The Impact of Family Circumstances on Work and Education: Evidence from Vietnam

Nga Thi Thuy Nguyen

University of Wollongong

Follow this and additional works at: https://ro.uow.edu.au/theses1

Recommended Citation
The Impact of Family Circumstances on Work and Education: Evidence from Vietnam

A thesis submitted in partial fulfilment of the requirements of award of the degree:

Doctor of Philosophy

University of Wollongong
School of Accounting, Economics and Finance

By

Nga Thi Thuy Nguyen

2017
Originality Statement

I, Nga Thi Thuy Nguyen, hereby declare that this thesis, submitted to the Faculty of Business, University of Wollongong, is my own work unless otherwise referenced or acknowledged. The work contained in this thesis has not been previously submitted for a degree or other qualification at UOW or any other educational institution. To the best of my knowledge, this thesis contains no materials previously published or written by another person, except where due reference is made. Any contribution made to the research by others is explicitly acknowledged in the thesis.

Signed.................................................................

Date........................................................................
Acknowledgements

I would like to express my greatest appreciation to my supervisors, Dr Oleg Yerokhin, Dr Silvia Mendolia and Professor Amnon Levy for their guidance, support, in-depth feedback and encouragement throughout the development of this thesis. Thank you for being so generous with your time, knowledge and enthusiasm.

Oleg, Silvia and Amnon continually provided motivation and inspiration and helped me to develop my own ideas and research expertise. I am very grateful to them for being my mentors throughout my thesis. I am very lucky to have had such dedicated and helpful supervisors.

I am also grateful to Dr Khorshed Chowdhury for supporting me in the first year of my degree and to my friends at UOW for their help and for the time we spent together talking about our research.

I would like to thank my extended family, especially my mother and mother-in-law, for their love, support and patience in seeing me through this process. My special thanks to my husband Duc Dinh for his love, encouragement, support and patience over all stages of this long journey. I am also grateful to my son Duc Anh and my little son Bao Chau for making my life whole.
Abstract

This thesis analyses the connection between family circumstances and work and educational outcomes of individuals in the family. It consists of a theoretical chapter developing a conceptual framework analysing various factors affecting female labour force participation and two empirical chapters analysing the relationships between family well-being and education and labour market outcomes.

The first chapter develops a broad conceptual framework encompassing both supply and demand aspects to explain women’s participation in the labour force. The supply side factors include fertility, partner’s income, potential childcare assistance from extended family members, the youngest child’s age and the female’s age and education. The demand side factors include the gender wage differential and the unemployment rate. From these different factors, this chapter derives a model of female labour supply which assumes that the female’s utility is a function of her family’s disposable income. This conceptual framework could lead to hypotheses on the participation of women in the labour market.

The second chapter is an empirical analysis testing one of the derived hypotheses from the conceptual framework on the determinants of female labour supply. This chapter focuses on analysing empirically the impact of fertility on female labour force participation in the less researched context of developing countries. Particularly, this thesis focuses on the effect of having an additional child on a female’s probability of labour force participation. This study is based on a sample of women aged 18-35 with at least two children from the Vietnam Population and Housing Census Survey 2009.

This study uses an instrumental variable methodology based on the widely observed phenomenon of parental son preference in Vietnamese society to address the endogeneity problem of fertility. Results suggest that the presence of an additional child among families with two or more children is likely to reduce the labour force participation probability of mothers and the effect of these children on female labour supply varies with the mother’s and father’s educational attainments.

The third chapter investigates the relationship between parental health shocks and children’s engagement in education and the labour market, using panel data from the Vietnam Household Living Standards Surveys between 2004 and 2008. While there is
substantial evidence showing the intergenerational transmission of health, the literature investigating the impact of parental health on children’s educational and labour market outcomes is limited, especially in developing countries. In this chapter, child fixed effects and control for a detailed set of household and local area characteristics are used. This study’s main findings show that maternal illness substantially decreases the chances of being enrolled in school for children between 10 and 23 years old and, at the same time, increases children’s likelihood of entering the labour market and working more hours for children aged 15-18 years old. The effect is particularly pronounced for girls who seem to experience the worst adverse consequences in terms of education and labour market engagement.
Table of Contents

Originality Statement .................................................................................................................. ii
Acknowledgements ........................................................................................................................ iii
Abstract ........................................................................................................................................ iv
List of Tables .................................................................................................................................. ix
List of Figures ................................................................................................................................. xi
List of Abbreviations ...................................................................................................................... xii

Chapter 1. Introduction ................................................................................................................ 13

Chapter 2. Literature review ....................................................................................................... 18
  2.1 Theoretical review ................................................................................................................... 18
  2.1.1 Fertility and female labour force participation ................................................................. 18
  2.1.2 Parental health and children’s education and work ......................................................... 29
  2.2 Empirical review .................................................................................................................... 31
    2.2.1 The effect of fertility and family size on female labour force participation ............... 31
        Vietnamese studies ................................................................................................................ 49
  2.2.2 The impact of parental health on children’s schooling and labour force participation 50

Chapter 3. Conceptual framework ............................................................................................... 62
  3.1 Introduction ............................................................................................................................ 62
  3.2 Conceptual framework .......................................................................................................... 63
  3.3 A model of female labour supply under employment uncertainty and discrimination ... 64
  3.4 Propositions .......................................................................................................................... 70
    3.4.1 Proposition 1: The effect of the number of children ...................................................... 70
    3.4.2 Proposition 2: The effect of grandparents .................................................................... 74
    3.4.3 Proposition 3: The effect of wage discrimination .......................................................... 75
    3.4.4 Proposition 4: The effect of partner’s income ............................................................... 77
    3.4.5 Proposition 5: The effect of the youngest child’s age ................................................... 79
    3.4.6 Proposition 6: The effect of the female’s age ................................................................. 81
    3.4.7 Proposition 7: The effect of unemployment rate ......................................................... 82
    3.4.8 Proposition 8: The effect of female educational attainment ....................................... 83
  3.5 Conclusion and testable hypotheses ...................................................................................... 85

Chapter 4. The impact of fertility on female labour force participation: evidence from Vietnam .......................................................................................................................... 89
  4.1 Introduction ............................................................................................................................ 89
  4.2 Data ......................................................................................................................................... 91
3. Completed empirical results ........................................................................................................ 166
4. 2SLS estimate results of using IVs *SameSex*, *Twogirl* and *Twoboy* .................................. 168
Appendix C ................................................................................................................................... 169
References .................................................................................................................................... 174
List of Tables

Table 2.1 - Fertility and female labour participation relationship literature ........................................... 37
Table 2.2 – Parental health and children’s education and work relationship literature .......................... 55
Table 4.1 - Labour force participation rates by sex, urban/rural areas, 2009 (%) .................................... 93
Table 4.2 - Fertility and female labour force participation rate ................................................................. 96
Table 4.3 - The sample of all women and married women with two or more children in Population and Housing Census 2009 ................................................................................. 97
Table 4.4 - Sample descriptive statistics - Independent variables (families with two or more children) ............................................................................................................................. 99
Table 4.5 - Proportion of families with another child by the sex of existing children ........................... 105
Table 4.6 - Differences in means (standard errors) of maternal demographic characteristics by Twogirl and Twin2 ........................................................................................................ 109
Table 4.7 - Sex ratio by birth order ............................................................................................................ 109
Table 4.8 - First stage regression - Impact of two girls and twin birth on probability of having more than two children (women aged 18-35 with two or more children) ........................................... 113
Table 4.9 - Effect of having more than two children on female labour force participation (women aged 18-35 with two or more children) ............................................................................ 118
Table 4.10 - Heterogeneity in the impact of children on female labour supply by interaction terms ................................................................................................................................. 122
Table 5.1 – Net enrolment rate by education level (%) ........................................................................... 136
Table 5.2 – Children labour supply ......................................................................................................... 138
Table 5.3 - Children’s school enrolment rates by parental health status ................................................. 141
Table 5.4 - Children’s labour force participation by parental health status ............................................. 142
Table 5.5 - Sample descriptive statistics - independent variables .......................................................... 142
Table 5.6 - Effect of parental health shocks on child school enrolment (children aged 10-23) 147
Table 5.7 - Effect of maternal and parental health shocks on child school enrolment (children aged 10-23) .................................................................................................................. 148
Table 5.8 - Effect of parental health shocks on child school enrolment by age group and gender of the child .............................................................................................................................. 149
Table 5.9 - Sensitivity analyses for the impact of parental health shocks on child’s education (Estimation with FE) .................................................................................................................... 149
Table 5.10 - Effect of parental health shocks on children’s labour force participation (children aged 10-23) .................................................................................................................. 151
Table 5.11 - Effect of maternal and parental health shocks on children’s labour force participation (children aged 10-23) .................................................................................................... 152
Table 5.12 - Effect of parental health shocks on children’s labour force participation, by age group and gender of the child ........................................................................................................... 152
Table B.1 - Effect of having more than two children on female labour force participation (2SLS estimation with all women sample aged 18-35 with two or more children) ........................................ 166
Table B.2 - Effect of having more than two children on female labour force participation (2SLS estimation with all women sample aged 18-35 with two or more children) ........................................ 167
Table B.3 - 2SLS estimates of FLFP(Women aged 18-35 with two or more children) .......................... 168
Table C.1 - Effect of parental health shocks on child school enrolment (children aged 10-23) 169
Table C.2 - Effect of maternal and paternal health shocks on child school enrolment (children aged 10-23) .................................................................................................................. 170
Table C.3 - Effect of parental health shocks on children’s labour force participation (children aged 10-23) ..................................................................................................................................... 171
Table C.4 - Effect of maternal and paternal health shocks on children’s labour force participation (children aged 10-23) ........................................................................................................ 172
List of Figures

Figure 2-1 An indifferent curve ........................................................................................................ 19
Figure 2-2 An individual’s Decision Not to Participate in Labour Market (Corner Solution) ........ 21
Figure 2-3 An individual’s Decision to Participate in Labour Market ......................................... 22
Figure 2-4 The effect of an increase in non-labour income on labour force participation ............. 23
Figure 3-1 Supply and demand considerations in female participation in out-of-home work in a developing and culturally changing economy: The case of Vietnam ........................................... 63
Figure 3-2 Marginal utility function of female labour force participation .................................... 78
Figure 4.1: Female labour force participation of selected countries (%) ...................................... 92
Figure 4.2: Total fertility rate in Vietnam, 1999-2009 .................................................................. 94
List of Abbreviations

GSO: General Statistics Office of Vietnam

ILO: International Labour Organization

OLS: Ordinary Least Squares

2SLS: Two Stage Least Squares

FLFP: Female Labour Force Participation

MPI: Ministry of Planning and Investment

VHLSS: Vietnam Household Living Standards Survey

VPHC: Vietnam Population and Housing Census

WHO: World Health Organization
Chapter 1. Introduction

It is well understood that current family circumstances have a significant effect on an individual’s choices and wellbeing. Modern economic thought recognises this fact by emphasising that many economic decisions are made by the household and not on the individual level, and that the wellbeing levels of individual members of a family are interrelated. Over the past few decades, a large body of literature has developed studying theoretically and empirically how decisions are made within families and how various shocks affecting some members of a family affect the economic outcomes of other members. This thesis contributes to this literature by studying how family decisions and circumstances affect two key economic outcomes, labour force participation and education level, in the context of Vietnam, a developing country which has undergone a very rapid economic transition in recent decades.

With respect to the first outcome, labour force participation, the thesis looks at the relationship between family fertility decisions and female labour force participation (FLFP). It does this by first developing a conceptual framework which discusses various factors that can in principle affect FLFP. Subsequently, it addresses this question empirically following on from the large and established literature on this topic in developed countries.

Empirical evidence shows that female labour force participation rates have increased dramatically, especially in the recent decades, accompanied by the increase in women’s education level and decline in fertility. While approximately 25% of married women participated in the labour market in the 1950s worldwide, this reached 60% by 2000 (Del Boca and Pasqua 2005). Female labour force participation has made important contributions to economic growth and economic development all over the world. However, FLFP has been constant or declining during the last three decades in some African and most Asian countries, despite a continuously declining trend in fertility and a remarkable increase in education levels (Priebe 2011).

This trend is also seen in Vietnam which has experienced a continuous decrease in fertility rates, especially in the last 10 years, while GDP per capita increased from 1161.5 to 1770.3 USD per capita during the period from 2007 to 2016 (Worldbank 2017) and female education has achieved remarkable progress in both quantity and
quality. While Vietnam has a relatively high FLFP rate by international standards, with about 72% of women participating in the labour force (ILO 2013), the FLFP rate has been decreasing during the last two decades and according to the ILO’s project on Vietnam labour market trends, this rate will continue to decline further (ILO 2008).

The determinants of female labour supply have changed dramatically over the course of economic transition. These include positive impacts of improved healthcare services, increased life expectancy, decreased fertility rates, increased educational attainment and diffusion of household appliances. Traditional family structures have started to change with nuclear families becoming more common and intergenerational co-residence in the context of extended families decreasing. On the other hand, various determinants of female labour supply arising from the labour market have also rapidly changed effecting FLFP negatively. The unemployment rate has increased and the gender pay gap has grown in the Vietnamese labour market during the transition.

The first part of the thesis is motivated by the importance of Vietnamese women’s contribution to economic development and the paucity of prior Vietnamese research on female labour force participation. First, a broad conceptual framework will be constructed analysing various factors affecting FLFP from both supply and demand sides (chapter 3). The supply side includes factors such as fertility, partner’s income, potential childcare assistance from extended family members, the youngest child’s age and the female’s age and education. The demand side includes factors such as the gender wage differential and the unemployment rate. From all of these different factors, this chapter introduces a model of female labour supply and uses it to study the comparative statistics of the FLFP problem. This conceptual framework will be used to study empirically the effect of fertility, which is one of the determinants of female labour supply.

Guided by the conceptual framework in chapter 3, this thesis uses data from the Vietnam Population and Housing Census (VPHC) 2009, which is a census sample survey of fifteen percent of the population covering 3,692,042 households, to study the effect of fertility on FLFP in Vietnam (chapter 4). The instrumental variable (IV) methodology is applied in order to address the endogeneity of fertility in order to understand more clearly the causal mechanisms linking fertility and female labour supply. Previous studies have used a variety of identification strategies to address the
endogeneity problem of fertility and most studies suggest that fertility decline is a key factor in the explanation of women’s labour supply decisions in developed countries. In the context of developing countries, however, this question has received less attention. By studying the relationship between fertility and female labour supply in Vietnam, this study contributes to the existing literature on the determinants of FLFP in developing countries.

Motivated by influential earlier literature (Angrist and Evans 1998), this study addresses the endogeneity problem by exploiting the exogenous variation in family size based on the sex composition of the first two children. In particular, this study hypothesises that the gender composition of the first two children will have an effect on the probability of having a third child due to the prevalence of gender bias in parental preferences, namely the preference for sons which exists in Vietnam. It is known that Vietnamese families prefer a male dominated composition of children to a mixed sex composition. This means that if parents have two girls consecutively, they are more likely to have a third child than those having two boys or those having one girl and one boy. As a robustness check, multiple births (twins at second birth) is also used as an instrument for moving from the second child to the third child.

This study demonstrates that having an additional child has a negative impact on a female’s probability of participation in the labour market. Using the first two children being girls twogirl as an IV, the Two Stage Least Square (2SLS) estimates show that the presence of the third child decreases the labour force participation probability of women by about 3 percentage points. The study also found that the effect of fertility on female labour supply may vary with the different educational attainments of both parents. The 2SLS estimates suggest that the impact of children on labour supply increases with the educational attainment of mothers and the more highly educated women experience the largest impacts of childbearing on labour supply. These results are consistent with theoretical predictions about household time allocation developed in the conceptual framework.

The second economic outcome related to the family circumstances that is investigated in this thesis is the education and labour force participation of children. This part of the thesis analyses the link between parental health shocks and children’s educational and labour market outcomes (chapter 5). Specifically, the study focuses on the effect of
parental illness on children’s school enrolment and engagement in the labour market in Vietnam. In the context of recent literature, the analysis of connections between parental health shocks and children’s education and child labour is particularly meaningful in the context of a developing country and Vietnam is an interesting case to look at, given its poor system of social protection and the limited access to the social health insurance system and health expenditure heavily depending on direct out-of-pocket payments. Further, the study period (2004-2008) corresponds to a critical phase of changes in the Vietnamese health insurance system.

Health shocks have a high degree of unpredictability and may have severe impacts on families’ well-being, especially for rural and low income households. On the one hand, parental illness may decrease their working hours as well as their productivity and may cause a decrease in household income. This loss of income along with the increased health expenditures may affect their ability to finance a child’s education and make children more likely to drop out of school. On the other hand, the parental illness may reduce family labour supply, so children may be required to leave school and start working to substitute for their parents’ labour. Therefore, it is particularly interesting to analyse the effect of parental health shocks on both children’s education and child labour because these two elements are closely related.

This chapter addresses this question using panel data from three waves of the Vietnam Household Living Standards Survey (VHLSS) from 2004-2008, which is conducted every two years by the General Statistics Office of Vietnam (GSO). This is an ongoing longitudinal survey of the Vietnamese population and conducted over a nationally representative sample. A model with individual child fixed effects is used to estimate the impact of parental illness on child schooling, which allows the study to take into account unobserved time-invariant characteristics that may affect both parental health and child education. With few exceptions, most studies in the prior literature look at the relationship between parental death and child education and are based on cross-sectional data or longitudinal data and try to limit the unobservable heterogeneity by using a rich set of simultaneous or lagged observable child and parent characteristics or by including matching techniques. This research contributes to the existing literature on parental health shocks and children’s education by providing new evidence on the relationship between parental health and child schooling in a developing Asian country, using a
credible identification strategy. Findings from Vietnam can be easily generalised to other Asian developing countries with similar education and health system. The results of this study can also provide important policy implications for the on-going reform process of the social health insurance system in Vietnam.

This study documents strong evidence suggesting that maternal health shocks can cause considerable and significant changes in a child’s likelihood to be enrolled in education. Children aged 10-23 with unhealthy mothers are 3 percentage points less likely to be enrolled in education than children with healthy mothers. The results also show that parental illness significantly increases children’s probability of entering employment as well as the number of hours of work. The negative effects of parental health shocks are particularly pronounced for girls and for children aged 15-18.

Taken together, the results of the three main chapters of this thesis (the conceptual framework, and an empirical investigation of the relationship between fertility and FLFP, and the chapter on the effects of parental health shocks on children’s educational enrolment and labour) contribute to the limited literature on the effect of family circumstances on various outcomes in Vietnam. They can be used as a stepping stone in subsequent literature on these topics and also provide empirical evidence which can be used by policymakers to implement social security programs and to design and implement other economic policies in the context of the Vietnamese economy.
Chapter 2. Literature review

2.1 Theoretical background

2.1.1 Fertility and female labour force participation

This chapter reviews most important theoretical models to explain the labour supply decision of females.

Female labour force participation has been first studied in seminal contributions by Mincer (1962), Becker (1965), Cain and Dooley (1976). Following these authors, this issue has raised the interest of many scholars. They further analysed women’s behaviours in the labour market by applying different econometric models to contribute to a great body of literature in this area. Mincer (1962) attempts to analyse one period model of the labour supply in which this period is a life-time and indicated that one of the most important factors affecting females’ lifetime labour supply curve is the number of children. Becker (1965) generalises the role of time in economic activities, and time is a central factor affecting decisions of fertility, health, and others. This theory laid a foundation of household production model that later became the standard approach to modelling labour supply decisions. According to economic theory, labour force participation decisions should be analysed in the context of the family (Mincer 1962; Cain 1966). This implies that individuals will compare benefits and costs of participating in market work with benefits and costs of activities at home.

2.1.1.1 The neoclassical theory of labour supply

The theory is typically used by economists to study labour supply decisions. The framework is an extension of consumer theory in which a consumer makes a choice between goods and leisure consumption. The model analyses how individuals allocate their limited amount of time to the labour market and leisure activities. Individuals have to decide whether they work and if so, how many hours for working and how many hours for other activities.

In a simple model of labour supply, an individual has two choices: spending his/her time on working in the labour market and consuming leisure. Assuming that each individual has possibility to trade-off between the consumption of goods and leisure,
individuals maximise their utility by consuming goods and services (C) in the market and consuming time for leisure (L).

The individual’s utility function is \( U = U(C, L) \)

The utility of an individual is represented by an indifferent curve that reflects the trade-off between the consumption of goods (C) and the consumption of leisure (L). Indifferent curve that is farther out from the origin reflects a higher level of utility. Consumers prefer indifferent curve that is farther out from the origin (O). Thus, a higher level of utility means that more C, more L, and more satisfaction.

**Figure 2-1 An indifferent curve**

![Indifferent Curve](image)

Individually maximise their utility, they are bound by two constraints: time is the first constraint that they must allocate to working for pay (h) and leisure (L) under the available time of the day. Besides that, they also face with income constraint that they need to buy goods and services in the market. The income derives from their labour wages (w) and nonlabour activities (R).

Individually maximise the utility \( U = U(C, L) \) subject to time and income constraints:

\[
T = h + L
\]

\[
C = wh + R
\]

\( U \) is a utility function that measures the individual’s well-being; \( T \) is total time available; \( h \) is working time for pay; \( L \) is leisure; \( w \) is hourly wage; \( R \) is non-labour income.
Setting up the Lagrangian equation to represent the individual’s utility maximization problem yields

\[ \ell = U(C, L) + \lambda [w(T - L) + R - C] \]

The first order condition for maximum are

\[ \frac{\partial \ell}{\partial C} = \frac{\partial U}{\partial C} - \lambda = 0 \]
\[ \frac{\partial \ell}{\partial L} = \frac{\partial U}{\partial L} - \lambda w = 0 \]

\[ \Rightarrow \text{Equilibrium condition is } \frac{MU_L}{MU_C} = w. \] This can be interpreted that a given wage rate, in order to maximize her utility, the individual will choose the number of working hours at which the marginal rate of substitution between leisure and a composite consumption good is equal to the market wage rate (Nicholson 1992).

An interior solution of the model exists will explain how many work hours to be devoted. With an increase in wage rate, the substitution effect will increase the opportunity cost of leisure because leisure becomes more expensive. The more hours of leisure are consumed, the more forgone income will be given up. On the other hand, assuming that leisure is a normal good, the income effect will induce individuals to consume more on leisure because the increase in wages makes people to become better off. According to this approach, an increase in the wage rate may increase or decrease the number of work hours depending on whether substitution effect or income effect dominates. If the substitution effect dominates, the number of work hours supplied will increase. In contrast, the result will be a decrease in the number of hours for working.

If there is an increase in non-labour income under the assumption of leisure as a normal good, leisure time rises and the hours of working falls.

A corner solution of the model exists as the individual decides not to enter the labour force. It is explained by reservation wage rate (w^*). The reservation wage rate is the wage level that would make a worker indifferent between not working and working that first hour (Borjas 2000). The marginal rate of substitution at the zero work point is called the reservation wage (Thornton et al. 2013). The reservation wage depends on the form of the utility function at the zero work point, non-labour income level, and factors influencing their home time value (number of children, marital status).
Borjas (2000) uses the utility function and income line to explain the decision of an individual to participate or not in the labour market. Figure 2.2 describes the labour supply decision of an individual through the utility maximization problem.

**Figure 2-2 An individual’s Decision Not to Participate in Labour Market (Corner Solution)**

Time is described on the x axis and the value of goods and services purchased is on y axis. The number of hours spent at work is calculated from right to left and hours spent on leisure is calculated from left to right. A set of indifferent curves $U_1, U_2, U_3$ indicates the different levels of utility. Consumers prefer the indifferent curve that is farther out from the origin. These indifferent curves have negative slopes that are the marginal rate of substitution between leisure and consumption ($\text{MU}_L/\text{MU}_X$). TAB line represents the individual’s budget constraint with non-labour income of TA and the market wage of AB with a negative slope. The slope of indifferent curve at point A is the reservation wage ($w^*$). As can be seen from Figure 2.2, the individual will not participate in the labour market because the market wage is lower than the reservation wage ($w < w^*$).

If the market wage is higher than the reservation wage ($w > w^*$), the individual will enter the labour force as shown in Figure 2.3. In this case, she will obtain a higher
utility level ($U_3$), and she will maximise her utility at point E where the highest indifferent curve is tangent with the market wage ($w$) line.

It can be seen from this analysis that when market wage or the value of market time increases, individuals are more likely to participate in the labour force because the opportunity cost of time out of labour force is more expensive. However, factors that increase the value of non-market time or reservation wage such as children, marital status have negative impacts on labour participation probability.

**Figure 2-3 An individual’s Decision to Participate in Labour Market**

Now consider the effect of higher non-labour income ($V$) on the female’s labour participation decision. If leisure is a normal good, the demand for leisure increases as income increases. As a result, she reduces the number of work hours or even out of the labour force. As can be seen from Figure 2.4, the non-labour income increases from TA to TC, the market wage rate is unchanged (CD line is parallel with AB line). With the higher non-labour income, the female can maximise her utility at point F that lies on the CD line and right of point E, so she can reach a higher utility level. In this case, she will work less and consume more time for leisure compared to the case in Figure 2.3. She may even be out of labour market by choosing point C.
2.1.1.2 The theory of allocation of time

The simplest model of labour supply studies an individual maximises his/her utilities through the rational allocation of time between working for pay and leisure. According to this approach, the wage rate is an important factor to determine the labour supply of an individual, subject to time and income constraints (Killingsworth 1983). However, this model is restricted by an analysis of individual framework and ignores the family context, non-market time, and unpaid non-market work. In a household, labour supply of in individual can be affected by other family members. For example, the labour supply of a married woman can be affected by her husband’s income, so it is difficult to analyse women’s labour supply because of the productive uses of non-market time. Researchers recognise that it is necessary to extend the basic labour supply model to capture the complication of women’s labour supply behaviours. The utility maximization problem of a single individual in this labour supply model has been extended to study in the context of family. This new approach has been used by Mincer (1962), Ashenfelter and Heckman (1974), Smith (1977), Killingsworth (1983), and
some other scholars. This approach assumes that the time allocation decision of an individual needs to be put in family context. It allows the interdependence about labour supply decisions of family members (Smith 1977). This means that family is considered as a decision making unit of analysis and an individual’s decision of labour supply should be analysed and takes into account other members of family. According to this new approach, when analysing female labour supply, especially married women, besides the women’s own wage rate, the income of other family members such as husband’s income also affects the women’s labour supply decisions.

A further extension of the model for the analysis of labour supply is proposed by Becker (1965). He introduces a theory of the allocation of time in which the allocation of time between work and leisure is extended to the trichotomy of time: work in the market, work at home (cooking, childcare, and so on), and leisure. Becker (1965) and Mincer (1962) recognise women’s non market commodities production which was ignored by previous scholars in the basic labour supply model. Becker tries to incorporate the concept of non-market time or non-market work into the labour supply theory. He suggests that time can be efficiently allocated between market work and household work and any changes of one family member to be more effective in the market will lead to the reallocation of time of other family members. According to the model of household production of Becker (1965), time not spending at work is considered as an input of household commodities production. Households are both producers and consumers of goods: “Households will be assumed to combine time and market goods to produce more basic commodities that directly enter their utility functions” (Becker 1965, p. 495). They combine market goods and time as inputs of a production process to produce household commodities such as cooked meals, health, entertainment, and childrearing. Then these commodities are consumed and yield utility for family members. In order to maximise their utility, families try to combine effectively inputs of production process, so family members need to incorporate in the effective allocation of time, income, and goods consumption.

The family’s utility function now is

\[ U = U(Z_1, \ldots, Z_n) \]  

where \( Z_i = f(x_i, t_i) \), household production function

Subject to income constraint: \( \sum p_i x_i = Y = wt + R \)  \hspace{1cm} (1)
Each household maximises its utility by consuming these commodities \(Z_i\) subject to time and income budget constraints.

\[
\sum t_i = t_c = t - t_w
\]  

(2)

where \(w\): hourly wage

\(t_w\): time spent at working for pay

\(t_c\): total time spent at consumption

\(t\): total time available

\(R\): non-labour income

\(p_i\): price of market goods \(x_i\)

From (1) & (2): \[
\sum p_i x_i + w \sum t_i = wt + R
\]  

(3)

The right hand side of (3) is “full income” = non labour income + total earnings if all time of the household spent on working. The left hand side is total cost on market goods \(\sum p_i x_i\) and time \(w \sum t_i\) to produce commodity \(Z_i\). Full income is divided into 2 parts: spending directly on goods and indirectly through foregoing income (time spent at home rather than at work) (Becker 1965).

Setting up the Lagrangian equation to represent the individual’s utility maximization problem yields

\[
\ell = U(x_i, t_i) + \lambda (wt + R - \sum p_i x_i - w \sum t_i)
\]

The first order condition for maximum are

\[
\frac{\partial \ell}{\partial x_i} = \frac{\partial U}{\partial x_i} - \lambda p = 0
\]

\[
\frac{\partial \ell}{\partial t_i} = \frac{\partial U}{\partial t_i} - \lambda w = 0
\]
The equilibrium condition: \[
\frac{MU_{t_i}}{MU_{x_i}} = \frac{w}{p}
\]

This expression can be interpreted that the marginal rate of substitution between time and market goods equals real wage rate. The necessary conditions for maximum propose equality between the marginal rate of substitution between goods and the ratio of their full prices, where the full prices include the forgone incomes associated with time spent on the consumption of the respective goods.

Consider an increase in wage rate, according to this approach, an increase in the wage rate changes the relative price of different commodities (the price here is full price that includes the price of market goods and of time used to produce a unit of commodity). The prices of time-intensive commodities will rise more quickly in comparison with those of goods-intensive commodities. As a result, households will adjust their consumption behaviour in the way of substituting from time-intensive commodities towards goods-intensive commodities. Thus, total time for consumption reduces and the time spent working increases. In other word, a wage rise causes a reallocation of time toward work for pay.

Since the ratio of time and market goods to produce household commodities can be varied in the production process of these commodities, a rise in wage increases the opportunity cost of time. An increase in the opportunity cost of time induces households reduce the ratio of time to market goods in the production process of each commodity to maximize the use of their time.

If non-labour income increases and no change in the relative price of commodities, the consumption of most commodities will increase or in other word, the total time spent on consumption will increase, so time spent at work will decrease (Becker 1965, p. 501).

The time allocation model permits more detailed analyses of non-market work of family members. It also provides a useful framework for analysing the variety of different factors of labour supply. This model is capable of explaining both participation decision and the decision on how much time to devote to work conditional on participating in the labour market. It provides explanations of women’s decisions to take part in the labour market or not by comparing the value of time in the labour market to the value of time
spent at home and if she decides to enter the labour force, how she allocates her limited amount of time to market work and other activities such as child care, and leisure. When market wage or the value of market time increases, individuals are more likely to participate in the labour force because the opportunity cost of time out of the labour force is more expensive. However, factors that increase the value of non-market time such as children, marital status have negative impacts on the probability of female labour participation.

2.1.1.3 The theory of human capital investment

Another theoretical approach to support the study of female labour force participation is human capital investment theory. This theory emphasizes the relationship between the return of investment in education and training and the time individuals expect to work during their lives. In other words, acquired knowledge and skills through education and training are related to labour market participation outcomes. Labour supply decisions require a substantial investment. An individual who is planning to enter in the labour market as a full time worker is more likely to invest more in education and training and vice versa (Altonji and Blank, 1999; Blau et al., 2006). As in many other investment decisions, rational investment in education and training should also be based on a comparison of the costs to the expected returns. The basic model of human capital investment explains how individuals maximize their utility in making the decision about education and training. In order to maximize the utility, individuals will compare the present value of expected future benefits with the current costs. If the present value of future benefits is greater than costs, this investment decision is attractive. The present value of future benefits is calculated as follows:

$$\text{Present value} = \frac{B_1}{(1+r)} + \frac{B_2}{(1+r)^2} + .... + \frac{B_i}{(1+r)^i} + .... + \frac{B_T}{(1+r)^T}$$

where $B_i$ is the benefit year $i$; the period of time is from 1 to $T$; $r$ is discount rate.

The human capital theory also provides an explanation of why women may choose to invest less in the formal education and spend a smaller amount of time during their lives in the labour market than men. For example, many studies stressed the traditional role of women in the household in which childbearing is one of the most important factors
because they are aware that childbearing and other family responsibilities may force them to drop out of market work during childbearing years (Mincer and Ofek 1982; Becker 1985). According to the human capital theory, if a woman has a plan to interrupt her market work life at a point of time in her lifetime, she is likely to reduce her amount of investment in education and training. In addition, the expectation of work discontinuity also may reduce her incentive for investment in the types of jobs that require the high level of commitment or demanding. The reason is that benefits from her investment in additional education and training will be lower because the time period spent in market work is expected to be shorter and the investment depreciates rapidly during the time out of the labour market.

Becker (1985) used the model of effort allocation to explain why women have a lower investment in market human capital. He stated that housework or childcare is larger effort intensive activities, so women tend to spend less effort on each hour of market work. As a result, this would reduce their hourly earnings. The lower benefit derived from human capital investment in the market work discourages their investment in human capital even in the case men and women work the same number of hour. This fall of human capital investment in the market work reduces their incentive to participate in the labour force and their lower participation in the labour market in turn lower their investment in the market human capital.

Thus, women’s responsibilities of childcare and other household work affect adversely their hourly earnings that are below men’s earnings. These family responsibilities also explain for occupational segregation between men and women. This is the reason why women prefer to choose less demanding jobs than men. Women also tend to avoid fields of study or occupations that make their skills easily depreciated during the time of market work interruption.

On-the-job training is another kind of human capital investment. It includes all form of training, both formal training and informal training. The human capital theory shows that women who adhere to their traditional gender role in the family are less likely to pursue higher education as well as acquire opportunities to engage in on-the-job training because they have weaker attachment to the labour market (Becker 1985). Accordingly,

---

their earnings are lower than men. The human capital explains why women’s earnings are less than men and they have less attachment to the labour market. According to Orazem et al. (2003) and Frazis et al. (1998), women often acquire less levels of on-the-job training. This emanates from both employer and employee sides. Since family responsibilities, women who are expected to interrupt their market work will have fewer opportunities to be offered on-the-job training. Women will be less likely to receive job training if their employers expect that women have shorter work lives than men, so they prefer invest in job training for male workers to female workers. Furthermore, employers do not want to employ women for jobs requiring much level of on-the-job training. On the other hand, if women themselves expect discontinuities of experience in the future because of childbearing or other household responsibilities, so having shorter work lives, they will choose jobs that require less level of training (Waldfogel 1998b; Lundberg and Rose 2000; Spivey 2005). Thus, childbearing responsibilities of women discourage such training investments.

The human capital theory also predicts recent changes in women’s labour force participation, especially married women in their childbearing years. This increasing trend of women’s participation in the labour market is leading to dramatic changes in the acquisition of human capital, including both education and training by women. The theory predicts that women’s incentives to acquire training will change. The formal education received by women also shows dramatic changes in recent years that reflect a rise in returns to human capital investments as a result of the longer expected work lives of women (Ehrenberg and Smith 2008, p. 300).

2.1.2 Parental health and children’s education and work

The study analyses the impacts of parental health on children education, so it relates to a large body of literature on the intergenerational transmission of socioeconomic status across generations, such as education, income, and earnings (Solon 1999; Black and Devereux 2011; Cobb and Deborah 2012). Low socioeconomic parents are more likely to experience a higher probability of negative health shocks than their counterparts who are more advantaged conditions. Both economic and other social science research have emphasized the importance of parental or family circumstances and choices on

---

children’s outcomes such as educational attainment, future income or earnings. Economists have begun to consider the process of children’s attainment to be an aspect of the theory of family behaviour (Becker and Tomes 1979, 1986; Browning 1992). Family is viewed as a production unit in which inputs are used to generate utility for family members. Parents (or adults) in the family make decisions to generate family economic resources and also determine the use of these resources (consumption, saving, investment in children). The amount of family resources and the timing of distribution of these resources depend on socioeconomic statement of the family and influence the attainments of children. Children are also affected by a variety of choices made by parents such as number of siblings, residence location, and family structure changes. According to the framework of Becker and Tomes (1986), children begin life with a natural transmission of endowment from their parents regardless of any decisions of parents. On average, if parents with high level of educational attainment will produce children who can also attain high level of education. Apart from endowment, parents can care about their children’s attainments and influence children’s human capital by making decisions on expenditures on health, learning, and skills. These expenditures depend on parental socioeconomic statements, and preferences. As noted in Solon (2004), children of wealthy parents can earn higher incomes in part because their parents invest more in human capital and have more education.

The work of Becker and Tomes provides theoretical background and testable hypotheses on the effects of family-based factors on investment in children but it has little guidance for empirical work (Goldberger 1989). Basing on this general framework, Leibowitz (1974) introduces an economic model of the determinants of children’s attainment and includes additional implications for empirical research. In this model, on the one hand, children inherit genetic endowments from their parents via heredity. On the other hand, the ability and education of parents determine family income as well as the quantity and quality of time and goods inputs that parents invest in their children. Children’s ability and family income and home investments in time and goods inputs (both quantity and quality), in turn, determine school attainment of children. And then schooling level and post school investment will affect children’s earnings and income in later life. Thus, apart from natural endowments, family background/circumstance has significant effects on children’s outcomes.
2.2 Empirical review

In this section, a review of empirical studies will be presented analysing the determinants of female labour force participation, (and, in particular, the effect of fertility on FLFP) and the impact of parental health shocks on children’s education and labour force participation.

2.2.1 The effect of fertility and family size on female labour force participation

Most developed countries have experienced dramatic increases in FLFP since 1960. For example, a tremendous increasing trend in FLFP can be seen in the United States. In 1900, only 20% of all women older than 15 took part in paid work in which approximately 6% of married women worked for pay. This number increased to 60% of all women and 62% of married women by the end of this century (Costa 2000). Similar trends have also been observed in other developed countries such as Canada, Germany, Sweden and France (Ehrenberg and Smith 2008). Ehrenberg and Smith (2008) used data from the Organization for Economic Co-operation and Development, Labour Force Statistics and found that in these countries the participation of females in the labour market, on average, increased to approximately two-thirds or more between 1965 and 2000.

Given the great social and political significance of the increase in FLFP, especially among married women, this issue has stimulated scholarly interest in women’s labour force participation behaviour. Researchers began to search for factors which would explain this behaviour. The studies of Mincer (1962) and Cain (1966) in the US have laid the theoretical and empirical foundation for many studies of FLFP in different countries.

A large body of literature has examined the effects of various factors on FLFP in the US as well as other developed countries. Some studies suggest fertility decline as a key factor in the explanation of women’s labour supply decisions. Other studies emphasize the role of female education (Michael 1985; Smith and Ward 1985; Goldin 1994; Lam and Duryea 1999; Bratti 2003; Maria and Elyce 2003) or the availability of part time jobs (Gregory et al. 1985; Blank 1989; Meulders et al. 1994; Falzone 2000; Del Boca 2002), while Greenwood et al. (2005) argue that the diffusion of home technological progress plays an important role as “engines of liberation” for women from household
chores and allows them to devote more time to market work. Many studies focus on the
direct effects of wage rate and husband’s income (Mincer 1962; Cain 1966; Ben-Porath
1973; Ashenfelter and Heckman 1974; Heckman 1974; Gronau 1977; Smith 1977),
while some others emphasize the role of childcare services and costs (Blau and Robins
1988; Ermisch 1989; Connelly 1992; Gustafsson 1995; Kimmel 1998; Chiuri 2000;

In general, the effects of factors on FLFP vary across countries in accordance with
traditional, social and economic characteristics and there is no consistent international
pattern in female labour force behaviour. While developed countries have experienced
an enormous rise in the participation of females in the labour force from the 20th
century, in developing countries, despite a continuously declining trend in fertility and a
remarkable increase in education levels, FLFP has been constant and even declined in
some African and most Asian countries over the last three decades (Priebe 2011). This
can be seen in several empirical studies as outlined in the next paragraphs.

There are few empirical studies in economics about female labour force participation in
developing countries. In Latin American countries, women’s participation rates in the
labour market overall remain at a low level, with an average of only one third of
working age females involved in the labour market. However, an increasing trend of
women’s incorporation into the Latin American labour market has been recognised over
recent decades. It increased from 21% in the early 1960s to 38% in the 1980s
(Psacharopoulos and Winter 1992). Scholars stated that the reason for this trend could
have been due to the economic crisis of the 1980s, but the increases have continued to
remain after the crisis.

In a study of factors affecting women’s behaviours of participation in 14 Latin
American and Caribbean countries, Psacharopoulos and Winter (1992) found that
education is one of the most important factors affecting women’s decisions to work
along with women’s earnings in this area. For example, in Argentina, the probability of
participation is only 28% among women with less than primary education, but the
probability increases to 58% if they have a university degree.

In addition, family characteristics such as marital status, young children and the head of
the household also have strong effects on women’s labour force participation decisions.
In most of the Latin American countries, married women are less likely to enter the labour market than single women. This seems to be a common characteristic in developing countries where antidiscrimination laws have not been strongly enforced. The study points out that discrimination in the labour market plays an important role affecting women’s decisions to participate or remain in market work. Econometric techniques are used to investigate how much the wage gap can be attributed to gender differences and how much can be attributed to human capital endowments. They found that human capital differentials only explain about 20% of differences in male-female wages. They also argued that the proportion of the wage gap attributable to discrimination in Latin America is significantly higher than that in developed countries (Psacharopoulos and Winter 1992).

Unlike developed and Latin American countries, Priebe (2011) shows that FLFP in Indonesia has been constant over the last two decades. The author focuses on the impact of fertility and child care and child raising costs on women’s participation in the labour market to explain this observed trend. In contrast to findings for developed countries, children play a dual role in women’s decisions in relation to labour supply. Children can not only decrease the probability of labour force participation by women but also conversely create an incentive to engage in the labour market, especially in developing countries where the women’s share in the labour force is essential to finance basic needs. This could explain why the FLFP rate has stagnated in Indonesia over the last decades. The study also indicates the very different effects of child care costs on female labour supply between developed and developing countries in the sense that childcare costs in Indonesia appear to have much less impact in comparison to developed countries.

In general, in economics literature, many theoretical and empirical studies have investigated FLFP as being determined by various factors such as age, marital status, number and age of children, education, work experience, income and non-labour income. This thesis however mainly focuses on empirical studies on the effect of fertility on female labour force participation.

As mentioned above, the investigation of the determinants of FLFP in developing countries is particularly interesting because they differ significantly from developed countries. Although there is substantial evidence of the effect of fertility on FLFP in the
existing economics literature, the existing studies on the subject provide contradictory results and the relationship between fertility and FLFP in developing countries is still ambiguous (see for example Priebe (2011), among others).

The relationship between fertility and FLFP in developed countries has attracted many scholars. A major factor that reduces the FLFP rate is that females tend to concentrate more on providing services for their family after they get married. Studies in economic literature focus on the effects of fertility and demonstrate fertility decline over time as a key factor in the explanation of women’s labour supply decisions. However, decisions about childbearing and allocation of time for participating in out-of-home work and housework are jointly determined by the maximisation of household utility. Thus, the endogeneity of fertility is the biggest challenge that researchers have to face\(^3\). It is likely due to omitted factors such as heterogeneity in parents’ preference that influence both decisions of childbearing and participation in the labour market. It may also be due to the impact of exogenous variables on fertility such as mother’s age, wage and schooling in the labour force participation equation. The presence of endogeneity of fertility in the labour force participation equation can cause the Ordinary Least Square (OLS) results to be biased or inconsistent.

Studies on this topic can be categorised by how they deal with the endogeneity of fertility. Some early studies treated fertility as an exogenous variable in the labour supply estimation (Mincer 1962; Cain 1966; Gronau 1973; Heckman 1974; Heckman and Willis 1977). Other studies recognised fertility as an endogenous variable and attempted to find methods to solve the endogeneity problem such as simultaneous equation estimates and instrumental variable methods.

Simultaneous equations to estimate the effect of fertility on FLFP have been used in some studies (Cain and Dooley 1976; Schultz 1978; Fleisher and Rhodes 1979; Moffitt 1984; Hotz and Miller 1988)\(^4\). However, one of the problems with simultaneous equation models is over-identification. Implementing this method is complicated

---

\(^3\) Browning (1992) notes in his literature review that the main problem researchers face is that fertility decisions may be endogenous and influenced by female labour supply. Thus, we cannot interpret the negative relationship found between different measures of fertility and female labour supply as causal effects.

\(^4\) Although most of these papers find parameters are considerably different when fertility is assumed as an endogenous variable, Mroz (1987) shows that basically the extensive specification tests cannot reject the exogeneity of fertility.
because it is difficult to find plausible identifying restrictions. Since there is more input information for solving the system than needed, certain restrictions are imposed on coefficients of the reduced-form equation and test hypothesis. Unbiased and consistent estimates can only be obtained from just-identified systems but inconsistent estimators come from over-identified systems.

Ideally, labour supply responses and earnings should be observed through an exogenous variation in the number of children within a family. Therefore, the IV method has been used to deal with the endogeneity of fertility. In this approach, IVs can generate exogenous variations in fertility and these exogenous fertility variations have been used to estimate its effect on FLFP. In other word, IVs are used to isolate the part of fertility that is uncorrelated with the error term in the labour participation equation. However, the choice of valid IVs is very complicated and not always easy to obtain. A valid instrument must satisfy two conditions: the relevance condition and the exogeneity condition. So first, it must be correlated with the endogenous regressor; second, it must be uncorrelated with the error term (Stock and Watson 2012).

Many IVs have been applied to generate the exogenous variation of fertility. Some studies use a twin at first birth (Rosenzweig and Wolpin 1980b; Bronars and Grogger 1994; Gangadharan et al. 1996; Jacobsen et al. 1999). Hotz et al. (2005) uses miscarriage in the first pregnancy of a woman, while Agüero and Marks (2008) and Cristia (2008) use events of infertility. Some other studies use the sex of the first child as an IV (Chun and Oh 2002; Lundberg and Rose 2002). Angrist and Evans (1998) and Cruces and Galiani (2007) use the sex of the first two children. These IVs can satisfy the two requirements of a valid instrument which are known as instrument relevance and instrument exogeneity. First, these instruments are randomly assigned so they are exogenous and not correlated with any unobserved variables in the participation equation. Second, they are highly correlated with fertility. For example, Angrist and Evans (1998) use an instrument which explores the widely observed phenomenon of parental preferences for a mixed sibling sex composition. Parents who have same sex siblings are more likely to have the third child. Since gender mix is randomly assigned, so whether the sex of the first two children matched will be a plausible instrument for having an additional child among families with at least two children (Angrist and Evans 1998).
Empirical studies on the effect of fertility on female labour supply in the existing literature have usually focused on the US and other European countries; only a few studies have focused on developing countries. Most of these studies in the literature found a significant negative effect of having more children on FLFP (see Table 2.1).
<table>
<thead>
<tr>
<th>Study</th>
<th>Country (Data)</th>
<th>Relevant Research Topics</th>
<th>Method</th>
<th>Main Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cross Section Studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosenzweig and Wolpin, 1980b</td>
<td>US (The Office of Population Research (1965) and Department of Health, Education, and Welfare (1973))</td>
<td>Estimate and explain effect of fertility on FLFP</td>
<td>Pooled sample, OLS, ML logit DV: FLFP IV: twin-first births</td>
<td>-10.2% to -37.1% probability of FLFP</td>
</tr>
<tr>
<td>Bronars and Grogger, 1994</td>
<td>US (1970 and 1980 census micro data)</td>
<td>Estimate and explain effect of fertility shocks on FLFP, earned income, poverty status, welfare for unwed mothers</td>
<td>OLS, IV estimates IV: twin-first births</td>
<td>-2.6% overall effect of unplanned child on FLFP for unwed mothers -11.4% to -12.3% effect of unplanned child aged 0-3 on FLFP for unwed mothers</td>
</tr>
<tr>
<td>Jacobsen et al. 1999</td>
<td>US (1970 and 1980 census micro data)</td>
<td>Estimate and explain effect of unplanned births on married women’s labour supply and earnings</td>
<td>OLS, probit DV: female labour supply IV: twin-first births</td>
<td>-1.4% to -1.6% overall effect of unplanned child on FLFP of married women -11.5% to -15.7% effect of unplanned child aged 0-3 on FLFP of married women</td>
</tr>
<tr>
<td>Angrist and Evan (1998)</td>
<td>US (Census public use micro samples 1980, 1990)</td>
<td>Estimate and explain effect of a third child on parents’ labour supply</td>
<td>Wald estimates, OLS, 2SLS DV: parents’ labour supply Key variable: more than two children Controls: age, age at first birth, boy 1st, boy 2nd, Black, Hispanic, other race IV: the first two children’s sex</td>
<td>OLS: -16.7% to -17.6 % for married-all women sample 1980 -14.7% to -15.5% for married-all 1990 2SLS: -12% for all-married 1980 -9.2% for all-married 1990</td>
</tr>
<tr>
<td>Author</td>
<td>Study Details</td>
<td>Methodology</td>
<td>Key Variable</td>
<td>Controls</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------------------------------------</td>
<td>-------------------------------</td>
<td>--------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Iacovou (2001)</td>
<td>British National Child Development Study and British Household Panel Study 1992</td>
<td>Estimate and explain effect of a third child on FLFP</td>
<td>OLS, 2SLS DV: FLFP Key variable: third child Controls: age, age², education, non-labour income IV: the first two children’s sex; multiple second birth</td>
<td>-6% to -15% in FLFP under OLS Positive and not statistically significant under 2SLS</td>
</tr>
<tr>
<td>Cristia (2008)</td>
<td>US (NSFG 1995)</td>
<td>Estimate and explain effect of a first child on female labour supply</td>
<td>Linear probability estimates, Multinomial logit estimates DV: employed21 Key variable: anychildren21 Controls: age, year0, educ, smoke, exper, Black, Hispanic, married, insurance covered IV: infertility</td>
<td>OLS: -26% in female employment Multinomial logit estimates: -45.8% in full time employment -5.7% in part time employment</td>
</tr>
<tr>
<td>Agüero and Marks (2008)</td>
<td>Latin America (Demographic and Health Surveys in Peru, 1996; Guatemala, 1998; Colombia, 1995; Bolivia, 1994 &amp; 1998, Nicaragua, 1998; the Dominican Republic, 1996)</td>
<td>Estimate and explain effect of additional children on FLFP</td>
<td>OLS, IV, GMM estimates DV: FLFP Key variable: number of children Controls: age, educ, age*educ, country fixed effect, married, location, BMI indicator IV: infertility shocks</td>
<td>OLS: -2.4% to -3.2% in FLFP IV; and GMM estimates: no effect</td>
</tr>
<tr>
<td>Chun and Oh (2002)</td>
<td>Korea (NSFIE, wave1996)</td>
<td>Estimate the effect of an additional child on FLFP of married women</td>
<td>Probit, 2SLS probit IV: first child’s sex</td>
<td>-27.5% in FLFP probability</td>
</tr>
<tr>
<td>Authors</td>
<td>Country/Source</td>
<td>Research Question</td>
<td>Methodology</td>
<td>Results</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------------------------------------</td>
<td>------------------------------------------------------------------------------------</td>
<td>-------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
Key variable: more than two children  
Controls: age, age at first birth, first child’s sex  
IV: sex of the first two children  
OLS: -8.3% to -9.7% for all-married women sample in Argentina  
-8.1% to -9% for married-all women in Mexico  
2SLS: -8.2% to 9.6% for all-married in Argentina  
-6.3% to -8.6% for all-married in Mexico |
Key variables: fertility  
Controls: age, age at first birth, boy 1st, boy 2nd, foreign  
IV: the first two children’s sex  
OLS: -8.3% to -10.5% (1991 and 2001)  
2SLS: -10% significant in 1991 -13.6% insignificant in 2001  
IV: same-sex is weak |
| Orbeta (2005)     | Philippines (2002 Annual Poverty Indicators Survey) | Estimate and explain the impact of the number of children on labour participation and earnings of parents | Two stage probit, FIML, simple probit  
DV: LFP  
Key variables: number of children  
Controls: age, age^2, years of educ, father’s income, non-labour income, unemp rate, regional dummy.  
IV: the first two children’s sex  
-5.7% in FLFP in simple probit  
Two stage probit and FIML failed (no significant instruments) |
IV: miscarriage in the first pregnancy of teenage mothers (aged 13-17)  
OLS: -21% hours of work for teen mothers compared to their delayed childbearing until adulthood |

**Longitudinal**
<table>
<thead>
<tr>
<th>Studies</th>
<th>Country/Period</th>
<th>Estimation Method</th>
<th>Key Variables</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>De la Rica and Ferrero (2003)</td>
<td>Spain (European Household Panel for Spain (1994-1998))</td>
<td>Estimate and explain effect of a newborn on married women’s labour participation</td>
<td>Pooling data from 5 waves, Probit estimates DV: FLFP Key variables: fertility Controls: lagged participation, age, educ, child status, spouse’s characteristics, year dummy</td>
<td>-17% in FLFP for women having newborn with exogenous fertility -39% in FLFP for women having newborn with endogenous fertility</td>
</tr>
<tr>
<td>Moschion (2013)</td>
<td>Australia (2006 Australian Census)</td>
<td>Estimate effect of fertility on mother’s labour supply</td>
<td>OLS, IV estimates IV: twins, sex mix of siblings</td>
<td>-12% in FLFP and 4 hours worked/week if having more than one child -12% in FLFP and 3 hours/week if having more than two children</td>
</tr>
</tbody>
</table>

FIML: full information maximum likelihood; Employed21 (=1 if employed in month21); Anychildren21 (=1 if the woman gave birth by month21 and 0 if she did not); Year0 is year in which the women sought fertility service for the first time; BMI is an indicator of health status.
Rosenzweig and Wolpin (1980a) is one of the first studies to try to estimate the relationship between fertility and FLFP using twins at first birth as an IV for fertility. The authors note that the occurrence of twins in the first birth can allow a close approximation. Twins in the first birth is a natural, unplanned event that simulates having an extra child and represents a time failure; twins is an unexpected event. Therefore, the occurrence of twins is a good instrument to measure the effects of exogenous variations in fertility by simple statistical techniques\(^5\). Rosenzweig and Wolpin (1980a) use 25 twin pairs from India to estimate the effect of family size on school attainment. Rosenzweig and Wolpin (1980b) use a sample of 87 twin pairs from the US to examine the effect of fertility on FLFP. They find that twins at first birth has a negative effect on FLFP by 37 percentage points for women aged 15-24 and by 10 percentage points for the 25-34 age group. Moreover, they show that the true effect of fertility on women’s labour force participation is significantly underestimated without applying an IV for fertility. Although twins in the first birth is a good IV for fertility, occurrences of twins are rarely observed. The study was only based on a very small sample of 87 twins so, as a result, estimated coefficients may not be precise.

In order to address the data limitations of Rosenzweig and Wolpin (1980a and 1980b), in their study, Bronars and Grogger (1994) firstly use census data to investigate the consequences of unplanned births for both marital and non-marital mothers on various outcome variables including labour force participation, earnings, poverty and welfare recipiency. The authors demonstrated that larger samples of the occurrence of twins at the first birth are possible to obtain by using data from the 1970 and 1980 censuses in the US\(^6\). The study investigated the impact of an additional child on these outcome variables, in which they mostly focused on unwed mothers, and did not exploit much the labour supply responses of married women. By comparing twins’ mothers and single child mothers at first birth, the presence of an extra child has a significant negative effect on the participation probability of unwed mothers. The study also

---

\(^5\) One obvious limitation of using the twins-first method is that although it can be used to estimate the marginal effect of an additional child among mothers with one child, it does not allow measuring the effect of the transition from no children to one child. However, if the majority of married women in the sample have one or more than one child, it is useful to measure the effect of an additional child on the labour supply of mothers.

\(^6\) Angrist and Evans (1998) used data on twin births from the 1980 census in the US to estimate the effect of fertility on female labour supply. In addition, they also used another possible instrument: the sex mix component of the first two children.
suggests that the effects of multiple births are different by race and across cohorts of unwed mothers. They estimate an 11% to 12% decline in labour participation probability for unwed mothers having an unplanned child aged 0-3. These negative effects are larger and more persistent among black unwed mothers.

Another study in the US by Jacobsen et al. (1999) uses the above dataset. However, unlike Bronars and Grogger (1994) who mainly examine the effect of unplanned births on labour supply for unwed mothers, Jacobsen et al. (1999) expand upon Bronars and Grogger’s finding by examining the impact of twin births at first birth on FLFP as well as the hours of work of married women. Therefore, they were able to use a larger sample size of married women with twin births for the analysis of their labour supply responses. They conclude that the overall impact of an unplanned birth on labour supply and earnings of married women is small but statistically significant, which is generally consistent with the results of Bronars and Grogger (1994). When controlling for years since the first birth, the presence of an unplanned child has a large and statistically significant effect on labour supply during the first few years immediately following the unanticipated birth and the effects diminish when the child reaches four years old or more.

One of the other contributions to address the endogeneity problem of fertility is the work of Cristia (2008). Similar findings are exhibited in the work of Cristia (2008) with infertility shocks as instruments for fertility. Unlike the previous studies using a twin first birth (Rosenzweig and Wolpin 1980b; Bronars and Grogger 1994; Gangadharan et al. 1996; Jacobsen et al. 1999) or mixed sex siblings (Angrist and Evans 1998), the author used infertility as an IV. In order to overcome the problems of unobserved heterogeneity and endogeneity that Browning (1992) mentions⁷, this paper is restricted to a sample of women from the National Survey of Family Growth 1995 (NSFG) in the US who did not have a child and sought assistance from a fertility service to have the first pregnancy. After a certain time of seeking help to achieve pregnancy, some of these women then experienced childbirth. The study examines the effect of the first child on female labour supply. The result shows that the presence of the first child under one

---

⁷ Unobserved heterogeneity: some researchers state that fertility behaviours may be endogenous so the relationship between measures of fertility and FLFP cannot be interpreted as causal effects.
year old reduces the working probability of women giving birth by 26 percentage points compared with women who did not give birth.

The negative effect of children on the FLFP rate is also found in another empirical study in the US by Hotz et al. (2005). The approach of Hotz et al. (2005) is different from Bronars and Grogger (1994) who used twins at first birth as an IV to identify effects of an extra child between teen mothers having twins in the first birth and teen mothers having a single child. Hotz et al. (2005) used the miscarriage of teenagers as an IV to investigate the effect of teenage childbearing on their socioeconomic outcomes. The study uses National Longitudinal Survey of Youth, 1979 in America and the estimates find that the negative effect of childbearing in teens is much smaller than previous findings and that delaying childbirth did not greatly raise their educational level or later earnings. However, in general, in terms of many socioeconomic outcomes, the paper’s findings are consistent with previous studies in the US such as Geronimus and Korenman (1992), Bronars and Grogger (1994) about the extent and nature of teenage childbearing as a social problem in the country. However, using miscarriage as an IV makes it difficult to attain a sufficiently large sample size and miscarriage as well as abortion events that may not be randomly assigned at times and unexposed in surveys. This problem may create imprecise findings.

The impact of fertility on FLFP is also approached in other studies using panel data. Carrasco (2001) developed a switching binary panel data model to estimate the causal effect of fertility on FLFP using the Panel Study of Income Dynamics 1986-1989. The model controlled for unobserved heterogeneity and predetermined existing children. The study found that the negative effect of fertility on participation is downward biased with the exogeneity assumption of the fertility variable. The effect is upward biased if unobserved heterogeneity (unobserved individual effects) such as parents’ preferences for children or for participation is not taken into account. The effect of fertility is smaller when existing children are treated as strictly exogenous variables.

A Spanish study by De la Rica and Ferrero (2003) also uses panel data from five waves of the European Household Panel from 1994 to 1998 for a sample of Spanish women allowing for the existence of unobserved factors. The lag of participation is included in

---

8 Hotz et al. (1997) developed a test for assessing the validity of instruments “miscarriages”. Test results show that it cannot reject the validity of miscarriages as an IV.
estimating because the authors said that “past participation is likely to affect the current participation positively” (De la Rica and Ferrero 2003, p. 165). The results of this paper are consistent with previous studies by Rosenzweig and Wolpin (1980a and 1980b) and Carrasco (2001) where the true impact of fertility on participation of women is downward biased and, when the endogeneity of fertility is taken into account, fertility has a stronger and negative effect on participation. The authors found a woman with a newborn has a 17% lower participation probability in the labour market with an assumption of exogenous fertility. In contrast, if fertility is assumed to be endogenous, a woman with a newborn has a 39% lower participation in working than a woman without a newborn.

In general, previous studies attempted to apply IVs such as twin first birth, miscarriage and infertility shocks (Rosenzweig and Wolpin 1980b; Bronars and Grogger 1994; Hotz et al. 2005; Cristia 2008). These events occur randomly and so they can be used as good IVs to generate exogenous variations of fertility. However, the challenge of using these IVs is that it is difficult to obtain large sample sizes as such events are rarely observed in the survey sample. As a result, the small sample may produce estimated results that affect precision.

Another approach to the endogeneity of fertility problem was introduced by Angrist and Evans (1998). Angrist and Evans (1998) are the first authors to introduce exogenous variability in family size by using American parental preference for a mixed sibling sex composition to establish IVs for fertility. The study used data from the Census Public Use Micro Samples in the years 1980 and 1990 to investigate empirically the effect of having more than two children on parents’ labour supply in the US. The sex of the first two children is applied as an IV for the “more than two children” variable in the sense that couples with two same sex children are more likely to have a third one than families with one boy and one girl. The study found that the first child’s sex has a very small impact on fertility or, in other words, there is no significant difference between families where the first child is a girl or a boy on the probability of having a second one. Moreover, families whose first two children are of the same sex (two boys or two girls) are more likely to have additional children than families with a sex mix (one boy and one girl). The OLS result showed that having a third child reduces the probability of working for women (for both all and married 21-35 year old women samples) by 17
percentage points, while this effect on husbands is small (by less than 1%). Under the 2SLS method, having more than two children reduces female labour supply by 12% which is smaller than corresponding OLS results. The 2SLS estimates showed that the coefficients of labour supply are not very much different when decomposing the two components of same sex into two boys and two girls in both the 1980 and 1990 data.

Following Angrist and Evans (1998), many empirical studies also provide similar evidence of the negative impact of fertility on women’s participation in the labour market (Carrasco 2001; Iacovou 2001; De la Rica and Ferrero 2003; Moschion 2013). Iacovou (2001) shares the same ideas as Angrist and Evans (1998), based on the idea that parents prefer to have a balanced sex composition of children than a same sex composition of children. Thus, mothers are more likely to have a third child if their first two children are of the same sex. The study used the sex of the first two children and multiple second birth as IVs for the “having a third child” variable because these two instruments are highly correlated with fertility but uncorrelated with labour supply. Using two datasets from the British National Child Development Study (all women were at the same age - 33) and the British Household Panel Study 1992 (women aged 21-49), Iacovou (2001) found a significant negative effect of having a third child on FLFP under OLS and this impact is not very much different from the estimates obtained by Angrist and Evans (1998). But the 2SLS results showed that this effect becomes positive and not statistically significant and it tends to approach close to zero as more variables are included in the estimation. The author also mentioned other possible reasons such as policies for the labour market or the welfare system to explain the different effects of a third child on labour supply between the US and the UK. The limitation of both studies is that the analyses only focus on a sample of women with at least two children.

The effect of children on mothers’ labour participation is different across families because of other factors such as age, age at first birth, education, marital status, race, etc. Bloom et al. (2009) demonstrate that the effect of fertility on labour participation differs by age. The authors used unbalanced five year panel data for 1960-2000 for 97 countries. They stated that a birth reduces the labour supply of a woman by nearly two years. They also found that the probability of maternal labour participation decreases by 10 to 15 percentage points if a woman has an additional child when she is in the 25-39
age group and 5 to 10 percentage points in the 40–49 age group. These differences in the size of the effect of fertility on maternal labour participation are also shown in many other studies (Cain and Dooley 1976; Lehrer 1992; Chun and Oh 2002; Assaad and Zouari 2003; De la Rica and Ferrero 2003; Cristia 2008).

Few studies focus on the relationship between fertility and labour participation for developing countries. Cruces and Galiani (2007) used the IV approach of Angrist and Evans (1998) to examine the effect of fertility on FLFP of all women aged 21-35 and married women samples in Argentina and Mexico. They show the presence of a sex mix preference in these countries. The study finds that the impact of having more than two children on Argentinian and Mexican women’s labour participation is similar to the findings by Angrist and Evans (1998) in the US. They stated that Angrist and Evans (1998) results can be qualitatively and quantitatively generalised. In contrast, Agüero and Marks (2008) used infertility shocks as an IV for childbearing to identify the effect of children on female participation in the labour market. Their identification method is basically similar to Cristia (2008) approach who compared employment outcomes between two groups of women: women who gave birth and women who did not after seeking fertility assistance. But the difference is that the infertility in this study includes both primary and secondary infertility, while Cristia (2008) includes only primary infertility. The authors employ this IV for a sample of women in six Latin American countries. They find a significant negative effect of an additional child on women’s labour force participation from OLS estimates. Moreover, the probability of participation for working among women with children is about 9% lower than that of childless women. However, they could not find any evidence for this relationship in the IV and GMM models. In general, the result does not support the inverse relationship between family size and female labour supply. A weak identification strategy is one of the limitations of this study because the measure of infertility is based on women’s health status (poor health), while it is possible that their health status directly affects their participation in the labour market.

In a study of Cameroon, Nanfosso and Zamo-Akono (2010) examine the relationship between fertility, health and labour supply using cross-section data from a population survey in urban Cameroon. They demonstrate that fertility and health status are interrelated and the result will be misleading if we separately estimate fertility or health
status on women’s labour force participation. However, the authors show a positive effect of fertility on labour participation of Cameroon women. Besides the inverse relationship between fertility and women’s labour participation, some countries, such as Denmark and France, have high fertility rates as well as high labour participation, while both low fertility and labour participation can be seen in Italy and Spain (Del Boca and Pasqua 2005).

The negative relationship between fertility and female labour participation might be a general trend in developed countries. In contrast, in developing countries, despite a continuously declining trend in fertility and a remarkable increase in education levels, FLFP has been constant or declining in most Asian countries over the last three decades (ILO 2008). Studies of FLFP in Asian countries show mixed evidence; the effect of fertility on FLFP is still ambiguous. Several studies show negative relationships between fertility and FLFP (Chun and Oh 2002; Orbeta 2005; Ebenstein 2007). For example, in the Philippines, Orbeta (2005) examines the impact of the number of children on labour force participation and earnings of parents using data from the 2002 Annual Poverty Indicator Survey. The empirical results from the probit model indicated that the presence of children below six reduces the probability of mothers’ participation in paid work by 5.7 percentage points. Using census data from Taiwan and the US in his study, Ebenstein (2007) also points out that the average causal effect of the third child’s presence on reducing the probability of women’s participation in Taiwan is 10% among women of the 34-36 age group. In general, both the above studies point out that fertility reduces the probability of women’s participation but these effects are relatively small.

Some other studies have shown that the presence of children has positive impacts on FLFP. For example, the Indonesian study by Priebe (2011) is considered as an example of the ambiguous relationship between fertility and FLFP. Priebe (2011) found that children play a dual role in women’s decisions about labour supply. Children not only decrease the probability of labour force participation of women but also create an incentive to engage in the labour market, especially in developing countries where women’s participation in the labour force is often an essential financial decision in order to maintain basic needs (Priebe 2011). The study suggests that fertility has a positive effect on female labour supply.
However, the empirical results for the Philippines and Taiwan are different from Indonesia, even though the observed trend is similar. Both studies show that fertility reduces the probability of women’s participation in work outside the home but the effect is relatively small. In the Philippines, Orbeta (2005) examines the impact of the number of children on labour force participation and earnings of parents using data from the 2002 Annual Poverty Indicator Survey. Following the IV strategy of Angrist and Evans (1998), the author uses the sex of the first two children as an IV for the number of children. Results from the two stage probit and full information maximum likelihood (FIML) estimates show that instruments such as same sex or two boys or two girls are invalid for the endogeneity of the number of children. However, when the study applies the simple probit model, the presence of children below six reduces the probability of mothers’ participation in paid work by 5.7 percentage points.

Unlike the Philippines where there is an absence of son preference, son preference is dominant in Taiwan. Using census data from Taiwan and the US in his study, Ebenstein (2007) shows that the effect of children on mothers’ labour supply under 2SLS estimates is greater than under OLS in Taiwan but the effect under 2SLS estimates is smaller than that under OLS in the US. The study estimated a structural model that points out that the average causal effect of the third child’s presence on reducing the probability of women’s participation is 10% in Taiwan and 12% in the US among women between 34-36 years of age. The effect under IV estimate in the US (-7%) is lower than the average causal effect and this effect in Taiwan (-11%) is close to the average causal effect. The results suggest that IV estimates are affected by the strength of instruments in both Taiwan and the US. IV estimates will be lower than the average causal effect with weak instruments and the result will be closer with strong instruments.

While the presence of children presents a comparatively small effect on FLFP in the Philippines and Taiwan, the effect of fertility on FLFP can be seen more clearly in Korea in several studies (Park and Cho 1995; Chun and Oh 2002). Chun and Oh (2002) examine the relationship between fertility and FLFP, where the preference for a son is a dominant value among Korean families. It is different to Angrist and Evans (1998) who used the sex composition of the two first children as a good IV for having more children among women having at least two children in the US where parents’ preference is for a
balanced sex composition. In contrast, Korean families prefer sons to daughters, and so the author used the sex of the first child as an IV to generate exogenous variation for fertility. Using data from the Korean National Survey of Family Income and Expenditure 1996, they point out that mothers having a son in their first birth are less likely to have other children and that having an additional child reduces the probability of labour force participation among Korean married women by 27.5% compared with mothers with at least one child. This rate is 37.7% in families having two or more children. It shows that the effect of the presence of children on labour force participation of Korean women increases with the increase in the number of children that a family has.

**Vietnamese studies**

In Vietnam, to the best of my knowledge, there is only one existing study about FLFP by Le (2009) that uses the research methodology of the thesis. Le (2009) analyses the effect of fertility on parents’ labour supply. He uses the sex of the first child and the same sex of the first two children as IVs for dealing with the endogeneity of fertility. Le (2009) uses a sample of 3,935 households from the VHLSS 2004. He finds that the probability of FLFP declined by 26.4% when families have one more child. When investigating urban and rural areas separately, the effect of fertility is not significant on labour participation decisions of urban women, whereas mothers in rural area have a reduction of 29.3%.

The findings by Le (2009) have certain significant contributions to the literature in the research field to some extent but this study also has some major limitations. First, the author mentions the existence of a son preference in Vietnam but he uses the same sex of the first two children as an IV for having more children. This means that he believed a combination of children with different sexes is the Vietnamese parents’ preference. In the research by Angrist and Evans (1998), Iacovou (2001) and others, parents prefered a balanced sex combination of children, but Vietnamese families in fact have a strong preference for sons (Haughton and Haughton 1995; Bélanger 2002). Like China, India, Korea and other Asian countries, Vietnamese families have a strong preference for sons because of economic, social, symbolic and cultural reasons (Haughton and Haughton 1995; Bélanger 2002); hence the instrument “same sex of the first two children” might not be appropriate in the case of Vietnam. Second, the sample for his study includes
parents aged between 18 and 60 with at least one child. As a result, the analysis fails to recognise the gender composition of the first two children because some adult children who are older than 18 may get married and have separate families or other reasons. Moreover, the analysed sample includes women who are outside of the age group of fertility, so it may not lead to precise results. Third, Le (2009)’s analysis is based on data for a single year with a small sample size of 3,935 households from the VHLSS 2004.

In order to address these limitations, this thesis will use census data from 2009 with a much larger sample size of 3,692,042 households to analyse the effect of fertility on FLFP. The study also restricts the analytical sample to households with children living with their parents. More importantly, this study will use IVs that are consistent with son preference in Vietnamese households. In addition, the study also uses unplanned fertility (twins at second birth) as an instrument and then compares results from these two instrumental methods. Thus, it can generate more reliable results.

Given the lack of empirical evidence about FLFP, the study will fill a gap in Vietnamese literature by investigating determinants of FLFP, focusing on the effect of fertility using a large dataset and accounting for the endogeneity of fertility with instrument technology. This study will contribute to the existing literature by providing new evidence on the effect of having more than two children on FLFP, using instrument variable “twogirl” (appropriate given parent’s son preference in Vietnam) as an identification strategy.

2.2.2 The impact of parental health on children's schooling and labour force participation

Family and parental characteristics not only affect the current generation but also have transmission effects between generations from parents to children such as health and education. Children are more likely to be healthy and well educated if their parents have good health and a good educational background. In contrast, unhealthy and less educated parents may have negative impacts on children’s outcomes such as children’s psychological health, behavioural development, educational attainments and employment.
This study relates to the substantial literature on intergenerational transmission of education, health and income which has been of interest to many economists (Black et al. 2003; Black and Devereux 2010; Fleury and Gilles 2015) and, more generally, on the effect of parental characteristics on children’s educational attainments because of the substantial impact of an individual’s education on her/his future life and achievements (see Currie (2009) for a survey and Case and Paxson (2011) among many others).

The topic in this study relates to a smaller and currently growing area of literature that explores the relationship between parental health and children’s education. The effect of parental health on children’s education has received rather limited attention in the economic empirical literature (Haveman and Wolfe 1995). According to the well-known survey by Haveman and Wolfe (1995) on the determinants of children’s attainment, the authors show that this research area severely lacks data on the health status of parents and children and, until now, empirical studies of the relationship between parental health and children’s education is still very limited. Most studies focus on related literature on the effect of parental death on a child’s education. Recently, scholars have begun to explore the consequences of poor parental health on children’s education.

The study of consequences of parental health shocks is particularly important in developing countries where there is an absence of adequate systems of social protection and people have restricted access to formal health insurance (Gertler and Gruber 2002; Wagstaff 2007). Major illnesses are unpredictable and can severely affect families, both in the rural and urban context (Wagstaff 2007; Sun and Yao 2010). Adverse health events can impose monetary costs on family members due to costs of health care services, the reduction in family labour supply, productivity and income by ill family members (Gertler and Gruber 2002; Wagstaff 2005, 2007; Somi et al. 2009; Wagstaff and Lindelow 2014).

Also, these health shocks can cause non-monetary costs in terms of emotional and psychological distress which lead to negative impacts on children’s educational attainments (Sieh et al. 2010). Parental health shocks can also have negative impacts on the development of children’s behavioural skills which is a very important component of human capital and closely related to educational attainment and labour market outcomes (see for example Cunha and Heckman (2008)). Using panel data from
Germany, Mühlenweg et al. (2016) analyse the importance of parental health on the development of child behaviour and find that maternal health shocks have negative effects on emotional symptoms, conduct problems and hyperactivity among children aged 6.

Furthermore, in the production of child quality, parental inputs play a very important role. Apart from provision of financial inputs, the time that parents spend with children is one very important channel for intergenerational transmission of human capital. A decrease in the amount of time spent with children, as well as a lack of involvement by parents in their children’s educational activities, may negatively affect a child’s quality in general and educational achievements in particular (Leibowitz 1977; Greenwood and Hickman 1991; Morefield 2010; Gayle et al. 2012).

All of these factors can have detrimental consequences on the family in both the short and long term because of the reduction in consumption as well as the decrease in accumulation of productive assets including education (Sun and Yao 2010). The effect of adverse health events has been found to be more pronounced on poor and less educated families (Dercon and Krishnan 2000; Genoni 2012) and even may lead households into a chronic poverty trap (Wagstaff 2007; Sun and Yao 2010).

When analysing the economic consequences of health shocks in Vietnam, Wagstaff (2007) shows that the financial impact of adverse health events can be catastrophic because it puts pressure on households to significantly reduce consumption, even necessary consumption such as enough food to maintain healthy calorie intake, and at the same time increase other costs such as housing, electricity and transportation. When household income is decreased, families must smooth consumption through a set of coping mechanisms including borrowing, selling valuable assets, reducing human capital investment and so on.

Similarly, Mitra et al. (2015) use fixed effects to analyse the economic effect of health shocks on working-age adults in Vietnam and show that health shocks in the working-age population led to serious economic consequences for families over the period 2004-2008 because of the large increase in healthcare costs. They also show that female-headed and rural households are the most vulnerable to these adverse health events and, in general, households tend to reduce non-health related expenditures such as food and
education. Households also adopt a number of coping mechanisms to smooth the effect of current consumption such as borrowing, selling assets and cutting education expenditures, and such mechanisms may compromise future family welfare.

The economic consequences of parental illness not only affect the current generation but also transmit to future generations (Hamoudi and Sachs 1999). In the context of poor and less developed countries with a lack of adequate social protection systems and where access to credit funds is often limited, an adverse health event occurring for family members causes a rise in out-of-pocket health expenditures and decreases income. In the short term, families may respond and eliminate these health events using the coping mechanisms mentioned above. However, in the long term, families may be compelled to choose other coping mechanisms that can affect the human capital development of future generations (Strauss and Thomas 1995). Labour substitution is one of these coping mechanisms. This means that when a health shock happens, households may use child labour to substitute for adult labour. This is especially true for less developed and developing countries where the majority of households rely on agriculture and the main income is derived from agriculture. In some cases, children may also be required to help their family by taking care of the sick family member(s). As a result, children may have to reduce attendance time at school or drop out altogether.

Related literature focuses on the impact of parental death on children’s outcomes (Case et al. 2004; Gertler et al. 2004; Yamano and Jayne 2005; Case and Ardington 2006; Evans and Miguel 2007; Chen et al. 2009; Adda et al. 2011; Cas et al. 2014; Senne 2014). Some studies find that the deaths of both mothers and fathers have negative effects on child schooling (Case et al. 2004; Gertler et al. 2004; Beegle et al. 2006), while some others argue that maternal death has a much larger impact on children’s education outcomes than paternal death (Ainsworth et al. 2005; Case and Ardington 2006; Evans and Miguel 2007; Chen et al. 2009). The first empirical evidence came from studies on the impact of adult mortality on household income and consumption in the era of the HIV/AIDS epidemic (Yamano and Jayne 2004; Naidu and Harris 2005). Since then, studies have focused more on consequences of a parent’s death on health and schooling outcomes but most of them are based on cross-sectional data, which has provided mixed results (see for example Case et al. 2004; Ainsworth and Filmer
One explanation for this is that parental death can be highly correlated with household socioeconomic conditions which also influence the demand for children’s schooling, and therefore when analysing the impact of orphanhood, these cross-sectional studies are limited because they only control for current household and child characteristics after parental death has occurred. Another reason is that due to the existence of unobserved factors that can influence both death and schooling, such estimations may be biased (Senne 2014).

Some recent influential studies analyse the impact of parental death. Gertler et al. (2004) use three repeated cross-sectional data of households from Indonesia. They show that a recent parental death has a substantial effect on children’s schooling enrolment and that these effects are little different by the gender of the child or the parents. Some other studies use longitudinal data to address endogeneity and omitted variable problems. Two studies in Africa by Case and Ardington (2006) and Evans and Miguel (2007) use two year and five year panels respectively to analyse the effect of parental death on children’s education outcomes. They show that maternal death has an important and much larger impact on children’s education than paternal death. Chen et al. (2009) also find similar evidence on the importance of maternal death in Taiwan. The role of mothers is also confirmed in the research by Adda et al. (2011) who find that Swedish mothers are somewhat more important for the cognitive skills of children and fathers for non-cognitive skills. Similarly, using a three year panel from rural Kenyan households, Yamano and Jayne (2005) find that working-age adult mortality has an adverse impact on children’s primary school attendance among the poor, and while girls are less likely to remain at school before death, boys are less likely to be at school after death.

The more recent study by Senne (2014) uses longitudinal data and difference-in-differences methods to investigate the effect of adult death in rural households in Madagascar. These results suggest that children are less likely to attend school following adult death than their counterparts, and these negative effects are particularly pronounced for girls and young children from poorer families.

In general, all studies mentioned above recognise that major health problems may predate parental death and if all these health problems occur before parental death, the negative effects of parental death can be seriously biased. The effects that predate
parental death are probably due to the demand for caring for ill family members, so the estimates in cross-sectional studies tend to be biased by omitted variables (Evans and Miguel 2007; Bratti and Mendola 2014). In such cases, parental health can be considered as a compounding factor.

While there is substantial literature investigating the effect of parental death on child education, the economics literature on the impact of parental health on children’s education is very limited (only some studies) and gives unclear results, especially with respect to the impact of paternal versus maternal health (see Table 2.2).

Table 2.2 – Parental health and children’s education and work relationship literature

<table>
<thead>
<tr>
<th>Study</th>
<th>Country (Data)</th>
<th>Relevant Research Topics</th>
<th>Method</th>
<th>Main findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sun and Yao (2010)</td>
<td>Rural China (National Fixed-point survey 1987-2003)</td>
<td>Examine effect of major health shocks of adults on children’s school attainment (children aged 13-21)</td>
<td>Linear probability model (LPM) for baseline scenario with village fixed effects added</td>
<td>-9.9 percentage points chances for primary children to enter middle school</td>
</tr>
<tr>
<td>Choi (2011)</td>
<td>Russia (Russia Longitudinal Monitoring Survey 1994-2004)</td>
<td>Estimate effect of parental self-reported health problems on a child’s educational attainment and employment during their subsequent years (sample is children aged 13-29)</td>
<td>Probit model</td>
<td>Lower educational attainment and lower probabilities of participation in the labour market during subsequent years of daughters</td>
</tr>
<tr>
<td>Bratti and Mendola (2014)</td>
<td>Bosnia and Herzegovina (Living standards measurement)</td>
<td>Effect of parental health shocks on investment in education of children aged 15-24</td>
<td>Child fixed effect</td>
<td>-7 percentage points of likelihood of children’s school enrolment if mothers have poor health.</td>
</tr>
</tbody>
</table>
Sun and Yao (2010) examine the effect of major illnesses of working age adults in households on a child’s chances to enter or finish middle school in rural China. Their paper uses two sources of data: one is the National Fixed-point Survey started in 1987 which is a longitudinal survey; the other is a retrospective survey in 2003 that collects information on major illness of family members for the period from 1987 to 2002. They find that primary school aged children (and girls in particular) are the most vulnerable to family health shocks. A child’s chances to enter middle school drop by 9.9 percentage points if he/she experiences a family health shock at primary school age. However, middle school aged children are not affected by these health shocks.

Similarly, Choi (2011) explores the effect of parental self-reported health problems on children’s educational attainment and children’s employment during their subsequent years in Russia using longitudinal data from 1994-2004. She estimates the impact of self-reported health issues of parents on their child’s probability of completing at least 15 years of schooling. Results show that the probability of obtaining higher education as well as engaging in the labour market during the subsequent years for children is significantly lower for daughters if they have an unhealthy father.

Another study that relates to a broader measure of children’s outcomes in cognitive and non-cognitive skill development is developed by Morefield (2010). This study investigates the effect of parental health events that are reported to limit parents’ daily activities or the type of work they can do on children’s cognitive and non-cognitive skill development in the US. The two measures of health events used include specific health conditions and unspecific physical conditions. The specific health conditions are measured by the diagnosis of one of specific health conditions (stroke, heart attack,
heart disease, hypertension, asthma, lung cancer, diabetes, arthritis, memory loss, learning disorders, cancer and psychiatric problems) and non-specific physical conditions. His results suggest that the parental health shocks have no significant effect on children’s cognitive skills but have a small negative effect on their non-cognitive skills and the effect of health events related to vascular or cancerous conditions are more negative. The paper also identifies that the consequences of health events are more pronounced for a son’s behaviour than a daughter’s.

Unlike Morefield (2010), Johnson and Reynolds (2013) use a narrower measure of health status, long hospitalisations of family members, to reflect health shocks. Their study uses data from the National Longitudinal Survey of Youth 1997 to explore the effect of long hospitalisations of family members on the educational attainment of children at the ages of 12-18 in 1997. They find that household hospitalisations reduce children’s educational attainment in terms of years of schooling, the likelihood of completing high school, and of attending and finishing further education. The study also shows that the negative effect of household hospitalisations is different across children within a family. The burden of household health shocks appears to be more negative and concentrated among male respondents and oldest children.

Parental mental health is also considered as an important influential factor to children’s outcomes. Some studies investigate the effect of the mental health of parents on children’s schooling or human capital development (Farahati et al. 2003; Frank and Meara 2009; Bratti and Mendola 2014). Farahati et al. (2003) investigate the effect of parents’ psychiatric disorders on the high school dropout rate of children using the National Comorbidity Survey in the US 1990-1992. They show that parents’ mental illnesses increase a child’s probability of dropping out of high school. The study also finds that parents’ mental illnesses appear to be more negative for girls than boys. Similarly, Frank and Meara (2009) also explore the impact of mental health on cognitive and behavioural development and educational outcomes during children’s early school years, focusing on mothers’ mental health using data from the National Longitudinal Survey of Youth. They find that maternal depression and substance abuse have a larger effect on children’s behavioural outcomes than cognitive outcomes.

Mont and Nguyen (2013) use cross-sectional data from the VHLSS 2006 to show that parental disability significantly decreases the likelihood of child school enrolment and
the impact of maternal disability is the strongest. This effect is greater among children aged 15-17 and boys are more negatively affected than girls. Children of less educated parents tend to experience greater impacts. The study also suggests that parents with disability can lower the number of completed grades for children.

However, all of these studies use cross-sectional data or longitudinal data but the dependent variable *child schooling* is observed only at one particular point in time. For example, Sun and Yao (2010) examine the effect of adults’ illness on child schooling chances in rural China. While information on children’s education is collected in the 2003 retrospective survey, information of major illnesses is obtained over the period of 1987-2002. Moreover, the data from Sun and Yao’s study may suffer from recall errors because the 2003 retrospective survey asks family members to remember information pertaining to illnesses that happened during the prior 15 years. Therefore, these studies are not able to exploit the variation over time in children’s education and to control for individual characteristics that do not vary over time.

It may be due to the lack of suitable data and instruments that studies in the literature have had to use two main identification strategies to address the potential endogeneity concerns in estimating the impact of parental health on child schooling. The first strategy attempts to limit the effect of unobservable heterogeneity by using a rich set of simultaneous or lagged observable child and parent characteristics or by including matching techniques (Morefield 2010; Sun and Yao 2010; Choi 2011; Johnston et al. 2013). The second approach uses panel data and controls for unobserved time-invariant factors by household fixed effects (Chen et al. 2009; Adda et al. 2011)⁹ to exploit differences in educational achievement between siblings because it is assumed that in order to remove unobservable heterogeneity (parental characteristics or family background), the best control group for a child is his/her siblings. However, children of the same parents, in fact, may differ in both cognitive and non-cognitive abilities, and so parents may adjust their resource investment in children’s human capital (Ermisch and Francesconi 2000; Figlio et al. 2014).

The most recent approach to address the above concerns, is the use of the child fixed effect to remove differences in individual characteristics of children.

⁹ Moreover, the weakness of this approach is that identification is only used for families with at least two children (Adda et al., 2011)
The studies that are closest to ours, in terms of methodology, are Bratti and Mendola (2014) and Alam (2015). Both papers use child fixed effects to analyse the impact of parental health on child education, and also analyse the effect on child labour. However, these two papers provide different results with respect to the role of maternal and paternal illness. Further, these two studies consider a European and an African country with very different socioeconomic structures and challenges, and the study by Bratti and Mendola (2014) focuses on the effect of parental illness on child enrolment, while Alam (2015) focuses on the effect of parental illness on attendance at school.

Bratti and Mendola (2014) is the first study to investigate a causal relationship between parental illness and child schooling in Bosnia and Herzegovina. They use longitudinal panel data and child fixed effects as an identification strategy to address the potential endogeneity bias of parental health status on child schooling. Results using parents’ self-reported health measures suggest that a mother’s poor health status decreases the likelihood of school enrolment of children aged 15-24 by 7 percentage points and increases a child’s likelihood to enter employment due to the large increase in health expenditures rather than due to forgone income. However, a father’s poor health status appears to be not statistically significant on children’s education. The authors also employ several alternative health indicators - limitations in activities of daily living indicators and mental health indicators - to check for measurement errors of self-reported health status. Results suggest that the effect of a mother’s poor health on child school enrolment is robust with these more precise measures.

Following Bratti and Mendola (2014), Alam (2015) employs a longitudinal panel data survey of four waves during the period of 1991-1994 with child fixed effects. This paper investigates the effect of parental illness on children’s education in Tanzania. While Bratti and Mendola (2014) focus on older children aged 15-24, Alam (2015) focuses on primary and middle school aged children (aged 7-15) and only includes children who have already started primary school. In other words, it focuses on the school attendance of children with the condition of already having started primary school. The health measure of illness used includes individuals who are reported to be ill and unable to do their usual activities for at least one day due to the illness. The study finds that only a father’s illness has a significantly negative effect on children’s school attendance. The father’s illness also has a long-term impact on children’s education by
reducing the likelihood of finishing primary school and resulting in fewer years of schooling. However, there is no evidence for the effect of a father’s sickness on the reallocation of children’s time from school to work. Also, there is no differential impact of parental illness on child attendance or child labour by age or gender. Results show that the effects of a father’s illness on children’s education is through income channels by decreasing household income which leads to a decrease in the household’s ability to afford children’s education. In contrast, a mother’s illness or another family member’s illness have no effect on child schooling.

Economic literature on the relationship between parental health and child labour is also very limited. The effect of an individual’s poor health status on the labour supply of other family members is considered as an “added worker effect” (Bazen and Salmon 2010). A few studies show that parental illness can increase a child’s likelihood to enter employment. Bazen and Salmon (2010) investigate the impact of parental health on child labour in Bangladesh. They show that paternal illness increases a child’s probability of labour force participation. Furthermore, they also show that while a father’s permanent or chronic illness increases a mother’s labour supply, a child’s likelihood of participation tends to increase if the father’s illness is short term. Similarly, Dillon (2012) uses cross-sectional data from a household survey in Northern Mali and finds that parental health shocks increase children’s number of working hours in both household enterprises and child care. The paper also shows evidence that households adjust child labour in response to unexpected health events, and that children may not only be required to work as a complement for adult labour in agriculture, but also as a substitution for adult labour in child care. As stated by Edmonds (2005) in child labour literature, the adjustment of child labour to substitute for adult labour depends on labour market conditions and technology used in production. Thus, in response to parental health shocks, adjustment of child labour to substitute for adult labour depends on specific tasks (Dillon 2012).

Most studies in the literature on the effects of parental health shocks on child labour have considered various transmission channels of health shocks to offspring’s educational outcomes. First, parental health shocks may reduce children’s education by increasing the need for children to substitute for adult labour supply. In some cases, children may be required to work for pay to relieve their family’s financial difficulties
due to the consequences of the health shocks and, as a result, children are required to leave school. In some other cases, children need to increase labour supply for their family by decreasing school attendance or school work, and this leads to reallocation of children’s time from school to work (Choi 2011; Bratti and Mendola 2014; Alam 2015). Bratti and Mendola (2014) state that higher health expenditures may be the main reason children are pushed to enter the labour market. They find evidence that maternal health shocks increase employment probability of children and these effects are most likely due to the need to cover healthcare costs. In contrast, Alam (2015) shows that although paternal illness substantially reduces family income, this illness does not increase child labour nor cause a reallocation of children’s time from school to work.

Second, children may be required to care for a sick parent in the household therefore reducing the time available for school attention and school work and, in some cases, children may have to drop out of school (Mont and Nguyen 2013; Bratti and Mendola 2014). In this case, girls are often more affected than boys due to the responsibilities of caring for family members and housework typically being a traditional role for females (Pitt and Rosenzweig 1990; Strauss and Thomas 1995; Rose 2000). Further, parental illness may reduce parental engagement with the child’s schooling, as well as increasing the child’s emotional distress, and both factors are likely to have adverse effects on educational achievements (Guryan et al. 2008; Bratti and Mendola 2014).

In Vietnam, to my best knowledge, there is no study on the effect of parental illness on child schooling and child labour using child fixed effect as an identification strategy. There is only one study by Mont and Nguyen (2013) using cross-sectional data from the VHLSS 2006 to explore the effect of parental disability on educational attainments of children aged 6-17. Mont and Nguyen (2013) find that having a disabled parent reduces the likelihood of child school enrolment by 16 percentage points. This effect is greater among children aged 15-17 and boys are more negatively affected than girls. Children of less educated parents also tend to experience greater impacts. The study also suggests that parents with a disability can lower the number of completed grades of children. However, this study only uses one wave of the VHLSS dataset due to the fact that the measure of the health variable parental disability is only available in 2006 and this work only focuses on one health measure - parental disability.
Chapter 3. Conceptual framework

3.1 Introduction

The objective of this chapter is to review theoretical studies on female labour supply and then develop a broad conceptual framework encompassing both supply and demand aspects to explain women’s participation in the labour market.

Female labour force participation can be affected by many factors. Some factors come from labour market conditions such as the unemployment rate and gender discrimination. These disadvantages in the labour market can cause women to face employment uncertainty and lowers their probability of getting a job in the labour market. Other factors come from individual and family characteristics such as health, age, education, preference, fertility, family gender responsibilities and intergenerational co-residence, all of which can directly or indirectly affect FLFP. All these factors can affect female labour supply and they have also multiple effects through transmission channels.

Given the importance of Vietnamese women’s contribution to economic growth and development, issues related to gender equality policies in the labour market and women’s advancement, and the lack of prior Vietnamese research on FLFP, it is valuable to study the determinants of FLFP. If there is a decrease in the female labour participation rate, investigation of various factors affecting this participation is necessary in order to make appropriate policy adjustments. The effective design of labour market policies relating to female labour supply requires studying the various factors affecting FLFP from both the supply and demand sides. Assessing the determinants of women’s labour force behaviour in the specific context of a country is necessary for policy design aimed at equal opportunity in the labour market.

This study will contribute to the existing economic literature on FLFP by developing a broad conceptual framework analysing various factors affecting the labour force participation of Vietnamese women from both the supply and demand sides. The conceptual framework will then be used to analyse empirically the impact of various factors on female labour supply. However, since the availability of data is limited, this framework can only empirically analyse the effect of fertility on FLFP in Vietnam. Further, this conceptual framework also could serve as a basis for the formulation of
labour market policies towards gender equality in the labour market and women’s advancement.

3.2 Conceptual framework

This section will develop a conceptual framework by analysing the various factors affecting the labour force participation of women from both supply and demand aspects. The supply side includes factors such as fertility, partner’s income, potential childcare assistance from extended family members, the youngest child’s age, female age and education. The demand side includes factors such as gender wage differentials and the unemployment rate. The different ways in which FLFP is affected are summarised in Figure 3.1. The effects result from both the supply side and demand side of the labour market.

A Flowchart of the Conceptual Framework

Figure 3-1 Supply and demand considerations in female participation in out-of-home work in a developing and culturally changing economy: The case of Vietnam
3.3 A model of female labour supply under employment uncertainty and discrimination

Since the late 1970s, there have been dramatic changes in the business climate such as economic recessions, technological change, economic crises and global competition. In particular, intensified global competition has pushed firms to reduce production costs and become more flexible. New technologies have been applied in production with more demand for skilled workers and less demand for less skilled workers. All of these changes have forced companies to implement restructuring and downsizing in their attempts to improve their organisational effectiveness and competitive ability to adapt to a changing business environment. As a consequence, many workers have become unemployed, involuntarily part time unemployed or hired on temporary and seasonal employment contracts.

When facing employment uncertainty, women are often more vulnerable than men due to expected gender responsibilities in the family, employment discontinuity arising from motherhood roles and gender discrimination in the labour market (Becker 1957; Mincer and Ofek 1982; Thornton et al. 2013). While the female has to face all these factors that may affect her labour supply from both supply and demand sides, in order to maximise her utility, she must decide whether she should participate in the labour market and if yes, how much time she should devote to outside work. This study introduces a model of female labour supply under employment uncertainty and discrimination in the labour market. The model assumes that, if she participates in market work, the female’s utility is then a function of her family’s disposable income, which includes her partner’s income, her own income and her leisure. Otherwise, if she does not participate in the labour market, the female’s utility is a function only of her partner’s income. The model will make a prediction in comparative statics about the female’s optimal decision of the level of participation in the labour market, particularly the amount of time supplied by the female to market work in order to maximise her utility based on the given factors from both supply and demand sides as mentioned above.

A model of female labour supply

Facing uncertainty about employment, a female’s utility is random (binomial, for simplicity):

The model will assume that:
\[ \tilde{u} = \begin{cases} 
    u([y_p + (1-\delta)w_m(e)L - c(L, n, z, s)](1-L)), & p(L, n, z, z, v) \\
    u(y_p, 1) & 1 - p(L, n, z, z, v)
\end{cases} \]

Where

\( u \) is the female’s utility from the household’s disposable income \([y_p + (1-\delta)w_m(e)L - c(L, n, z, s)]\) and her leisure \((1-L)\) is displaying positive but diminishing marginal utilities with respect to both arguments;

\( 0 \leq L < 1 \) is the time supplied by the female to out-of-home work (a control variable);

\( y_p \geq 0 \) is the female’s partner’s income (taken to be exogenous);

\( w_m > 0 \) is the male wage rate (taken to be exogenous);

\( 0 \leq \delta < 1 \) is the rate of deviation of the female wage rate from the male wage rate (exogenous, reflecting discrimination);

\( e \) is the female’s educational attainment level or schooling years;

\( n \) is the female’s number of children (may be endogenous under rationality);

\( z_j \) is the age of the youngest child;

\( z_i \) is the age of the observed female;

\( v \) is the unemployment rate;

\( s \) is parents’ and other extended family members’ potential care provision for the female’s children (measured, say, by proximity to the household);

\( c(L, n, z, s) \) is the cost associated with the female’s absence from the household during outside work and where \( \frac{\partial c(L, n, z, s)}{\partial z} > 0, \frac{\partial c(L, n, z, s)}{\partial n} > 0 \),

\( \frac{\partial c(L, n, z, s)}{\partial z} < 0 \) and \( \frac{\partial c(L, n, z, s)}{\partial s} < 0 \); and

\( p(L, n, z, z, v) \) is the female’s probability of obtaining an out-of-home job and where

\( \frac{\partial p(L, n, z, z, v)}{\partial v} < 0, \quad \frac{\partial p(L, n, z, z, v)}{\partial L} < 0, \quad \frac{\partial p(L, n, z, z, v)}{\partial n} < 0, \)
\( \partial p(L, n, z_j, z_i, v) / \partial z_j > 0 \) and \( \partial p(L, n, z_j, z_i, v) / \partial z_i = 0 \) as \( z_j \leq z_i \); reflecting employer’s tastes and statistical discrimination against females, mothers in particular, with a possible sensitivity to their age and the age of their dependents.

The probability of obtaining outside work is affected by many different factors including both supply and demand sides such as the number of children, the time supplied by the female to outside work, the youngest child’s age, the female’s age and the unemployment rate.

While a woman has to face uncertainty about employment along with many other factors that may affect her labour supply, in order to maximise her utility she will have to decide whether she should participate in the labour market and if yes, how much time she should devote to outside work. According to Becker (1965), time allocation decisions of an individual need to be put in family context. This means that family is considered as a decision making unit of analysis and an individual’s decision of labour supply should be analysed and takes into account other members of family. Therefore, the female’s decisions of labour supply need to be considered in her family context and taken into account her family members such as her partner and her utility also depends on her family’s utility. The model will identify the optional time to be supplied by the female to the labour market work to maximise her utility.

Assuming von Neumann-Morgenstern utility theorem holds, a decision maker faced with different risks from his/her choices will behave as if he/she is maximising expected utility. Therefore the female’s decision problem is:

\[
\max_{(L,n)} \{ E\tilde{u} = p(L, n, z_j, z_i, v)u([y_p + (1-\delta)w_m(e)L - c(L, n, z_j, s)], (1-L)) \\
+ [1 - p(L, n, z_j, z_i, v)]u(y_p, 1) \} 
\]

For simplicity, the study will start with the case where the number of children is predetermined, hence exogenous. Namely:

\[
\max_L \{ E\tilde{u} = p(L, n, z_j, z_i, v)u([y_p + (1-\delta)w_m(e)L - c(L, n, z_j, s)], (1-L)) \\
+ [1 - p(L, n, z_j, z_i, v)]u(y_p, 1) \} \tag{\*
\]

Set \( y = [y_p + (1-\delta)w_m(e)L - c(L, n, z_j, s)] \); \( l = 1 - L \)
The necessary condition for maximum expected utility is:

\[
\frac{\partial E_u}{\partial L} = \frac{\partial p(.)}{\partial L} \left[ u(y, l) - u(y_p, 1) \right] + p(.) \left[ \frac{\partial u(.)}{\partial y} \left( (1 - \delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] = 0 \quad (3.1)
\]

In the special case where the probability of being employed is independent of the level of participation by women, \(dp/dL = 0\). This means that the woman’s probability of obtaining outside work is not dependent on how much time the woman can work. The first order condition for maximum expected utility becomes:

\[
\frac{\partial E_u}{\partial L} = p(.) \left[ \frac{\partial u(.)}{\partial y} \left( (1 - \delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] = 0
\]

The second order condition (SOC) for maximum expected utility (*) is:

\[
\frac{\partial^2 E_u}{\partial L^2} = \frac{\partial^2 p(.)}{\partial L^2} \left[ u(y, l) - u(y_p, 1) \right] + \frac{\partial^2 p(.)}{\partial L^2} \left[ \frac{\partial u(.)}{\partial y} \left( (1 - \delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] + \frac{\partial p(.)}{\partial L} \left[ \frac{\partial u(.)}{\partial y} \left( (1 - \delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] + p(L, n, z_j, z_v) \left[ \frac{\partial^2 u(.)}{\partial y^2} \left( (1 - \delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right)^2 - \frac{\partial u(.)}{\partial y} \frac{\partial^2 c(.)}{\partial L^2} \right] - \frac{\partial^2 u(.)}{\partial l^2} \right]
\]

\[
= \frac{\partial^2 p(.)}{\partial L^2} \left[ u(y, l) - u(y_p, 1) \right] + 2 \frac{\partial p(.)}{\partial L} \left[ \frac{\partial u(.)}{\partial y} \left( (1 - \delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] + p(L, n, z_j, z_v) \left[ \frac{\partial^2 u(.)}{\partial y^2} \left( (1 - \delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right)^2 - \frac{\partial u(.)}{\partial y} \frac{\partial^2 c(.)}{\partial L^2} \right] + \frac{\partial^2 u(.)}{\partial l^2} \right]
\]

SOC: \( \frac{\partial^2 E_u}{\partial L^2} < 0 \) if

\[
\left\{ \frac{\partial^2 p(.)}{\partial L^2} \left[ u(y, l) - u(y_p, 1) \right] + 2 \frac{\partial p(.)}{\partial L} \left[ \frac{\partial u(.)}{\partial y} \left( (1 - \delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] + p(L, n, z_j, z_v) \left[ \frac{\partial^2 u(.)}{\partial y^2} \left( (1 - \delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right)^2 - \frac{\partial u(.)}{\partial y} \frac{\partial^2 c(.)}{\partial L^2} \right] + \frac{\partial^2 u(.)}{\partial l^2} \right\} < 0
\]

(3.2)
By assumption, marginal utility is positive but diminishing with respect to disposable income and female’s leisure, so $\frac{\partial u(.)}{\partial y} > 0$, $\frac{\partial^2 u(.)}{\partial y^2} < 0$, $\frac{\partial u(.)}{\partial l} > 0$, $\frac{\partial^2 u(.)}{\partial l^2} < 0$, $\frac{\partial p(L,n,z_j,z_l,v)}{\partial L} < 0$ and $\frac{\partial c(L,n,z_j,s)}{\partial L} > 0$. The first term of equation (3.2) can be positive, negative or equal to zero if $\left[u(y,l) - u(y_p,1)\right] \leq 0 \Rightarrow u(y,l) = u(y_p,1)$. However, in fact, a woman only participates in the labour force if her utility obtained from participating in the labour market is greater than or at least equal to the utility from not entering into market work, meaning $u(y,l) \geq u(y_p,1)$, then the first term is negative. On the other hand, if the utility obtained from participating in the labour force is less than that of not participating $u(y,l) < u(y_p,1)$, the female will choose to not work in the labour market and choose to stay at home.

Since $\frac{\partial p(L,n,z_j,z_l,v)}{\partial L} < 0$, $(1-\delta)w_m(e) \geq \frac{\partial c(.)}{\partial L}$, a woman only chooses to participate in the labour market if the marginal return of being away from home is greater than or at least equal to the marginal cost of being away from home. If $(1-\delta)w_m(e) < \frac{\partial c(.)}{\partial L}$, the female will not be interested in allocating her time between market work and household work. The second term of the equation is negative if $\left[(1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L}\right] > \frac{\partial u(.)}{\partial l}/\frac{\partial u(.)}{\partial y}$. This means that the female’s wage rate after subtracting the marginal cost of the time of her absence from home is greater than the marginal rate of substitution between her leisure and disposable income.

The third term in equation (3.2) is also negative because the utility function is positive and diminishing with respect to disposable income and the female’s leisure, so $\frac{\partial u(.)}{\partial y} > 0$, $\frac{\partial^2 u(.)}{\partial y^2} < 0$ and $\frac{\partial^2 u(.)}{\partial l^2} < 0$. Furthermore, the cost function associated with the female’s absence from the household during outside work $c(L,n,z_j,s)$ is convex to the time supplied to market work $L$, so $\frac{\partial^2 c(.)}{\partial L^2} > 0$. The female’s probability of obtaining an outside job $p(L,n,z_j,z_l,v)$ is between 0 and 1. Consequently, term 3 is negative.
Assuming that SOC<0 is satisfied, an interior solution exists: the female is interested in allocating a part of her time to out-of-home work and a part of her time to household work. The two following conditions must be satisfied: 

\[ (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} > \frac{\partial u(.)}{\partial l} / \frac{\partial u(.)}{\partial y} \] and \( u(y, l) > u(y_p, 1) \). In other words, the female’s wage rate after subtracting the marginal cost of the time of her absence from home is greater than the marginal rate of substitution between her leisure and disposable income, and her utility from participating in market work is greater than her utility from not participating in labour market.

A violation of the second order condition for maximum results in a corner solution: either work full time or devote no time to work. If \( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} < \frac{\partial u(.)}{\partial l} / \frac{\partial u(.)}{\partial y} \) and \( u(y, l) < u(y_p, 1) \), the first and second terms are positive. If the sum of the first and second terms is greater than the third term, or the first term is positive and the second term is negative and the first term is larger than the sum of the second and third terms, or the first term is negative and the second term is positive and the sum of the first and third terms is smaller than the second term, then SOC for maximum is positive and not satisfied. In this case, the female decides not to participate or work full time in the labour market.

In the special case, the model assumes that the probability of being employed is independent of the level of participation by women \( dp/dL = 0 \), the first and second terms of SOC are equal to zero, so SOC becomes:

\[
p(L, n, z_j, z_i, v) \left[ \frac{\partial^2 u(.)}{\partial y^2} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right)^2 - \frac{\partial u(.) \partial^2 c(.)}{\partial y \partial L^2} \right] + \frac{\partial^2 u(.)}{\partial l^2} < 0
\]

In this case, SOC is always satisfied and an interior solution exists. The female is interested in allocating her time to market work and household work.
In the general case where \( dp / dL < 0 \), assuming that the SOC for utility maximisation is satisfied, so an interior solution exists. Comparative statics of the interior solution obtained by total differentiation of the first order condition leads to the following propositions.

### 3.4 Propositions

#### 3.4.1 Proposition 1: The effect of the number of children

The number of children negatively or positively affect female labour supply.

**Proof:**

The first order condition for maximum expected utility is:

\[
\frac{\partial E\bar{u}}{\partial L} = \frac{\partial p(.)}{\partial L} \left[ u(y,l) - u(y_p,1) \right] + p(L,n,z_j,z_i,v) \left[ \frac{\partial u(.)}{\partial y} \left( 1 - \delta \right) w_m(e) - \frac{\partial c(.)}{\partial L} \right] - \frac{\partial u(.)}{\partial l} = 0
\]

Total differentiation of the first order condition (1) with respect to \( n \) generates comparative statics of the interior solution, so we have:

\[
\frac{\partial^2 E\bar{u}}{\partial L^2} dL + \frac{\partial^2 E\bar{u}}{\partial n^2} dn = 0 \Rightarrow \frac{dL}{dn} = -\frac{\partial^2 E\bar{u}}{\partial n^2} \frac{\partial^2 E\bar{u}}{\partial L^2}
\]

where

\[
\frac{\partial^2 E\bar{u}}{\partial n^2} = \left\{ \frac{\partial p(.)}{\partial L} \left[ -\frac{\partial u(.)}{\partial y} \frac{\partial c(.)}{\partial n} + \frac{\partial p(.)}{\partial n} \left[ \frac{\partial u(.)}{\partial y} \left( 1 - \delta \right) w_m(e) - \frac{\partial c(.)}{\partial L} \right] - \frac{\partial u(.)}{\partial l} \right] \right\} + p(L,n,z_j,z_i,v) \left[ -\frac{\partial^2 u(.)}{\partial y^2} \frac{\partial c(.)}{\partial n} \left( 1 - \delta \right) w_m(e) - \frac{\partial c(.)}{\partial L} \right], \text{ and}
\]

\[
\frac{\partial^2 E\bar{u}}{\partial L^2} = SOC = \left\{ \frac{\partial^2 p(.)}{\partial L^2} \left[ u(y,l) - u(y_p,1) \right] + \frac{\partial p(.)}{\partial L} \left[ \frac{\partial u(.)}{\partial y} \left( 1 - \delta \right) w_m(e) - \frac{\partial c(.)}{\partial L} \right] - \frac{\partial u(.)}{\partial l} \right\}
\]

\[
+ p(L,n,z_j,z_i,v) \left[ \frac{\partial^2 u(.)}{\partial y^2} \left( 1 - \delta \right) w_m(e) - \frac{\partial c(.)}{\partial L} \right] - \frac{\partial u(.)}{\partial l} \left[ \frac{\partial c(.)}{\partial L} \right] - \frac{\partial^2 u(.)}{\partial l^2} \right\}
\]
Since SOC is held by the assumption $\frac{\partial^2 E\hat{u}}{\partial L^2} < 0$, so

$$
dL > 0 \text{ if } \frac{\partial p(.)}{\partial n} \left[ \frac{\partial u(.)}{\partial y} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] + p(L,n,z_j,z_i,v) \left[ -\frac{\partial^2 u(.)}{\partial y^2} \frac{\partial c(.)}{\partial n} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) \right] - \frac{\partial p(.)}{\partial L} \frac{\partial u(.)}{\partial y} \frac{\partial c(.)}{\partial n} \right] \frac{\partial^2 u(.)}{\partial y} \frac{\partial c(.)}{\partial L} > 0 \quad (3.3)
$$

(See Appendix A for proof of this equation).

By assumption, SOC is negative so $\left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) > \frac{\partial u(.)}{\partial l} / \frac{\partial u(.)}{\partial y}$ hold, $\frac{\partial p(.)}{\partial n} < 0$, so the sign of the first term in equation (3.3) is negative.

According to the assumption, $\frac{\partial u(.)}{\partial y} > 0$, $\frac{\partial^2 u(.)}{\partial y^2} < 0$, $\frac{\partial u(.)}{\partial l} > 0$, $\frac{\partial c(L,n,z_j,z_i,s)}{\partial n} > 0$, $\frac{p(L,n,z_j,z_i,v)}{\partial L} < 0$ and $p(L,n,z_j,z_i,v)$, the third term is positive. The sign of the second term depends on the sign of $\left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right)$. However, as mentioned above, a woman only chooses to participate in the labour market if the marginal return of being away from home is greater than or at least equal to the marginal cost of being away from home: $(1-\delta)w_m(e) \geq \frac{\partial c(.)}{\partial L}$. Moreover, by assumption, the SOC is satisfied that $\left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) > \frac{\partial u(.)}{\partial l} / \frac{\partial u(.)}{\partial y}$, and we have $\frac{\partial u(.)}{\partial y} > 0$; $\frac{\partial u(.)}{\partial l} > 0$, so $(1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} > 0$. Thus, the second term is positive. As a result, the relationship between the number of children that a woman has and her hours of work depends on the relative magnitudes of the three terms in equation (3.3). Thus, the number of children will have a negative or positive impact on the time that a mother devotes to market work.

A violation of SOC that $\left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) < \frac{\partial u(.)}{\partial l} / \frac{\partial u(.)}{\partial y}$ and $u(y,1) < u(y,1)$ will result into a corner solution: either work full time or devote no time to work. The
woman is not interested in the allocation of her time between outside work and housework.

Under the simplifying assumption that probability of being employed is not dependent on the amount of time worked by women $dp/dL = 0$, the third term of equation (3.3) is zero. Equation (3.3) becomes:

$$\frac{dL}{dn} = 0 \text{ if } \left\{ \frac{\partial p(.)}{\partial n} \left[ \frac{\partial u(.)}{\partial y} \left( (1-\delta) w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] \right\} > 0$$

$$(3.3')$$

If equation (3.3’) is positive or negative or equal to zero, then the relationship between the number of children and the level of participation by women will be positive or negative or equal to zero.

As stated in the general case, the first term of equation (3.3’) is negative and the second term is positive. This means that if the magnitude of the first term is greater than the second term, then the effect of children on the time supplied by the mother to market work is negative. In contrast, if the magnitude of the second term is larger than the first term, then this effect is positive.

It is easy to see a negative relationship between the number of children and the female’s participation, as well as the level of participation, in the labour market. The more children the female has, the less time she can devote to work. The more children the female has, the lower her probability of labour force participation because of the opportunity cost of women’s labour force participation and the time needed for caring for children increases. Moreover, having more children negatively affects the female’s probability of obtaining a job in the labour market as well as her hours of work.

However, the number of children does not necessarily have a negative effect on the level of FLFP. On the one hand, the presence of an additional child generates a variety of demands on household resources that increase household expenditures such as accommodation, consumption of food and the many other items needed for a new family member, and this leads to parents reducing their consumption of goods. This may push mothers to increase their labour supply to the labour market. On the other
hand, the increased number of children also increases time needed for childrearing, so women participate less in the labour market. Furthermore, according to the utility function, holding other factors constant, the number of children increases the value of non-market time, so the cost associated with the female’s absence from home during outside work increases. As a result, it reduces the family’s disposable income and so affects the female’s utility. A decrease in disposable income may lead the female to increase her supply of labour to support her family’s expenses. Therefore, the total effect of the number of children on labour supply depends on whether the magnitude of the substitution effect is larger than the income effect. However, a woman can also choose to stay at home and take care of her children and do housework. In this case, she will not have to pay for the cost of her absence from home but her family’s disposable income will only include her husband’s income, so her utility function then depends on her husband’s income.

The effect of the number of children on female labour supply in developing countries may be different from developed countries (Priebe 2011). In developed countries where families often have better financial conditions and better support from social welfare by governments, the income effect of an increase in the number of children may be smaller than that in developing countries. However, the substitution effect of the presence of an additional child may be greater than that in developing countries because most of them are nuclear families and so mothers have to spend time taking care of their children, especially young children. Further, childcare services are often more expensive and not available for very young children. Therefore, when the number of children increases, the substitution effect of time needed for childrearing may be dominant in comparison to the positive effect of increase of living costs. Consequently, the number of children has a negative impact on women’s level of employment due to the relative magnitude of the substitution effect being greater than the income effect.

However, this story may be different in developing countries. The number of children does not necessarily have a negative effect on the level of FLFP. This is especially true for developing countries where people still struggle to meet basic needs such as food, accommodation and clothes. The presence of an additional child may encourage mothers to work more to support their family’s financial needs and this leads to greater labour supply, plus women in developing countries may find childcare support from
their parents easier to obtain, and so the time needed for taking care of their children may be decreased. In this case, the income effect of having more children on labour supply may be greater than the substitution effect. Therefore, the number of children may have a positive or negative impact on women’s level of employment depending on the relative magnitude of the substitution and income effects.

3.4.2 Proposition 2: The effect of grandparents
Potential childcare provision by other extended family members have a positive effect on female labour supply.

Proof:

\[
\frac{\partial^2 E\tilde{u}}{\partial L^2} dL + \frac{\partial^2 E\tilde{u}}{\partial s^2} ds = 0 \Rightarrow \frac{dL}{ds} = -\frac{\partial^2 E\tilde{u}}{\partial s^2} \frac{\partial s}{\partial L^2}
\]

Since SOC is held by the assumption, \(\frac{\partial^2 E\tilde{u}}{\partial s^2} \frac{\partial s}{\partial L^2} < 0\)

\[
\Rightarrow \frac{dL}{ds} > 0 \text{ if } \left\{ -\frac{\partial p(\cdot)}{\partial L} \frac{\partial u(\cdot)}{\partial y} \frac{\partial c(\cdot)}{\partial s} - p(L,n,z_j,z_i,v) \frac{\partial^2 u(\cdot)}{\partial y^2} \frac{\partial c(\cdot)}{\partial s} \left( (1-\bar{\delta})w_m(e) - \frac{\partial c(\cdot)}{\partial L} \right) \right\} > 0
\]

\(\text{(3.4)}\)

(See Appendix A for proof of this equation).

According to the assumption, \(\frac{\partial p(L,n,z_j,z_i,v)}{\partial L} < 0, \frac{\partial u(\cdot)}{\partial y} > 0, \frac{\partial^2 u(\cdot)}{\partial y^2} < 0, \) and \(\frac{\partial c(L,n,z_j,s)}{\partial s} < 0\), the first term of equation (3.4) is negative and the second term is positive if \( (1-\bar{\delta})w_m(e) - \frac{\partial c(\cdot)}{\partial L} < 0 \). When the second term is positive and greater than the first term of equation (3.4), the other family members’ childcare support, grandparents in particular, positively affects female labour supply.

For the special case where the probability of being employed is independent of the level of participation by women \( dp/dL = 0 \), we can see the relationship between potential parental childcare provision and the level of participation in the labour market by women only depends on the sign of the second term in equation (3.4).
\[
\frac{dL}{ds} > 0 \text{ if } \left\{ -p(L,n,z_j,z_i,v) \frac{\partial^2 u(.)}{\partial y^2} \frac{\partial c(.)}{\partial s} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) \right\} > 0
\]

\[
\Rightarrow \frac{dL}{ds} < 0 \text{ if } \left\{ (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right\} < 0 \Rightarrow \frac{dL}{ds} > 0 \text{ if } (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} > 0.
\]

It means that only when the marginal cost of being away exceeds the marginal return of being away from home, does the support given by parents have a positive effect on female labour supply.

In this case, the parental childcare support would have a direct effect on the cost associated with the female’s absence from the household during out-of-home work. If the marginal cost of being away from home exceeds the marginal return of being away from home, that is \( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} < 0 \), the optimal condition for the maximum of the female’s utility is violated. However, since the parental care provision for the female’s children would have a direct effect on the marginal cost of being away from home, the parental support directly reduces the marginal cost of being away. Therefore, the parental support would encourage the female to devote more time to outside work.

It is clear that the help of grandparents or other extended family members in providing childcare significantly reduces the cost associated with the female’s absence from the household during work in the labour market. In addition, this help in taking care of children and housework can reduce the burden on women, so the female will have more time to devote to work. Hence, the family’s disposable income and the female’s utility increases. The female will have a greater probability of participating in work outside the home as well as potentially increasing her working hours.

3.4.3 Proposition 3: The effect of wage discrimination
Gender wage discrimination has a negative effect on female labour supply.

Proof:

\[
\frac{dL}{d\delta} = -\frac{\partial^2 E\tilde{u} / \partial \delta^2}{\partial^2 E\tilde{u} / \partial L^2}
\]
The gender wage differential comes from the demand side of the economy. It arises from the statistical and taste discrimination of employers. Females are often offered lower wage rates compared to males because employers assume that females have lower productivity and an expected period of childbearing and childrearing. Also, employers often prefer not to hire mothers with young children because they assume that these mothers will have reduced efforts in the workplace and so will lower productivity.

From the utility function, as there is wage discrimination against females \( \delta > 0 \), when an increase in wage discrimination occurs, the female’s earnings will be lower than males. The female may feel that her participation in the labour market is not as efficient as her husband’s participation. In order to obtain the highest efficiency in the household’s resource allocation, in this case the mother may reduce her working hours.
or choose to be out of the labour market, especially if her husband has a good job and high income which can support her family. As a result, her time to devote to outside work will be decreased or she may spend no time in the labour market. Therefore, gender wage discrimination has a negative effect on female labour supply.

Wage discrimination is more widespread among developing countries where anti-discrimination laws have not been enforced, especially in Asian countries. For example, according to the International Labour Organisation, Vietnam is among the few countries in the world where the gender pay gap has been widening. Data from the GSO in 2011(GSO 2011) showed that women’s earnings are 13% less than men.

3.4.4 Proposition 4: The effect of partner’s income
A partner’s income have a negative effect on the time supplied by the female to outside work.

Proof:

\[
\frac{dL}{dy_p} = -\frac{\partial^2 E\bar{u} / \partial y_p^2}{\partial^2 E\bar{u} / \partial L^2}
\]

\[
\frac{dL}{dy_p} > 0 \text{ if } \frac{\partial p(.)}{\partial L} \left( \frac{\partial u(y,l)}{\partial y} - \frac{\partial u(y_p,1)}{\partial y} \right) + p(L,n,z_j,z_v,v) \left[ \frac{\partial^2 u(.)}{\partial y^2} \left( 1 - \delta \right) w_m(e) - \frac{\partial c(.)}{\partial L} \right] > 0 \quad (3.6)
\]

(See Appendix A for proof of this equation).

Since the family’s disposable income including female income is greater than that not including female income \( y > y_p \) and the marginal utility function is positive but diminishing with respect to income and female’s leisure, the marginal utility of participating in the labour market is smaller than the marginal utility of not participating in market work: \( \frac{\partial u(y,l)}{\partial y} < \frac{\partial u(y_p,1)}{\partial y} \). Hence, the first term of (3.6) is positive.
Figure 3-2 Marginal utility function of female labour force participation

Since \( p(L, n, z, j, z_i, v) \) is between 0 and 1, \( \frac{\partial^2 u(.)}{\partial y^2} < 0 \) and, as mentioned above, the female is not concerned about the allocation of her time in the labour market if the marginal cost from being away home exceeds the marginal return of being away from home, the female only chooses to work in the labour market if the marginal return from being away home is greater than the marginal cost of being away from home, so \( \frac{(1-\delta)w_m(d) - \hat{c}(j)}{\partial L} > 0 \). The second term is negative. The sign of \( \frac{dL}{dy_p} \) depends on the relative magnitude of the first and second terms. If the magnitude of the second term is greater than the first term, then the husband’s income will negatively affect the time supplied by the female to outside work.

An increase in the husband’s income will increase the family’s disposable income, holding other factors constant, and therefore this will increase utility. In some cases, the husband prefer his wife to have more time for housework and taking care of their family and to reduce her time for outside work, particularly if these families do not face financial difficulties. The wife also feel that it is better and more efficient for her family if she spends more time at home and takes care of her family instead of working outside. As a result, the female’s labour supply will decrease. However, some women like participating in activities outside of the home and she may prefer working outside to doing housework even though her husband may have high earnings. In these cases,
the female does not reduce her time to market work and so the partner’s income does not affect the labour supply of the female.

In the special case, the model assumes that the probability of being employed is independent of the level of participation by women, the first term will be zero, so the sign of $dL/dy_p$ depends only on the sign of the second term:

$$
\frac{dL}{dy_p} > 0 \quad \text{if} \quad p(L,n,z_j,z_i,v) \left[ \frac{\partial^2 u(.)}{\partial y^2} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) \right] > 0
$$

Since $p(L,n,z_j,z_i,v)$ is between 0 and 1, $\frac{\partial^2 u(.)}{\partial y^2} < 0$, and $\left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) > 0$, the sign of $dL/dy_p$ is negative. In this special case, the partner’s income always discourages the female’s labour supply. In other words, an increase in partner’s income decreases the time that a female supplies to outside work.

3.4.5 Proposition 5: The effect of the youngest child’s age
A greater age of the youngest child have a positive impact on the labour supply of mothers.

Proof:

$$
\frac{dL}{dz_j} = \frac{\partial^2 E\hat{u}}{\partial z_j \partial L} - \frac{\partial^2 E\hat{u}}{\partial L^2}
$$

$$
\frac{dL}{dz_j} > 0 \quad \text{if} \quad \left[ \frac{\partial p(.)}{\partial L} \left[ \frac{\partial u(.)}{\partial y} - \frac{\partial c(.)}{\partial z_j} \right] + \frac{\partial p(.)}{\partial y} \left[ \frac{\partial u(.)}{\partial y} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] \right] > 0
$$

$$
\left[ \frac{\partial^2 u(.)}{\partial y^2} \left( - \frac{\partial c(.)}{\partial z_j} \right) \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) \right] < 0
$$

(3.7)

(See Appendix A for proof of this equation).

The first term of equation (3.7) is negative because $\frac{\partial p(L,n,z_j,z_i,v)}{\partial L} < 0$, $\frac{\partial u(.)}{\partial y} > 0$ and $\frac{\partial c(L,n,z_j,s)}{\partial z_j} < 0$. The second term is positive because
\[ \frac{\partial p(L,n,z_j,z_i,v)}{\partial z_j} > 0 \text{ and } \left[ \frac{\partial u(.)}{\partial y} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] > 0 \] by the assumption of the existence of a utility maximisation problem (SOC is negative). The third term is negative because \[ (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} > 0; \] \[ p(L,n,z_j,z_i,v) \text{ is between 0 and 1; and } \frac{\partial^2 u(.)}{\partial y^2} < 0. \] Therefore, the sign of \[ \frac{dL}{dz_j} \] depends on the magnitude of these three terms.

Under the simplifying assumption \( dp/dL = 0 \), the sign of \( dL/dz_j \) only depends on the second and third terms:

\[
\frac{dL}{dz_j} = \left[ \frac{\partial p(.)}{\partial z_j} \left[ \frac{\partial u(.)}{\partial y} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] \right] + p(L,n,z_j,z_i,v) \left[ \frac{\partial^2 u(.)}{\partial y^2} \left( \frac{\partial c(.)}{\partial z_j} \right) \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) \right]
\]

(3.7’)

The sign of \( dL/dz_j \) depends on the magnitude of these two terms. As mentioned in the general case, the second term is positive and the third term is negative, if the second term is greater than the third term in equation (3.7’), then \( dL/dz_j \) is positive. It means that the effect of the youngest child’s age on the labour supply of the mother is positive if the second term is greater than the third term.

According to the model, the age of the youngest child affects the labour supply of mothers on both the supply and demand sides, holding all other factors constant. On the one hand, the age of the youngest child affects the cost associated with the absence of mothers from home during the time she is working in the labour market. The older the youngest child, the lower the cost associated with the female’s absence from the household during out-of-home work, and so the lower the opportunity costs of participating in the labour market. Subsequently, the total disposable income of the family increases. Furthermore, as the youngest child becomes older, the mother is freed up from spending time taking care of them. In this case, the mother may spend more time on market work, reducing her time with the children, so her labour supply increases. In addition, as children become older, expenditures for children such as
schooling, food and health also increase and the mother may be encouraged to work more to support her family expenditures, especially in the case where her partner’s income cannot cover the family’s needs.

On the other hand, from the demand aspect, the greater age of the youngest child also positively affect the mother’s probability of obtaining an outside full-time job. When the youngest child becomes older, the mother may also be out of the average reproductive age range of women. She has more opportunities to seek paid work as well as having more time to devote to outside work. Thus, the older the youngest child’s age, the more time the female can supply to the labour market in terms of both supply and demand aspects of the economy.

3.4.6 Proposition 6: The effect of the female’s age
The female’s age have positive or negative effects on the time supplied by the female to outside work depending on whether her age is above or below the average reproductive age range.

Proof:

\[
\frac{dL}{dz_i} = -\frac{\partial^2 E\hat{u}}{\partial z_i^2} - \frac{\partial^2 E\hat{u}}{\partial L^2}
\]

\[
\frac{dL}{dz_i} \geq 0 \text{ if } \left\{ \frac{\partial p(.)}{\partial z_i} \left[ \frac{\partial u(.)}{\partial y} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] \right\} \geq 0 \tag{3.8}
\]

(See Appendix A for proof of this equation).

By assumption, we have \[ \left[ \frac{\hat{u}(.)}{\partial y} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\hat{u}(.)}{\partial l} \right] > 0 \]. The sign of \( \frac{dL}{dz_i} \) depends on \( \frac{\partial p(.)}{\partial z_i} \), that is if \( \frac{\partial p(.)}{\partial z_i} > 0 \), then \( \frac{dL}{dz_i} > 0 \). According to the assumption, \( \frac{\partial p(L,n,z_j,z_i,\nu)}{\partial z_i} > 0 \) as \( z_i > z_j \), when the female reaches an age that is higher than the average reproductive age range of women, her age has a positive effect on her probability of obtaining an outside job: \( \frac{\partial p(L,n,z_j,z_i,\nu)}{\partial z_i} > 0 \) as \( z_i > z_j \), then the
sign of \( \frac{dL}{dz_i} \) is positive. At this time, the female would be freed from her children and she is more likely to devote more time to outside work as her age is far from the average reproductive age range of women. Thus, her time to devote to outside work increases with her age.

However, it is also assumed that when the age of the female is below the the average reproductive age range, i.e. the older the female’s age is, the lower the probability of obtaining an out-of-home job is: \( \frac{\partial p(L, n, z_j, z_i, v)}{\partial z_i} < 0 \) as \( z_i < z_r \), then the sign of \( \frac{dL}{dz_i} \) is negative. This implies that when the female is expected to be busier with her children, she is likely to devote less time to outside work and more time to housework as her age approaches the average reproductive age range of women. Thus, her time to devote to outside work decreases with her age.

The special case where the probability of being employed is independent of the level of participation by women is not different from the general case.

**3.4.7 Proposition 7: The effect of unemployment rate**

The unemployment rate has a negative effect on the amount of time that the female devotes to the labour market.

Proof:

\[
\frac{dL}{dv} > 0 \text{ if } \frac{\partial p(.)}{\partial v} \left[ \frac{\partial u(.)}{\partial y} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] < 0 \quad (3.9)
\]

(See Appendix A for proof of this equation).

According to the assumption of the model \( \frac{\partial p(.)}{\partial v} < 0 \), and

\[
\left[ \frac{\partial u(.)}{\partial y} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] > 0 \text{ is held by the assumption of SOC for maximum expected utility, then } \Rightarrow \frac{dL}{dv} < 0 . \text{ This means that the unemployment rate has negative effects on the time supplied by a female to outside work.}
The special case where the probability of being employed is independent of the level of participation by women \((dp/dL = 0)\) is not different from the general case.

An economic recession or economic crisis can cause high unemployment rates. Demand for labour decreases while many people are looking for a job. This reduces individual employment opportunities as well as the amount of working time available. Furthermore, when the economy is at a high unemployment rate, wage rates are offered at a lower level. As the expected wage falls, individuals who are looking for a job may find that spending time at home is more productive than spending time looking for a job. Conditional on the desire for labour force participation, individuals also find that employment opportunities are limited in the stage of high unemployment. Individuals who would like to enter the labour market may become discouraged by the low wage rate.

On the other hand, in the same way, a high unemployment rate is often caused by an economic recession or economic crisis. In order to face economic recession, employers need to reduce production costs in some way. Employers may cut costs by decreasing the wages of existing workers but this is difficult because wage rate are often set at the time of signing a labour contract. Employers may choose to reduce production costs by requiring existing employees to reduce their working number of hours due to a decrease in business activities and labour demand in the stage of economic recession. As a consequence, many workers are faced with an involuntary unemployed status or their working hours are reduced. Thus, again, the high unemployment rate has a discouraging effect on labour supply.

### 3.4.8 Proposition 8: The effect of female educational attainment

Female education may have a positive or negative effect on women’s participation level in the labour market.

Proof:

\[
\frac{dL}{de} = -\frac{\partial^2 E u / \partial e^2}{\partial^2 E u / \partial L^2}
\]
\[
\frac{dL}{de} > 0 \quad \text{if} \quad \left\{ \begin{array}{l}
\frac{\partial p(.)}{\partial L} \frac{\partial u(.)}{\partial y} (1-\delta)w_m'(e)L \\
+ p(.) (1-\delta)w_m'(e) \left[ \frac{\partial^2 u(.)}{\partial y^2} L \left( (1-\delta)w_m'(e) - \frac{\partial c(.)}{\partial L} \right) + \frac{\partial u(.)}{\partial y} \right] \end{array} \right\} > 0
\] 

(3.10)

(See Appendix A for proof of this equation).

Since \(0 \leq \delta < 1, w_m > 0, 0 \leq L < 1\), the first term is negative. The second term depends on the sign of \( \left[ \frac{\partial^2 u(.)}{\partial y^2} L \left( (1-\delta)w_m'(e) - \frac{\partial c(.)}{\partial L} \right) + \frac{\partial u(.)}{\partial y} \right] \). The second term is positive if

\[
\left( (1-\delta)w_m'(e) - \frac{\partial c(.)}{\partial L} \right) > - \frac{\partial u(.)}{\partial y} / L \frac{\partial^2 u(.)}{\partial y^2}.
\]

If the second term is positive and greater than the first one, then \( \frac{dL}{de} > 0 \). This means that female education has positive impacts on the level of their participation in the labour market. If the second term is positive and smaller than the first term, then \( \frac{dL}{de} < 0 \) meaning that female educational attainment has negative impacts on labour supply.

In the special case, the model assumes that \( dp/dL = 0 \), so the first term is equal to 0 and the sign of \( \frac{dL}{de} \) only depends on the second term:

\[
\frac{dL}{de} > 0 \quad \text{if} \quad \left\{ \frac{\partial^2 u(.)}{\partial y^2} L \left( (1-\delta)w_m'(e) - \frac{\partial c(.)}{\partial L} \right) + \frac{\partial u(.)}{\partial y} \right\} > 0
\]

If \( \left( (1-\delta)w_m'(e) - \frac{\partial c(.)}{\partial L} \right) < 0 \), then \( \left[ \frac{\partial^2 u(.)}{\partial y^2} L \left( (1-\delta)w_m'(e) - \frac{\partial c(.)}{\partial L} \right) + \frac{\partial u(.)}{\partial y} \right] > 0 \)

\[\Rightarrow \frac{dL}{de} > 0\]

This means that if the marginal cost of being away from home is greater than the marginal return of being away from home, female education has a positive effect on participation level in the labour market. However, this violates the SOC of the utility maximisation problem that requires the marginal return of being away from home to be greater than the marginal cost. Thus, the condition for a woman to be concerned about the allocation of her time between outside work and housework is that SOC must be
satisfied, so the relationship between education and the time supplied to out-of-home work is positive if
\[
(1 - \delta)w_m(e) - \frac{\partial c(.)}{\partial L} > \frac{\partial u(.)}{\partial y} / L \frac{\partial^2 u(.)}{\partial y^2}
\]

A woman’s educational level can directly affect the probability of participation and the amount of time devoted by the woman to outside work as well as her earnings. A female with a university degree will have more opportunities to be employed and can get a job with a good salary. As a result, her family’s total income and her utility will increase, holding other factors unchanged. Moreover, according to human capital investment theory, when there is more investment in education, a woman can expect to spend a longer period in the labour market during her life. In other words, female educational attainment has a positive impact on a female’s decision to work and therefore her labour supply (substitution effect).

In contrast, female education may have a negative effect on women’s participation levels in outside work. As women gain greater educational attainment, they can access higher incomes in comparison to women with lower educational levels. Therefore, they may not need to work as much and may prefer to spend more time with their family to balance their work and family life (income effect). In contrast, women with no education or lower education may have to work harder and for a longer time because of their restricted opportunities in employment. Furthermore, more highly educated women are more likely to get marriage with higher educated and income men. So these women also tend to be less likely to work and more likely to spend more time for children because of income effects. Therefore, the total effect of the educational attainment on labour supply depends on whether the magnitude of the substitution effect is larger than the income effect.

3.5 Conclusion and testable hypotheses

*Conclusion

In summary, this chapter developed a broad conceptual framework analysing various determinants of FLFP from both supply and demand sides. The conceptual framework would lead to hypotheses about the participation of women in out-of-home work in both categories quantitatively. The list of such derived hypotheses on determinants of FLFP include supply-side factors such as number and age of dependent children, partner’s
income, potential assistance from extended family members, the youngest child’s age and the female’s age and educational attainments on female labour supply, as well as demand-side factors such as gender-wage differentials and unemployment rates.

The theoretical framework predicts eight hypotheses, including six derived hypotheses on determinants of FLFP from supply side and two hypotheses from demand side. From supply side, first, the conceptual framework predicts that the number of children have positive or negative impact on women’s level of participation in market work depending on the relative magnitude of substitution and income effect. Second, the partner’s income is predicted to have both positive and negative impacts on the time devoted by the female to labour market. Third, parents’ and other extended family members’ potential care provision for the female’s children increases female labour supply. Forth, the age of the youngest child also affects positively and negatively the labour supply of mothers. Fifth, the age of the female herself is predicted as an important factor affecting her labour supply. The female is likely to devote less time to outside work and more time housework as her age approaches the average reproductive age range of women. In contrast, she is more likely to devote more time to market work as her age is far away from the average reproductive age range of women. The last hypothesis from supply side is that female education plays a very important role determining women’s labour supply. the female’s educational attainment has both positive and negative impacts on the time supplied by her to the labour market depending on whether substitution effect is larger than income effect.

From demand side, first, the theoretical framework predicts that gender wage discrimination has negative effects on the level of participation of women in the out-of-home work. Second, the unemployment rate in the economy also plays an significant role in reducing the amount of time that the female devotes to the labour market.

Because the date is not available, the study cannot test all these hypotheses derived from the conceptual framework, only the hypothesis on the relationship between fertility and female labour supply will be tested in the next empirical chapter.

*Testable hypotheses
In view of data availability, observations on the amount of time spent at work outside of the household are not available. It is only observed whether the female was involved in work outside the home. In this case, the empirical analyses only focus on the choice of working or not working in the labour market, which is a categorised variable taking the value 1 or 0. Furthermore, the study also cannot test all the hypotheses derived from the conceptual framework due to the limited availability of data, so only the hypothesis on the relationship between fertility and FLFP will be tested in the next empirical chapter. This study will focus on the neglected context of developing countries, Vietnam in particular.

Using the new conceptual framework, the next chapter will conduct an empirical study on the effect of fertility on FLFP in Vietnam. The analysis will be based on a sample of women aged 18-35 who have two or more children in Vietnam Population and Housing Census Survey. The empirical study will test the hypothesis as to whether having a third child has a negative or positive effect on FLFP among Vietnamese families with two or more children, and whether this effect is different between different groups of women.

The number of children (having a third child) is an endogenous variable because a mother’s labour supply decisions and the number of children are jointly determined, and therefore the number of children may be influenced by labour supply decisions or by other variables in the labour supply equation. It may also be affected by omitted variables (e.g. different preferences for different groups of women) that affect both fertility and FLFP. Therefore, the IV methodology will be used to address the potential endogeneity problem arising from the causal relationship between fertility and FLFP of this testable hypothesis. In order to test this causal relationship, I need to use instrumental variables that are correlated with having more than two children, but uncorrelated with error terms in the labour force participation equation. To seek appropriate instruments, the analysis is based on a common observed symptom among Vietnamese families, that is the dominant son preference (Haughton and Haughton 1995; Bélanger 2002). This means that couples are more likely to have a third child if their first two children are girls. Therefore, the first two girls can be a valid instrument for having more than two children. An unplanned fertility (twins at second birth) is also used as an IV because the birth of twins is a fertility shock and a natural event.
The study in the next chapter will empirically test the causal effect of fertility on FLFP using the IV methodology and the expected result will be a negative effect of the presence of a third child on the female’s probability of labour force participation.
Chapter 4. The impact of fertility on female labour force participation: evidence from Vietnam

4.1 Introduction

The relationship between fertility and female labour supply is of longstanding interest to many scholars in both theoretical and empirical literature. However, the way an increase in the number of children affects the labour supply behaviour of women is difficult to predict. On one hand, the presence of an additional child will increase household work and opportunity costs of participating in the labour market, so this may reduce the net benefit from female labour supply. On the other hand, an increase in the number of children may increase financial difficulties for households because of increased need of consumption, and this can motivate the labour force participation of women. The magnitude of the effect of fertility on female labour supply is different across different countries. An analysis and understanding of the effect of fertility on labour force participation rates may help explain the post-war increasing trends in the FLFP rate in most countries in the last century. It may also help to explain different trends in labour supply in different countries.

The majority of empirical studies of this subject find a negative impact of fertility on FLFP. However, as shown in the survey of previous literature, Willis (1987) notes that the interpretation of this relationship is still not clear, it is complicated and ambiguous due to the endogeneity of fertility. The author states “… it has proven difficult to find enough well-measured exogenous variables to permit cause and effect relationships to be extracted from correlations among factors such as the delay of marriage, decline of childbearing, growth of divorce, and increased female labour force participation with aggregate or even micro level data” (p.74). In theory, it is believed that fertility and labour supply are jointly determined because fertility is a choice variable that may be influenced by FLFP decisions. In addition, female labour supply is likely due to be affected by omitted factors such as heterogeneous preferences among different groups of women that may also influence fertility (Jacobsen et al. 1999). Thus, observed negative relationships between fertility and FLFP may be spurious and the negative

---

10 Browning (1992) also writes when analysing household economic behaviour: “we have a number of robust correlations but there are very few credible inferences that can be drawn from them”.

11 Also see Schultz (1978) and Goldin (1994).
correlations found may be misleading and cannot be interpreted as evidence of causal impacts (Jacobsen et al. 1999; Cristia 2008).

Many studies have attempted to disentangle the causal mechanisms linking fertility and FLFP. They have exploited exogenous variations in family size to identify the causal relationship between the number of children and the labour supply. For example, several studies use twins at first birth (Rosenzweig and Wolpin 1980b; Bronars and Grogger 1994; Gangadharan et al. 1996; Jacobsen et al. 1999). These studies find that twins at first birth have a negative effect on labour force participation of women. Although twins at first birth is a good IV for fertility, occurrences of twins are rarely observed and studies are often based only on very small samples.

Angrist and Evans (1998) first propose a source of exogenous variations based on the sex composition of the first two children. They exploit a widely observed phenomenon of parental preference for a mixed sex composition of children in Western society and, in particular, in the United States. Parents whose the first two children being of same sex are more likely to have a third child than those of mixed sex siblings. The authors also use twins at second birth as an instrument and compare this with the case of a mixed sex composition to estimate the consequences of moving from the second to the third child on female labour supply and both show that children have negative effects on female labour supply.

This chapter will exploit the identification strategy of Angrist and Evans (1998) that uses the sex composition of the first two children to study the effect of fertility on FLFP in Vietnam but it is different to Angrist and Evans (1998) is that this study uses another instrument which is suitable with Vietnamese parents’ preference of sex composition. The study will analyse the causal impact of having more than two children on women’s labour force participation among families with two or more children in Vietnam, a developing country where the socioeconomic environment is very different from Western countries. The choice of mixed sex siblings as an IV in Angrist and Evans (1998) study is motivated by the fact that in the US, parents prefer a balanced sex composition, so an unbalanced sex composition of children is a good instrument for having more than two children. However, Vietnam is an Asian society where a son preference exists and parents prefer a male dominated sex composition of children to a mixed sex composition (Haughton and Haughton 1995). Thus, the present study
hypothesizes that the first two children being girls should be a good instrument to predict the probability of having a third child, or the total number of children, better than a balanced sex composition. If parents have two girls consecutively, they are more likely to have a third child than those with two boys or those with one girl and one boy. The study also uses twins at second birth as an instrument for having a third child. Moreover, this chapter also investigates heterogeneity in the impact of children on women’s labour force participation, depending on different educational levels of mothers and fathers, living areas and access to childcare support from grandparents.

The rest of this chapter is organised as follows: section 2 describes the data, sample construction and descriptive statistics; section 3 presents construction of key variables; section 4 analyses the effect of sex composition of children on fertility and validity of instrumental variables; section 5 describes the econometric methodology; section 6 presents empirical results and discussion; section 7 analyses heterogeneity in the impact of children on female labour supply; and section 8 provides concluding remarks.

4.2 Data

4.2.1 Vietnam population, female labour force, and fertility rate

Along with the increase in education level and the worldwide decline in fertility, female labour force participation (FLFP) rates have increased dramatically, especially in the last few decades. Although there has been a continuous decline trend in fertility and a remarkable increase in education levels in developing countries, female labour force participation has been constant or declined in some African and most Asian countries over the last three decades (Priebe, 2011). The trend is also seen in Vietnam (see Figure 4.1).
Vietnam is a developing country in Southeast Asia with a population of 86 million in 2009 and 89 million in 2011 (OECD 2013). At the point of the 2009 census, the country had 49.2 million people aged 15 and older who belonged to the labour force, accounting for 57.3% of the entire population, in which women shared a smaller proportion than men (48% female and 52% male). By 2009, 29.6 percent of the population was living in urban areas compared to 23.7 percent in 1999. Labour force participation rate is defined as the proportion of the population accounted for by labour force to the population aged 15 and over. In 2009, 76.5% of total of 64.3 million people aged 15 and over participated in the labour force, in which 81.8% men compared to 71.4% women. Labour force participation rate varies between urban and rural areas as well as across regions with a higher participation rate for the population in rural areas. The labour force participation rate in 2009 in rural areas was 14 percentage points higher than for urban areas (80.6% compared to 67.1%), in which the differential is larger among women than men (see Table 4.1). While 60.4 percent of urban women participated in the labour force, this figure for women in rural areas was 76.3 percent.
Women’s participation in the labour force has improved and one of the key national targets is to reduce gender gaps in labour market, and employment. The females’ participation rate in the labour market has accounted for nearly a half of the labour force and contributed considerably to economic growth and development in Vietnam. However, as with many other Asian countries, female labour force participation rate has been decreasing during the last two decades 1990-2010 (see Figure 4.1). In addition, the gender pay gap has been widening in the Vietnamese labour market.

Over past years, many programmes creating job opportunities for women have been implemented such as the project named “Providing vocational training, creating jobs for women during the period 2010-2015” has achieved certain results in improving women’s competitiveness in the labour market, helped them find better paid jobs. However, gender inequality still persists in the labour market with the labour force participation of men is higher than women because men often have better access to decent employment opportunities. According to the millennium development goals report 2013 of Ministry of Planning and Investment (MPI), in 2011, women’s labour participation rate was at 72.6 percent, which is 9.1 percent lower than men. Women account for a high proportion in the informal sector with limited access to skill development and training. During the period 2009-2011, women’s participation in the wage employment labour force is 20 percent lower than men’s participation (MPI 2013).

Unlike the case of developed countries, tradition based expectations, nonexistence of enforced antidiscrimination laws, and the continuing importance of extended family ties still significantly affect women’s behaviour of labour force participation, especially in Asian countries such as China, Thailand, Korea. These characteristics are also present in Vietnam where tradition based expectations such as son preference (Haughton and

---

**Table 4.1 - Labour force participation rates by sex, urban/rural areas, 2009 (%)**

<table>
<thead>
<tr>
<th>Residence region</th>
<th>Total</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entire country</td>
<td>76.5</td>
<td>81.8</td>
<td>71.4</td>
</tr>
<tr>
<td>Urban</td>
<td>67.1</td>
<td>74.4</td>
<td>60.4</td>
</tr>
<tr>
<td>Rural</td>
<td>80.6</td>
<td>85.0</td>
<td>76.3</td>
</tr>
</tbody>
</table>

Source: VPHC, 2009
Haughton, 1995; Bélanger, 2002) and extended family as a source of social and emotional support (Hirschman and Vu, 1996) still plays a role.

Vietnam has experienced a continuous decrease in fertility rates, especially in the period of 1999-2009 and female education has achieved remarkable progress in both quantity and quality. Total fertility rate (TFR) decreased rapidly from 2.33 children per women in 1999 to 2.03 children per women in 2009 (see Figure 4.2). In general, TFR in urban areas was lower than in rural areas. In 2009, TFR in urban areas was 1.81 children per women, while the figure for rural areas was 2.14 children per women. This differential may be due to urban areas have easier access to information, medical facilities that help prevent unwanted pregnancies. It may also be due to urban areas have better living conditions compared to rural areas. The rapid fall in TFR had important contributions to reducing population growth.

A decline in fertility also accompanied an increase in income. While fertility rate dramatically dropped, GDP per capita increased from 1161.5 to 1770.3 USD per capita during the period from 2007 to 2016 (Worldbank 2017). This pattern is consistent with microeconomic predictions: higher income leads to a reduction in fertility and the inverse relationship of fertility and labour force participation (Becker and Lewis 1973; Willis 1973).

Figure 4.2: Total fertility rate in Vietnam, 1999-2009

Sources:
Taken together all factors affecting women’s participation in the labour force, some factors have positive effects, while other factors have negative impacts on their participation. One question is that while fertility rate in Vietnam dropped rapidly and female education also has achieved remarkable progress during past years, why women’s participation in the labour force has not changed much, even slightly decreased. Whether the decrease in fertility rate had no impacts on the participation or these impacts have been offset by other factors such as gender inequality in the labour market, gender-biased perception. Thus, assessing the determinants of women’s labour force behaviour in the specific context of a country is necessary for policy design aimed at equal opportunities in the labour market. In the chapter, the author focuses on empirical analysis the impacts of fertility on women’s participation behaviour in the labour market. The thesis uses micro data from Vietnamese Census of Population and Housing 2009.

4.2.2 Vietnam Population and Housing Census (VPHC)

The Population and Housing Censuses were conducted by the Vietnamese GSO with financial and technical support by the United Nations Population Fund (UNFPA) and other United Nations organizations. Among the United Nations agencies, the United Nations Population Fund (UNFPA) played a lead role in supporting the censuses. UNFPA provided assistance in planning the censuses, design of the questionnaires and piloting, training and monitoring processes. When the census was finished, UNFPA assisted in data analysis and dissemination.

The Population and Housing Censuses in Vietnam are conducted every 10 years for all households, with the objective of collecting basic data on population and housing for research and analysis of socioeconomic development plans. Up to now, Vietnam has implemented four Population and Housing Censuses in total since reunification in 1975.

The 2009 Vietnamese Population and Housing Census is the fourth and most recent population census and was conducted on April 1st, 2009. Compared with previous censuses, the 2009 Census had a much larger scope. Beside contents similar to previous censuses on sex, age, ethnicity, education, marital status and demographic features, etc.,
this census surveyed additional indicators on disabilities, cause of death, maternal mortality and basic amenities in the household. The size of the 2009 census sample survey was increased to 15% of the population to allow expansion of the survey contents, while the 1989 and 1999 Census sample survey sizes were only 5% and 3% respectively.

Data from the fifteen per cent Census sample survey of the national population with a national probability sample of 3,692,042 households consisting of 14,177,590 individuals was included to expand the census coverage. The size of the 2009 census sample provides data representing the district level, while the sample sizes of the 1989 and 1999 censuses provide data representing the provincial level only. In addition, Intelligent Character Recognition (ICR) technology was used for data capture instead of traditional keyboard data entry (Nakamura et al. 1979). This study will use the 2009 VPHC to utilise a large sample.

The sample is randomly selected to ensure representation of the whole population. The census is implemented taking the household as a survey unit. The household head is the main respondent and face-to-face interviews are utilised. The VPHC includes detailed information on the characteristics of individuals and households, such as demographics of household members, ethnicity, area of residence, educational background, employment status and birth history of women (having given birth, number of children ever born, alive, died, etc.) as well as information on mortality and housing (owners, source of water, main facility and appliances, etc.).

4.2.3 Sample construction
The trends in fertility and labour supply of women in different age groups can be seen in Table 4.2.

Table 4.2- Fertility and female labour force participation rate

<table>
<thead>
<tr>
<th>Sample</th>
<th>1989</th>
<th>1999</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women aged 18-35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean children ever born</td>
<td>1.53</td>
<td>1.23</td>
<td>1.12</td>
</tr>
<tr>
<td>Per cent of women with 2 or more children</td>
<td>43.56</td>
<td>37.86</td>
<td>36.43</td>
</tr>
<tr>
<td>Per cent of women worked last year</td>
<td>77.07</td>
<td>74.73</td>
<td>80.45</td>
</tr>
<tr>
<td>Number of observations</td>
<td>409,899</td>
<td>373,275</td>
<td>2,188,962</td>
</tr>
</tbody>
</table>
Women aged 36-50

<table>
<thead>
<tr>
<th></th>
<th>1999</th>
<th>2009</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean children ever born</td>
<td>4.31</td>
<td>3.19</td>
<td>2.58</td>
</tr>
<tr>
<td>Per cent with 2 or more</td>
<td>87.18</td>
<td>84.98</td>
<td>83.21</td>
</tr>
<tr>
<td>children</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per cent of women worked</td>
<td>77.66</td>
<td>78.66</td>
<td>88.56</td>
</tr>
<tr>
<td>last year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>151,447</td>
<td>202,164</td>
<td>1,374,027</td>
</tr>
</tbody>
</table>

Women aged 18-35 with 2 or more children

<table>
<thead>
<tr>
<th></th>
<th>1999</th>
<th>2009</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean children ever born</td>
<td>3.07</td>
<td>2.60</td>
<td>2.35</td>
</tr>
<tr>
<td>Per cent with more than 2</td>
<td>56.16</td>
<td>37.91</td>
<td>24.98</td>
</tr>
<tr>
<td>children</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per cent of women worked</td>
<td>81.53</td>
<td>80.80</td>
<td>88.58</td>
</tr>
<tr>
<td>last year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>178,569</td>
<td>141,316</td>
<td>797,544</td>
</tr>
</tbody>
</table>

Married women aged 18-35 with 2 or more children

<table>
<thead>
<tr>
<th></th>
<th>1999</th>
<th>2009</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean children ever born</td>
<td>3.08</td>
<td>2.60</td>
<td>2.35</td>
</tr>
<tr>
<td>Per cent with more than 2</td>
<td>56.65</td>
<td>38.08</td>
<td>25.04</td>
</tr>
<tr>
<td>children</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per cent of women worked</td>
<td>81.53</td>
<td>80.62</td>
<td>88.49</td>
</tr>
<tr>
<td>last year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>171,293</td>
<td>137,379</td>
<td>777,981</td>
</tr>
</tbody>
</table>

Notes: The 1989 and 1999 census data are from 5% and 3% of the population respectively. The 2009 data are from 15% of the population. Married samples include women married at the time of the census. It does not include single/never married, separate/divorced/spouse absent, and widowed. Percentage of women worked in the last year applied for 1989 and 1999 not for 2009. In 2009, women asked whether they worked in the last 7 days.

This table reports on labour force participation rates, the mean number of children ever have born to a woman, and the probability of having one more child among women at the ages of 18-35 and 36-50 in the VPHC 2009. The table shows substantial declines in fertility and increases in labour supply in both age groups. A similar trend can also be seen among all and married women aged 18-35 with two or more children.

The study will use a sample from the VPHC 2009 which is detailed in Table 4.3. A number of sample restrictions are applied and are shown in this table.

Table 4.3 - The sample of all women and married women with two or more children in Population and Housing Census 2009

<table>
<thead>
<tr>
<th>Restrictions</th>
<th>Remaining Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sample</td>
<td>14,177,590</td>
</tr>
<tr>
<td>Women</td>
<td>7,213,415</td>
</tr>
<tr>
<td>Aged between 18 and 35</td>
<td>2,197,393</td>
</tr>
</tbody>
</table>
Women aged 18-35 with two or more children 800,266
Consistent number of children ever born and number of children observed in the household 721,315
Not missing employment status 720,119
Not missing ethnicity status 720,119
Not missing marital status 720,096
Not missing education data 720,035
Consistent age at first birth 719,127
Not eldest children >18 714,777
Not missing relationship to householder data 714,776
Not matching error mothers to children 706,602

Married women aged 18-35 with two or more children 693,244
Not missing employment status of husbands 669,322
Not missing marital status 669,315
Not missing education data of husbands 669,296
Consistent age of husbands at first birth 669,062

Source: General Statistics Office, Vietnam
Notes: Age at first birth is the age of a woman in years when her first child was born, so age at first birth is inconsistent and dropped if a woman’s age is less than 12 at first birth. Children aged 18 and over are dropped, and women whose the number of children ever born is not the same as the number of children observed in households are also dropped.

The total sample size is 14,177,590 observations including both males and females. The sample is restricted to mothers aged 18-35, which results in 2,259,011 women aged less than 18 years old being dropped and 2,757,011 women aged older than 35 also being excluded from the sample. The reason for choosing the research sample restricted to women aged from 18 to 35 is to reduce the need to drop observations because children may have left home and lived separately from their parents if they are older than 18 and so they cannot be linked and their sex is unknown. Since the research sample includes mothers aged 18-35 with at least two children, all women with fewer than two children are dropped from the sample, resulting in 1,397,127 women being dropped. In some cases, the number of children observed in households is fewer than the number of reported births because children have died or are not living with the mother. In other cases, the number of children observed in the household may exceed the number of children ever born because adopted or step children are not identified in the dataset. So a further 78,951 observations are dropped due to inconsistent child data where mothers
whose number of children observed in the household does not match the reported number of children ever born.

The FLFP is a dummy variable which takes 1 if a woman is employed or looking for a job and 0 if she is inactive. 1,196 women with unknown or missing employment status used to construct the labour force participation variable are also excluded.

Also, observations with missing data on marital status (23), education status (61) and relationship to household head (1) are excluded. Nine hundred and eight observations of women with an inconsistent age at first birth are also dropped. In addition, 4,350 women with an eldest child aged 18 or over are dropped because children over 17 are more likely to live separately from their parents.

Finally, the study excludes 8,174 mothers who could not be matched with their children because these mothers were classified in the same subfamily numbers in households containing multiple families (see Appendix B). After applying all restrictions, a final sample of 706,602 women aged 18-35 with two or more children was obtained.

4.2.3 Descriptive statistics
Definitions of variables, instruments, means and standard deviations of these variables are provided in Table 4.4.

Table 4.4 - Sample descriptive statistics - Independent variables (families with two or more children)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean (standard deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All women sample</td>
</tr>
<tr>
<td></td>
<td>All women</td>
</tr>
<tr>
<td>Children ever born</td>
<td>2.287</td>
</tr>
<tr>
<td></td>
<td>(0.618)</td>
</tr>
<tr>
<td>More than two children (=1 if woman had more than two children)</td>
<td>0.221</td>
</tr>
<tr>
<td></td>
<td>(0.415)</td>
</tr>
<tr>
<td>Girl 1\textsuperscript{st} (=1 if first child was a girl)</td>
<td>0.502</td>
</tr>
<tr>
<td></td>
<td>(0.499)</td>
</tr>
<tr>
<td>Girl 2\textsuperscript{nd} (=1 if second child was a girl)</td>
<td>0.485</td>
</tr>
<tr>
<td></td>
<td>(0.499)</td>
</tr>
<tr>
<td>Two girls (=1 if the first two children were girls)</td>
<td>0.243</td>
</tr>
<tr>
<td></td>
<td>(0.429)</td>
</tr>
<tr>
<td>Two boys (=1 if the first two children were boys)</td>
<td>0.255</td>
</tr>
<tr>
<td></td>
<td>(0.436)</td>
</tr>
<tr>
<td>Same sex (=1 if the first two children were the</td>
<td>0.499</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
same sex) & (0.499) & (0.499) & (0.499) \\
Twin2 (=1 if second birth was twins) & 0.005 & 0.005 & 0.005 \\
& (0.070) & (0.070) & (0.070) \\
Age & 29.912 & 29.875 & 33.052 \\
& (3.740) & (3.746) & (4.806) \\
Age at first birth (mother and father’s age in years when first child was born) & 21.113 & 21.076 & 24.253 \\
& (2.838) & (2.814) & (3.993) \\
Ethnicity (=1 if the woman is Kinh, =0 if other ethnic group (minority)) & 72.1% & 71.7% & 71.7% \\
& (0.448) & (0.450) & (0.450) \\
Urban & 0.192 & 0.184 & 0.184 \\
& (0.394) & (0.388) & (0.388) \\
Co-residence with parents/in-laws & 0.095 & 0.096 & 0.096 \\
& (0.293) & (0.295) & (0.295) \\

<table>
<thead>
<tr>
<th>Regions</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>0.149</td>
<td>0.149</td>
<td>0.149</td>
</tr>
<tr>
<td>&amp; (0.357)</td>
<td>(0.356)</td>
<td>(0.356)</td>
<td></td>
</tr>
<tr>
<td>North East</td>
<td>0.209</td>
<td>0.211</td>
<td>0.211</td>
</tr>
<tr>
<td>&amp; (0.407)</td>
<td>(0.408)</td>
<td>(0.408)</td>
<td></td>
</tr>
<tr>
<td>North West</td>
<td>0.068</td>
<td>0.068</td>
<td>0.068</td>
</tr>
<tr>
<td>&amp; (0.253)</td>
<td>(0.253)</td>
<td>(0.253)</td>
<td></td>
</tr>
<tr>
<td>North Central</td>
<td>0.123</td>
<td>0.121</td>
<td>0.121</td>
</tr>
<tr>
<td>&amp; (0.328)</td>
<td>(0.327)</td>
<td>(0.327)</td>
<td></td>
</tr>
<tr>
<td>Central Coast</td>
<td>0.085</td>
<td>0.084</td>
<td>0.084</td>
</tr>
<tr>
<td>&amp; (0.278)</td>
<td>(0.278)</td>
<td>(0.278)</td>
<td></td>
</tr>
<tr>
<td>Central High</td>
<td>0.078</td>
<td>0.078</td>
<td>0.078</td>
</tr>
<tr>
<td>&amp; (0.268)</td>
<td>(0.269)</td>
<td>(0.269)</td>
<td></td>
</tr>
<tr>
<td>South East</td>
<td>0.123</td>
<td>0.121</td>
<td>0.121</td>
</tr>
<tr>
<td>&amp; (0.328)</td>
<td>(0.326)</td>
<td>(0.326)</td>
<td></td>
</tr>
<tr>
<td>Mekong</td>
<td>0.163</td>
<td>0.164</td>
<td>0.164</td>
</tr>
<tr>
<td>&amp; (0.369)</td>
<td>(0.370)</td>
<td>(0.370)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parent’s Education</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than primary</td>
<td>0.270</td>
<td>0.273</td>
<td>0.233</td>
</tr>
<tr>
<td>&amp; (0.444)</td>
<td>(0.445)</td>
<td>(0.423)</td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>0.623</td>
<td>0.627</td>
<td>0.644</td>
</tr>
<tr>
<td>&amp; (0.484)</td>
<td>(0.483)</td>
<td>(0.478)</td>
<td></td>
</tr>
<tr>
<td>Secondary</td>
<td>0.066</td>
<td>0.064</td>
<td>0.086</td>
</tr>
<tr>
<td>&amp; (0.248)</td>
<td>(0.244)</td>
<td>(0.280)</td>
<td></td>
</tr>
<tr>
<td>University</td>
<td>0.041</td>
<td>0.036</td>
<td>0.036</td>
</tr>
<tr>
<td>&amp; (0.198)</td>
<td>(0.186)</td>
<td>(0.187)</td>
<td></td>
</tr>
</tbody>
</table>

| Number of observations | 706,602 | 669,155 | 669,155 |

Notes: The all women samples include women aged 18-35 with at least two children. Married women samples include women who were married at the time of the Census.

The dependent variable \( FLFP \) is a binary variable that indicates whether women participate in the labour market. Among families with two or more children, the number of children ever born on average to a woman in 2009 is 2.28. \( Moretwo \) is an endogenous variable of interest in the labour supply equation. As can be seen from the table, 22\% of
women with two children had an additional child. IVs for having more than two children are Twogirl and Twin2. Twogirl is constructed based on the sex components of the first two children in the population of women with at least two children. About 24% of these families had children of the same sex.

Multiple second births are also used to generate instruments for measuring the movement from two children to three children. A multiple second birth Twin2 is defined as children having the same mother, the same age and the same month of birth. The mean of multiple second births in the full sample is 0.0049 for the 2009 census.

The table also describes demographic and geographic location variables, including indicators of mothers’ age, age at first birth, ethnic background, an urban indicator, geographic region and educational attainment. The average age of mothers in both the all women and married women samples is approximately 30 and the age at first birth mean is 21. There is little difference on the age mean between the all women and married women samples. This could be due to married women accounting for a high proportion of the full sample. The sample includes more than 50 ethnic groups. The majority ethnic group is Kinh, accounting for 72% in the 2009 sample. The majority of women in the sample live in rural areas, only nineteen per cent of women live in urban areas.

With respect to geographic location, the sample includes eight regional variables corresponding to eight areas: the Red River Delta, North East, North West, North Central, Central Coast, Central Highlands, South East and the Mekong River Delta. The majority of women live in the North East, the Red River Delta and the Mekong River Delta, with the largest proportion of women in the North East area.

The majority of women in the sample have a low educational attainment (primary education or lower). Over half of all women aged 18-35 with at least 2 children completed primary education. Women with secondary education level or above account for a very low proportion. Only 10.7% of women completed secondary and university levels, in which 6.6% of whom completed secondary education and only 4.1% of whom achieved university level. The proportion of women with secondary and university education is slightly lower among the married women sample.
4.3 Construction of key variables

In order to construct instruments, firstly, children need to be matched to mothers within households. There is no retrospective fertility information in the 2009 census data, only the information on the number of children ever born is available. The process of matching children to mothers is described in Appendix B. According to this method, in households with a single family, individuals who are labelled “child” in the primary relationship code are directly attached to the woman who is the household head or the spouse of the household head. In households including multiple families, matching children and mothers is more complicated. Besides using the primary relationship and detailed relationship codes, subfamily numbers (family unit number) are also used to match children and mothers.

The number of children matched to mothers (or the number of children observed in the household) may be smaller than the number of children ever born because some children might not be living with their mother but are living elsewhere or they are no longer alive at the time of the census. The number of children observed in a household may be greater than the number of given births because adopted or step children are also included and not separately identified in the survey. Thus, the study will drop all mothers where the observed number of children in the household does not match the number of children ever born.

Female labour force participation

Women’s labour force participation is a binary variable that takes two values: 1 if women take part in the labour force and 0 if not. People who are working or looking for a job are identified as being in the labour force. Based on the available data on employment status in general and employment status details, women who are employed and unemployed are classified as being in the labour force and take a value of 1. Working in agriculture, fishing and forestry or being self-employed are also included as being in the labour force. Women who are studying in school, doing housework or unable to work/disabled are not included in the labour force and data on their employment status is labelled as inactive, so the variable $FLFP$ is 0 for these women.

More than two children in the household

This variable indicates whether women have more than two children. It is constructed from the sample of families with at least two children who are currently residing with
their mothers (see Appendix B for more details of the construction of this variable). Mothers in the sample are defined as women where the reported number of children ever born is equal to the observed number of children in the household. Thus, the variable Moretwo takes a value of 1 if the mother has more than two children who are living with her who are her biological children.

Two girls in the household
This is an IV for Moretwo and is a binary variable. After matching children to mothers, the variable Twogirl (where a mother’s first two children are daughters) can be constructed from the sex of the first two children. If the mother has two daughters consecutively, Twogirl is equal to 1. Otherwise, if only one of the first two children is a girl or both of them are sons, then Twogirl is zero.

Twins at second birth
Twin2 is a dummy variable that is constructed from the year of birth and the month of birth of children in the second birth of a mother. Twins at second birth include children who have the same mother location code, the same year of birth and the same month of birth in the dataset. So if a mother has two children with the same year of birth and the same month of birth, Twin2 is equal to 1, otherwise it is 0.

Maternal age
Since Vietnamese women give birth at a younger age in comparison to American women, the study is restricted to women aged 18-35. Angrist and Evans (1998) is limited to mothers aged 21-35.

Age at first birth
Age at first birth Ageatfirst is the age of a mother when she gives birth for the first time. A number of studies have examined whether the age of a mother at first birth affects her decision to participate in the labour force. An increase in the age at first birth often relates to a decrease in fertility rates and an increase in the FLFP rate (Killingsworth and Heckman 1986; Dercon and Krishnan 2000; Genoni 2012).

In the data survey, the age at first birth of a mother is constructed by subtracting the age of her eldest child from her current age. The age of a father at first birth Ageatfirst_sp is calculated by subtracting the age of the eldest child from their father’s age. There may be some measurement errors due to inaccurate reports by respondents. In order to avoid
inconsistencies in the age of the mother at first birth or where age differences between mothers and children are implausible, all mothers and fathers with the age of less than twelve at first birth are excluded from the study.

4.4 The effect of children’s sex on fertility and validity of instrumental variables

4.4.1 The effect of children’s sex on fertility

There have been studies in the literature on parents’ preference on the sex composition of their children (Angrist and Evans 1998; Carrasco 2001; Iacovou 2001; Chun and Oh 2002; Cruces and Galiani 2007). This preference may affect parents’ childbearing decisions and the number of children they wish to have and therefore their labour supply decisions. This parental preference may vary between cultures, nations or over time. According to previous studies, a preference for mixed sex siblings is mainly found in developed countries. Families who have a mixed sex composition of children are less likely to have more children. The effect of a mixed sex composition on childbearing decisions of couples has been examined in a number of studies. Angrist and Evans (1998) examine the 1980 US census data and found that 43.2% of women with two girls or two boys had a third child, while only 37.2% of women with one boy and one girl had a third child. They also showed that the probability of having a third birth among households where the first two children have the same sex (two girls or two boys) is 5-7% higher than households with different sexes in their children. Also, Iacovou (2001) finds families where the first two children have the same sex have an 8-11% higher probability of having a third child than families with a mixed sex component in Britain. Both studies showed that there is no significant difference in the probability of having a third child between couples with two daughters and couples with two sons. This implies that a balanced sex combination is preferred in the US and Britain.

However, the story is different in Asian countries in general and Vietnam in particular. The main finding here is for a male dominated preference. Table 4.5 reports the impact of children’s sex composition on the probability of having an additional child. In the first part of the table, the study describes the fraction of women who have a second birth based on the sex of the first child among families with one or more children. The first two rows show whether a woman’s first child is a girl or boy. The third row shows the difference in the fraction of women having a second child with the first child as a girl
and the first child as a boy. The table shows that the proportion of having a second child among the couples with the firstborn child as a girl is higher than among the couples with the first child as a boy by 5%. In Angrist and Evans’ (1998) study, the likelihood of having a second child is not different between families with a daughter and families with a son as the first child.

The second part of the table describes the proportion of women who have a third birth, conditional on the sex of the first two children in families with two or more children. The first three rows describe the combination of the sexes of the first two children: one boy, one girl; two girls; or two boys. The next line shows the differences in the fraction of women having a third child between the women with two girls and with two boys. The fifth line describes differences between women with two girls and women with one girl, one boy. Data suggests that women with two girls are much more likely to have a third child in comparison to women with two boys as well as to those whose first two children are of mixed sex (one girl, one boy). The impact of two girls on the probability of mothers having more than two children is considerably large in the 2009 census. For example, about 37.7% of women with two girls have a third child, while only 15.5% of women with two boys and 16.7% of women with one girl, one boy have a third child. Women with two girls have a higher probability of having a third birth than women with two boys or women with one boy and one girl, by about 22 and 21 percentage points respectively. The results of the t test show that these differences are statistically significant at the 1 percent level.

Table 4.5 - Proportion of families with another child by the sex of existing children

<table>
<thead>
<tr>
<th>Families with one or more children</th>
<th>Sex of first child</th>
<th>Fraction of families that had another child</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>One girl (1)</td>
<td>0.624</td>
<td>0.645</td>
</tr>
<tr>
<td></td>
<td>(0.0000239)</td>
<td>(0.0000246)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>One boy (2)</td>
<td>0.573</td>
<td>0.594</td>
</tr>
<tr>
<td></td>
<td>(0.0000219)</td>
<td>(0.0000232)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Difference (1) - (2)</td>
<td>0.051</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>(0.0000057)</td>
<td>(0.0000060)</td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>1,213,230</td>
<td>1,105,653</td>
<td></td>
</tr>
</tbody>
</table>

All women | Married women |
<table>
<thead>
<tr>
<th>Families with two or more children</th>
<th>Sex of first two children</th>
<th>Fraction of families that had another child</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>One girl, one boy (1)</td>
<td>0.167 (.0000347) 0.170 (.0000359)</td>
</tr>
<tr>
<td></td>
<td>Two girls (2)</td>
<td>0.377 (.0000489) 0.386 (.0000498)</td>
</tr>
<tr>
<td></td>
<td>Two boys (3)</td>
<td>0.155 (.0000409) 0.158 (.0000430)</td>
</tr>
<tr>
<td>Difference (2) - (3)</td>
<td>0.223 (.0000357) 0.228 (.0000368)</td>
<td></td>
</tr>
<tr>
<td>Difference (2) - (1)</td>
<td>0.210 (.0000288) 0.216 (.0000296)</td>
<td></td>
</tr>
<tr>
<td>Difference (3) - (1)</td>
<td>-0.120 (.0000115) -0.125 (.0000124)</td>
<td></td>
</tr>
</tbody>
</table>

Number of observations: 706,721 669,155

Note: The samples are the same as in Table 4.4. Standard errors are reported in parentheses. The statistical significance of differences in the table is based on t-test results.

Like China, India, Korea and some other East and South Asian countries where son preference has been so persistent (Gu and Li 1994; Park and Cho 1995; Poston Jr et al. 1997; Das Gupta et al. 2003), Vietnamese families have a strong preference for sons rather than a balanced sex composition of children for economic, social, symbolic and cultural reasons, and this is different to the environments of the US in the research by Angrist and Evans (1998) and in Britain by Iacovou (2001). The relationship between the number of girls and the probability of having an additional child is also confirmed in the works of Haughton and Haughton (1995) and Bélanger (2002) in regards to Vietnam. Vietnamese families are much more likely to have a second child if the first child is a daughter. Moreover, if the first two children are daughters, households will try to have at least one son (Haughton and Haughton 1995; Bélanger 2002). However, they may not have the motivation to have additional children if their first two children are sons. Haughton and Haughton (1995) stated that the probability of having more children is very high among families with no sons yet, regardless of the given number of children, while the probability is lower among families with no daughters.

Hence, the probability of having more children increases along with the rise in the number of daughters. For families with one or more children, a first girl may be a good predictor for having a second child. For families with two or more children, the first two girls is a good instrument for whether families will have more than two children. However, the study will use two girls *Twogirl* as an instrument for an identification
strategy of having a third child in the sense that parents with the first two children being girls are more likely to have another child than those who had two sons or those who had one son and one daughter.

The study also uses variables related to unplanned fertility (twins at second birth) as an IV. Since twins are fertility shocks, these events occur randomly and so they can be used as good IVs to generate exogenous variations in fertility. It needs to be noted that the challenge of using these IVs is that it is difficult to obtain a large sample size due to occurrences rarely being observed in the survey sample. As a result, a small sample may produce estimated results that may not be very accurate. However, this study uses a large sample size from the VPHC, so limitations due to small sample size should not be an issue.

4.4.2 Validity of instrumental variable

A natural concern is that why I focus on the third child, but not on the second child or in other words why does this study not exploit the outcome of moving from the first child to the second child and use only the first child’s sex as an instrument. Although there is a difference in the likelihood of having a second child between families with a daughter and families with a son as the first child, this difference is not too large. However, there is a big difference in the probability of having a third child between families with two girls and families with two boys or families with one boy and one girl. Thus, the thesis focuses on the movement from two children to three children. If the sex composition of the first two children is randomly assigned, it will be a good predictor for having a third child.

If the sex of existing children was randomly determined by nature, the sex ratio (the number of boys born per one hundred girls) should stay around the normal level (about 106). If the sex of existing children is not randomly assigned, we cannot use the sex components of the first two children as instruments for fertility because these instruments will be correlated with error terms in the labour supply equation. This leads to violation of the exogeneity condition of a valid instrument. In order to become good instruments for having more children, the instruments created from the sex composition of children must be correlated with having more children and have no impact on the labour force participation of mothers. This can be attained if the sex of children is randomly assigned, irrespective of the preference of parents.
In some developing countries, there are concerns that the presence of son preference may cause the sex composition to no longer be randomly assigned. For example, in Korea or China, the sex ratio by birth order provides evidence of sex selection against female embryos by parental actions (Chun and Oh 2002). It may be claimed that *Twogirl* may be not randomly determined because, with current modern ultrasonic technologies, it is easy for parents to learn the sex of a fetus. If parents use the sex identification of a fetus to have sex selective abortions, the sex composition of children is no longer randomly assigned. Prenatal sex identification may cause a sex ratio imbalance at birth, as has occurred in recent years in some developing countries. However, the documents of law aspects promulgated as well as the sample data in the study rule out this concern.

First, there may be claims that son preference in Vietnam may lead parents to seek prenatal sex selection. However, a law prohibiting prenatal sex selection was introduced in 2003 by the National Assembly Standing Committee and strengthened by Government Decree No.114 issued in 2006, so sex selective abortions are illegal behaviour under Vietnamese law. Therefore, if this is a case, families who carry out such illegal activities only account for a very small proportion of the entire population.

Second, one can claim that the virtual random assignment of *Twogirl* may lead to regression estimates of fertility and labour supply on these instruments that have a causal interpretation. The random assignment of *Twogirl* can be checked by comparing the demographic characteristics of women who have girls as the first two children with those have two boys or those have one girl and one boy. Table 4.6 describes the difference in mothers’ age, age at first birth, ethnicity, urban indicator and years of schooling by *Twogirl* for the 2009 sample. In general, the differences of magnitude in means (standard errors) for mothers’ age at first birth, ethnicity, living areas and years of schooling in the 2009 sample by *Twogirl* are too small and insignificant. There is also no contrast of *Twogirl* for mothers’ demographic characteristics. None of these differences are significant at the 5% level, despite the large sample size of the 2009 census. This means that there is no difference in means or standard errors for demographic characteristics between the groups of women who have two girls and those who do not have two girls as their first children.

---

12 For example, older women are more likely to have twins (Waterhouse, 1950).
Table 4.6 - Differences in means (standard errors) of maternal demographic characteristics by Twogirl and Twin2

<table>
<thead>
<tr>
<th>Variables</th>
<th>Difference in means (standard errors)</th>
<th>By Twogirl</th>
<th>By Twin2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.0100</td>
<td>0.2820</td>
<td>(0.0103)</td>
</tr>
<tr>
<td>Age at first birth</td>
<td>-0.0096</td>
<td>0.248</td>
<td>(0.0078)</td>
</tr>
<tr>
<td>Kinh</td>
<td>-0.00074</td>
<td>0.062</td>
<td>(0.0012)</td>
</tr>
<tr>
<td>Urban</td>
<td>-0.0011</td>
<td>0.0238</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Years of schooling</td>
<td>0.082</td>
<td>-2.993</td>
<td>(0.083)</td>
</tr>
</tbody>
</table>

Notes: The samples are the same as in Table 4.4. Standard errors are reported in brackets. Kinh is the main ethnic group in Vietnam.

However, there are traditional differences among mothers with twins. The average age as well as age at first birth of mothers with twins is higher than those with no twins by about one quarter of a year. These results are also consistent with a previous well-known study by Waterhouse (1950) which reveals that twins are more likely for older women.

Third, the study also checks the sex ratio of children in the used data sample. Table 4.7 reports the sex ratio of children by birth order for the data used in this study. The sex ratio (the number of boys divided by the number of girls times 100) of the first child and second child are around 104 and 105 in the 1989 and 1999 censuses, respectively. This ratio is 106 for the 2009 census. In general, the sex ratio of children from data used in the sample is close to the natural sex ratio, and therefore there are no signs of sex selection or parental control that may affect the sex ratio.

Table 4.7 - Sex ratio by birth order

<table>
<thead>
<tr>
<th>Birth order</th>
<th>1989</th>
<th>1999</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>First birth</td>
<td>104.49</td>
<td>105.14</td>
<td>106.56</td>
</tr>
<tr>
<td>Second birth</td>
<td>104.35</td>
<td>105.16</td>
<td>105.92</td>
</tr>
<tr>
<td>Sex ratio of children under 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First birth</td>
<td>106.25</td>
<td>107.12</td>
<td>106.96</td>
</tr>
<tr>
<td>Second birth</td>
<td>103.68</td>
<td>106.03</td>
<td>106.13</td>
</tr>
</tbody>
</table>


Last, the primary school enrolment rates in Vietnam are approximately the same for boys and girls. According to Vietnam Women’s Union of Ho Chi Minh City, the primary school enrolment rate for girls is 91.5% and that for boys is 92.3% (Union 2014). This shows that there is no discrimination by parents against girls, and the son preference mainly reflects the cultural preference in Vietnam. Son preference does not lead parents to use abortion methods to control the sex of their children. Son preference does not cause sex selective abortion as is seen in Korea. In general, there is no sign of sex selective abortions despite of the existence of a son preference.

In summary, all the evidence shows that son preference mainly reflects cultural preferences among Vietnamese people. It does not affect the sex selection, or the sex ratio, of children. At least the first and second order births in this study are randomly determined, and therefore this study assumes that Twogirl is randomly assigned so it meets the exogeneity condition of a valid instrument. Further, as mentioned above, the first two girls are correlated with having more than two children because of the existence of the parental son preference in Vietnam, and therefore this instrument also meets the relevance condition of a valid instrument. This relevance condition of Twogirl continues to be considered in the next section.

4.5 Empirical methodology and estimation

This section describes econometric models and the estimation method. This study can only observe whether the female was involved in work outside the home. In this case, the empirical analyses only focus on the choice of working or not working in the labour market which is a categorical variable taking the value 1 or 0.

4.5.1 OLS models

To investigate the direct impact of fertility on FLFP, this study will start with the simplest model by treating fertility as an exogenous variable. This can be obtained by estimating the OLS model in the following equation:

\[ Y_i = \alpha + \beta X_i + \gamma W_i + u_i \]  

(4.1)
where $Y_i$ is the dependent variable for FLFP for individual $i$ ($i = 1, 2, \ldots, n$); it takes the value of 1 if the woman participates in the labour market and 0 if not;

$X_i$ is the independent variable that indicates whether mother $i$ has more than two children; it is equal to 1 if the woman has more than two children;

$u_i$ is a random error term representing measurement error or unobserved factors on $Y_i$;

Parameter $\beta$ captures the effect of having a third child on FLFP. A significantly negative or positive $\beta$ shows whether the presence of a third child reduces or increases the mother’s probability of working in the labour market; and

$W$ is a vector of control variables, namely, age, age at first birth, ethnicity, the urban indicator, regional dummy variables, co-residence with parents or parents-in-law and education of mothers.

4.5.2 Two stage least squares
The OLS models consider the fertility variable as an exogenous determinant of female labour supply behaviour. However, fertility (the number of children) is an endogenous variable. An important challenge in the estimation is the possible endogeneity of fertility. First, labour supply decisions of women and the number of children as well as the time of childbearing are jointly determined in the household’s utility maximisation problem. The number of children may be influenced by FLFP decisions or by other variables in the labour supply equation such as mothers’ education, mothers’ income or her husband’s income. Second, there may be some omitted variables such as different groups of women with heterogenous tastes (that are unable to measured) that affect both fertility and labour force participation decisions. Consequently, estimates of the effect of fertility and other exogenous variables on labour supply behaviour could bias estimation results if we fail to account for the endogeneity of fertility. The OLS results may be biased and unreliable and not reflective of true relations between fertility and labour supply. Thus, in exploring the relationship between fertility and women’s labour supply, the endogeneity problem needs to be taken into account. To address the endogeneity problem, this study estimates the model using 2SLS methods.
In the first stage, the endogenous variable “fertility” $X_i$ is decomposed into two parts: one part is uncorrelated with error term $\varepsilon_i$ and will be used in the second stage; the other part is correlated with the error term (Stock and Watson 2012). The part of $X$ that is uncorrelated with error term is isolated by regressing $X_i$ on instruments $Z_i$ and exogenous variables $W_i$ to obtain the predicted values $\hat{X}_i$

$$X_i = \lambda + \delta Z_i + \pi W_i + \varepsilon_i \quad (4.2)$$

where $Z$ is a vector of IVs that are correlated with $X_i$ but uncorrelated with error term $\varepsilon_i$; and the parameter $\delta$ is unknown coefficients that represent the correlation between instruments and fertility.

In the second stage, we regress $Y_i$ on the predicted value of the endogenous variable $X_i$ obtained from the first stage and exogenous variables $W_i$ as in equation (4.3). The coefficient of $\hat{X}_i$ obtained from the second stage regression is the 2SLS estimator $\hat{\beta}^{T\text{SLS}}$ and $\hat{\beta}^{T\text{SLS}}$ is a consistent estimator of $\beta$.

$$Y_i = \alpha + \beta X_i + \gamma W_i + u_i \quad (4.3)$$

As explained in section 4.4, the sex of the first two children is randomly determined and the instruments used in this study Twogirls and Twin2 are strongly correlated with having a third child because of the son preference mentioned above and as checked in the next section.

This thesis uses the 2SLS method to estimate the effect of having more than two children on the labour force participation decisions of mothers. The IVs used are Twogirl and Twin2. In the first stage, the endogenous variable $X_i$ “Moretwo” is regressed on IVs $Z_i$, Twogirl and Twin2 separately and include exogenous variables $W_i$ as in equation (4.2). In the second stage, the dependent variable FLFP is estimated on the fitted value of the Moretwo variable $\hat{X}_i$ that is obtained from the first stage and other exogenous variables as in equation (4.3). The coefficient obtained on $\hat{X}_i$ shows the effect of the existence of a third child on FLFP.
4.6 Empirical results

This study mainly focuses on the empirical analysis of the 2009 census sample. The sample is restricted to mothers aged 18-35. After applying all the restrictions, a final sample of 706,602 women aged 18-35 with two or more children was obtained.

4.6.1 OLS and 2SLS results

This chapter reports the empirical results. Only the coefficients of the variable of interest are presented. Completed estimates can be found in Appendix B. Firstly the results of the first stage of 2SLS estimation are presented, and then the results of OLS and the second stage of 2SLS estimation are reported.

Table 4.8 reports the results of the first stage of 2SLS estimation for the all women and married women samples. The first two columns report the results for the all women sample; the last two columns report the results for married women. The choice of controlling variables follows the literature and, in particular, Angrist and Evans (1998). I begin by estimating model (A) with inclusion of the sex of the first two children, age, age at first birth, and ethnicity. Model (B) additionally controls for the indicator of whether the family lives in an urban area. This is the preferred specification. Model (C) further controls for the regional variables. In model (D), the indicator of whether the woman resides with her parents or parents-in-law is included, and finally model (E) controls for the educational variables of mothers. The models use two different instruments (the first two children being girls $Twogirl$ and twins at second birth $Twin2$) separately.

<table>
<thead>
<tr>
<th>Models</th>
<th>All women</th>
<th>Married women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Twogirl</td>
<td>Twin2</td>
</tr>
<tr>
<td>(A)</td>
<td>0.1771***</td>
<td>0.7887***</td>
</tr>
<tr>
<td></td>
<td>(.0013)</td>
<td>(.0064)</td>
</tr>
<tr>
<td>(B) *</td>
<td>0.1771***</td>
<td>0.7892***</td>
</tr>
<tr>
<td></td>
<td>(.0012)</td>
<td>(.0064)</td>
</tr>
<tr>
<td>(C)</td>
<td>0.1768***</td>
<td>0.7876***</td>
</tr>
<tr>
<td></td>
<td>(.0013)</td>
<td>(.0064)</td>
</tr>
<tr>
<td>(D)</td>
<td>0.1873***</td>
<td>0.7782***</td>
</tr>
</tbody>
</table>

Table 4.8 - First stage regression - Impact of two girls and twin birth on probability of having more than two children (women aged 18-35 with two or more children)
<table>
<thead>
<tr>
<th></th>
<th>(E)</th>
<th>(F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(.0014)</td>
<td>0.1867***</td>
<td>0.7793***</td>
</tr>
<tr>
<td>(.0014)</td>
<td>0.1903***</td>
<td>0.7753***</td>
</tr>
<tr>
<td>F-test for</td>
<td>F = 19360.50***</td>
<td>F = 15257.24***</td>
</tr>
<tr>
<td>significance of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IVs (for model C)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No of obs</td>
<td>706,602</td>
<td>706,602</td>
</tr>
<tr>
<td></td>
<td>669,062</td>
<td>669,062</td>
</tr>
</tbody>
</table>

Notes: All models control for firstgirl, secondgirl, age, age at first birth and ethnicity. Model (B) controls for the urban dummy variable. Model (C) additionally controls for regional variables. Model (D) = Model (C) + Coresidence variable. Model (E) = Model (D) + Education variables. Standard errors are reported in brackets. * denotes statistical significance at 10 percent; ** denotes significance at 5 percent; *** denotes significance at 1 percent. Regional variables include the Red River Delta, North East, North West, North Central, Central Coast, Central Highlands, South East and the Mekong River Delta. Co-residence indicates whether the woman resides with her parents or parents-in-law.

As can be seen from the first column, women with two girls consecutively have 18 percentage points higher probability of having a third child than women with two boys or one boy and one girl. The results are very stable and consistent across the different specifications of the models from (A) to (E). Similar figures are also reported for the married women sample (column 3) because married women account for 93% of the all women sample.

Noticeably, the coefficients of Twogirl for the 2009 data in all models (A-E) are very high and much higher than those of Samesex (the coefficients of Samesex are reported in Appendix B). All estimates are statistically significant at the 1% level. Moreover, we can use the first stage F-statistic to test whether given instruments are weak. The results of the first stage F-statistic for the joint significance of IVs, as can be seen in Table 4.8, is very high (F-statistic>10 is in the safety zone). If first stage F-statistic is less than 10, the given instruments are weak. These F-statistic results demonstrate the statistical significance of this instrument. Twogirl is correlated with Moretwo, so it meets the relevance condition of a valid instrument. Moreover, as shown in the previous section, this instrument also meets the exogeneity condition of a valid instrument. Thus, it will be used as a good instrument for having more than two children.

I also check instruments such as Samesex, Twogirl and Twoboy that Angrist and Evans (1998) used. The effect of Samesex on having more than two children is also significant but its coefficient is smaller than the coefficient of Twogirl. Furthermore, using Twogirl as an instrument for having a third child is appropriate with the son preference of
parents in Vietnam, and therefore I choose *Twogirl* as a preferred instrument. for this study and *Samesex* is only used for comparison purposes against the study of Angrist and Evans (1998).

In addition, *Twogirl* and *Twoboy* are instruments which are decomposed from the *Samesex* instrument by Angrist and Evans (1998). I also examine the correlation of two instruments with endogenous variable ‘Moretwo’. Only *Twogirl* has a significant effect on *Moretwo* and the estimate is statistically significant at the 1 percent level, while the coefficient of *Twoboy* is very small, approaching zero and statistically insignificant. This implies that there are almost no incentives for parents of two boys to have a third child and they are also more likely to stop additional childbirth if they have two boys. Therefore, this study does not use *Twoboy* as an IV.

I also run several regressions using the instruments *Samesex, Twogirl* and *Twoboy* as in Angrist and Evans (1998) for the purpose of comparison, and results are reported in Appendix B. The results show that the effect of *Samesex* on having more than two children in this study is similar to the results in Angrist and Evans (1998) study. The results of the 2SLS estimates show that coefficients of *Moretwo* in the 2SLS estimates using *Twogirl* and *Twoboy* as IVs and the estimates using *Twogirl* are nearly the same. This is due to the effect of *Twoboy* being insignificant and approaching zero in the first stage of the 2SLS estimates and *Twogirl* being the main factor affecting the *Moretwo* variable (having more than two children). These effects of *Twogirl* and *Twoboy* are very different to those in Angrist and Evans (1998). In their study, both *Twogirl* and *Twoboy* have significant effects on having more than two children when *Samesex* is decomposed into *Twogirl* and *Twoboy*. However, in this study, since *Twoboy* has no significant effect on having more than two children, the effect of *Samesex* is mainly due to *Twogirl*, and therefore *Twogirl* is the main factor for having a third child among Vietnamese parents.

Like *Twogirl*, twins at second birth *Twin2* is also used to measure the consequences of moving from the second child to the third child. This study also uses twins at second birth as an instrument for having a third child. Apparently, twins are events that occur naturally and unplanned, so twins at second birth is also a good instrument for having a third child. Results from columns (2) and (4) of Table 4.8 show that *Twin2* has
significant effects on having more than two children and results are unchanged across all the different specifications of the models.

In this part, the results of the OLS estimates and the second stage of 2SLS estimations are presented. *Twogirl* and *Twin2* are used as IVs to estimate the effect of having more than two children on women’s labour force participation in 2SLS regressions. A set of control variables are included: the sex of the first two children; female’s age; age at first birth; educational attainment; ethnicity; whether the family lives in an urban area; residential region variables; and whether the female resides with her parents or parents-in-law, (for other coefficients not reported here, see Appendix B).

Table 4.9 summarises the results of OLS regression and the second stage regression of 2SLS estimates for the women aged from 18 to 35 with two or more children sample. The first three columns summarise the results for the all women sample, the last three columns summarise the results for the married women sample. The OLS results for the all women and married women samples are represented in the first columns respectively. The OLS estimates show that having a third child decreases a woman’s probability of labour force participation by about 3 percentage points. These impacts are not changed much when additional controls for other covariates are added. Similar results are also shown for the married women sample. All results of the OLS estimates are statistically significant. These negative effects are similar to the results reported in the analysis of the US data by Angrist and Evans (1998), Iacovou (2001) for Britain and Cruces and Galiani (2007) for two developing countries (Mexico and Argentina).

The 2SLS results for the all women sample using *Twogirl* and *Twin2* as instruments are presented in the second and third columns of Table 4.9. Column (2) of Table 4.9 reports the results using *Twogirl* in the 2SLS estimates of mothers’ labour supply responses. Using *Twogirl* as an instrument is totally appropriate with the existence of the son preference in Vietnam. The estimate results indicate the negative and significant effects of having a third child on FLFP from models (A) to (E) for both samples. For example, in the model (B), the results suggest that the presence of a third child reduces the probability of participating in the labour force by about 2.8 percentage points for the all women sample (or by 2.5%) and 2.9 percentage points for married women (or by 2.58%). The mother’s labour supply response is slightly smaller when controlling for regional variables, co-residence and education variables as in models (C), (D) and (E).
for both samples. Particularly, in model (C), the effect of a third child on mothers’ participation probability in the labour market reduces by 2.5 percentage points for the full sample and by 2.7 for the married women sample when controlling for regional dummy variables. All coefficients of Moretwo are statistically significant at the 1 percent level.

Thus, the 2SLS estimates using Twogirl as an instrument demonstrate the significantly negative effects of the Moretwo variable on women’s labour force participation. Noticeably, the results of the 2SLS estimates using Twogirl as an instrument are smaller than the corresponding OLS results for both the all and married women samples. It appears that OLS estimates overstate the causal effects of fertility on labour supply. This finding is similar to the findings of Angrist and Evans (1998).

Multiple births are considered to be one of the most important exogenous variations in the economic literature. Most studies focus on twins at first birth such as Bronars and Grogger (1994), Jacobsen et al. (1999) and Rosenzweig and Wolpin (1980b). Some studies use twins at second birth such as Angrist and Evans (1998) and Iacovou (2001). Like Samesex or Twogirl, twins at second birth Twin2 is used to measure the consequences of moving from two to three children. Although a twin is randomly assigned and a good instrument for the additional childbearing variable, using twins also has some difficulties because the number of observations is often limited and it can be difficult to attain a sufficiently large sample size.

In order to address the problem of data limitations, this study uses the population census survey with a much larger sample size than previous studies including Angrist and Evans (1998) study. The results of 2SLS estimates using Twin2 are reported in the third and sixth columns of Table 4.9. The results of 2SLS estimates using Twin2 as an instrument show statistical significance at the 1% level. Having an additional child decreases a woman’s probability of labour force participation by about 4.9 percentage points with including other covariates such as sex of the first two children, mother’s age, age at first birth, ethnicity and the indicator for whether families are living in an urban area as in model (B). The estimate results also show that this impact slightly reduces if regional variables, co-residence or education variables are added in the regression. The results are similar for the married women sample.
Table 4.9 - Effect of having more than two children on female labour force participation (women aged 18-35 with two or more children)

<table>
<thead>
<tr>
<th>Estimation method</th>
<th>All women</th>
<th>Married women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>2SLS</td>
</tr>
<tr>
<td>IVs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model (A)</td>
<td>-0.0277***</td>
<td>-0.0270***</td>
</tr>
<tr>
<td></td>
<td>(0.0009)</td>
<td>(0.0058)</td>
</tr>
<tr>
<td>Model (B)</td>
<td>-0.0340***</td>
<td>-0.0275***</td>
</tr>
<tr>
<td></td>
<td>(0.0009)</td>
<td>(0.0057)</td>
</tr>
<tr>
<td>Model (C)</td>
<td>-0.0320***</td>
<td>-0.0246***</td>
</tr>
<tr>
<td></td>
<td>(0.0009)</td>
<td>(0.0055)</td>
</tr>
<tr>
<td>Model (D)</td>
<td>-0.0310***</td>
<td>-0.0247***</td>
</tr>
<tr>
<td></td>
<td>(0.0009)</td>
<td>(0.0055)</td>
</tr>
<tr>
<td>Model (E)</td>
<td>-0.0280***</td>
<td>-0.0241***</td>
</tr>
<tr>
<td></td>
<td>(0.0009)</td>
<td>(0.0055)</td>
</tr>
<tr>
<td>Obs</td>
<td>706,602</td>
<td>706,602</td>
</tr>
</tbody>
</table>

Notes: All models control for firstgirl, secondgirl, age, age at first birth and ethnicity. Model (B) additionally controls for the urban dummy variable. Model (C) additionally controls for regional variables. Model (D) = Model (C) + Coresidence variable. Model (E) = Model (D) + Education variables. Standard errors are reported in brackets. * denotes statistical significance at 10 percent; ** denotes significance at 5 percent; *** denotes significance at 1 percent.

4.6.2 Discussion

Overall, both OLS and 2SLS estimates in this study show that the presence of the third child has significantly negative effects on female labour supply. Noticeably, results of the 2SLS estimates constructed using twins at second birth Twin2 are greater than the estimates using Twogirl. As noted in Angrist and Pischke (2008), different instruments do not necessarily generate the same estimate results even in the case of all valid instruments. This point can be explained by Local Average Treatment Effects (LATEs).

LATEs only provide an estimate of the average treatment effect on a population of compliers (subgroups). It does not provide the treatment effect for everyone. According to Angrist and Imbens (1995), the LATEs is equivalent to the estimator of the IV estimate (βIV) if instruments are binary variables. LATEs is calculated by dividing the difference in the means of the outcome variable (y) by the difference in the means of treatment (d) at the two different instruments (z):
\[ \beta_{iv} = \frac{(\bar{y}_1 - \bar{y}_0)}{(\bar{d}_1 - \bar{d}_0)} \]

where \( \bar{y}_1 \) is the mean of outcomes of women’s labour supply for those observations with instruments (such as Twogirl and Twin2) = 1;

\( \bar{y}_0 \) is the mean of outcome of women’s labour supply for those observations with the value of instruments = 0;

\( \bar{d}_1 \) is the mean of women having more than two children when the instruments Twogirl or Twin2 are equal to 1; and

\( \bar{d}_0 \) is the mean of women having more than two children if these instruments are 0.

Since an instrument may have stronger impacts on one or several subgroups of the population (the compliers), different instruments may lead to different LATEs. This is the reason why the causal effects of the presence of a third child on the labour supply of women (coefficients of Moretwo are different) are different when using different instruments such as Twogirl or Twin2. In this case, the effect of the third child on labour supply using twins as an IV is bigger than the third child due to two girls. One explanation for the bigger twins effects is that a third child who was born due to twins has no space with the second child and is an unplanned birth, so two young children at the same time may require more attention than an older child and younger child. Moreover, mothers with twins may spend more economic sources as well as time in comparison with mothers with the older child and younger child.

In general, both OLS and 2SLS estimates in this study show negative effects of the existence of a third child on female labour supply. This is similar to the findings in the study in the US by Angrist and Evans (1998) or in Britain by Iacovou (2001) or in Mexico by Cruces and Galiani (2007). However, the magnitude of these impacts on female labour supply is smaller than the findings by Angrist and Evans (1998) and Iacovou (2001), but similar to the findings by Cruces and Galiani (2007). These smaller effects may be due to supportive policies from governments such as childcare services, maternity leave, part time job availability and so on that vary between the different countries. For example, childcare services in developing countries in general and Vietnam in particular are more available and cheaper than in developed countries like.
the US or Britain. According to a report from the University of Pennsylvania in America, the US does not have a national policy on child care and the cost for child care is very high (UniversityofPennsylvania 2014). Also, a study by Child Care Aware America in 2013 reports that the day care cost per year for an infant ($15,000) exceeds the average tuition costs at public colleges in 31 of the states of America (University of Pennsylvania, http://knowledge.wharton.upenn.edu/article/economic-impact-of-child-care/).

Maternity leave policy is also very different between countries. According to ILO (2014), the US provides new mothers with 12 weeks of maternity leave and women in France are entitled to 16 weeks of paid leave. Meanwhile women in Vietnam received 18 weeks of paid maternity leave with the policy up to 2014, and with the current policy, now get 26 weeks.

The effect of children on FLFP decisions are also relatively small in comparison to the effect of other determinants of FLFP such as education or ethnicity in Vietnam. This may be due to the availability of childcare support from grandparents or childcare services as well as the lower costs of childcare services in Vietnam. It also may be that a large proportion of women are working in agricultural areas. With the characteristics of agricultural field work, they can work part time. All of these factors may reduce the tensions of having more children, so the effect of having additional children may not significantly affect the labour supply decisions of mothers. These may also be reasons why the Vietnamese women’s labour force participation rate has stagnated, even though the country has experienced a continuous decline in fertility rates and an increase in female education levels over the last two decades. This suggests that there may be other factors, such as discrimination in the labour market, that have negative effects on the FLFP rate. However, the study cannot investigate this factor because of the limitations of data availability. The next section will examine how the effect of children on female labour supply may vary with parents’ socioeconomic characteristics.

4.7 Heterogeneity in the impact of children on female labour supply

The impact of children on labour supply may vary with wages or educational attainments of mothers and fathers. This prediction has been shown in a number of theoretical models that describe how a mother’s labour supply response may vary with her wage rate or schooling as well as her husband’s. Gronau (1977) describes a
theoretical model that provides a unifying explanation of different behaviour patterns of people with different wages, income, education, and the impact of children on the allocation of time. Angrist and Evans (1996) also use a theoretical framework that incorporates Becker and Lewis (1973) child quantity/quality model into Gronau (1977) model to explore what factors affect the relationship between children and labour supply of parents. This model predicts that the labour supply response of childbearing will be exaggerated as the wages of parents are higher. This can be explained as, at equilibrium, higher wage earning women have greater marginal returns for an hour spent at home than lower wage earning women.

There are also many empirical studies that are consistent with this theoretical prediction. According to a survey conducted by Strauss and Thomas (1995), most economists find that better educated women spend more time with child care as well as housework than less educated women. Consequently, the labour supply of better educated mothers is more sensitive to the presence of children in comparison with less educated mothers (Gronau 1973; Leibowitz 1975; Hamoudi and Sachs 1999; Gayle et al. 2012). For example, Hamoudi and Sachs (1999) estimate the time parents allocate to preschool children for different socioeconomic groups and suggest that mothers who have high status and high potential wages tend to spend two or three times as much time with their children than mothers with lower status do. Gronau (1973) finds that the effect of a child on his/her mother’s value of time varies with the child’s age and the mother’s education. For a given age of a child, the effect of a child on his/her mother’s value of time increases with the mother’s education.

Since data about parents’ earnings is unavailable, the study explores how the impact of children on a mother’s labour supply varies with the level of schooling of the mothers and their husbands, schooling being an important predictor of potential earnings. The first part of Table 4.10 reports the OLS and 2SLS regression results of the impact of having more than two children on married women’s labour supply in the 2009 sample, conditional on the education levels of these women. In this section, Twogirl is used as an instrument for having more than two children. The second column shows results of OLS estimates, while the first and third columns present 2SLS estimates results. The OLS results show that the effect of having more than two children on labour supply increases with the mother’s education. For example, the effect of the presence of a third
child reduces his/her mother’s labour force participation by only 1.6 percentage points if the mother has not completed primary education. However, the mother’s probability of working decreases by 7.5 percentage points when the mother has at least completed secondary education.

The 2SLS estimates are presented in the first and third columns of panel 1. The first column reports the first stage results of the effect of Twogirl on Moretwo by the different education levels of mothers (such as less than primary education completed, primary education completed, secondary completed or university completed). The results show a strong and significant connection between Twogirl and Moretwo in all schooling groups. In families with two girls as the first two children, the probability of having an additional child among women with primary education or less is around 19 percentage points but the probability of having more than two children becomes smaller for mothers with higher education categories (secondary or university education completed) with 8.6 percentage points.

The second stage results of 2SLS estimates on labour supply equations (as seen in column 3) show that the impact of children on labour supply rises with the educational attainments of mothers. Having an additional child causes a decline of 8.7 percentage points in their probability of working if the women completed general universal education (year 12) or university or higher degrees, while the presence of an additional child decreases the participation by around 2 percentage points if the woman’s education is at primary level or less. In other word, the better educated women experience the largest impacts of childbearing on labour supply. This finding is consistent with the theoretical prediction and the OLS estimates but contradict the finding of Angrist and Evans (1998) that suggest that the effect of childbearing on female labour supply declines with women’s educational attainments.

Table 4.10 - Heterogeneity in the impact of children on female labour supply by interaction terms

<table>
<thead>
<tr>
<th>Dependent variables</th>
<th>More than 2 children</th>
<th>Labour supply response of mothers</th>
<th>No of obs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First stage</td>
<td>OLS</td>
<td>2SLS</td>
</tr>
<tr>
<td>1. Labour supply responses of mothers by educational level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than primary completed</td>
<td>0.198***</td>
<td>-0.016***</td>
<td>-0.020**</td>
</tr>
</tbody>
</table>

(0.003) (0.0014) (0.009)
<table>
<thead>
<tr>
<th>Educational Level</th>
<th>Less than primary</th>
<th>Primary completed</th>
<th>Secondary and/or university completed</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary completed</td>
<td>0.187***</td>
<td>0.190***</td>
<td>0.124***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.0016)</td>
<td>(0.0028)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.021***</td>
<td>-0.031***</td>
<td>-0.054***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0016)</td>
<td>(0.0012)</td>
<td>(0.0037)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.036***</td>
<td>-0.022***</td>
<td>-0.052***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.0068)</td>
<td>(0.025)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>156,060</td>
<td>431,215</td>
<td>81,787</td>
<td></td>
</tr>
</tbody>
</table>

2. Labour supply responses of mothers by father’s educational levels

<table>
<thead>
<tr>
<th>Educational Level</th>
<th>Less than primary</th>
<th>Primary completed</th>
<th>Secondary and/or university completed</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary completed</td>
<td>0.193***</td>
<td>0.190***</td>
<td>0.124***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0017)</td>
<td>(0.0016)</td>
<td>(0.0028)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.027***</td>
<td>-0.031***</td>
<td>-0.054***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0014)</td>
<td>(0.0012)</td>
<td>(0.0037)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.018***</td>
<td>-0.022***</td>
<td>-0.052***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0076)</td>
<td>(0.0068)</td>
<td>(0.025)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>327,824</td>
<td>431,215</td>
<td>81,787</td>
<td></td>
</tr>
</tbody>
</table>

3. Labour supply responses of mothers by educational level of mothers whose husbands completed primary education.

<table>
<thead>
<tr>
<th>Educational Level</th>
<th>Less than primary</th>
<th>Primary completed</th>
<th>Secondary and/or university completed</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary completed</td>
<td>0.193***</td>
<td>0.190***</td>
<td>0.124***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0017)</td>
<td>(0.0016)</td>
<td>(0.0028)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.027***</td>
<td>-0.031***</td>
<td>-0.054***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0014)</td>
<td>(0.0012)</td>
<td>(0.0037)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.018***</td>
<td>-0.022***</td>
<td>-0.052***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0076)</td>
<td>(0.0068)</td>
<td>(0.025)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>327,824</td>
<td>431,215</td>
<td>81,787</td>
<td></td>
</tr>
</tbody>
</table>

4. Regions

<table>
<thead>
<tr>
<th>Region</th>
<th>Less than primary</th>
<th>Primary completed</th>
<th>Secondary and/or university completed</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban</td>
<td>0.116***</td>
<td>0.116***</td>
<td>0.105***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0026)</td>
<td>(0.0026)</td>
<td>(0.0032)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.063***</td>
<td>-0.063***</td>
<td>-0.070***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0034)</td>
<td>(0.0034)</td>
<td>(0.0057)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.055**</td>
<td>-0.055**</td>
<td>-0.156***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.028)</td>
<td>(0.051)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>123,608</td>
<td>327,824</td>
<td>29,516</td>
<td></td>
</tr>
</tbody>
</table>

5. Co-residence with parents

<table>
<thead>
<tr>
<th>Co-residence</th>
<th>Less than primary</th>
<th>Primary completed</th>
<th>Secondary and/or university completed</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-residence</td>
<td>0.257***</td>
<td>0.257***</td>
<td>0.105***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0048)</td>
<td>(0.0048)</td>
<td>(0.0032)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0115***</td>
<td>-0.0115***</td>
<td>-0.070***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.0057)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0014</td>
<td>-0.0014</td>
<td>-0.156***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0094)</td>
<td>(0.0094)</td>
<td>(0.051)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>55,893</td>
<td>527,824</td>
<td>29,516</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Co-residence</th>
<th>Less than primary</th>
<th>Primary completed</th>
<th>Secondary and/or university completed</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non co-residence</td>
<td>0.183***</td>
<td>0.183***</td>
<td>0.105***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0015)</td>
<td>(0.0015)</td>
<td>(0.0032)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.034***</td>
<td>-0.034***</td>
<td>-0.070***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.0057)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.030***</td>
<td>-0.030***</td>
<td>-0.156***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0065)</td>
<td>(0.0065)</td>
<td>(0.051)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>523,063</td>
<td>545,518</td>
<td>29,516</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Standard errors are reported in brackets. * denotes statistical significance at 10 percent; ** denotes significance at 5 percent; *** denotes significance at 1 percent. Data are from the 2009 married women and husband samples. Other covariates in the models are as listed in the notes of Table 4.8. Twogirl is used as an IV in the 2SLS estimation.
In economic theory, the income of husbands is an important factor affecting wives’ labour supply. Since there is no available data on husbands’ wages or incomes to analyse female labour supply effects conditional on husbands’ wages, the study uses the condition of the schooling of husbands which is a good prediction of potential earnings. Panel 2 of Table 4.10 describes the labour supply consequences for married women with at least two children, conditional on their husbands’ education. The OLS estimates in the second column show that the effect of childbearing on women’s labour supply increases as their husband’s education level is higher. The first stage results of the 2SLS estimates suggest that the effect of Twogirl on having more than two children reduces with their husbands’ education. For women with husbands in the low education categories, the probability of having an additional child is about 19 percentage points, while this probability is only 12 percentage points for women with better educated husbands. The 2SLS estimates of the labour supply equation show the labour supply of women whose husbands are better educated is more sensitive to the presence of an additional child.

In general, the variation in 2SLS results by both mother’s and father’s educational levels are similar to the variation in OLS estimates which implies that the effect of children on women’s labour supply rises as education levels increase.

In order to distinguish more clearly how labour supply responses to childbearing vary with educational levels of women or their husbands, the sample is then restricted to women whose husbands completed primary education. This is interesting because with this restriction, we can see how women’s labour supply may change in response to an increase in the number of children by their different education levels. Panel 3 of Table 4.10 reports on the labour supply consequences by women’s education levels with the restricted sample including married women with husbands who completed primary education. The 2SLS estimates show that the effect of childbearing on women’s labour supply increases with their education levels. These findings are consistent with the theoretical prediction and the OLS results which suggest that the labour supply of better educated women is more sensitive to childbearing.

The effect of childbearing on women’s labour supply also varies between urban and rural areas. As can be seen from Table 4.10, mothers who live in urban areas are more sensitive to the presence of additional children than mothers living in rural areas. The
2SLS estimates show that the probability of working reduces by 5.5 percentage points when a woman residing in an urban area has an additional child, while the figure is only 2.5 percentage points for a woman living in a rural area. This may be explained by the majority of women in rural areas working in agricultural field work (about 72%), while only 21% of women living in urban areas work in agricultural areas. They can work part time because of the nature and characteristics of agricultural jobs in the field, so the presence of a child may not affect their labour supply as much as women working in other fields. Moreover, in rural areas, women often have support from their parents who can help to take care of their children. In contrast, in urban areas, women often work in full time paid jobs such as accountants, doctors, teachers and so on. They also have better economic conditions. In addition, urban areas mainly include nuclear families, so it is more difficult to get support from grandparents for childcare. Therefore, having more children will have more influence on their labour supply decisions.

It is interesting to note that the reduced form first stage results in column (1) of panel 4 suggest that the effect of TwoGirl on fertility is greater among rural families than urban families. Again this can be explained as the son preference is stronger in rural areas where the traditional idea is that a boy belongs to us and a daughter belongs to someone else (Mont and Nguyen 2013, pp.61) and where sons are needed to maintain the family line according to the Confucian tradition which remains strongly influential among many Vietnamese people. Moreover, sons are often responsible for financially supporting their parents’ old age. Therefore, sons are highly desirable for social, symbolic and economic reasons.

The final panel of Table 4.10 reports on the labour supply responses of married women in families with two or more children and whether the women reside with their parents or in-laws as an interaction term. The OLS results suggest that the presence of a third child has a greater significant negative impact on women’s participation in the labour market when they do not live with their parents or in-laws. In other words, when there is no help from parents and parents-in-law, the effect of having more children on the labour force participation of women is greater than on women who do receive such help. Similar results can also be seen in the 2SLS estimates for women who do not reside with their parents/in-laws, where having a third child reduces their probability of
working by 3 percentage points. However, this effect is small and insignificant in the case of women who co-reside with their parents or parents-in-law.

4.8 Conclusion and policy implications

This study explores the issue of causal effects of fertility on FLFP in Vietnam. Understanding the linkage between variations in fertility and female labour supply is important for predicting changes in the structure of the labour force in the future as well as for identifying potential sources of gender wage differentials. The endogeneity problem of the relationship between fertility and female labour supply is a challenge. To deal with this problem, the study uses an IV strategy which exploits exogenous variations in family size of moving from two to three children to investigate the causal effect of both planned and unplanned births on women’s labour force participation decisions.

This study is the first attempt to investigate the effect of fertility on the labour supply of Vietnamese women using the IV method, based on the existence of a son preference among families in Vietnam. It has been shown that the first two children being girls is a good instrument for having a third child. The study also uses twins at second birth as an IV.

Results show that having a third child decreases the probability of labour force participation of women by about 3-5 percentage points following the presence of the additional child.

The study also finds that the impact of fertility on labour supply may vary with the educational attainments of mothers and fathers. The effect of children on labour supply rises as educational attainments of mothers rise. It suggests that the better educated women experience the greatest impacts of childbearing on labour supply. This finding is consistent with predictions of some theories of household time allocation and OLS estimates. The 2SLS estimates also show that the labour supply responses of women whose husbands are better educated are also more sensitive to the presence of additional children.

These findings suggest that reduction in fertility is one of channels to increase the labour supply of women in Vietnam. However, this is not the only channel. Other
factors such as better education, reductions in gender discrimination in the workplace and women’s empowerment could be driving both trends. As mentioned above, although fertility rate in Vietnam dropped rapidly and female education also has achieved remarkable progress during past years, labour force participation rate of women has not changed much, even slightly decreased. This may be due to these impacts of fertility have been offset by other factors such as gender inequality in higher education, in employment. Because there is a fact that even though Vietnam has achieved universal primary education, enrolment rates for girls in higher education and among disadvantaged groups such as rural and remote areas, and among ethnic minority communities is still lower than for boys. Policies to aim improvement and better access to higher education for girls, especially for disadvantaged groups will help girls attain higher levels of education and this is also an important factor to enhance the role of women in society. In order to implement this target, the government can provide financial solutions such as school facilities, school fees to support for students in disadvantaged socioeconomic areas.

Moreover, higher level of education achieved by girls can also be a useful channel to reduce gender inequality in the labour market. In Vietnam, gender inequality still persists in the labour market with the rate of participation in the labour market is higher for men and women account for a high rate in vulnerable types of jobs such as informal sector, or jobs in the short term and without social security. Therefore, in order to improve and reduce gender disparity in the labour market, the government can issue anti-discrimination decrees in the workplace. However, this is difficult because discrimination is intangible and hard to measure. Therefore, this study suggests policies should aim to improve female education. As women with better level of education can have more opportunities to choose good jobs as well as better positions at work with higher income. Their knowledge also helps them have the desired family size as well as doing well other roles in the family.

Although this study focuses on the population of Vietnamese women, the approach and method of this study may also be applicable to other developing countries where cultural and social traditions are similar to Vietnam and the existence of a son preference is still a dominant theme.
Chapter 5. The impact of parental health on children's schooling and labour force participation: evidence from Vietnam

5.1 Introduction

The objective of this paper is to investigate the relationship between parental health shocks, and children's schooling and participation in the labour market using a large, rich longitudinal dataset of Vietnamese households.

The analysis of investments in human capital is a critical issue in the context of developing economies and has received increasing attention among economists and policy makers in recent years. The achievement of universal primary education is specifically mentioned as one of the Millennium Development Goals and, in developing regions, children from the poorest households are more likely than others to not be in school (Unesco and Unicef 2010). Further, substantial gender disparities still exist with respect to secondary and tertiary education and enrolment rates still favour males both in high school and university in several developing regions, including South Asia.

The Vietnamese education system has outperformed many other Asian and developed countries in standardised tests in recent years, thanks to substantial government investments in education, and Vietnam has almost achieved universal primary education (WorldBank 2011). However, important gaps exist, especially for children from ethnic minorities and low socioeconomic backgrounds, and educational attainments significantly vary across the different groups ((Unicef) https://www.unicef.org/vietnam/girls_education.html#_ftn1). Similarly, enrolments and completion rates in lower and/or upper secondary schools (ages 11-15 and 15-18 respectively) have improved over the recent years, but are still substantially lower for children who come from ethnic minorities and/or from households at the bottom of income distribution (WorldBank 2011; MOET 2015). As reported in Badiani et al. (2013), 40% of people who are 21 years and older in the richest quintile have completed a university degree compared to less than 2% in the poorest quintile. In fact, more than a quarter of individuals in the poorest quintile had not even completed primary school by 2010.

In addition, completion rates in Vietnamese rural areas are about two thirds of what they are in urban areas and these figures are a particular source of concern because of their impact on overall inequality, especially considering that the urban population only
accounts for 20% of the total population (WorldBank 2011). Lastly, learning outcomes and educational achievements substantially differ across socio economic groups, with students from ethnic minorities, rural areas and poor households, showing to be about three grades behind their counterparts from the better advantaged urban families (WorldBank 2011). All of these data raise substantial concerns for the potential to perpetuate and aggravate the intergenerational transmission of poverty in Vietnam (Badiani et al. 2013).

The evidence on child labour in Vietnam shows similar trends and concerns. The incidence and intensity of child labour has substantially decreased over the last two decades (O’Donnell et al. 2005). However, children from ethnic minorities and children who live in rural and deprived areas still remain particularly vulnerable and seem to have missed out on the improvements in this context (Edmonds and Turk 2004). Beegle et al. (2009) analyse the impact of child labour in rural areas of Vietnam and show that, as expected, child labour significantly reduces the chance of being in school and gaining educational attainments.

The analysis of children’s education and child labour in conjunction with parental health shocks is particularly meaningful in the context of a developing country like Vietnam. Health shocks have a high degree of unpredictability and may have severe and sometimes devastating effects on families’ well-being, especially in rural households who have limited resources, are mostly employed in the agricultural sector and are often not covered by adequate health insurance. Health shocks can suddenly increase medical expenditures and decrease available income for all family members. Families may need to adjust their consumption patterns and may have to use accumulated savings, sell valuable items or borrow from relatives or financial institutions (Wagstaff 2007; Bales 2013). The impact of health shocks on labour supply in the family is unclear. On one hand, labour supply may be reduced both because one family member is ill and because others need to look after them. On the other hand, rural households are more likely to compensate for the sudden decrease in working hours by increasing the labour supply of other family members (Wagstaff 2007). In the Vietnamese context, caring for ill family members is typically a woman’s responsibility (WorldBank 2011) and therefore girls are particularly likely to be affected by the illness of a family member.
This study explores the relationship between parental health shocks and children’s education and engagement in the labour market in Vietnam. The connection between parental illness and children’s education is twofold. First, parental illness may reduce the number of working hours as well as productivity and consequently may cause a reduction in available family income and this can in turn reduce available resources to support children in education. Second, parental illness may decrease the overall family labour supply and children may need to leave education and start working (Alam 2015). For this reason, it is particularly interesting to analyse the impact of parental health shocks on both children’s education and child labour as the two elements are strongly related.

While there is substantial evidence on the intergenerational transmission of health from parents to children (see for example Currie and Moretti (2007), among others), the economics literature analysing the relationship between parental health shocks and children’s education is limited. In this framework, Vietnam is a particularly interesting case to look at, given its poor system of social protection, the limited access to the social health insurance system, and health expenditure heavily depending on direct out-of-pocket payments. Further, the study period (2004-2008) corresponds to a critical phase of changes for the Vietnamese health insurance system.

The study uses longitudinal data and child fixed effects to control for child’s time-invariant characteristics that might be associated with parental health and children’s schooling and the impact of parental illness is separately analysed by gender and age of the child.

The results suggest that children (ages 10-23) of unhealthy mothers are 3 percentage points less likely to be enrolled in school than children with healthy mothers, and girls are particularly strongly affected by parental illness. Results also show that that maternal illness significantly increases the probability of children entering employment, as well as increasing their number of working hours. In contrast, a father’s illness has smaller impact and is only significant at the 10% level on school enrolment of children in the sample but moderately increases the probability of children working. These results are consistent with existing literature on the impact of parental health in developing countries, in particular with Bratti and Mendola (2014).
The rest of the paper is organised as follows: section 2 discusses the existing literature on the relationship between parental health shocks and children’s education; section 3 presents Vietnam’s context in the period of analysis, the data and the descriptive statistics; section 4 describes the econometric methodology; section 5 presents the results; and section 6 provides concluding remarks.

5.2 Literature review

Economists have been interested in the intergenerational transmission of health, income and education and, more generally, in the effect of parental characteristics on children’s educational attainments, because of the substantial impact of an individual’s education on her/his future life and achievements (see for example Currie (2009), Case and Paxson (2011), and among many others). There is also a large literature documenting on the effects of parental health on the outcomes of children later in life such as human capital development, income, and health (Haveman and Wolfe 1995; Currie 2009). In general, the existing literature shows that poor health of parents has negative impacts on socioeconomic and health outcomes of children in the later life. However, the effect of parental health on children’s education has received rather limited attention in the economics empirical literature (Haveman and Wolfe 1995).

The study of the impact of parental health shocks is particularly important in developing countries, where social protection is limited and people have restricted access to formal health insurance (Gertler and Gruber 2002; Wagstaff 2007). Major illnesses are unpredictable and can represent very serious shocks for families, both in the rural and urban context (Wagstaff 2007; Sun and Yao 2010). In these circumstances, negative health shocks can impose monetary and non-monetary costs on families, due to costs of accessing health care and reductions in labour supply, income and productivity, along with emotional and psychological distress (Gertler and Gruber 2002; Wagstaff 2007; Somi et al. 2009; Sieh et al. 2010; Wagstaff and Lindelow 2014). These factors can have short and long run consequences on the family, because of the reductions in consumption and decreased accumulation of productive assets including education (Sun and Yao 2010). The effect of adverse health events has been found to be particularly pronounced for poor and less educated families (Dercon and Krishnan 2000; Genoni 2012) and may lead households into chronic poverty (Wagstaff 2007; Sun and Yao 2010).
Wagstaff (2007) analyses the economic consequences of health shocks in Vietnam and shows that, as a consequence of health shocks, families are likely to significantly reduce per-capita food consumption, and at the same time increase other household expenses, such as housing and electricity. Rural households are also more likely to adjust their labour supply, and compensate for the reduced activity of the sick member of the family. Similarly, Mitra et al. (2015) use fixed effects to show that Vietnamese female-headed and rural households are the most vulnerable to health shocks and, in general, families tend to reduce their non-health related expenditures on items such as food and education.

A separate strand of literature has analysed the impact of parental death on children’s outcomes and generally shows that a parent’s recent death has a large and significant negative effect on children’s enrolment into education (see for example Gertler et al. (2004) for evidence on Indonesia and Case and Ardington (2006) for evidence on South Africa).

The limited literature on parental health and child labour in developing countries has considered various transmission channels of health shocks to children’s educational attainments. First, parental health shocks may reduce children’s education by increasing the need for children to substitute for adult labour supply and therefore decrease school attendance (Choi 2011; Bratti and Mendola 2014; Alam 2015). Second, children may be required to care for a sick parent in the household, again reducing the time available for school work (Mont and Nguyen 2013; Bratti and Mendola 2014). Further, parental illness may reduce parental engagement with the child’s schooling as well as increasing the child’s emotional distress, and both factors are likely to have adverse effects on educational achievements (Guryan et al. 2008; Bratti and Mendola 2014).

The economics literature on the impact of parental health on children’s education in developing countries is limited and delivers unclear results, especially with respect to the impact of paternal versus maternal health.

Sun and Yao (2010) examine the effect of major illnesses of working age adults in households on a child’s chances of entering or finishing middle school in rural China and they show that primary school age children (and girls in particular) are the most vulnerable to family health shocks. Similarly, Choi (2011) explores the effect of
parental self-reported health problems on children’s education and employment during subsequent years in Russia and shows that a father’s poor health substantially reduces a daughter’s educational attainment and labour market engagement. Mont and Nguyen (2013) use cross-sectional data from the VHLSS 2006 to show that parental disability significantly decreases the likelihood of child school enrolment and the impact of maternal disability is the strongest. However, all of these studies rely on data that only records children’s schooling at one particular point in time and therefore they are not able to exploit the variation over time in children’s education and to control for individual characteristics that do not vary over time.

The studies that are closest to this one in terms of methodology are Bratti and Mendola (2014) and Alam (2015). Both papers use child fixed effects to analyse the impact of parental health on children’s education, and also to analyse the effect on child labour. However, these two papers provide different results with respect to the impacts of maternal and paternal illness. Further, these two studies consider a European and an African country, which have very different socioeconomic structures and challenges.

Bratti and Mendola (2014) use longitudinal panel data and child fixed effects to investigate the causal relationship between parental illness and child schooling in Bosnia and Herzegovina. Their results show that a mother’s poor self-reported health decreases the chance of school enrolment of children aged 15-24 and increases a child’s likelihood of working due to the large increase in health expenditures. However, a father’s poor health does not have such significant effects on children’s education.

On the other hand, Alam (2015) employs a panel data survey and child fixed effects to investigate the effect of parental illness on the education of children aged 7-15 in Tanzania and shows that only the father’s illness has a significantly negative effect on children’s school attendance but there is no evidence of reallocation of children’s time from school to work.

This study contributes to this existing literature in three main ways. First, the study provides new evidence on the relationship between parental health and child schooling in an Asian developing country, using a credible identification strategy, and distinguishing between paternal and maternal health status. The institutional context of Vietnam is very different from the one of both Bosnia and Tanzania, and findings from
Vietnam can be more easily generalised to other Asian developing countries with similar education and health systems. Further, in the period of the study, Vietnam undertook a substantial process of reform of its social protection health insurance system and therefore it offers a very interesting case to analyse and the results of this study could have important policy implications for the ongoing reforms of the system.

Second, I use longitudinal data with child fixed effects to explore the effect of parental illness on child education and therefore the study is able to take into consideration time-invariant unobserved factors that may affect parental health and children’s education.

Lastly, the study explores the impact of parental illness on child labour by focusing on the intensity of the child’s engagement in the labour market captured by the number of hours of work rather than simple labour force participation.

5.3 Data
5.3.1 Vietnam's institutional background
Vietnam is a developing country in Southeast Asia with a population of 89 million in 2011 (OECD 2013). Around 67% of the total population lives in rural areas (World Bank, http://data.worldbank.org/indicator/SP.RUR.TOTL.ZS) and 47% of the overall population is employed in the agricultural sector (World Bank, http://data.worldbank.org/indicator/SL.AGR.EMPL.ZS?locations=VN). Recent economic reforms have generated substantial economic progress in the Vietnamese agricultural sector, with rapid increases in farming production, lifting rural incomes, reducing rural area poverty and under-nourishment (OECD 2015).

Vietnam has achieved a remarkable level of economic growth in recent years and has reduced the poverty rate from 58% in 1992 to 10% in 2010 (WorldBank 2012). Further, life expectancy and the average level of education in the population have substantially increased (OECD 2013). However, even though many Vietnamese families have risen out of poverty, many still have incomes very close to the poverty line and therefore are still highly vulnerable to unexpected shocks, such as sudden changes in rainfall, animal diseases and health shocks (Badiani et al. 2013). Further, the remaining poor in the population experience substantial challenges, such as isolation, poor health, low education, and limited assets. Ethnic poverty remains as a strong problem, as minority groups accounted for 47% of the poor in 2010, even though they make up less than 15%
of the overall population (Badiani et al. 2013). Overall inequality is rising, especially between rural and urban areas, and disparities between different socioeconomic groups are widening in rural areas such as the increasing gap between the rich and the poor, inequality in income, employment as well as access to other social services such as education and healthcare services. The country’s rapid economic growth and modernisation have pushed many people out of rural areas and into new city-based jobs which often are in the informal sector and lack security and social protection (Badiani et al. 2013).

The Vietnamese Government introduced a social health insurance scheme in 1992 called Vietnam Health Insurance, initially to provide coverage to civil servants and state enterprise workers. The Vietnamese Health Insurance system has undergone substantial changes after the period of 1992-1998. In 2002, the non-contributory health insurance scheme was expanded to target vulnerable groups, such as poor individuals, ethnic minorities and families experiencing socio-economic disadvantage through the new Health Care Fund for the Poor. Since 2005, the non-contributory health insurance program has also included children under 6 years old (Mitra et al. 2015). As noted in Lieberman and Wagstaff (2009), the health insurance coverage rate increased from 14% to 42% between 2002 and 2006. In 2009, the Vietnamese Government approved the Law on Social Health Insurance to create a Social Health Insurance (SHI) system to be used as the main vehicle to achieve universal coverage (Somanathan et al. 2014).

However, there still are significant challenges for the Vietnamese health system, especially related to improving equity, financial protection and enrolment rates. Further, out-of-pocket health-related expenses remain high and can leave households exposed to substantial risks, as well as deterring utilisation and reinforcing health inequalities (Somanathan et al. 2014). At the current stage, the financial protection provided by the health insurance scheme is still too low, also because of informal payments and other indirect costs (Sepehri et al. 2006).

Further, expanding the breadth of coverage remains an important priority for the Vietnamese health system, with a particular focus on reaching the large faction of the population working in the informal sector. These individuals constitute a large proportion of the 31.9 million Vietnamese not enrolled in SHI, who mostly represent the so-called “missing middle” problem, a problem typical of developing countries, where
enrolments in the health system are high among lower and higher income groups, but low among individuals from the middle income groups (Somanathan et al. 2013).

The national education system in Vietnam has 5 levels of education: pre-primary, primary, lower secondary, upper secondary and higher education. Pre-primary education includes preschool and kindergarten for ages 3-6. Children start primary school at age 6. Primary school lasts for 5 years and is free and compulsory for all Vietnamese children. Secondary education is divided into lower secondary, which lasts for 4 years from grades 6 to 9 (ages 11-15), and upper secondary education, which includes grades 10-12 for children aged 15-18. An alternative to upper secondary education is the vocational or technical track that varies from 6 months to 3 years in length. Students graduating from secondary education could enrol in any university by passing the qualification examination organised by the institution. Higher education includes universities, senior colleges and research institutes. A typical undergraduate course in the university last from four to five years and 3 years in the college (WorldBank 2011).

Government support for education has dramatically increased over the past 25 years and Vietnam spent over 5% of its GDP on education in 2009 (WorldBank 2011). Important recent achievements include almost universal primary education and a substantial increase of the population’s literacy rates. The major current challenges for the education system include targeting inequalities in grade attainments, attendance and completion across the different ethnic and socioeconomic groups (WorldBank 2011).

5.3.2 Parental health, children labour and education

Vietnam has also made remarkable progress in multiple aspects of gender equality, particularly in education. Gender disparity in term of access to primary education has been eliminated and mostly obtained gender equality in access to lower secondary education. The country has succeeded in making primary education universal and is moving towards universal lower secondary education. By 2012, net enrolment rate in primary education reached 97.7 percent. Net enrolment rate for lower secondary education had risen considerably over the period of 2006-2012, from 81 percent in 2007 to 87.2 percent in 2012 (Table 5.1).

| Table 5.1 – Net enrolment rate by education level (%) |
Although the country has achieved universal primary education, reducing school dropout rates has been still challenging. In 2012, students who did not complete primary education accounted for 7.9 percent and about 20 percent of students who enrolled in lower secondary education did not complete. A variety of reasons for dropping out of school such as financial problems, learning capacity of students, language barriers for ethnic minority, low quality teaching and learning, particularly in highly remote and disadvantaged areas.

Moreover, inequality in access to education and quality of education still exists across different socioeconomic groups, between ethnic majorities and ethnic minorities, between rural and urban areas, and among geographical regions. Over the period 2004-2008, enrolment rate in urban areas for students aged 10-23 accounted for 70.6 percent, while this figure for rural areas was 65.4 percent. Particularly, at primary education level, enrolment rates are nearly the same between urban and rural areas. However, inequality in access to education between rural and urban areas has become larger at higher education levels (76.8 percent of children in urban areas enrolled in lower secondary education compared to 65.7 percent in rural areas; 63.2 percent in urban and 48.7 percent in rural areas for upper secondary education) (VHLSS, 2004-2008). By region, completion rates at primary education level in 2012 were especially low in some certain regions such as the Northern Midland and Mountainous Areas, Central Highland regions (at 89.4% and 83.6% respectively) (VHLSS, 2012). Literacy rate for ethnic minorities still remains at low level and even lower among women and girls.

Gender disparities in higher education and among disadvantaged groups still remain. Although there was no significant difference in the net enrolment rate between boys and girls in primary education, enrolment rates for girls in higher education levels in rural, mountainous and remote areas, and among ethnic minority groups is much lower than for boys. A significant gap in literacy rates still exists between boys and girls among ethnic minority groups over the period 2004-2012. In 2012, ethnic minority girls aged 10 and over have 10.8 percent lower level of literacy than men. In some highly remote
areas such as the Northern Midland and Mountainous Areas, boys demonstrated a higher level of literacy than girls and remained at 7.8 percent in 2012.

The huge discrepancy in living conditions, access to healthcare facilities and quality of healthcare service shows substantial inequality. Although maternal mortality rate in Vietnam considerably dropped by about two thirds over the period of 1990-2012, from 233 per 100,000 live births in 1990 to 64 per 100,000 in 2012, disparities in the health status of women still exists across geographical regions, ethnicities and income quintiles (GSO, 2012). Every maternal health indicator such as maternal mortality rate, antenatal care, and adolescent birth shows a far more disadvantaged situation for women living in remote, rural and mountainous areas, minority communities, as well as women with lower education levels. Women belong to ethnic minority groups, the poorest quintiles and rural areas have lower rate of antenatal care and are less likely to give birth with the assistance of skilled health staffs than those of majority groups and in urban areas. While 98 percent of majority women have a birth with attendance of skilled staff, only 63 percent of their minority counterparts have this health assistance (GSO, 2006, 2011).

Child labour in Vietnam is predominant in agriculture and in the household unit. In 1998, around 90 percent of child workers in both urban and rural areas participated in this activity. Children working in this sector account for a substantial proportion and even greater in rural areas. Child labour supply in term of both labour force participation and weekly working hours in rural areas shows a considerably higher rate than in urban areas (see Table 5.2). The dominance of children’s work on household farm or agricultural fields mainly reflects the general structure of the Vietnamese economy with a large proportion of population living on agriculture.

### Table 5.2 – Children labour supply

<table>
<thead>
<tr>
<th></th>
<th>2004</th>
<th></th>
<th>2006</th>
<th></th>
<th>2008</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Urban</td>
<td>Rural</td>
<td>Urban</td>
<td>Rural</td>
<td>Urban</td>
<td>Rural</td>
</tr>
<tr>
<td>Labour force participation (%)</td>
<td>22.39</td>
<td>41.65</td>
<td>23.28</td>
<td>36.34</td>
<td>24.61</td>
<td>39.69</td>
</tr>
<tr>
<td>Weekly work hours</td>
<td>11.56</td>
<td>17.12</td>
<td>12.20</td>
<td>16.23</td>
<td>12.78</td>
<td>18.02</td>
</tr>
</tbody>
</table>


Along with the rapid economic growth in Vietnam during the 1990s, there was a considerable decrease in both incidence and intensity of child labour. The decline in child labour in the 1990s was linked with significant improvement in living standards. However, the rate of child labour is still high in poor and developing countries like Vietnam, particularly in rural areas and ethnic minority groups and its outcome is a trade-off between child labour and schooling as demonstrated by many studies in the existing literature. Children from ethnic minority communities and disadvantaged groups appear to remain particularly vulnerable. In 2008, the proportion of children from ethnic minority communities participating in the labour force accounted for 56 percent in comparison with 30 percent for children from ethnic majority group (VHLSS, 2008).

Overall, all indicators on children education and labour show that children from rural and remote areas, disadvantaged socioeconomic groups are particularly vulnerable in comparison with children from urban areas and ethnic majority groups. Thus, whether household financial problems are main reasons and critical factors that affecting dropping out of school, low enrolment rates and engagement in the labour force of children in these areas or other factors such as family members’ health shocks, their parents’ discouragement for attending school. This thesis focuses on the impacts of parental health shocks on school enrolment and labour of children.

5.3.3 Data and descriptive statistics

In this study, data is used from three waves of the VHLSS covering the years 2004, 2006 and 2008. Only half of households interviewed in the 2004 wave are recalled in 2006, and also only half of households in 2006 are reinterviewed in the 2008 wave. Therefore, an unbalanced panel of 13,930 observations has been created from these three waves. Any one individual in this panel data can follow for at least two waves. The VHLSS dataset is implemented by the GSO of Vietnam with technical support from the World Bank, with the objective of monitoring and evaluating the implementation of the Comprehensive Poverty Alleviation and Growth Strategy. This is an ongoing longitudinal survey of the Vietnamese population and is conducted on a nationwide scale. The sample is not only representative at the national level but also of urban and rural areas, and eight geographical regions. The VHLSS includes detailed information on the characteristics of individuals, households and communities, such as
demographics of household members, ethnicity, area of residence, educational background, employment status, income, expenditures, housing, household assets, utilities, etc.

The survey also incorporates detailed information on individual health status, including morbidity (and occurrence of symptoms such as diarrhoea, nausea, fever, and cough); and physical difficulties and limitations in daily activities, which is only available in 2006. This study follows Mitra et al. (2015) and uses a different measure of health shocks. In each wave, all household members were asked whether they spent any days in bed due to illness or injury and whether they needed someone to take care of them. Similarly, they were asked whether there were any days on which they were unable to carry out regular activities due to illness/injury. The survey includes two questions for each of these indicators where individuals are asked to report any major illness events in the last 4 weeks and then in the last 12 months. Unfortunately, there is no way of exactly defining when the episodes happened, as individuals who were sick in the last 4 weeks could also report that they were sick in the last 12 months. If respondents answered positively to any of these questions, they are asked to report how long the episode lasted. These measures allow capturing individuals’ functioning levels that affect their ability to work, and are close to the objective measures of health defined under the International Classification of Functioning Disability and Health (WHO 2001). Lastly, individuals are asked whether they were hospitalised in the previous 12 months and how many times they were hospitalised.

The study defines a binary variable for parental illness equal to 1 if the individual:

- Spent at least 1 day in bed and needed someone to take care of her/him; or
- Spent at least 1 day unable to carry out regular activities because of illness; or
- Was hospitalised at least one time.

I also run several sensitivity analyses, by varying the number of days (for example choosing 2 or 3 days as a threshold to define illness) and excluding hospitalisation, to show that these do not affect the main results. The average number of days spent in bed for parents in the estimation sample is 4, while the average number of days unable to carry out regular activities because of illness is 12.
In general, health measures that refer to the ability to perform daily activities have been considered “more objective” and less prone to measurement error than measures based on self-reported health status, which are likely to be affected by the subjectivity of individual response scales (Bratti and Mendola (2014); Gertler and Gruber (2002); Dow et al. (1997), among many others).

This study follows Bratti and Mendola (2014) and includes children aged 10-23 who are living with both parents in the sample. However, I tested the validity of my estimates by including children living with a single parent and the main results are unchanged. The sample of children aged 10-23 includes those who attend secondary school (lower or upper) and university. This study decided to focus on this group of children (excluding primary school aged children) given that primary attendance in Vietnam is free and almost universal, and therefore less likely to be affected by parental illness.

Table 5.3 presents descriptive statistics of enrolment rates in the sample, by age and parental health status. As expected, enrolment rates are higher for younger children (around 92% for children aged 10-14) and decrease for children in upper secondary school (around 63% in the age group 15-18). In general, the school enrolment rate is higher for girls than for boys and children with a sick parent seem less likely to be enrolled in education.

<p>| Table 5.3 - Children’s school enrolment rates by parental health status |
|---------------------------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th></th>
<th>Enrolment</th>
<th>No parent ill</th>
<th>Any parent ill</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full sample</td>
<td>0.665</td>
<td>0.690</td>
<td>0.640</td>
</tr>
<tr>
<td></td>
<td>(0.472)</td>
<td>(0.463)</td>
<td>(0.479)</td>
</tr>
<tr>
<td>Boys</td>
<td>0.636</td>
<td>0.661</td>
<td>0.612</td>
</tr>
<tr>
<td></td>
<td>(0.481)</td>
<td>(0.473)</td>
<td>(0.487)</td>
</tr>
<tr>
<td>Girls</td>
<td>0.698</td>
<td>0.722</td>
<td>0.674</td>
</tr>
<tr>
<td></td>
<td>(0.459)</td>
<td>(0.448)</td>
<td>(0.468)</td>
</tr>
<tr>
<td>Ages 10-14</td>
<td>0.924</td>
<td>0.931</td>
<td>0.915</td>
</tr>
<tr>
<td></td>
<td>(0.265)</td>
<td>(0.252)</td>
<td>(0.278)</td>
</tr>
<tr>
<td>Ages 15-18</td>
<td>0.633</td>
<td>0.665</td>
<td>0.601</td>
</tr>
<tr>
<td></td>
<td>(0.482)</td>
<td>(0.472)</td>
<td>(0.489)</td>
</tr>
<tr>
<td>Ages 19-23</td>
<td>0.309</td>
<td>0.319</td>
<td>0.300</td>
</tr>
<tr>
<td></td>
<td>(0.462)</td>
<td>(0.466)</td>
<td>(0.458)</td>
</tr>
</tbody>
</table>

Bratti and Mendola (2014) also examine the effect of parental illness on children living with both parents currently alive but their sample is comprised only of older children aged 15-24. The purpose is to focus on parental health only and avoid the effect of parents’ death and single parents, so only households with both parents alive are included.
Notes: The table provides the children’s school enrolment rate over three survey rounds. The full sample includes children aged 10-23 and 13,930 observations over all three rounds of survey. Standard deviations are in parentheses.

Table 5.4 presents the proportion of children in the sample who are working and the average number of hours of work per week by parental health status. As expected, children who have a sick parent are more likely to work and tend to work longer hours. Interestingly, boys seem more likely to work than girls. However, this can partially be due to the fact that girls are more likely to work in the household and therefore their work is not adequately captured in the data.

### Table 5.4 - Children’s labour force participation by parental health status

<table>
<thead>
<tr>
<th>Labour force participation rate</th>
<th>Total</th>
<th>No parent ill</th>
<th>Any parent ill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full sample</td>
<td>35.53</td>
<td>31.53</td>
<td>39.55</td>
</tr>
<tr>
<td>Boys</td>
<td>38.70</td>
<td>34.81</td>
<td>42.53</td>
</tr>
<tr>
<td>Girls</td>
<td>31.80</td>
<td>27.75</td>
<td>35.95</td>
</tr>
<tr>
<td>Ages 10-14</td>
<td>13.95</td>
<td>11.24</td>
<td>16.90</td>
</tr>
<tr>
<td>Ages 15-18</td>
<td>39.20</td>
<td>34.14</td>
<td>44.19</td>
</tr>
<tr>
<td>Ages 19-23</td>
<td>63.95</td>
<td>61.90</td>
<td>65.86</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Weekly working hours</th>
<th>Total</th>
<th>No parent ill</th>
<th>Any parent ill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full sample</td>
<td>15.96 (23.42)</td>
<td>14.38 (22.86)</td>
<td>17.54 (23.87)</td>
</tr>
<tr>
<td>Boys</td>
<td>17.79 (24.32)</td>
<td>16.21 (23.95)</td>
<td>19.35 (25.59)</td>
</tr>
<tr>
<td>Girls</td>
<td>13.79 (22.12)</td>
<td>12.28 (21.36)</td>
<td>15.36 (22.77)</td>
</tr>
<tr>
<td>Ages 10-14</td>
<td>4.43 (12.23)</td>
<td>3.67 (11.46)</td>
<td>5.26 (12.97)</td>
</tr>
<tr>
<td>Ages 15-18</td>
<td>16.79 (23.07)</td>
<td>14.84 (22.44)</td>
<td>18.72 (23.52)</td>
</tr>
<tr>
<td>Ages 19-23</td>
<td>32.59 (26.48)</td>
<td>31.63 (26.69)</td>
<td>33.47 (26.27)</td>
</tr>
</tbody>
</table>

Notes: The table provides children’s labour force participation over three survey rounds. The full sample includes children aged 10-23 and 13,930 observations over all three rounds of survey. Standard deviations are in parentheses.

### Table 5.5 - Sample descriptive statistics- independent variables

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Sample Average</th>
<th>Sample standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mother ill only</td>
<td>0.179</td>
<td>(0.383)</td>
</tr>
<tr>
<td>Father ill only</td>
<td>0.114</td>
<td>(0.318)</td>
</tr>
<tr>
<td>Both parents ill</td>
<td>0.204</td>
<td>(0.403)</td>
</tr>
<tr>
<td>Age</td>
<td>15.822</td>
<td>(3.704)</td>
</tr>
<tr>
<td>Male</td>
<td>0.541</td>
<td>(0.498)</td>
</tr>
<tr>
<td>Kinh (main ethnic group)</td>
<td>0.789</td>
<td>(0.408)</td>
</tr>
<tr>
<td>Urban</td>
<td>0.210</td>
<td>(0.407)</td>
</tr>
<tr>
<td>Age of mother</td>
<td>43.108</td>
<td>(7.073)</td>
</tr>
</tbody>
</table>
Table 5.5 presents the descriptive statistics of all independent variables included in the model. Around 18% of mothers and 11% of fathers are classified as ill in the estimation sample, according to the definition above. Not surprisingly, the average level of parental education is quite low in the estimation sample, and there are a substantial proportion of parents (more than 20%) without any educational qualification. The major ethnic group (Kinh) includes about 80% of the estimation sample, and about 20% of the sample lives in urban areas.

5.4 Methodology and Estimation

The analysis begins with the impact of parental health on children’s outcomes by estimating a simple linear model:
\[ c_{ijt} = \beta_0 + \beta_1 H_{ijt} + \beta_2 X'_{ijt} + \varepsilon_{it} \] (5.1)

Where \( c_{ijt} \) represents the outcome for child \( i \) (school enrolment; labour force participation or working hours) living in household \( j \) at time \( t \);

\( H_{ijt} \) represents the parental health shock;

\( X'_{ijt} \) represents a vector of control variables including individual and household characteristics; and

\( \varepsilon_{it} \) is an idiosyncratic individual error term normally distributed with variance normalised to be equal to 1.

The parameter of interest is \( \beta_1 \) that captures the effect of parental health shock on children’s outcomes.

An important challenge in the estimation is the possible endogeneity of parental health shocks due to unobserved variables that might affect both parental health status and children’s outcomes, and therefore could bias estimation results. For example, parents with poor living conditions are more likely to have poor health and, at the same time, invest less time and money in their children’s health and education. To address the endogeneity problem, this study estimates the model using child fixed effects and therefore takes into account unobserved time-invariant factors that might affect both parental health and children’s outcomes:

\[ c_{ijt} = \beta_0 + \beta_1 M_{ijt} + \beta_2 F_{ijt} + \beta_3 P_{ijt} + \beta_4 X'_{ijt} + \gamma_i + \varepsilon_{it} \] (5.2)

In order to capture the specific differences between the effect of maternal and parental illness, this study follows Bratti and Mendola (2014) and separately includes in the model each parent’s health status:

\[ c_{ijt} = \beta_0 + \beta_1 M_{ijt} + \beta_2 F_{ijt} + \beta_3 P_{ijt} + \beta_4 X'_{ijt} + \gamma_i + \varepsilon_{it} \] (5.3)

Where \( M_{ijt} \) represents a mother’s poor health status;

\( F_{ijt} \) represents a father’s illness;

\( P_{ijt} \) equals 1 if both parents report poor health; and
\( \gamma_i \) is a child fixed effect.

The model includes an extensive set of control variables, such as the demographic characteristics of parents and children (child’s age and ethnicity, mother’s and father’s age and educational attainments), the demographic structure of the household (household size, number of children), proxies for household wealth (house ownership, number of houses, living area, availability of safe water, computers), per capita income of the household, region of current residence and time fixed effects\(^{15}\). These time and region fixed effects will capture macroeconomic and local conditions, such as the local provision of education, health services and labour market conditions.

As already mentioned, the child fixed effect \( \gamma_i \) will capture the effect of all observable and unobservable time-invariant characteristics associated with both parents’ health and children’s outcomes, such as birth order, gender, religion, personality traits and time preferences, plus any other unobserved time-invariant factors. Therefore, the time-invariant controls are excluded in fixed effects estimations because they are perfectly absorbed by the child fixed effects.

In the fixed effects model, the causal interpretation of results relies on the assumption that the time-dependent error term is independent of changes in parental health, conditional on the regressors included in the model and on the individual fixed effect. This assumption will not hold if there are unobserved yearly random shocks that affect parental health and children’s outcomes at the same time. For this reason, the study controls for a wide set of individual and family characteristics, such as the child’s health, and several indicators of socio-economic status.

In other words, identification comes from time variations in parental health for the same individual, and the underlying assumption of the model is that these variations are exogenous, conditional on the set of independent variables. In practice, children living in specific conditions may be systematically more likely to have sick parents, but the time variation in parental health should not be related to children’s outcomes (Bratti and Mendola 2014).

---

\(^{15}\) The study also estimates models that include regions by time fixed effects to capture time-varying region-level shocks (see Table 5.9 in the next section).
5.5 Results

5.5.1 Parental health and children education
This section presents the main empirical findings of the investigation of the impact of parental health on children’s outcomes. Only coefficients of main regressors in regressions are presented; completed estimates can be found in Appendix C. I begin by presenting results of the impact of parental illness using a binary indicator equal to 1 if any parent reported illness, and then distinguishing between maternal and paternal health shocks.

Table 5.6 reports estimate results of the impact of parental health shocks on children’s enrolment in education, using the full sample of children aged 10-23. In the first column, the OLS results are presented. Columns 2-5 present results from the model including child fixed effects, with a progressively more extended set of independent variables. The choice of regressors follows the literature and, in particular, Bratti and Mendola (2014). Initially a simple model is estimated including child’s illness, and the month and year of the survey (model 2). In model 3, child’s age, parental education and age are also included. Model 4 further controls for other demographic characteristics of the family, such as the number of children and family members, as well as some proxies for household wealth (home ownership, size of the house, access to clean water, computer ownership) and an indicator of whether the family lives in an urban area. This is the preferred specification. Model 5 also controls for per capita household income in the previous year. The study follows Bratti and Mendola (2014) and does not control for parental employment in the main specification because this variable is potentially endogenous and likely to be affected by parental health. Therefore, the estimates will capture the pecuniary and non-pecuniary effects of parental health shocks on children’s outcomes. However, a sensitivity test was run including binary indicators of parental employment and the main results are unchanged.

Results are very stable and consistent across the different specifications of the model. Children with an unhealthy parent are around 2 percentage points less likely to be enrolled in school than children having healthy parents (on average, 66% of children in the sample are enrolled in education) and this result is stable when additional control variables are included. Interestingly, estimates from model 5 show that the effect of
parental illness on child school enrolment is not affected by inclusion of per capita household income.

Table 5.6 - Effect of parental health shocks on child school enrolment (children aged 10-23)

<table>
<thead>
<tr>
<th></th>
<th>OLS (1)</th>
<th>FE (2)</th>
<th>FE (3)</th>
<th>FE (4)</th>
<th>FE (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parental health shock (any parent)</td>
<td>-0.0165</td>
<td>-0.0223</td>
<td>-0.0225</td>
<td>-0.0232</td>
<td>-0.0232</td>
</tr>
<tr>
<td></td>
<td>(0.0066)</td>
<td>(0.0089)</td>
<td>(0.0089)</td>
<td>(0.0089)</td>
<td>(0.0089)</td>
</tr>
<tr>
<td>No. of obs</td>
<td>13,930</td>
<td>13,930</td>
<td>13,930</td>
<td>13,930</td>
<td>13,930</td>
</tr>
<tr>
<td>No. of children</td>
<td>7,380</td>
<td>7,380</td>
<td>7,380</td>
<td>7,380</td>
<td>7,380</td>
</tr>
<tr>
<td>No. of households</td>
<td>3,720</td>
<td>3,720</td>
<td>3,720</td>
<td>3,720</td>
<td>3,720</td>
</tr>
</tbody>
</table>

Notes: Column 1 reports the OLS results; columns 2-5 represent child fixed effect estimations of the effect of parental health shocks on the probability of children’s schooling. All models control for children’s illness, month of interview and the round of the survey. Model (3) additionally controls for child’s age, age and education of parents. Model (4) controls for the household’s demographic structure as well as household’s wealth indicators. Model (5) adds control for per capita income in the household. Standard errors are in brackets. * denotes statistical significance at 10 percent; ** denotes significance at 5 percent; *** denotes significance at 1 percent. In additional, model (1) also controls for sex, ethnicity and region fixed effects. The coefficient of other control variables is reported in Appendix C.

In Table 5.7, I use separate indicators variables for maternal and paternal illness (maternal health shock, paternal health shock and both parents’ health shock) in order to further investigate the impact of parental health shocks (see equation 5.3 in the previous section) and to allow for non-linearity in the impact of parental health (Bratti and Mendola 2014). Interestingly, results show that the impact of maternal poor health is stronger than the effect of paternal health. The coefficient of mother’s and father’s illness remains unchanged when including the set of control variables from columns 2 to 5. Children living with an unhealthy mother are over 3 percentage points less likely to be enrolled in education, and the effect is statistically significant at 1%. However, the effect of paternal health shock is smaller (-2%) and only significant at 10%.

This finding is consistent with results from the existing literature on parental health shocks in general and children’s outcomes, such as Bratti and Mendola (2014) and Chen et al. (2009), as well as with findings from economics literature showing that investments in children are primarily made by mothers (see for example Case and Paxson (2001)). The negative effect can be driven by a decrease in maternal time and inputs in the child’s education, as well as reallocation of the child’s time from education to work, either in the labour market or at home to care for the sick mother.
The impact of having both parents ill is not significantly different from zero in the fixed effects estimation. This result is similar to findings from Bratti and Mendola (2014) and could be explained by the lack of variation in the binary variable indicating that both parents were sick, as well as by the high level of measurement error, and therefore a higher attenuation bias for the estimator.

Table 5.8 further explores the impact of parental illness on child schooling by gender and age group of the children. For reasons of parsimony, I focus on the preferred specification, which is the fixed effects model with controls for individual and parental demographics, as well as proxies for household wealth (other results are available on request). The first three columns report the results on the effect of parental illness on child school enrolment by the age of the child, while the two remaining columns report the results by gender. Girls are clearly more affected than boys by parental health shocks, and their chances to be enrolled in education decrease by about 4-5 percentage points because of parental illness. Further, children aged 15-18 (in upper secondary school) are the most vulnerable to maternal poor health status and their probability to be enrolled in education is reduced by 5.5 percentage points. As already discussed, these results can be driven by several transmission channels for children in this age group, such as the negative impact of maternal absence and lack of supervision, as well as reduced resources to fund education.

Table 5.7 - Effect of maternal and parental health shocks on child school enrolment (children aged 10-23)

<table>
<thead>
<tr>
<th></th>
<th>OLS (1)</th>
<th>FE (2)</th>
<th>FE (3)</th>
<th>FE (4)</th>
<th>FE (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal health shock</td>
<td>-0.0351*</td>
<td>-0.0313**</td>
<td>-0.0315**</td>
<td>-0.0320***</td>
<td>-0.0320***</td>
</tr>
<tr>
<td></td>
<td>(0.0087)</td>
<td>(0.0112)</td>
<td>(0.0112)</td>
<td>(0.0112)</td>
<td>(0.0112)</td>
</tr>
<tr>
<td>Paternal health shock</td>
<td>-0.0260**</td>
<td>-0.0254*</td>
<td>-0.0264*</td>
<td>-0.0259*</td>
<td>-0.0259*</td>
</tr>
<tr>
<td></td>
<td>(0.0104)</td>
<td>(0.0137)</td>
<td>(0.0137)</td>
<td>(0.0137)</td>
<td>(0.0137)</td>
</tr>
<tr>
<td>Both parents’ health shocks</td>
<td>0.0124</td>
<td>-0.0060</td>
<td>-0.0055</td>
<td>-0.0075</td>
<td>-0.0075</td>
</tr>
<tr>
<td></td>
<td>(0.0092)</td>
<td>(0.0125)</td>
<td>(0.0125)</td>
<td>(0.0125)</td>
<td>(0.0126)</td>
</tr>
<tr>
<td>N. of obs</td>
<td>13,930</td>
<td>13,930</td>
<td>13,930</td>
<td>13,930</td>
<td>13,930</td>
</tr>
</tbody>
</table>

Notes: Control variables are reported in Table 5.6. Standard errors are in brackets. * denotes statistical significance at 10 percent; ** denotes significance at 5 percent; *** denotes significance at 1 percent.
Table 5.8 - Effect of parental health shocks on child school enrolment by age group and gender of the child

<table>
<thead>
<tr>
<th></th>
<th>Age 10-14</th>
<th>Age 15-18</th>
<th>Age 19-23</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal health shock</td>
<td>-0.0035</td>
<td>-0.0545**</td>
<td>-0.0381</td>
<td>-0.0466***</td>
<td>-0.0171</td>
</tr>
<tr>
<td></td>
<td>(0.0129)</td>
<td>(0.0276)</td>
<td>(0.0318)</td>
<td>(0.0161)</td>
<td>(0.0157)</td>
</tr>
<tr>
<td>Paternal health shock</td>
<td>0.0090</td>
<td>-0.0370</td>
<td>-0.0486</td>
<td>-0.0398**</td>
<td>-0.0148</td>
</tr>
<tr>
<td></td>
<td>(0.0164)</td>
<td>(0.0348)</td>
<td>(0.0397)</td>
<td>(0.0201)</td>
<td>(0.0191)</td>
</tr>
<tr>
<td>Both parents’ health shocks</td>
<td>0.0090</td>
<td>-0.0381</td>
<td>-0.0230</td>
<td>-0.0111</td>
<td>-0.0063</td>
</tr>
<tr>
<td></td>
<td>(0.0152)</td>
<td>(0.0311)</td>
<td>(0.0347)</td>
<td>(0.0183)</td>
<td>(0.0174)</td>
</tr>
<tr>
<td>No of obs</td>
<td>5,604</td>
<td>4,678</td>
<td>3,648</td>
<td>6,393</td>
<td>7,537</td>
</tr>
</tbody>
</table>

Notes: Control variables are reported in Table 5.6. Standard errors are in brackets. * denotes statistical significance at 10 percent; ** denotes significance at 5 percent; *** denotes significance at 1 percent.

This section runs several sensitivity tests to verify the stability of my results. In particular, since child fixed effects only can capture the impact of all time-invariant characteristics, fixed effects results will not be valid if there are any time-varying yearly shocks to parental health that also affect the child schooling at the same time. For this reason, the model is tested by controlling for other family members’ illnesses (specifically for grandparents and young siblings’ health shocks, and other members in the household), as well as controlling for a series of interactions between survey wave and region of residence. These time-region fixed effects should capture time-specific local shocks as well as local provision of health and education facilities. Lastly, the estimation sample also is expanded to include single parent families and it includes one indicator variable for single parent status. All results are reported in Table 5.9 and are consistent with findings presented in the previous tables. There is no significant change in the effect of parental illness on child school enrolment with the inclusion of any of these controls.

Table 5.9 - Sensitivity analyses for the impact of parental health shocks on child’s education (Estimation with FE)

<table>
<thead>
<tr>
<th></th>
<th>Ctrl for illness of children</th>
<th>Ctrl for grandparents’ illness</th>
<th>Ctrl for other family members’ illness</th>
<th>Ctrl for time-region fixed effects</th>
<th>Incl. single parent indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal health shock</td>
<td>-0.0324***</td>
<td>-0.0318***</td>
<td>-0.0317***</td>
<td>-0.0330***</td>
<td>-0.0305***</td>
</tr>
<tr>
<td></td>
<td>(0.0112)</td>
<td>(0.0112)</td>
<td>(0.0112)</td>
<td>(0.0112)</td>
<td>(0.0108)</td>
</tr>
</tbody>
</table>
I also investigate whether the impact of parental illness on children’s schooling varies when I control for parental health insurance status (results are available on request). The impact of parental health on children’s schooling remains stable and negative, irrespective of insurance status. Previous literature has documented the limited effect of health insurance in protecting families from the negative consequences of health shocks in Vietnam. For example, the findings of Xu et al. (2003) have showed that when families experience a health shock, out-of-pocket health expenditures are very high, even for insured individuals. This may be because the public health insurance system is not efficient and many insured individuals who are covered by the public health scheme still have to use private healthcare services, or because the level of financial protection of the public health insurance scheme is low and individuals have to spend a substantial amount of money on informal payments and other indirect costs (Sepehri et al. 2006). This issue has been further investigated by estimating the effect of parental health shocks on household health expenditure in the sample. Results show a substantial increase in health expenditure as a consequence of negative health shocks, in line with findings in earlier studies in Vietnam (Wagstaff 2007; Mitra et al. 2015). And (Xu et al. 2003) also show that Vietnam is a country that experiences among the highest levels of catastrophic health expenditures in the world, with out-of-pocket health expenditures being the main financial source for health care. This increase in health expenditures could have a direct negative effect on child school enrolment, through the need to reduce educational expenditures. Similarly, parental health shocks (and the related increase in health expenditures) can have an indirect effect on child’s education through child labour channels by creating an incentive for the child to leave education and enter the labour market. For this reason, in the next section, the impact of parental health shocks on children’s work is explored.

<table>
<thead>
<tr>
<th></th>
<th>Paternal health shock</th>
<th>Both parents’ health shocks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.0261*</td>
<td>-0.0080</td>
</tr>
<tr>
<td></td>
<td>(0.0137)</td>
<td>(0.0098)</td>
</tr>
<tr>
<td></td>
<td>-0.0258*</td>
<td>-0.0069</td>
</tr>
<tr>
<td></td>
<td>(0.0137)</td>
<td>(0.0126)</td>
</tr>
<tr>
<td></td>
<td>-0.0259*</td>
<td>-0.0072</td>
</tr>
<tr>
<td></td>
<td>(0.0145)</td>
<td>(0.0126)</td>
</tr>
<tr>
<td></td>
<td>-0.0266*</td>
<td>-0.0095</td>
</tr>
<tr>
<td></td>
<td>(0.0137)</td>
<td>(0.0126)</td>
</tr>
<tr>
<td></td>
<td>-0.0204</td>
<td>-0.0035</td>
</tr>
<tr>
<td></td>
<td>(0.0131)</td>
<td>(0.0120)</td>
</tr>
</tbody>
</table>

Notes: Other independent variables are reported in Table 5.6. Standard errors are in brackets. * denotes statistical significance at 10 percent; ** denotes significance at 5 percent; *** denotes significance at 1 percent.
5.5.2 Parental health and children work

The VHLSS data include available information on employment status and number of working hours of all family members. I extend the existing literature on parental health and children outcomes and use both measures of child labour to explore whether parental health shocks affect child labour. Therefore, this study is able to investigate the effect of parental health shocks on labour force participation as well as on the intensity of engagement in the labour market.

The estimates are reported in Tables 5.10-5.12 and show that children living with a sick parent are significantly more likely to enter into employment (+4%, where the sample average is 36%) and work longer hours (+1 hour or 4.5% of a standard deviation). Table 5.11 shows that the impact of maternal illness is slightly stronger than the effect of a paternal health shock, both on labour force participation and the number of working hours. Maternal illness significantly increases the likelihood of entering the labour market with an increase of over 5 percentage points and the number of weekly working hours increase by 1.6 (around 7% of a standard deviation). These effects remain unchanged with inclusion of different control variables as in models (2) and (3).

When the results are analysed by age and gender of the child in Table 5.12, results show that girls are more likely to enter the labour market after a health shock for both parents, while boys are significantly affected only in the case of maternal illness.

Further, as expected, the effect is stronger for children in upper secondary school. Interestingly, paternal illness seems to have a stronger effect on child labour than on child education, especially for girls and children aged 15-18.

Table 5.10 - Effect of parental health shocks on children’s labour force participation (children aged 10-23)

<table>
<thead>
<tr>
<th>Parental health shock (any parent)</th>
<th>Labour force participation</th>
<th>Weekly working hours</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS (1)</td>
<td>FE (2)</td>
</tr>
<tr>
<td></td>
<td>0.0413***</td>
<td>0.0382***</td>
</tr>
<tr>
<td></td>
<td>(0.0071)</td>
<td>(0.0094)</td>
</tr>
<tr>
<td>No. of obs</td>
<td>13,930</td>
<td>13,930</td>
</tr>
</tbody>
</table>

Notes: All models control for children’s illness, month of interview and the round of survey. Model (1) also controls for sex, ethnicity and region fixed effects. Model (2) additionally controls for child’s age, age and education of parents. Model (3) additionally controls for household demographic structure and household wealth indicators. Standard errors are in brackets. * denotes statistical significance at 10 percent; ** denotes significance at 5 percent; *** denotes significance at 1 percent.
Table 5.11 - Effect of maternal and parental health shocks on children’s labour force participation (children aged 10-23)

<table>
<thead>
<tr>
<th>Labour force participation</th>
<th>Weekly working hours</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS (1)</td>
</tr>
<tr>
<td>Maternal health shock</td>
<td>0.0520***</td>
</tr>
<tr>
<td></td>
<td>(0.0093)</td>
</tr>
<tr>
<td>Paternal health shock</td>
<td>0.0337***</td>
</tr>
<tr>
<td></td>
<td>(0.0111)</td>
</tr>
<tr>
<td>Both parents’ health shocks</td>
<td>0.0346***</td>
</tr>
<tr>
<td></td>
<td>(0.0098)</td>
</tr>
<tr>
<td>No. of obs</td>
<td>13,930</td>
</tr>
</tbody>
</table>

Notes: All models control for children’s illness, month of interview and the round of survey. Model (1) also controls for sex, ethnicity and region fixed effects. Model (2) additionally controls for child’s age, age and education of parents. Model (3) additionally controls for household demographic structure and household wealth indicators. Standard errors are in brackets. * denotes statistical significance at 10 percent; ** denotes significance at 5 percent; *** denotes significance at 1 percent.

Table 5.12 - Effect of parental health shocks on children’s labour force participation, by age group and gender of the child

<table>
<thead>
<tr>
<th>Labour force participation</th>
<th>Ages 10-14</th>
<th>Ages 15-18</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal health shock</td>
<td>0.0508***</td>
<td>0.0556</td>
<td>0.0453***</td>
<td>0.0595***</td>
</tr>
<tr>
<td></td>
<td>(0.0190)</td>
<td>(0.0292)</td>
<td>(0.0174)</td>
<td>(0.0164)</td>
</tr>
<tr>
<td>Paternal health shock</td>
<td>0.0215</td>
<td>0.0805**</td>
<td>0.0598***</td>
<td>0.0240</td>
</tr>
<tr>
<td></td>
<td>(0.0237)</td>
<td>(0.0367)</td>
<td>(0.0216)</td>
<td>(0.0199)</td>
</tr>
<tr>
<td>Both parents’ health shocks</td>
<td>0.0302</td>
<td>0.0140</td>
<td>0.0072</td>
<td>0.0314*</td>
</tr>
<tr>
<td></td>
<td>(0.0219)</td>
<td>(0.0328)</td>
<td>(0.0197)</td>
<td>(0.0181)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Weekly working hours</th>
<th>Ages 10-14</th>
<th>Ages 15-18</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal health shock</td>
<td>0.631</td>
<td>2.399</td>
<td>1.483**</td>
<td>1.832**</td>
</tr>
<tr>
<td></td>
<td>(0.656)</td>
<td>(1.338)</td>
<td>(0.796)</td>
<td>(0.767)</td>
</tr>
<tr>
<td>Paternal health shock</td>
<td>-0.201</td>
<td>3.787**</td>
<td>2.316**</td>
<td>0.867</td>
</tr>
<tr>
<td></td>
<td>(0.819)</td>
<td>(1.684)</td>
<td>(0.992)</td>
<td>(0.930)</td>
</tr>
<tr>
<td>Both parents’ health shocks</td>
<td>0.246</td>
<td>-0.087</td>
<td>-0.717</td>
<td>0.805</td>
</tr>
<tr>
<td></td>
<td>(0.759)</td>
<td>(1.504)</td>
<td>(0.905)</td>
<td>(0.847)</td>
</tr>
<tr>
<td>No. of obs</td>
<td>5,604</td>
<td>4,678</td>
<td>6,393</td>
<td>7,537</td>
</tr>
</tbody>
</table>

Notes: Control variables are reported in Table 5.6. Standard errors are in brackets. * denotes statistical significance at 10 percent; ** denotes significance at 5 percent; *** denotes significance at 1 percent.
Therefore, results of this study show that while parental illness has a significant negative effect on schooling, the study also finds strong evidence that parental illness increases child labour, both the likelihood of entering the labour market and the working hours. These findings are broadly consistent with Bratti and Mendola (2014), but are in contrast to Alam (2015) who shows that a father’s illness has a significant impact on school attendance of children and that this effect is not due to increased child labour. However, it is very hard to disentangle the reasons for these differences which can be attributed to the different educational and institutional contexts of the two countries (Bosnia and Tanzania).

Furthermore, the study also provides evidence that parental illness has different effects on child labour as well as schooling by gender and age of the child. Children at ages 15-18 are the most vulnerable group and girls are more susceptible to an unexpected parental health shock than boys. This is consistent with the literature on gender differences, especially studies in developing countries that show that girls often have to work more and receive less educational opportunities than boys. This may be due to increased responsibilities in caring for family members which are typically traditional roles for females. It may also be due to resource allocation within the household that favours boys over girls (Pitt and Rosenzweig 1990; Strauss and Thomas 1995; Rose 2000).

Overall, these results confirm that parental illness has a very strong negative effect on children and substantially decreases enrolment in education while increasing child labour. These results are particularly pronounced for girls who seem to experience the most harmful effects from parental health shocks. The negative effect of parental illness can be mediated through an increase in health expenditure, that reduces family income and funds available for education and at the same time may increase the need for the children to be in employment to replace a sick parent.

The impact of poor parental health on children’s outcomes could have a long term impact, larger and is even more serious if the rate of return to schooling is high. In Vietnam, empirical studies are generally consistent and show a significant increasing trend in the rates of return to education over time (Gallup 2002; Doan and Gibson 2010; Phan and Coxhead 2013). The rate of return on one-year increase in education almost double from 2.9 per cent in 1993 to 5.0 per cent in 1998 (Gallup 2002). The rates of
return to education increased significantly and consistently over the period of 1998-2008 from 3.8 per cent to 10 per cent in 2008 (Doan and Gibson 2010). However, it is argued that both domestic policies and foreign trade liberalization have an impact on wages and after taking into account factors such as wage effects and employment structure effects of open door policies and foreign trade liberalization, results show a increasing trend in rates of return to education but generally still low (Oostendorp and Doan 2013).

5.6 Conclusion
Health shocks are important unpredictable events that can undermine the emotional and financial well-being of families, especially in developing countries where citizens have limited access to social protection and health insurance. However, the impact of parental illness on children’s outcomes has received very little attention in economics literature.

This study explores the impact of parental health shocks on child school enrolment and labour force participation in Vietnam, using unbalanced panel data from the VHLSS. Child fixed effects are used in order to address potential endogeneity arising from the relationship between parental health shocks and children’s outcomes and control for unobserved characteristics that do not vary over time. This study contributes to the recent limited economic literature on the effect of parental health shocks on child education in developing countries, by providing new evidence based on Vietnam, which can be particularly meaningful for analysing several other Asian developing countries, with similar education and health systems. Vietnam is a middle income developing country which has experienced substantial economic growth in recent years, and has reached important development milestones, but it still lacks provision of adequate education and social protection to some certain vulnerable groups in its population, especially ethnic minorities and rural families at the bottom end of the income distribution.

The results show that maternal health shocks substantially decrease a child’s likelihood to be enrolled in education by about 3 percentage points. Further, parental health shocks increase children’s employment probabilities as well as the number of weekly hours of work. The results are very stable across different specifications of the model and several sensitivity tests, including a very detailed set of independent variables. The negative effects are particularly pronounced for girls and for children aged 15-18.
It is clear that parental investment on both time and money inputs are especially important for children’s education attainment and that poor parental health has significant negative impacts, and even it could have a long term impact on children’s outcome. The findings of this study have important policy implications. It may suggest that the policy implications should target at improving women’s health - a crucial source of children’s educational achievement. In particular, the implementation of specific policies aimed at improving: access to health care for women; the social insurance system; and, more broadly, maternal health, could all have substantial spillover benefits for children’s education and labour market outcomes. In addition, the implementation of an adequate system of social protection and an improved access to healthcare services for women as well as an improved social insurance system with better efficiency may be suggestive solutions to reduce the intergeneration transmission costs from parental health shocks to offspring. Further, specific care could be intensified for rural households or families with middle to low income who are particularly vulnerable to health shocks and whose children are at a high risk of dropping out of education and entering the labour market. The focus of policies should aim at the rural, remote, mountainous areas and disadvantaged socioeconomic groups where living condition, health centers, healthcare facilities, as well as healthcare skilled staffs are still poor, lack and backward. The government can enhance more investment in healthcare and education, especially health and education facilities in these areas. Important policies to encourage skilled people to willing to work in remote, mountainous and disadvantaged socioeconomic areas are especially needed. In some cases, the government may consider providing financial assistance such as reduction or exemption school fees, providing vouchers of buying books and uniforms for children of households having severe health shocks or providing educational loans for these households. These educational loans will bring positive impacts on both children as well as the society.
Chapter 6. Conclusions

This thesis studies the effect of family circumstances on outcomes of individual family members in Vietnam. It first develops a conceptual framework of analysing various factors affecting FLFP and examining the empirical test on the derived hypothesis on the relationship between fertility and female labour supply with instrumental variable strategy based on the existence of a parental son preference. In the second part of the thesis, the impact of parental health shocks on education and engagement in the labour market of Vietnamese children is explored.

The theoretical chapter develops a conceptual framework and introduces a model of female labour supply. Hypotheses on determinants of female labour supply are obtained from analysing a model of female utility maximisation. These hypotheses represent the impacts of various factors from both the individual, their family characteristics and labour market conditions on the female’s level of labour force participation.

Next, the two empirical chapters use two different econometric methodologies and two different datasets. The relationship between fertility and FLFP is analysed in the first empirical chapter. The IV methodology is applied in order to estimate the caused effect of fertility on female labour supply. This study exploits the parental preference for sons in Vietnam and uses the sex composition of the first two children as an IV for having a third child among families with two or more children. Parents whose first two children are girls are more likely to have a third child than families whose first two children are boys or who have one boy and one girl.

The empirical results show that the presence of an additional child has a negative effect on FLFP and this effect may vary with the educational attainments of mothers and fathers. Findings suggest that reductions in fertility is one of the important channels to increase female labour supply but it is not the only channel. Other factors such as greater educational attainments, reductions in gender discrimination in the workplace and empowerment of women could be driving both trends. Thus, understanding the linkage between variations in fertility and female labour supply is important for predicting changes in the structure of the labour force in the future as well as identifying potential sources of gender wage differentials. Apart from attempts to decrease fertility,
policy implications should aim at investment into female education to improve their position and empowerment in society.

The second empirical chapter analyses the impact of parental health shocks on children’s educational and labour market outcomes in Vietnam, using panel data from the VHLSS. The study addresses the potential endogeneity problems arising from the relationship between parental health shocks and children’s engagement in education and the labour market by employing panel data with a child fixed effects estimator. This study contributes to the recent limited economic literature on the effect of parental health shocks on child education in developing countries by providing new evidence based on Vietnam which could be particularly meaningful for analysing other developing Asian countries with similar education and health systems.

The findings of this chapter indicate that maternal health shocks substantially decrease the probability of a child’s school enrolment by about 3 percentage points. Furthermore, parental health shocks may lead to an increase in children’s likelihood to enter employment as well as increasing the number of hours worked, leading to a reallocation of their time from school to work. The results are very stable across different specifications of the model and several sensitivity tests, including a very detailed set of independent variables. The study also suggests that the impact of parental illness on children’s education and labour market outcome varies with age groups and gender. Children aged 15-18 are the most vulnerable age group and girls appear to be more vulnerable to the negative effects of parental health events.

These findings suggest policy implications targeting at improvement of women’s health. Improved access to healthcare services for women as well as a more efficient social insurance system may be needed to reduce intergenerational transmission costs from parental health shocks to offspring. Additionally, policies should be aimed at low income households, individuals, who are particularly vulnerable to health shocks.

Overall this thesis has made a contribution to the empirical literature on the effects of family circumstances on individual outcomes in general and in particular in the case of Vietnam. At the same time it must be acknowledged that much more empirical evidence must be accumulated before well devised welfare policies can be implemented. This will remain an important area for future research on this topic.
Appendix A. Proof for propositions of female labour supply

The first order condition for the maximum expected utility is:

\[
\frac{\partial E\tilde{u}}{\partial L} = \frac{\partial p(.)}{\partial L} \left[u(y,l) - u(y_{p},l)\right] + p(L,n,z_{j},z_{i},v) \left[\frac{\partial u(.)}{\partial y} \left(1 - \delta\right)w_{m}(e) - \frac{\partial c(.)}{\partial L} - \frac{\partial u(.)}{\partial l}\right] = 0
\]  

(1)

1. The effect of the number of children

Total differentiation of the first order condition (1) with respect to \(n\), we have:

\[
\frac{\partial^{2} E\tilde{u}}{\partial L^{2}} dL + \frac{\partial p(.)}{\partial L} \left[u(y,l) - u(y_{p},l)\right] + p(L,n,z_{j},z_{i},v) \left[\frac{\partial^{2} u(.)}{\partial y^{2}} \left(1 - \delta\right)w_{m}(e) - \frac{\partial c(.)}{\partial L} - \frac{\partial u(.)}{\partial l}\right] = 0
\]

\[
\Rightarrow \frac{\partial^{2} E\tilde{u}}{\partial L^{2}} dL + \frac{\partial p(.)}{\partial L} \left[u(y,l) - u(y_{p},l)\right] + p(L,n,z_{j},z_{i},v) \left[\frac{\partial^{2} u(.)}{\partial y^{2}} \left(1 - \delta\right)w_{m}(e) - \frac{\partial c(.)}{\partial L} - \frac{\partial u(.)}{\partial l}\right] = 0
\]

Recalling that \(\frac{\partial^{2} E\tilde{u}}{\partial L^{2}} < 0\), this then leads to the following proposition:

\[
\frac{dL}{dn} > 0 \text{ if } \frac{dL}{dn} < 0
\]
\[
\begin{align*}
&\left\{ \frac{\partial p(.)}{\partial n} \left[ \frac{\partial u(.)}{\partial y} \left( 1-\delta \right) w_m(e) - \frac{\partial c(.)}{\partial L} \right] - \frac{\partial u(.)}{\partial L} \right\} \\
&+ p(L, n, z_j, z_i, v) \left[ - \frac{\partial^2 u(.)}{\partial y^2} \frac{\partial c(.)}{\partial n} \left( 1-\delta \right) w_m(e) - \frac{\partial c(.)}{\partial L} \right] - \frac{\partial p(.)}{\partial L} \frac{\partial u(.)}{\partial y} \frac{\partial c(.)}{\partial n} \right\} \geq 0 (3)
\end{align*}
\]

2. The effect of grandparents

Total differentiation of the first order condition (1) with respect to \( s \), we have:

\[
\frac{\partial^2 E\tilde{u}}{\partial L^2} dL + \frac{\partial^2 E\tilde{u}}{\partial s^2} ds = 0
\]

\[
\Leftrightarrow \frac{\partial^2 E\tilde{u}}{\partial L^2} dL + \left\{ \frac{\partial p(.)}{\partial L} \left( - \frac{\partial u(.)}{\partial \bar{y}} \right) \frac{\partial c(.)}{\partial s} + p(L, n, z_j, z_i, v) \frac{\partial^2 u(.)}{\partial y^2} \frac{\partial c(.)}{\partial s} \left( 1-\delta \right) w_m(e) - \frac{\partial c(.)}{\partial L} \right\} ds = 0
\]

\[
\Rightarrow \frac{dL}{ds} = - \left\{ \frac{\partial p(.)}{\partial L} \frac{\partial u(.)}{\partial \bar{y}} \frac{\partial c(.)}{\partial s} - p(L, n, z_j, z_i, v) \frac{\partial^2 u(.)}{\partial y^2} \frac{\partial c(.)}{\partial s} \left( 1-\delta \right) w_m(e) - \frac{\partial c(.)}{\partial L} \right\}
\]

Recalling that \( \frac{\partial^2 E\tilde{u}}{\partial L^2} < 0 \)

\[
\Rightarrow \frac{dL}{ds} \begin{cases} > 0 & \text{if } \left\{ \frac{\partial p(.)}{\partial L} \frac{\partial u(.)}{\partial \bar{y}} \frac{\partial c(.)}{\partial s} - p(L, n, z_j, z_i, v) \frac{\partial^2 u(.)}{\partial y^2} \frac{\partial c(.)}{\partial s} \left( 1-\delta \right) w_m(e) - \frac{\partial c(.)}{\partial L} \right\} > 0 \\ < 0 & \text{if } \left\{ \frac{\partial p(.)}{\partial L} \frac{\partial u(.)}{\partial \bar{y}} \frac{\partial c(.)}{\partial s} - p(L, n, z_j, z_i, v) \frac{\partial^2 u(.)}{\partial y^2} \frac{\partial c(.)}{\partial s} \left( 1-\delta \right) w_m(e) - \frac{\partial c(.)}{\partial L} \right\} < 0 \end{cases}
\]

(4)

3. The effect of wage discrimination

\[
\frac{\partial^2 E\tilde{u}}{\partial L^2} dL + \frac{\partial^2 E\tilde{u}}{\partial \delta^2} d\delta = 0
\]

\[
\frac{\partial^2 E\tilde{u}}{\partial L^2} dL + \left\{ \frac{\partial p(.)}{\partial L} \frac{\partial u(.)}{\partial \bar{y}} \left( -w_m(e)L \right) + p(.) \left[ \frac{\partial^2 u(.)}{\partial y^2} \left( -w_m(e)L \right) \left( 1-\delta \right) w_m(e) - \frac{\partial c(.)}{\partial L} \right] + \frac{\partial u(.)}{\partial \bar{y}} \left( -w_m(e) \right) \right\} d\delta = 0
\]

\[
\Rightarrow \frac{dL}{d\delta} = - \left\{ \frac{\partial p(.)}{\partial L} \frac{\partial u(.)}{\partial \bar{y}} \left( -w_m(e)L \right) + p(.) \left[ \frac{\partial^2 u(.)}{\partial y^2} \left( -w_m(e)L \right) \left( 1-\delta \right) w_m(e) - \frac{\partial c(.)}{\partial L} \right] + \frac{\partial u(.)}{\partial \bar{y}} \left( -w_m(e) \right) \right\}
\]

\[
\Rightarrow \frac{dL}{d\delta} \begin{cases} > 0 & \text{if } \left\{ \frac{\partial p(.)}{\partial L} \frac{\partial u(.)}{\partial \bar{y}} \left( -w_m(e)L \right) + p(.) \left[ \frac{\partial^2 u(.)}{\partial y^2} \left( -w_m(e)L \right) \left( 1-\delta \right) w_m(e) - \frac{\partial c(.)}{\partial L} \right] + \frac{\partial u(.)}{\partial \bar{y}} \left( -w_m(e) \right) \right\} > 0 \\ < 0 & \text{if } \left\{ \frac{\partial p(.)}{\partial L} \frac{\partial u(.)}{\partial \bar{y}} \left( -w_m(e)L \right) + p(.) \left[ \frac{\partial^2 u(.)}{\partial y^2} \left( -w_m(e)L \right) \left( 1-\delta \right) w_m(e) - \frac{\partial c(.)}{\partial L} \right] + \frac{\partial u(.)}{\partial \bar{y}} \left( -w_m(e) \right) \right\} < 0 \end{cases}
\]
\[ \left\{ \frac{\partial p(\cdot)}{\partial L} \frac{\partial u(\cdot)}{\partial y} (-w_m(e)L) + p(\cdot) \left[ \frac{\partial^2 u(\cdot)}{\partial^2 y} (-w_m(e)L) \left( 1 - \delta \right) w_m(e) - \frac{\partial c(\cdot)}{\partial L} \right] + \frac{\partial u(\cdot)}{\partial y} (-w_m(e)) \right\} > 0 \]

\( w_m(e) > 0 \)

\[ \Rightarrow \frac{dL}{d\delta} > 0 \text{ if } \left\{ \frac{\partial p(\cdot)}{\partial L} \frac{\partial u(\cdot)}{\partial y} (-L) + p(\cdot) \left[ \frac{\partial^2 u(\cdot)}{\partial^2 y} (-L) \left( 1 - \delta \right) w_m(e) - \frac{\partial c(\cdot)}{\partial L} \right] - \frac{\partial u(\cdot)}{\partial y} \right\} > 0 \] (5)

\[ \text{or } - \frac{\partial p(\cdot)}{\partial L} \frac{\partial u(\cdot)}{\partial y} L - p(\cdot) \left[ \frac{\partial^2 u(\cdot)}{\partial^2 y} L \left( 1 - \delta \right) w_m(e) - \frac{\partial c(\cdot)}{\partial L} \right] + \frac{\partial u(\cdot)}{\partial y} \right\} > 0 \]

4. The effect of partner’s income

The effect of partner’s income on a female’s labour supply may be derived as follows:

\[ \frac{\partial^2 E\bar{u}}{\partial L^2} dL + \frac{\partial \bar{p}(\cdot)}{\partial L} \left( \frac{\partial u(y, l)}{\partial y} \frac{\partial u(y, l)}{\partial y} \right) + p(L, n, z_j, z_v, v) \left[ \frac{\partial^2 u(\cdot)}{\partial^2 y} \left( 1 - \delta \right) w_m(e) - \frac{\partial c(\cdot)}{\partial L} \right] dy_p = 0 \]

\[ \Rightarrow \frac{dL}{dy_p} = - \frac{\frac{\partial \bar{p}(\cdot)}{\partial L} \left( \frac{\partial u(y, l)}{\partial y} \frac{\partial u(y, l)}{\partial y} \right) + p(L, n, z_j, z_v, v) \left[ \frac{\partial^2 u(\cdot)}{\partial^2 y} \left( 1 - \delta \right) w_m(e) - \frac{\partial c(\cdot)}{\partial L} \right]}{\frac{\partial^2 E\bar{u}}{\partial L^2}} \]

\[ \Rightarrow \frac{dL}{dy_p} > 0 \text{ if } \]

\[ \frac{\partial \bar{p}(\cdot)}{\partial L} \left( \frac{\partial u(y, l)}{\partial y} \frac{\partial u(y, l)}{\partial y} \right) + p(L, n, z_j, z_v, v) \left[ \frac{\partial^2 u(\cdot)}{\partial^2 y} \left( 1 - \delta \right) w_m(e) - \frac{\partial c(\cdot)}{\partial L} \right] = 0 \] (6)

5. The effect of the youngest child’s age

From the first order condition equation, we have:
\[
\frac{\partial^2 E\bar{u}}{\partial L^2} dL + \left[ \frac{\partial p(.)}{\partial L} \left[ \frac{\partial u(.)}{\partial y} \left( -\frac{\partial c(.)}{\partial z_j} \right) + \frac{\partial p(.)}{\partial z_j} \left[ \frac{\partial u(.)}{\partial y} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] \right] \\
+ p(L, n, z_j, z, v) \left[ \frac{\partial^2 u(.)}{\partial y^2} \left( -\frac{\partial c(.)}{\partial z_j} \right) \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] \right] d\bar{z}_j = 0
\]

\[
dL \quad \frac{dL}{dz_j} = - \frac{\partial^2 E\bar{u}}{\partial L^2}
\]

\[
\Rightarrow \frac{dL}{dz_j} > 0 \quad \text{if} \quad \left[ \frac{\partial p(.)}{\partial L} \left[ \frac{\partial u(.)}{\partial y} \left( -\frac{\partial c(.)}{\partial z_j} \right) + \frac{\partial p(.)}{\partial z_j} \left[ \frac{\partial u(.)}{\partial y} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] \right] \\
+ p(L, n, z_j, z, v) \left[ \frac{\partial^2 u(.)}{\partial y^2} \left( -\frac{\partial c(.)}{\partial z_j} \right) \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] \right] \quad > 0
\]

6. The effect of female’s age

\[
\frac{\partial^2 E\bar{u}}{\partial L^2} dL + \left[ \frac{\partial p(.)}{\partial z_i} \left[ \frac{\partial u(.)}{\partial y} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] \right] d\bar{z}_i = 0
\]

\[
dL \quad \frac{dL}{dz_i} = - \frac{\partial^2 E\bar{u}}{\partial L^2}
\]

\[
\Rightarrow \frac{dL}{dz_i} > 0 \quad \text{if} \quad \left[ \frac{\partial p(.)}{\partial z_i} \left[ \frac{\partial u(.)}{\partial y} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] \right] \quad > 0
\]

7. The effect of the unemployment rate

Total differentiation of the first order condition equation (1):

\[
\frac{\partial^2 E\bar{u}}{\partial L^2} dL + \frac{\partial p(.)}{\partial v} \left[ \frac{\partial u(.)}{\partial y} \left( (1-\delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) - \frac{\partial u(.)}{\partial l} \right] dv = 0
\]
\[
\Rightarrow \frac{dL}{dv} = - \frac{\partial p(.)}{\partial v} \left[ \frac{\partial u(.)}{\partial y} \left( (1 - \delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) \right] - \frac{\partial u(.)}{\partial l} - \frac{\partial E\tilde{u}}{\partial L^2} \\
\Rightarrow \frac{dL}{dv} > 0 \text{ if } \frac{\partial p(.)}{\partial v} \left[ \frac{\partial u(.)}{\partial y} \left( (1 - \delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) \right] > 0
\]

8. The effect of female educational attainments

\[
\frac{\partial^2 E\tilde{u}}{\partial L^2} \frac{dL}{de} + \left\{ \frac{\partial p(.)}{\partial L} \left( (1 - \delta)w'_m(e)L \right) + p(.) \left[ \frac{\partial^2 u(.)}{\partial y^2} (1 - \delta)w'_m(e)L \left( (1 - \delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) + \frac{\partial u(.)}{\partial y} (1 - \delta)w'_m(e) \right] \right\} de = 0
\]

\[
\frac{dL}{de} = - \frac{\left\{ \frac{\partial p(.)}{\partial L} \left( (1 - \delta)w'_m(e)L \right) + p(.) \left[ \frac{\partial^2 u(.)}{\partial y^2} (1 - \delta)w'_m(e)L \left( (1 - \delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) + \frac{\partial u(.)}{\partial y} (1 - \delta)w'_m(e) \right] \right\}}{\frac{\partial^2 E\tilde{u}}{\partial L^2}}
\]

\[
\Rightarrow \frac{dL}{de} > 0 \text{ if } \left\{ \frac{\partial p(.)}{\partial L} \left( (1 - \delta)w'_m(e)L \right) + p(.) \left[ \frac{\partial^2 u(.)}{\partial y^2} (1 - \delta)w'_m(e)L \left( (1 - \delta)w_m(e) - \frac{\partial c(.)}{\partial L} \right) + \frac{\partial u(.)}{\partial y} (1 - \delta)w'_m(e) \right] \right\} > 0
\]

(10)
Appendix B

1. Matching children to mothers in the 2009 Population and Housing Census

Codes for the primary relationship and the detailed relationship to householders are used to match children to mothers. These codes describe the relationship of each individual in the household to the person identified as the household head. In households containing a single family, children are matched straight to mothers by attaching all individuals labelled as “child” in the primary relationship code to the woman who is the household head or the spouse of the household head and is reported as having given birth.

In the households including multiple families, matching children to mothers is more complicated. Besides using the primary relationship and detailed relationship codes, subfamily numbers (family unit number) are also used to match children to mothers. If a woman who is reported as having given birth is a child of the household head (includes biological child or child in-laws), all children who are labelled as “grandchild” in the detailed relationship to the householder code will be attached to the woman who is labelled as “child” and has the same subfamily number. If a woman who is reported as having a child is labelled as ‘other relative and non-relative’ of the householder, all children who share the same subfamily code as the woman will be matched to the woman.

In the 2009 census, in general, multiple families in a household are classified with subfamily numbers. However, the multiple families in the household are not always classified with subfamily numbers. In some cases, multiple families in the household are all labelled with the same family code (as family unit “1”). In these cases, it is impossible to recognise and match children to mothers, so households with two or more mothers who are children of the householders and have the same subfamily codes are excluded.

The number of children matched to mothers (or the number of children observed in the household) may be smaller than the number of children ever born because some children might not be living with their mother but are living elsewhere or they may no longer be alive at the time of the census. The number of children observed in a
household may be greater than the number of given births because adopted or step children are also included and not separately identified in the survey. Ninety per cent of mothers in the sample of women aged 18-35 with at least two children had the number of children born reported the same as the number of children observed in the household. Ninety four percent of mothers in the sample have the number of children surviving equal to the number of children observed in the household.

2. Construction of the Moretwo variable

This variable is constructed from the sample of families with at least two children who are currently residing in Vietnam. Mothers in the sample are defined as women where the number of children ever born who are currently living with their mothers is equal to the number of children observed in the household. All children who died or are not living with their mothers and step or adopted children are excluded when constructing the variable Moretwo. This means that all mothers who have the reported number of children ever born is not the same as the number of children observed in the household at the time of the census are excluded.

There are appropriate reasons for the exclusion of step and adopted children in the study. Theoretically, it is argued that the effects of children on women’s labour force participation are heterogeneous by children’s age. On one hand, the smaller the child’s age is, the higher the opportunity costs of a woman participating in the labour market are, so mothers can choose to stay at home to take care of their children rather than work outside. On the other hand, the child’s age may also affect the mother’s probability of obtaining an outside job because it may be more difficult for mothers to get a job while their children are too small. Since the responsibility of mothers for step children may only start many years after they are born, the effects of step children on a woman’s probability of participating in the labour force may be different from the effects of biological children. However, the effect of an adopted child on a mother’s labour supply may not be different from that of a natural child because most children are adopted as infants (Lindelow and Wagstaff 2005). Although it may be appropriate to include adopted children in the study sample, it is not possible to distinguish between step and adopted children in the census data. So both step and adopted children are excluded from the sample.
In the census data, there are a number of questions that provide information to construct the variable *Moretwo*. The survey includes the variable “observed number of children” that provides a count of the total number of children currently living in the household with their parents. This variable includes both step and adopted children, so it represents an inappropriate measure to construct *Moretwo*. However, this variable lists all children currently living with their mothers at the time of the census. These children are actually matched to their mothers in the process of linking information on children and mothers. The variable “the number of children ever born” is also reported in the survey. It counts the total number of children born to each mother. It indicates how many children the mother has ever given birth to, which is obviously including any children who have died. Respondents are also asked to identify how many children were born alive but are no longer living at the time of the census (the number of children who died). It is possible to calculate the surviving children by subtracting the number of children who died from the number of children ever born. This means that surviving children include biological children who are currently residing with their mothers and those who are living elsewhere.

As mentioned above, it would be appropriate to include adopted children in the study sample so the total number of children each mother currently has (including both biological and adopted children) should represent an appropriate measure to construct *Moretwo*. However, since it is not possible to distinguish between step and adopted children. Moreover, all children listed in the survey data matched to their mothers are children who are currently living in the household with their parents. This means that it includes children who are natural, step or adopted children but does not includes children living elsewhere. In order to exclude the step and adopted children, the study will drop all mothers where the observed number of children in the household is greater than the number of children ever born and still living ‘chsurv’. The study also drops children living elsewhere as the survey does not list these children, so mothers who have a greater number of surviving children than the number of children observed in the households are also dropped.

Consequently, the sample will include all mothers whose number of natural children alive is equal to the number of natural children currently residing with them, with no children living elsewhere at the time of the census. Thus, the variable *chsurv* now is
equal to observed number of children, and we can use these variables to construct Moretwo. Moretwo takes a value of 1 if the number of children surviving and currently residing with a woman is greater than two, and it takes 0 otherwise.

3. Completed empirical results

Table B.1 - Effect of having more than two children on female labour force participation (2SLS estimation with all women sample aged 18-35 with two or more children)

Instrument: Twogirl

<table>
<thead>
<tr>
<th>Model</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moretwo</td>
<td>-0.0270</td>
<td>-0.0275</td>
<td>-0.0246</td>
<td>-0.0247</td>
<td>-0.0241</td>
</tr>
<tr>
<td>Firstgirl</td>
<td>0.0030***</td>
<td>0.0032***</td>
<td>0.0024***</td>
<td>0.0024***</td>
<td>0.0028***</td>
</tr>
<tr>
<td>Secondgirl</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.0097***</td>
<td>0.0100***</td>
<td>0.0095***</td>
<td>0.0087***</td>
<td>0.0087***</td>
</tr>
<tr>
<td>Ageatfirst</td>
<td>-0.0088***</td>
<td>-0.0068***</td>
<td>-0.0068***</td>
<td>-0.0065***</td>
<td>-0.0082***</td>
</tr>
<tr>
<td>Kinh</td>
<td>-0.1124***</td>
<td>-0.0966***</td>
<td>-0.0447***</td>
<td>-0.0422***</td>
<td>-0.0456***</td>
</tr>
<tr>
<td>Urban</td>
<td>-0.1058***</td>
<td>-0.1011***</td>
<td>-0.0990***</td>
<td>-0.1090***</td>
<td>-0.1080***</td>
</tr>
<tr>
<td>Northeast</td>
<td>0.0035***</td>
<td>0.0015</td>
<td>0.0030**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northwest</td>
<td>0.0035**</td>
<td>0.0023</td>
<td>0.0029</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northcentral</td>
<td>-0.0118***</td>
<td>-0.0110***</td>
<td>-0.0096***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Centralcoast</td>
<td>-0.0581***</td>
<td>-0.0547***</td>
<td>-0.0515***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Centralhigh</td>
<td>-0.0093***</td>
<td>-0.0076***</td>
<td>-0.0052***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Southeast</td>
<td>-0.1753***</td>
<td>-0.1713***</td>
<td>-0.1669***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mekong</td>
<td>-0.2031***</td>
<td>-0.1976***</td>
<td>-0.1918***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coresidence</td>
<td>0.0130***</td>
<td>0.0134***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td></td>
<td></td>
<td>0.0060***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary</td>
<td></td>
<td></td>
<td>-0.0181***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>University</td>
<td></td>
<td></td>
<td>0.1212***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.8720***</td>
<td>0.8289***</td>
<td>0.8659***</td>
<td>0.8817***</td>
<td>0.9107***</td>
</tr>
<tr>
<td>F-stat</td>
<td>4870.18</td>
<td>6145.64</td>
<td>7232.85</td>
<td>5534.25</td>
<td>4828.90</td>
</tr>
</tbody>
</table>
Table B.2 - Effect of having more than two children on female labour force participation (2SLS estimation with all women sample aged 18-35 with two or more children)

**Instrument: Twins**

<table>
<thead>
<tr>
<th>Model</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moretwo</td>
<td>-0.0500</td>
<td>-0.0493</td>
<td>-0.0440</td>
<td>-0.0417</td>
<td>-0.0413</td>
</tr>
<tr>
<td>Firstgirl</td>
<td>0.0050</td>
<td>0.0051</td>
<td>0.0041</td>
<td>0.0040</td>
<td>0.0044</td>
</tr>
<tr>
<td>Secondgirl</td>
<td>0.0042</td>
<td>0.0039</td>
<td>0.0032</td>
<td>0.0034</td>
<td>0.0034</td>
</tr>
<tr>
<td>Age</td>
<td>0.0104</td>
<td>0.0107</td>
<td>0.0102</td>
<td>0.0092</td>
<td>0.0092</td>
</tr>
<tr>
<td>Ageatfirst</td>
<td>-0.0097</td>
<td>-0.0076</td>
<td>-0.0075</td>
<td>-0.0072</td>
<td>-0.0088</td>
</tr>
<tr>
<td>Kinh</td>
<td>-0.1160</td>
<td>-0.0997</td>
<td>-0.0473</td>
<td>-0.0446</td>
<td>-0.0446</td>
</tr>
<tr>
<td>Urban</td>
<td>-0.1072</td>
<td>-0.1024</td>
<td>-0.1000</td>
<td>-0.1088</td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>0.0032</td>
<td>0.0012</td>
<td>0.0025</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northwest</td>
<td>0.0041</td>
<td>0.0028</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northcentral</td>
<td>-0.0103</td>
<td>-0.0097</td>
<td>-0.0084</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Centralcoast</td>
<td>-0.0562</td>
<td>-0.0529</td>
<td>-0.0503</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Centralhigh</td>
<td>-0.0061</td>
<td>-0.0046</td>
<td>-0.0027</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Southeast</td>
<td>-0.1735</td>
<td>-0.1696</td>
<td>-0.1659</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mekong</td>
<td>-0.2030</td>
<td>-0.1975</td>
<td>-0.1929</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coresidence</td>
<td>0.0131</td>
<td>0.0135</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>0.0030</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary</td>
<td>-0.0220</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>University</td>
<td>0.1169</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.8732</td>
<td>0.8294</td>
<td>0.8657</td>
<td>0.8814</td>
<td>0.9108</td>
</tr>
<tr>
<td>F-stat</td>
<td>4061.45</td>
<td>5271.13</td>
<td>6717.45</td>
<td>5165.79</td>
<td>4560.55</td>
</tr>
</tbody>
</table>
4. 2SLS estimate results of using IVs Samesex, Twogirl and Twoboy

Table B.3 - 2SLS estimates of FLFP(Women aged 18-35 with two or more children)

<table>
<thead>
<tr>
<th></th>
<th>All women</th>
<th>Married women</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First stage - dependent variable: Moretwo</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Samesex</td>
<td>0.0887*** (.0009)</td>
<td>0.0909*** (.0009)</td>
</tr>
<tr>
<td>Twogirl</td>
<td>0.1768*** (.0012)</td>
<td>0.1814*** (.0013)</td>
</tr>
<tr>
<td>Twoboy</td>
<td>0.0007 (.0012)</td>
<td>0.0004 (.0013)</td>
</tr>
</tbody>
</table>

| **Second stage - dependent variable: FLFP** |                    |                     |
| IV: Samesex          | -0.0420*** (0.0080) | -0.0440*** (0.0078) |
| IV: Twogirl and Twoboy| -0.0247*** (0.0055) | -0.0267*** (0.0055) |
| No. of Obs           | 706,602             | 706,602             |

Note: All models control for other covariates as in the preferred model.
### Appendix C

**Table C.1 - Effect of parental health shocks on child school enrolment (children aged 10-23)**

<table>
<thead>
<tr>
<th>Model</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent variables</td>
<td>OLS</td>
<td>FE</td>
<td>FE</td>
<td>FE</td>
<td>FE</td>
</tr>
<tr>
<td>Any parent ill</td>
<td>-0.0165**</td>
<td>-0.0223**</td>
<td>-0.0225**</td>
<td>-0.0232***</td>
<td>-0.0232***</td>
</tr>
<tr>
<td>Child ill</td>
<td>-0.0042</td>
<td>0.0145</td>
<td>0.0143</td>
<td>0.0137</td>
<td>0.0137</td>
</tr>
<tr>
<td>Month of interview</td>
<td>-0.0022</td>
<td>-0.0058</td>
<td>-0.0009</td>
<td>-0.0006</td>
<td>-0.0006</td>
</tr>
<tr>
<td>Year06</td>
<td>-0.0130*</td>
<td>-0.2665</td>
<td>-0.0248</td>
<td>-0.0211</td>
<td>-0.0213</td>
</tr>
<tr>
<td>Year08</td>
<td>-0.0363***</td>
<td>-0.0633*</td>
<td>-0.0633*</td>
<td>-0.0564</td>
<td>-0.0569</td>
</tr>
<tr>
<td>Age</td>
<td>-0.0696***</td>
<td>-0.0428***</td>
<td>-0.0446***</td>
<td>-0.0446***</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>-0.0224***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother’s age</td>
<td>-0.0023**</td>
<td>-0.0161**</td>
<td>-0.0166**</td>
<td>-0.0166**</td>
<td></td>
</tr>
<tr>
<td>Father’s age</td>
<td>0.0019**</td>
<td>0.0085</td>
<td>0.0076</td>
<td>0.0076</td>
<td></td>
</tr>
<tr>
<td>Mother’s primary</td>
<td>0.0711***</td>
<td>-0.0711***</td>
<td>-0.0754***</td>
<td>-0.0754***</td>
<td></td>
</tr>
<tr>
<td>Mother’s low secondary</td>
<td>0.1218***</td>
<td>-0.0604*</td>
<td>-0.0648*</td>
<td>-0.0648*</td>
<td></td>
</tr>
<tr>
<td>Mother’s upper secondary</td>
<td>0.1460***</td>
<td>-0.0906*</td>
<td>-0.0943*</td>
<td>-0.0943*</td>
<td></td>
</tr>
<tr>
<td>Father’s primary</td>
<td>0.1212***</td>
<td>-0.1359*</td>
<td>-0.1470*</td>
<td>-0.1474*</td>
<td></td>
</tr>
<tr>
<td>Father’s low secondary</td>
<td>0.0628***</td>
<td>-0.0299</td>
<td>-0.0318</td>
<td>-0.0318</td>
<td></td>
</tr>
<tr>
<td>Father’s upper secondary</td>
<td>0.1184***</td>
<td>-0.0438</td>
<td>-0.0441</td>
<td>-0.0442</td>
<td></td>
</tr>
<tr>
<td>Father’s postsecondary</td>
<td>0.1482***</td>
<td>-0.0739*</td>
<td>-0.0741*</td>
<td>-0.0741*</td>
<td></td>
</tr>
<tr>
<td>Household size</td>
<td>0.1670***</td>
<td>-0.0626</td>
<td>-0.0566</td>
<td>-0.0566</td>
<td></td>
</tr>
<tr>
<td>Number of children</td>
<td>-0.0151***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of houses</td>
<td>0.0136</td>
<td>0.0351</td>
<td>0.0154</td>
<td>0.0156</td>
<td></td>
</tr>
<tr>
<td>House own</td>
<td>0.0351</td>
<td>0.0007***</td>
<td>0.0004***</td>
<td>0.0004***</td>
<td></td>
</tr>
<tr>
<td>Living area</td>
<td>0.0051</td>
<td>0.0135</td>
<td>0.0113</td>
<td>0.0113</td>
<td></td>
</tr>
<tr>
<td>Safe water</td>
<td>0.1102***</td>
<td>0.0665***</td>
<td>0.0663***</td>
<td>0.0663***</td>
<td></td>
</tr>
<tr>
<td>Computer own</td>
<td>0.0075</td>
<td>0.0287</td>
<td>0.0287</td>
<td>0.0287</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>-0.0196*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kinh</td>
<td>0.0392**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per capital income</td>
<td>0.0379***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North West</td>
<td>0.0392**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North East</td>
<td>0.0134</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Central</td>
<td>0.0379***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: Significant levels: *, p < 0.1; **, p < 0.05; *** p < 0.01*
Table C.2 - Effect of maternal and paternal health shocks on child school enrolment (children aged 10-23)

<table>
<thead>
<tr>
<th>Model</th>
<th>Independent variables</th>
<th>OLS</th>
<th>FE</th>
<th>FE</th>
<th>FE</th>
<th>FE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mother ill only</td>
<td>-0.0351***</td>
<td>-0.0313***</td>
<td>-0.0315***</td>
<td>-0.0320***</td>
<td>-0.0321***</td>
<td></td>
</tr>
<tr>
<td>Father ill only</td>
<td>-0.0260**</td>
<td>-0.0254*</td>
<td>-0.0264*</td>
<td>-0.0259*</td>
<td>-0.0258*</td>
<td></td>
</tr>
<tr>
<td>Both parents ill</td>
<td>0.0124</td>
<td>-0.0060</td>
<td>-0.0055</td>
<td>-0.0075</td>
<td>-0.0072</td>
<td></td>
</tr>
<tr>
<td>Child ill</td>
<td>-0.0161*</td>
<td>0.0092</td>
<td>0.0087</td>
<td>0.0086</td>
<td>0.0086</td>
<td></td>
</tr>
<tr>
<td>Month of interview</td>
<td>-0.0022</td>
<td>-0.0060**</td>
<td>-0.0011</td>
<td>-0.0008</td>
<td>-0.0008</td>
<td></td>
</tr>
<tr>
<td>Year06</td>
<td>-0.0142*</td>
<td>-0.1277</td>
<td>-0.0251</td>
<td>-0.0216</td>
<td>-0.0217</td>
<td></td>
</tr>
<tr>
<td>Year08</td>
<td>-0.0366***</td>
<td>-0.2672</td>
<td>-0.0630*</td>
<td>-0.0563</td>
<td>-0.0568</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.0696***</td>
<td>-0.0427***</td>
<td>-0.0445***</td>
<td>-0.0445***</td>
<td>-0.0445***</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>-0.0228***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother’s age</td>
<td>-0.0023**</td>
<td>-0.0162**</td>
<td>-0.0167**</td>
<td>-0.0167**</td>
<td>-0.0167**</td>
<td></td>
</tr>
<tr>
<td>Father’s age</td>
<td>0.0017*</td>
<td>0.0085</td>
<td>0.0075</td>
<td>0.0075</td>
<td>0.0075</td>
<td></td>
</tr>
<tr>
<td>Mother’s primary</td>
<td>0.0713***</td>
<td>-0.0711***</td>
<td>-0.0753***</td>
<td>-0.0752***</td>
<td>-0.0752***</td>
<td></td>
</tr>
<tr>
<td>Mother’s low secondary</td>
<td>0.1221***</td>
<td>-0.0604*</td>
<td>-0.0647*</td>
<td>-0.0646*</td>
<td>-0.0646*</td>
<td></td>
</tr>
<tr>
<td>Mother’s upper secondary</td>
<td>0.1457***</td>
<td>-0.0892*</td>
<td>-0.0928*</td>
<td>-0.0927*</td>
<td>-0.0927*</td>
<td></td>
</tr>
<tr>
<td>Mother’s post-secondary</td>
<td>0.1223***</td>
<td>-0.1344*</td>
<td>-0.1456*</td>
<td>-0.1456*</td>
<td>-0.1456*</td>
<td></td>
</tr>
<tr>
<td>Father’s primary</td>
<td>0.0633***</td>
<td>-0.0292</td>
<td>-0.0312</td>
<td>-0.0312</td>
<td>-0.0312</td>
<td></td>
</tr>
<tr>
<td>Father’s low secondary</td>
<td>0.1193***</td>
<td>-0.0437</td>
<td>-0.0440</td>
<td>-0.0441</td>
<td>-0.0441</td>
<td></td>
</tr>
<tr>
<td>Father’s upper secondary</td>
<td>0.1476***</td>
<td>-0.0755*</td>
<td>-0.0756*</td>
<td>-0.0756*</td>
<td>-0.0756*</td>
<td></td>
</tr>
<tr>
<td>Father’s post-secondary</td>
<td>0.1693***</td>
<td>-0.0627</td>
<td>-0.0567</td>
<td>-0.0567</td>
<td>-0.0567</td>
<td></td>
</tr>
<tr>
<td>Household size</td>
<td>-0.0149***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of children</td>
<td>-0.0104**</td>
<td></td>
<td>0.0130</td>
<td>0.0130</td>
<td>0.0130</td>
<td></td>
</tr>
<tr>
<td>Number of houses</td>
<td>0.0132</td>
<td></td>
<td>-0.0315</td>
<td>-0.0315</td>
<td>-0.0315</td>
<td></td>
</tr>
<tr>
<td>House own</td>
<td>0.0357</td>
<td></td>
<td>-0.0127</td>
<td>-0.0129</td>
<td>-0.0129</td>
<td></td>
</tr>
<tr>
<td>Living area</td>
<td>0.0007***</td>
<td>0.0004***</td>
<td>0.0004***</td>
<td>0.0004***</td>
<td>0.0004***</td>
<td></td>
</tr>
<tr>
<td>Safe water</td>
<td>0.0058</td>
<td></td>
<td>0.0114</td>
<td>0.0114</td>
<td>0.0114</td>
<td></td>
</tr>
<tr>
<td>Independent variables</td>
<td>Labour force participation</td>
<td>Weekly working hours</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>----------------------------</td>
<td>----------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1) OLS</td>
<td>(2) FE</td>
<td>(3) FE</td>
<td>(4) OLS</td>
<td>(5) FE</td>
<td>(6) FE</td>
</tr>
<tr>
<td>Any parent ill</td>
<td>0.0413***</td>
<td>0.0382***</td>
<td>0.0383***</td>
<td>1.4199***</td>
<td>1.0652***</td>
<td>1.0658***</td>
</tr>
<tr>
<td>Child ill</td>
<td>0.0019</td>
<td>-0.0059</td>
<td>-0.0063</td>
<td>-0.2155</td>
<td>-0.6038</td>
<td>-0.6096</td>
</tr>
<tr>
<td>Month of interview</td>
<td>-0.0050***</td>
<td>-0.0077***</td>
<td>-0.0081***</td>
<td>-0.1235</td>
<td>-0.1532</td>
<td>-0.1671</td>
</tr>
<tr>
<td>Year06</td>
<td>-0.0540***</td>
<td>0.0164</td>
<td>0.0045</td>
<td>-1.3484***</td>
<td>1.4113</td>
<td>1.0039</td>
</tr>
<tr>
<td>Year08</td>
<td>-0.0390***</td>
<td>0.1050***</td>
<td>0.0829**</td>
<td>-0.6169</td>
<td>5.0328***</td>
<td>4.2755***</td>
</tr>
<tr>
<td>Age</td>
<td>0.0613***</td>
<td>0.0204**</td>
<td>0.0235**</td>
<td>3.3297***</td>
<td>1.7063***</td>
<td>1.8209***</td>
</tr>
<tr>
<td>Male</td>
<td>0.0348***</td>
<td></td>
<td></td>
<td>2.1684</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother’s age</td>
<td>-0.0001</td>
<td>-0.0050</td>
<td>-0.0038</td>
<td>0.0014</td>
<td>-0.1297</td>
<td>-0.0937</td>
</tr>
<tr>
<td>Father’s age</td>
<td>-0.0007</td>
<td>-0.0032</td>
<td>-0.0023</td>
<td>-0.0575</td>
<td>-0.3806</td>
<td>-0.3337</td>
</tr>
<tr>
<td>Mother’s primary</td>
<td>-0.0847***</td>
<td>0.0110</td>
<td>0.0153</td>
<td>-3.7747***</td>
<td>1.6920</td>
<td>1.9649</td>
</tr>
<tr>
<td>Mother’s low secondary</td>
<td>-0.1039***</td>
<td>0.0701*</td>
<td>0.0764**</td>
<td>-5.0031***</td>
<td>2.9618***</td>
<td>3.3314*</td>
</tr>
<tr>
<td>Mother’s upper secondary</td>
<td>-0.1477***</td>
<td>0.0499</td>
<td>0.0529</td>
<td>-6.9306***</td>
<td>1.7994</td>
<td>2.0196</td>
</tr>
<tr>
<td>Mother’s postsecondary</td>
<td>-0.1388***</td>
<td>0.0984</td>
<td>0.1069</td>
<td>-5.9701***</td>
<td>5.1194</td>
<td>5.6371</td>
</tr>
<tr>
<td>Father’s primary</td>
<td>-0.0446***</td>
<td>0.0276</td>
<td>0.0304</td>
<td>-2.5843***</td>
<td>0.9646</td>
<td>1.0201</td>
</tr>
<tr>
<td>Father’s low secondary</td>
<td>-0.0947***</td>
<td>0.0631**</td>
<td>0.0631**</td>
<td>-5.5804***</td>
<td>2.1609</td>
<td>2.1073</td>
</tr>
<tr>
<td>Father’s upper secondary</td>
<td>-0.1515***</td>
<td>0.0855**</td>
<td>0.0869**</td>
<td>-7.7898***</td>
<td>2.9629</td>
<td>3.0230</td>
</tr>
</tbody>
</table>

Table C.3 - Effect of parental health shocks on children’s labour force participation (children aged 10-23)
Table C.4 - Effect of maternal and paternal health shocks on children’s labour force participation (children aged 10-23)

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>Labour force participation</th>
<th>Weekly working hours</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>OLS</td>
<td>FE</td>
</tr>
<tr>
<td>Mother only ill</td>
<td>0.0520***</td>
<td>0.0534***</td>
</tr>
<tr>
<td>Father only ill</td>
<td>0.0337***</td>
<td>0.0377***</td>
</tr>
<tr>
<td>Both parents ill</td>
<td>0.0346***</td>
<td>0.0159</td>
</tr>
<tr>
<td>Child ill</td>
<td>0.0045</td>
<td>0.0013</td>
</tr>
<tr>
<td>Month of interview</td>
<td>-0.0051**</td>
<td>-0.0075**</td>
</tr>
<tr>
<td>Year06</td>
<td>-0.0537***</td>
<td>0.0167***</td>
</tr>
<tr>
<td>Year08</td>
<td>-0.0385***</td>
<td>0.1044***</td>
</tr>
<tr>
<td>Age</td>
<td>0.0613***</td>
<td>0.0203**</td>
</tr>
<tr>
<td>Male</td>
<td>0.0349***</td>
<td>2.1782***</td>
</tr>
<tr>
<td>Mother’s age</td>
<td>-0.0002</td>
<td>-0.0048</td>
</tr>
<tr>
<td>Father’s age</td>
<td>-0.0007</td>
<td>-0.0028</td>
</tr>
<tr>
<td></td>
<td>Coefficient 1</td>
<td>Coefficient 2</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Mother’s primary</td>
<td>-0.0846</td>
<td>0.0110</td>
</tr>
<tr>
<td>Mother’s low secondary</td>
<td>-0.1038</td>
<td>0.0700*</td>
</tr>
<tr>
<td>Mother’s upper secondary</td>
<td>-0.1471**</td>
<td>0.0476</td>
</tr>
<tr>
<td>Mother’s postsecondary</td>
<td>-0.1389**</td>
<td>0.0959</td>
</tr>
<tr>
<td>Father’s primary</td>
<td>-0.0447***</td>
<td>0.0269</td>
</tr>
<tr>
<td>Father’s low secondary</td>
<td>-0.0954**</td>
<td>0.0629**</td>
</tr>
<tr>
<td>Father’s upper secondary</td>
<td>-0.1517**</td>
<td>0.0876**</td>
</tr>
<tr>
<td>Father’s postsecondary</td>
<td>-0.1496**</td>
<td>0.0849</td>
</tr>
<tr>
<td>Household size</td>
<td>0.0083***</td>
<td>-0.0136*</td>
</tr>
<tr>
<td>Number of children</td>
<td>0.0188***</td>
<td>-0.0138</td>
</tr>
<tr>
<td>Number of houses</td>
<td>-0.0097</td>
<td>-0.0106</td>
</tr>
<tr>
<td>House own</td>
<td>-0.0136</td>
<td>-0.0128</td>
</tr>
<tr>
<td>Living area</td>
<td>-0.0005***</td>
<td>-0.0003*</td>
</tr>
<tr>
<td>Safe water</td>
<td>-0.0092</td>
<td>-0.0059</td>
</tr>
<tr>
<td>Computer own</td>
<td>-0.0831***</td>
<td>-0.0343</td>
</tr>
<tr>
<td>Urban</td>
<td>-0.0767***</td>
<td>-0.0231</td>
</tr>
<tr>
<td>Kinh</td>
<td>-0.0875</td>
<td></td>
</tr>
<tr>
<td>North West</td>
<td>0.0110</td>
<td></td>
</tr>
<tr>
<td>North East</td>
<td>0.0468***</td>
<td></td>
</tr>
<tr>
<td>North Central</td>
<td>-0.0209*</td>
<td></td>
</tr>
<tr>
<td>Central High</td>
<td>-0.0202</td>
<td></td>
</tr>
<tr>
<td>South Coast</td>
<td>-0.0951***</td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td></td>
<td></td>
</tr>
<tr>
<td>South East</td>
<td>-0.0382***</td>
<td></td>
</tr>
<tr>
<td>Mekong</td>
<td>-0.0205</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.3233***</td>
<td>0.3020</td>
</tr>
<tr>
<td>No. of obs</td>
<td>13,930</td>
<td>13,930</td>
</tr>
</tbody>
</table>
References

Becker, GS 1957, The economics of discrimination, University of Chicago press.


Cain, GG 1966, *Married women in the labor force*.


Cunha, F & Heckman, JJ 2008, ‘Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation’, *Journal of Human Resources*, vol. 43, no. 4, pp. 738-82.


Doan, T & Gibson, J 2010, ‘Return to schooling in Vietnam during economic transition: does the return reach its peak?’.

Ebenstein, A 2007, The causal effect of fertility on female labor supply: evidence from Taiwanese son preference, Harvard University; University of California, Berkeley - Department of Economics.


Fleury, N & Gilles, F 2015, ‘A meta-regression analysis on intergenerational transmission of education: publication bias and genuine empirical effect’.


Meulders, D, Plasman, O & Plasman, R 1994, Atypical Employment in the EC, ULB--Universite Libre de Bruxelles.


MPI 2013, *Millennium development goals full report 2013 "Achievement and challenges in the progress of reaching millennium development goals of Vietnam"*.


University of Pennsylvania 2014, Why child care is the economy’s 'invisible' driver.


Willis, RJ 1987, ‘What have we learned from the economics of the family?’, The American Economic Review, vol. 77, no. 2, pp. 68-81.


