# FAMILY STRUCTURE AND OUTCOMES IN ADOLESCENCE, YOUNG ADULTHOOD, AND ADULTHOOD

**by**

**NGA THI HANG NGUYEN**

**A Dissertation submitted to the**

**Graduate School - New Brunswick**

**Rutgers, The State University of New Jersey**

**in partial fulfillment of the requirements**

**for the degree of**

**Doctor of Philosophy**

**Graduate Program in Economics**

**Written under the direction of**

**Carolyn Moehling**

**and approved by**

____________________

____________________

____________________

____________________

New Brunswick, New Jersey
October 2014

**ABSTRACT OF THE DISSERTATION**


**Family Structure and Outcomes in Adolescence, Young Adulthood, and Adulthood**

**By NGA THI HANG NGUYEN**

**Dissertation Director:**

**Carolyn Moehling**

I use longitudinal data sets to investigate the impact of family structure on socioeconomic outcomes including educational attainment, labor market outcomes, and engagement in risky behaviors in the short-run and the long-run.

The first essay, "Teenage Childbearing and Socioeconomic Outcomes of Teen Mothers in Young Adulthood and Adulthood using NLSY79 data" uses the National Longitudinal Survey of Youth (NLSY79) data set from 1979 to 1998 to examine the causal effects of teenage childbearing on the socioeconomic outcomes of teen mothers in two stages of young adulthood (between 18 and 24 years old) and adulthood (between 25 and 32 years old). I find that teenage childbearing has a statistically significant negative impact on years of schooling and annual earnings in the short-run but no impact in the long-run. These findings support the hypothesis that the short-run disadvantage of teenage childbearing in terms of education and labor market diminishes in the long-run.

The second essay, "An Examination of the Persistence of Impact of Teenage Childbearing on Labor Market Outcomes Using the Add Health Data" extends the analysis in my first essay in two ways. First, I use the National Longitudinal Study of

Adolescent Health (Add Health) data-set to examine the short-run and the long-run impact of Teenage Childbearing on the mother's future labor outcome. The Add Health data set has an advantage over the NLSY79 data in analysis of teenage childbearing issue because it uses computer-assisted personal interview (CAPI) technology allowing the respondents to answer sensitive questions by computer rather than by an open verbal conversation. Therefore, using Add Health data helps reduce the bias in self-reports of pregnancy outcomes. I find that teenage childbearing does not affect the mother's labor outcome in long run contrary to its effect in the short run. Second, I extend my first essay as well as the existing literature by looking for reasons behind the transition of this effect from short run to long run. I test the hypothesis that the lack of long-run impact of teenage childbearing on teen mothers' annual earnings could be due to their earlier participation in the labor market. My results show that although the first age at which a teen mother starts working full-time is negatively correlated with her annual earnings after controlling for her education, the negative effect of teenage childbearing does not reappear in the long run. This finding does not support the hypothesis above.

The third essay, entitled "Birth Spacing and Outcomes in Adolescence, Young Adulthood, and Adulthood" uses the sub-sample of sibling-pairs from the restricted-use Add Health data-set to investigate the linkage between birth spacing and outcomes of siblings such as test scores, years of schooling, college attendance, college degree, wages, and engagement in risky behaviors in adolescence, young adulthood, and adulthood. My results show that birth spacing does not have an impact on siblings' percentile rankings on test scores and years of schooling in adolescence and young adulthood. However, greater birth spacing increases the likelihood of enrolling in college for siblings in young adulthood. This effect persists when siblings transition to adulthood by increasing the possibility of obtaining a college degree in adulthood. I also find that wider birth spacing will have greater impact on the likelihood of enrolling in a college for the younger sibling than for the older sibling. The findings suggest that the allocation of family resources to and across siblings plays an important role in the post-secondary schooling decisions. I find no effect of birth spacing, however, on annual earnings in adulthood or the probability of engaging in cigarette smoking in adolescence.

Acknowledgments

I am forever indebted to my dissertation advisor, Professor Carolyn Moehling, for her invaluable research guidance. Without her encouragement and support, it would be impossible for me to overcome difficult time and reach this stage of my graduate study. Professor Moehling took me as her student, helped me access to the restricted dataset - a crucial part of this dissertation - she guided me in every single step through it. With the unsurpassed insight, knowledge, positive energy, and caring, Professor Moehling becomes my role model as a researcher, a teacher and an advisor.

I am extremely grateful to Professor Thomas Prusa and Professor Anne Piehl for their advice and numerous discussions to improve my research. Professor Prusa was instrumental in helping me improve the chapter 4 of this dissertation and my job market. He has been so generous in his support in teaching and research opportunities since I worked for him as a teaching assistant. The discussion with him has always been motivational and inspiring. Professor Piehl showed so much enthusiasm in helping me improve my dissertation. The profound research knowledge that I learned from Professor Piehl will stay with me as an invaluable asset. I would like to express my gratitude to Professor Ayse Akincigil for being a member of my dissertation committee.

I would like to express my sincere thanks to Professor Hugh Rockoff, Professor Roger Klein, Professor Hiroki Tsurumi, Professor John Landon-Lane, and Professor Ira Gang for helpful comments as well as insightful discussions at various stages of my dissertation. I owe special thanks to Ryan Womack for his help in the arrangement of restricted facility for the dataset used in my dissertation.

I am also grateful to my husband for his support in my study and research.

Many thanks are owed to Dorothy Rinaldi, Donna Ghilino for their help. I also appreciate the financial support from Economics Department and Graduate School at Rutgers University.

I would also like to thank my friends, in particular Hang, Omayma, Evin, Giang, Hoa, Xi Wang, Costie, Xuelian, Georgia, Ritu for the support and time we spent together.

Finally, I would like to thank my parents, my parent-in-law, and my kids - Hali and Son for their immense sacrifices.

# Dedication

To my parent, my parent-in-law, my husband, my kids - Son Duong and Hali Duong.

# Table of Contents

# List of Tables

# Chapter 1

# Introduction

I use longitudinal data sets to investigate the impact of aspects of family structure on socioeconomic outcomes including educational attainment, labor market outcomes, engagement in risky behaviors in the short-run and the long-run.

The second chapter, "Teenage Childbearing and Socioeconomic Outcomes of Teen Mothers in Young Adulthood and Adulthood using NLSY79 data" adds to the literature by empirically examining the causal effects of teenage childbearing on socioeconomic outcomes of teen mothers in the short and long-run. In the first stage, I analyze outcomes in young adulthood when teen mothers are between 18 and 24 years old. In the second stage, these outcomes are examined when teen mothers are between 25 and 32 years old. Previous studies have shown that the effects of teenage childbearing vary on yearly basis. However, these studies could not provide a clear picture as to whether these effects persisted as teen mothers transition to adulthood. I improve the comparative analysis by aggregating teen mother's outcomes in two separate stages. I follow the literature by adopting an instrumental variable (IV) method with miscarriage as an instrumental variable for teenage childbearing and apply it to the National Longitudinal Survey of Youth (NLSY79) data set from year 1979 to year 1998. I find that teenage childbearing has a small but statistically significant negative impact on years of schooling in the short-run. I find no impact though in the long-run. I get similar results for labor market earnings. These findings support the hypothesis that the short-run disadvantage of teenage childbearing in terms of education and labor market is diminishing in the long-run. In addition, I find that having a teen birth increases the likelihood of the receipt of public assistance but has little effect on the mother's subsequent fertility. In addition, I examine the validity of this instrument by analyzing risky behaviors such as drinking, smoking and failure to use contraceptive methods. I find that the results are similar to each other after including the variables. This finding suggests the validity of miscarriage as an instrumental variable for teenage childbearing.

The third chapter, "An Examination of the Persistence of the Impact of Teenage Childbearing on Labor Market Outcome using the Add Health Data" extends the analysis in my second chapter and the literature on teenage childbearing in two ways. First, I use the National Longitudinal Study of Adolescent Health (Add Health) data-set up to 2008 to examine the short run and the long run impact of teenage childbearing on labor market outcomes. The Add Health data is very rich as it includes surveys of school, personnel, and parents as well as contextual data such as relationships, families, social networks, neighborhoods, schools, and states. In addition, computer-assisted personal interview (CAPI) technology used in surveys will likely yield more accurate information about miscarriage, the instrumental variable for teenage childbearing. My results indicate that the effect of teenage childbearing on labor market outcomes does not persist in the long run. These findings are consistent with those in the second chapter. Second, I extend my second chapter as well as the existing literature by looking for reasons behind the dynamic change in the effects of teenage childbearing on later labor market outcomes. My results show that the age at which a teen mother starts working full-time negatively affects her annual earnings in the long run after controlling for her education. However, the negative effect of teenage childbearing does not reappear after controlling for the age at first full-time job. This diagnostic finding suggests that this age variable could be used to reject the hypothesis that the lack of long-run impact of teenage childbearing on annual earnings might not be due to an earlier participation or more work experience of a teen mother in the labor market.

In the fourth chapter, entitled “Birth Spacing and Outcomes in Adolescence, Young Adulthood, and Adulthood", I investigate the linkage between birth spacing and various outcomes of siblings such as percentile ranking on test scores, years of schooling, college attendance, college degree, labor market earnings, and engagement in risky behaviors. The study uses the sub-sample of sibling-pairs from the restricted-use Add Health data-set. Birth spacing is an element of family structure and is widely viewed as an important determinant of human capital investment. Although previous research papers found birth spacing affects early life outcomes such as birth weight, neonatal mortality, pre-school cognitive development, and test scores at elementary school age, I have not been able to find any papers that have examined the impact of birth spacing on the later life outcomes such as labor earnings and engagement in risky behaviors in adolescence, young adulthood, and adulthood. There are two key contributions of the paper. First, my study is broader than previous studies, both in terms of time frame (adolescence, young adulthood, and adulthood) and also in terms of the variety of outcomes studied, such as educational outcomes (percentile ranking on test scores, years of schooling, college attendance, and college degree), labor market outcomes (annual earnings), and risky behavior outcomes (cigarette smoking). Second, I examine and test the persistence of birth spacing effects during the time when siblings transition from adolescence to young adulthood. I use a family fixed-effect estimation to control for the heterogeneity across families. I find that birth spacing does not have an impact on siblings' test scores and years of schooling in adolescence and young adulthood. However, I do find that wider birth spacing increases the likelihood of enrolling in college in young adulthood. This effect does persist in to adulthood. I find that greater birth spacing increases the likelihood of obtaining a college degree for siblings. In addition, I find that birth spacing has a larger effect on the younger sibling than it does on the older sibling in terms of the possibility of enrolling in college. These findings suggest that the allocation of family resources to and across siblings plays important role in pursuing higher education of siblings. I find no effect of birth spacing on annual earnings in adulthood or the probability of engaging in cigarettes smoking in adolescence.

# Chapter 2

# Teenage Childbearing and Socioeconomic Outcomes of Teen Mothers in Young Adulthoold and Adulthood using NLSY79 Data

## 2.1 Introduction

Teenage childbearing is an important socioeconomic issue in the United States over the last several decades. At its peak in 1990, the U.S. teenage pregnancy rate was 116.9 pregnancies per 1,000 women aged between 15 and 19 (Kost and Henshaw 2012). In its lowest point in 2008, still 733,000 teenage females aged from 15 to 19 were pregnant in the United States. The teenage pregnancy rate was almost 4.02%. The U.S. teenage pregnancy rate remains the highest in the developed world and more than twice as high as those of Canada and Sweden (Hoffman 2006). In 2006, 59% of pregnancies among teenagers between 15 and 19 ended in birth, 27% in abortion, and 14% in miscarriage. (Kost, Henshaw and Carlin 2010).

Previous studies have found that teen mothers face difficulties finishing high school as well as entering the labor force a couple of years after they give birth (Hoffman, Foster, and Furstenberg 1993). They tend to receive low earnings and have to rely on government support (Hoffman 2006). They are less likely to marry than women who do not give birth as a teenager (Hoffman 2008). These negative effects are well understood and are considered in this paper as the "short-run" effects of teenage childbearing.

However, several papers such as Geronimus and Korenman (1992) and Hotz, McElroy and Sanders (2005) examine the long-run effects of teenage childbearing. They have found little adverse effect of teenage childbearing on women in their late 20s. In particular, the effects of teenage childbearing on a woman's life 10 or 15 years after giving birth are minimal. These findings suggest that negative effects of teenage childbearing may not be permanent or may even disappear as teen mothers and their children age. I examine this issue in my research.

In this chapter, I re-visit the topic of the effects of teenage childbearing on socioeconomic outcomes using the National Longitudinal Survey of Youth cohort 1979 (NLSY79). Previous studies using this data set usually look at how these effects changed from year to year. Given sample attribution, it is difficult to interpret the statistical differences in the effect between two consecutive years due to the fact that the sample size of the teen mothers gets smaller as they get older. Instead of yearly comparisons, I aggregate teen mother's outcomes in two separate stages. In the first stage (short-run), I analyze the effects of teenage childbearing on mothers who are in young adulthood, between 18 to 24 years old. In the second stage (long-run), I look at the effects of teenage childbearing on the same mothers but when they are between 25 to 32 years old. I refer to this age period as "adulthood". By comparing the effects from two stages, I can test whether the effects of childbearing are permanent or only temporary as well as whether these effects increase in the long run.

Another motivation for this research also comes from econometric challenges of sample selection and missing variables that plague this literature. The selected sample is those women who have their first births as teenagers. The question of interest for researchers is "what would have happened to this woman had she not given a birth as a teen?" The comparison group for this selected population is actually unobserved.

This is an important issue that must be addressed.

Several studies use instrumental variable (IV) methods. One recent approach by Hotz, McElroy, and Sanders (2005) utilized miscarriage as a natural experiment instrument. They pointed out that women who miscarried are considered to have been willing to give birth but are unable to because of the random miscarriage. Socioeconomic outcomes of women who experienced miscarriages as teenagers then could be contrasted to the outcomes of women who have their first birth as teenagers. However, the assumption that miscarriage is a natural random event is a strong parametric restriction. In fact, miscarriage might depend on factors such as drinking, smoking, and environment before or during the first pregnancy. Miscarriage could also be linked to the characteristics of the community where teen mothers grew up.

Fletcher and Wolfe (2009) and Ashcraft and Lang (2006) casted doubt on the validity of miscarriage as a natural instrument. They argued that miscarriage is a valid instrument only in the absence of abortion. If a mother miscarries early, it is difficult to know whether she truly miscarried or she aborted on purpose. Fletcher and Wolfe (2009) used late miscarriage - defined as 8 weeks or later as an instrument which is more close to a natural experiment.

Based on these studies I implement an empirical analysis using miscarriage as instrumental variable. I add further evidence to the literature on this topic by checking the validity of the natural randomness of miscarriage. I carry out sensitivity analysis using factors associated with the random nature of miscarriage such as drinking, smoking, and the use of contraceptive methods.

I find that in young adulthood, teenage childbearing has statistically significant effects on annual earnings, the receipt of public assistance, and subsequent fertility. There is no statistically significant effects on mother's years of education. There is a big "positive" transition in the impact of having a birth as a teen on mother's outcomes as she ages from young adulthood to adulthood. In the adulthood, she appears to be able to overcome the hardship of giving birth early as a teen. This could be explained by the fact that accumulated work experience might help teen mothers later in their work history. My results also confirm that the negative impacts of teenage childbearing are over-estimated by OLS and the use of IV estimation can reduce this bias.

The rest of this chapter is organized as follows: section 2.2 summarizes literature and section 2.3 discusses models. In section 2.4, I describe the data and the estimation results are presented in section 2.5. Finally, section 2.6 concludes the chapter.

## 2.2 Literature Review

### 2.2.1 The Effects of Teenage Childbearing

As pointed out by many researchers, early childbearing causes substantial adverse effects on mothers. Teen motherhood imposes both health and economic consequences. The risk of death during childbirth for a mother under age 17 is two to four times greater than that for a mother aged over 20 (McCauley and Salter 1995). Teen mothers receive less prenatal care than other mothers. Lack of prenatal care subsequently creates pregnancy-related issues for them. From an economics perspective, teenage childbearing raises the opportunity costs of human capital accumulation. Early childbearing leads to lower levels of investment in education and labor market outcomes, which could result in depressed socioeconomic status (Ribar 1994; Geronimus and Korenman 1992; Geronimus and Korenman 1992; Hoffman et al. 1993; Klepinger et al. 1999). Due to the hardship at the early stage of their life, it is much more difficult for teen mothers to complete high school or obtain a college degree. This adversely affects their value in labor market.

In addition, teenage childbearing also has been shown to have adverse effects on the children born to teen mothers. The weight of children of young mothers is commonly below the average level. Those children tend to experience more childhood health problems and receive doctoral treatments more frequently than the other children do. Sullivan et al. (1994) found that the risk of a child dying within the year of birth is 30 percent higher if the mother is 15 to 19 than if she is 20 to 29. Unfortunately, the risks for children of teen mothers are not just limited to their health status. Children of teen mothers are themselves more likely to become teen parents (Kahn and Anderson 1992; Bane and Ellwood 1986) and have to deal with all the problems associated with teenage childbearing.

Finally, teenage childbearing is costly to the public sector such as federal, state, and local governments and the taxpayers who support them. For instance, Hoffman (2006) shows that the public costs are as much as \$9.1 billion in 2004 which are mostly related to the children of the teen mothers. The component of these costs are \$1.9 billion for increased public sector health care costs, \$2.3 billion for increased child welfare costs, \$2.1 billion for increased costs for state prison systems, and \$2.9 billion lost in tax revenue because of the lower tax revenue associated with the children of teen mothers.

On the contrary, some scholars such as Geronimus and Korenman (1992); Geronimus, Korenman, and Hillemeier (1994) consider early childbearing not as a hardship but actually as an advantage to the teen mothers. As discussed in Lee (2010), the key argument of such scholars relies on the observation that many teen mothers grow up in poor families that are often located in lower status neighborhoods. Therefore, a delay in childbearing might not necessarily ensure better future outcomes. Even more interestingly, these scholars point out that teenage childbearing might be a culturally rational response to poverty. They argue that pregnant teenagers could think that giving birth at this early stage in life is an adaptive strategy which would allow them to receive support from their families and their community. Such arguments are consistent with recent studies suggesting that the effects of early childbearing might not last in the long run, i.e. 10 to 15 years after their giving birth. Those effects will monotonically decrease as the mothers and their children age. Teen mothers might develop skills to adapt their lives better to the difficult conditions they have had to experience. This adaptation could come from either more working experiences or higher earnings than those they would have had postponed their motherhood (Hotz, McElroy, and Sander 2005). The impact of early childbearing on adult outcomes is therefore ambiguous.

### 2.2.2 Econometric Methods

There is a large literature studying the causal effect of teenage childbearing on the socioeconomic outcomes of teen mothers. A general research question related to teen mothers' outcomes is "What would have happened to the young woman had she not given a birth as a teen?"

In answering this question, note that it is not feasible to identify these mothers' outcomes as we only observe each teenager in one situation, either as she is a teen mother (treatment group) or as she is not (control group). Randomized assignment to treatment and control group is not feasible for the study of teenage childbearing and most other social science research.[1] In addition, teenage childbearing is not naturally random because teen mothers tend to be raised in the impoverished families and communities. They are not representative of the entire population of teenagers. For example, teen mother usually comes from a single-parent family, lives in poverty neighborhood, and has parents with limited education. Therefore, it is difficult to identify whether negative outcomes of teen mothers such as their low commitment to school, low educational achievement, and possibility of engaging in risky, criminal behaviors are causal effect of teenage childbearing or lower socioeconomic status of their families and poor community. This is an important selection bias problem that

[1] See Lee (2010) for further details.

needs be addressed. To account for both observed and unobserved differences between teen mothers and other young women, several approaches have been proposed.[2]

The first approach is to use regressions and control for observable factors associated with economic status of teen mothers in their later years and then examine the differences in outcomes between teen mothers and other women at time of the analysis. This approach requires a strong assumption that the dummy variable specifying whether a teen is mother or not conditioning on observable factors (intact family, mother's and father's education, race, etc.) be uncorrelated with other unobserved factors (Upchurch and McCarthy 1990). The estimates of the effect of teenage childbearing are typically negative and large from this approach. This over-estimation of teenage childbearing effects occurs because this approach does not count for unobserved disadvantage characteristics of teen mothers such as living in poor environment, porvety neighborhood, or having negative peer effects. These unobserved factors have positive effect on teenage childbearing but negative effect on teen mother's outcomes. Therefore, without controlling these variables in the model the effects of teenage childbearing are over-estimated.

The second approach is to use instrumental variables. Various instrumental variables have been used to estimate the causal effects. Different choices of instrumental variables identify different causal parameters. For example, Geronimus and Korenman (1992) or Hoffman, Foster and Furstenberg (1993) use the outcomes of an adolescent mother's sister who did not experience early childbearing as the hypothetical outcomes for the teenage mother. It has been pointed out that the estimates are biased because the woman who had a child as a teen and her sisters/siblings faces different socioeconomic conditions because family size, family status change over time. For example, the effects of teenage childbearing is negative and upward biased if teen mothers' lower level of achievement than their sisters could not be controlled for. On

[2] For example, see Hoffman (1998) and Lee (2010) for more review.

the other hand, the effects is negative and downward biased if the sister who has a birth gets more parent's resources than the non-childbearing sister. Moreover, births are not naturally random even across daughters within the same family.

Another IV approach used in identifying the causal effect of teenage childbearing is to contrast the outcomes of a teen mother who had twins as a teen to the teen mother who had only one child in her first birth. Grogger and Bronars (1993) argued that the estimated effect will be more negative in this analysis because the marginal effects of having two children as a teen compared to having one child will be larger than that of having one child as a teen relative to having no children as a teen. Therefore, the negative effects of teenage child bearing are overestimated under this approach.

Recently, miscarriage has been used as a "natural experiment" instrumental variable (for the absence of birth) in the estimation of causal effect of having a child as a teenage. This instrument was firstly proposed by Hotz, McElroy and Sanders (2005) who show that the incorporation of miscarriage as a natural instrument helps construct unbiased estimates of the causal effects. However, miscarriage could be non-random both from medical and from social perspective. The estimates hence might not be correct. Miscarriage becomes non-random if a teen pregnant is able purposely to cause spontaneous abortion. In fact, among the population of pregnant teenagers, abortion is more common among teenagers with high economic status. Miscarriage itself occurs more frequently in the group of disadvantaged teen mothers. Additionally, the epidemiological literature suggests smoking and the use of certain drugs such as cocaine, heroine, caffeine before or during pregnancy could be viewed as risk factors for miscarriage (Curtis, Savitz, and Arbuckle, 1997 or Augood, Duckitt, and Templeton 1998). The use of contraceptive methods during the early stage of pregnancy also increases the probability of miscarriage (Ford and MacCormac 1995)

The third approach is to use a matching method. For example, Sanders, Smith and Zhang (2007) apply semi-parametric kernel matching estimator and Lee (2010) applies propensity-score matching. In propensity score matching, two young women, one of whom is a teen mother and the other is not are matched on the same preexisting observed characteristics. This approach generally finds more modest but nontrivial adverse effects although some estimates remain large and some suggest no effect.

## 2.3 Models and Empirical Approach

As discussed in previous sections, my empirical strategy in this paper is to examine the effect of giving birth as a teen on mothers' outcomes in young adulthood and adulthood. The empirical model can be summarized as follows:

Let $Y_i$ be a socioeconomic outcome of woman $i$. This outcome in the data could be annual earnings, years of education, take up of public assistance, or woman fertility. Let $B_i$ be a dummy variable that indicates how the pregnancy of this woman ends: $B_i = 1$ for giving birth as a teen and $B_i = 0$ otherwise. I restrict my empirical approach to linear models in order to estimate the the causal effect of $B_i$ on $Y_i$,[3] (noted as Model 1 below) i.e.

$$Y_i = \alpha + \beta B_i + \theta X_i + \varepsilon_i$$

where $\alpha$, $\beta$, $\theta$ are model parameters, $X_i$ is a vector of covariates and $\varepsilon_i$ is an error term. In this equation, $\beta$ quantitatively measures the marginal impact of teenage childbearing on a mother's outcome.

In the above set-up, $B_i$ is not always observed. For an illustration of the estimation problem, let $Y_i^1$ be the outcome of woman $i$ who gives birth as a teen ($B_i = 1$) and let $Y_i^0$ be the outcome of the same woman otherwise ($B_i = 0$). In a more general setting, Angris and Imbens (1991) pointed out a way to express the average treatment effect

[3] For example, Fletcher, Jason and Barbara (2009) use this framework for estimation.

$$E(Y_i^1 - Y_i^0|X, B_i = 1) = E(Y_i^1|X, B_i = 1) - E(Y_i^0|X, B_i = 1)$$

The first part in the right hand side of the above equation, $E(Y_i^1|X, B_i = 1)$ can be easily identified directly from data. However, the second part, $E(Y_i^0|X, B_i = 1)$ is missing. One solution to this missing variable issue is to use an instrumental variable (IV) for the missing part. The choice and validity of the instrumental variable are crucial for an unbiased estimation. Grogger and Bronars (1993) used twin births. Geronimus and Korenman (1992) and Hoffman, Foster and Furstenberg (1993) used siblings. As pointed out in the previous sections, those instruments could lead to a biased estimation.

Instead of using twin births or siblings, I follow a recent empirical strategy introduced by Hotz, McElroy, and Sander (2005) by using miscarriage ($M_i$) to be an instrument for $B_i$ where $M_i = 1$ if a woman experience a miscarriage and $M_i = 0$ otherwise. Miscarriage could be seen as the closest to the natural experiment. The instrumental variable method will produce IV estimate for $\beta$ via the following model (noted as Model 2 below)

$$Y_i = \alpha + \beta M_i + \theta X_i + \varepsilon_i$$

For the validity of using miscarriage as an instrumental variable, Hotz, McElroy, and Sander (2005) showed that the following conditions should be met: (1) all miscarriages are random; (2) all fertility events are correctly reported; (3) having a miscarriage or an abortion has the same direct effect on $Y_i$. In application, these conditions are not always satisfied. For example, the epidemiological literature shows that smoking, using certain drugs (alcohol, cocaine, and heroine) at certain level, or using some contraceptive methods at early age and at early stage of pregnancy are believed to affect the likelihood of miscarriage.[4] In this paper, I also look at this issue through the Model 3:

$$Y_i = \alpha + \beta M_i + \theta X_i + \pi Z_i + \varepsilon_i$$

where $Z_i$ includes covariates such as smoking, drinking, and the use of contraceptive methods prior to the first pregnancy of mothers.

In particular, I implement three models. Model 1 is ordinary least square (OLS) in which I only control for teen birth by using a dummy variable to specify whether a woman gives birth as a teen and other covariates associated with disadvantaged backgrounds of teen mothers such as race, cognitive ability, parent's education, and family income. The control group in the OLS model includes all woman who reported a teen pregnancy.[5] Model 2 applies two-stage least squares using miscarriage as an instrumental variable for giving birth as teenagers. In this model, the control group consists of women whose teen pregnancy ended at miscarriage. And finally, in Model 3, I use the covariates in Model 2 and then incorporate other behavior variables considered to be correlated with the incidence of miscarriage such as smoking, drinking, and the use of contraceptive methods prior to the first pregnancy of mothers. By implementing Model 3, I am able to check the robustness of the natural randomness of miscarriage as a valid instrument. This is an important check because non-random miscarriage would lead to a biased estimate of the negative effects of early childbearing on mothers' outcome. Specifically, covariates such as smoking and drinking may be positively correlated with teenage childbearing and negatively correlated with teen outcomes. Therefore, the effects of teenage childbearing on mothers could be quantitatively negative but over-estimated.

Another important step of my empirical strategy is the specification of young

[4] See Hassan and Killick (2004) for details.

[5] I do not include all women in my OLS model following the approach in Hotz et al. (2005).

adulthood and adulthood to capture the short and long run effects of teenage childbearing. As discussed in the previous section, earlier studies on this topic examine the effects of teenage childbearing on a year to year basis, for instance estimating effect at each year of age 18 to 32. Several papers such as Hotz, McElroy, and Sander (2005) found early years result in the significantly negative impacts on the outcomes while the later years do not. However, they do not specify formally the short and the long run period of teen mothers's life cycle. For example, they make conclusions that teen mothers earn an average of 24 percent less per year during their early twenties, 43 percent during their late twenties, and 27 percent during the early thirties. In addition, their estimates could be affected by the limited size of the yearly data. In fact, their estimates show marginal differences across years. One of my contributions is to provide an aggregate analysis on the short run period (young adulthood) and long run (adulthood). Specifically, I analyze the effects of teen births on teen mothers in two aggregate separated stages, Stage 1 from 18 to 24 years old and Stage 2 from 25 to 32 years old, defined as the short-run and long-run samples, respectively.[6] By doing so, the sample size of each stage is much larger. For example, Stage 1 with 5736 observations including 4117 cases of ending at birth, 392 ending at miscarriage and 1227 ending at abortion. In stage 2, these numbers are 4928, 3623, 358 and 947, respectively.[7] However, in these samples there are multiple observations observed on the same individual. To take account of this fact, the standard errors are clustered by individual identification in all models. In addition, due to the varying ages and the year of birth, in the set up of the extended model, I add controls for ages and birth year in each stage.

[6] In the next chapter, the cutoff point of 24 year old between young adulthood and adulthood is consistent with the cutoff point of these two stages defined in the Add Health data.

[7] First, I group observations into groups by age (age18 to age 32). There are 15 groups. Then I pool all observations into two big groups, one includes women who are between 18 and 24 years old and the other includes women who are between 25 and 32 years old.

## 2.4 Data

My analysis focuses on the female sample from 1979 cohort of the National Longitudinal Survey of Youth (NLSY79). This is a nationally representative sample of 12,686 young men and women who were from 14 to 21 years old as of December 31, 1978. These individuals were interviewed annually till 1994 and then were biennially updated by the year 2002. As this paper looks at the causal effects of teenage childbearing in two stages from 18 to 32 years old, I use data sample to year 1998, as the youngest female in the cohort who was at age 14 in 1979 turned 32 by this time.[8] As in previous studies which use NLSY79, I exclude the economically disadvantaged white supplementary and military samples from our empirical analysis because most of females from military sample were dropped from the interviews after 1984 and all of the economic disadvantaged white supplementary have not been interviewed since 1990. The number of sample cases in 1979, excluding the discontinued military and nonblack/non-Hispanic samples, was 9,964 and this number in 1998 was 7,565.[9]

Table 2.1 presents summary statistics of teen pregnancy and non-teen pregnancy samples. A teen is included in the teen pregnancy sample if her first pregnancy is oc-

[8] Note that attrition is an important issue that needs to be addressed in the use of NLSY79 data set as it could potentially lead to attrition bias in the estimation of the effects of teenage childbearing in the short run and long run. The current chapter of this dissertation does not account for this attrition issue. I build on the results in Aughinbaugh (2014) which shows that effects of this attrition issue on results of analysis appear to be minimal. In particular, Aughinbaugh (2014) directly looks at the attrition problem in the NLSY79. This study reports rates of attrition and first-attrition for the full sample and the cross-section among women in the NLSY79. Attrition has been relatively low among the women of the NLSY79, although it has increased over time. The attrition rates range from 3 percent in the early years of the survey to about 13 percent in 1998 with the patterns being quite similar for the full-sample of women and the cross-section (see Table 1 in Aughinbaugh 2004). The study presents descriptive statistics of two different samples. One includes all person-year observations (full sample) and the other includes only annual observations who have never missed an interview (nonattritors). The comparison between two samples shows that nonattritors have family incomes that are slightly larger than those in the full sample. However, other characteristics (marriage, childbearing, poverty, and educational level) seem to be unaffected by the exclusion of the attrition (shown in Table 2a in Aughinbaugh 2004). In sum, these findings imply that althougth attrition may be nonrandom over the years the effects of this attrition on results of analysis appear to be minimal.

[9] Hotz, McElroy and Sanders (2005) use based-year weight for the combination of the cross-sectional data and supplemental samples to address attrition issue in the NLSY79. However, the information on this weight is not available in the public-released version of the NLSY79.

curred prior to her $18^{th}$ birthday. Non-teen pregnancy sample refers to women whose first pregnancy are delayed after their $18^{th}$ birthday.[10] By combining the information on date of a first pregnancy and birthday of respondents, I end up with a sample of 974 teen pregnancies and 2,443 of non-teen pregnancies.[11] Three sub-samples from the teen pregnancy sample are then constructed on the basis the outcome of that pregnancy. The data is coded into four outcomes of pregnancy: birth, abortion, miscarriage and stillbirth. I merge stillbirth and miscarriage into one outcome as there are only 6 cases in which a teen pregnancy ended at stillbirth. There are 706 teen pregnancies ended at birth, 68 ended at miscarriage and 200 ended at abortion. Background characteristics of these samples and sub-samples are reported in Table 2.1 and Table 2.2, respectively.

The first three rows of the Table 2.1 summarize the early maternal characteristics of women whose first pregnancy happened prior to the $18^{th}$ birthday. It contains two dummy variables of whether a woman was black or Hispanics. By doing this I can capture the pre-existing disadvantage of the black and Hispanic females, often seen to have higher risk of teenage childbearing, compared with the white population. On average, 42 percent of women in the teen pregnancy sample is black while that number for the non-teen pregnancy sample is 25 percent. This suggests that blacks are more likely than whites to become pregnant as teens. This, however does not hold for Hispanics with their shares being17 percent for the teen pregnancy sample versus 18 percent for the non-teen pregnancy sample.

The third row shows the Armed Forces Qualifying Test (AFQT) composite score variable - a proxy for the woman's cognitive ability.[12] Lower levels of cognitive ability

[10] Women who delay their first pregnancy after their 18th birthday tend to come from higher socioeconomic status and have higher test scores than women whose pregancies prior to their 18th birthday.

[11] The sum of teen pregnancies and non-teen pregnancies is less than the number of observations in each stage sample (5736 in Stage 1 and 4928 in Stage 2) because one individual can be observed more than once in each sample.

[12] Other papers such as Hotz, Mc Elroy, and Sander (2005) also use AFQT score as a proxy for the woman's cognitive ability.

could be associated increased risk of teenage childbearing. As seen from the table, women in the teen pregnancy sample had an average lower AFQT score (73 points) than those in the non-teen pregnancy sample (90 points).

The last rows of Table 2.1 provide variables on family environment which were collected when the subjects were 14 years old. In the data, I use two dummy variables for family structure, specifically whether the subject lived in an intact family at age 14 and whether she lived in a female-headed family at age 14. Variables reflecting family knowledge include: a dummy whether magazines were in the home at age 14; a dummy whether newspapers were in the home at age 14; a dummy whether a library card was in the home at age 14; mother's and father's education; and indicators for missing parent education. To measure family income, I use family income in 1978, a dummy for whether family received welfare in 1978, and an indicator for missing family income. As seen in the Table 2.1, women in the teen pregnancy sample come from more disadvantaged backgrounds than those in the non-teen pregnancy sample. For example, at the age of 14, family income in the teen pregnancy sample was an average of \$9,484 per year compared to \$13,696 in the non-teen pregnancy sample. The families of teens who got pregnant were more likely get welfare in 1978: 21 percent for teen pregnancy versus 2 percent for non-teen pregnancy. In addition, the parents of teens in the teen pregnancy sample had 10.0 years of education on average while those in the non-teen pregnancy sample had 10.8. In addition, at age 14 they were less likely to live with both their mother and father with the difference being 54 percent for the teen pregnancy sample versus 70 percent for the non-teen pregnancy sample. Chances of receiving public information from other resources were also lower for the teen pregnancy: 39 percent versus 58 percent for magazines, 65 percent versus 74 percent for newspapers, and 62 percent versus 73 percent for library cards.

Table 2.2 shows the background characteristics of the three teen pregnancy subsamples grouped by pregnancy outcome with same set of variables as Table 2.1.

Each group includes all women who actually give a birth as teen, whose pregnant outcome ended at miscarriage or ended at abortion, respectively. As shown in Table 2.2, the women who experienced an abortion at first pregnancy as a teen came from more advantaged backgrounds than those who actually gave a birth or experienced miscarriages. For example, the women who experienced abortions had an average AFQT score of 91 points compared to 80 points for women who gave birth and 67 points for women who experienced miscarriages. Their family income was higher, i.e. \$14,287 versus \$8,129 and \$9,417. They were less likely to be black or Hispanic and their parents had higher education.

## 2.5 Estimation Results

I explore the effects of teenage childbearing on teen mothers in two aggregate stages of young adulthood (Stage 1 - short- run) and adulthood (Stage 2 - long-run) and then look for the differences between them. Table 2.3 presents a summary of the key impacts of teenage childbearing on teen mother's outcomes at both Stage 1 and Stage 2 including years of schooling, annual earnings, public assistance income (annual amount of Food Stamps and AFDCF received), and the number of children (excluding a teen birth born). For an overview, the numerical values are coefficients of the teen birth variable extracted from Tables 2.4 - 2.11. OLS estimation is shown in Model 1 and IV estimations are in Model 2 and Model 3 for the two stages. In comparison with IV models, the estimates of the effects of teenage childbearing using OLS are largely negative compared to those of IV models for the education and labor market outcomes (Model 2 and 3). In addition, estimates of Model 2 and Model 3 are very close to one another and far from those of Model 1. This suggests the bias of the estimates of Model 1 as reported in earlier studies on this topic.[13] For instance,

[13] I also follow the literature by implementing estimation results by each age from 18 to 32. As the results by age are numerous and are not the main focus of the paper, for brevity, I do not present

estimates of the negative effects of teenage childbearing captured by Model 2 and 3 are -0.31 and -0.19 while the estimate of Model 1 is -1.79. The estimates of the effect of teenage childbearing on annual mother's annual earnings is -$1,470 for Model 1, compared to -$879.79 and -$885.98 for Model 2 and Model 3, respectively. For the public assistance income, the positive estimates in all 3 models ($1,007.29, 1212.90$, $1115.93 for Model 1, 2, and 3, respectively). I also look at the subsequent fertility outcomes of the mothers.

Besides, the insignificant difference between the estimates of the effects between Model 2 and Model 3 implies that the miscarriage is a good instrument variable.

In the following sections, I present the detailed implementation and empirical results for the four outcomes: educational attainment, annual earnings, receipts of public assistance and subsequent fertility. Those outcomes are important to teen mothers.[14]

### 2.5.1 Impact on Educational Attainment

There are several ways to measure the outcome of educational attainment of teen mothers. One way is to look at whether teen mothers earned high school diplomas or obtained General Educational Development (GED). Other researchers use indicators for whether the teen mothers had enrolled in a college by a certain age as a measure of the education attainment.

Different from the above approaches, I use years of schooling as a measure of education achievement. The rational for this choice is that most of the teen mothers in the sample obtained their high school diploma or GED by early their 20s in Stage 1 and some attended colleges in Stage 2. It is not possible to make a comparison of education outcomes in two stages if they are in different measures. By using the same measure (years of schooling) I can see whether a teen mother improve her education

tables and figures for estimation results by age. They are available upon request.

[14] See for example Hoffman (2006) for further discussion.

level in the future.

The key results for the estimates of teenage childbearing are summarized in Table 2.3 for both stages and the detailed model estimation is presented in Table 2.4 for Stage 1 and in Table 2.5 for Stage 2. As seen in Table 2.3, there is little negative impact of teen childbearing on years of schooling under all three models and in both stages. All coefficients show small negative impact which ranges from -0.11 to -1.78 years and are smaller in Stage 2 compared to Stage 1. As discussed in the previous sections, the OLS model will overstate the effects of the teenage childbearing. The introduction of miscarriage as an instrument variable helps resolve this issue. Model 2 is implemented without controlling for factors that might affect the natural randomness of miscarriage while in Model 3 I improve Model 2 by incorporating risky behaviors such as smoking before pregnancy, drinking before pregnancy, and the use of conceptive methods.[15] My results are consistent with the findings in the literature that the OLS estimator is overestimated toward an adverse effect. Estimated coefficients from OLS model are statistically significant at 1 percent in Stage 1 and 10 percent in Stage 2 while the estimates coefficients from IV models (Model 2 and Model 3) are not significant.

The fact that the coefficients are smaller in Stage 2 compared to Stage 1 suggests that teen births have smaller effect on education attainment in young adulthood than in adulthood. An explanation could be that in the long run teen mothers are able to achieve education levels similar to that they would have achieved if they had delayed childbearing.[16]

Now I turn to the other variables of interest introduced in Model 3. The use of contraceptive methods before first pregnancy does not have an impact on years of schooling at either of two stages. Drinking alcohol before first pregnancy is signifi-

[15] Hozt et al. (2005) use variables such as smoking before pregnancy, drinking before pregnancy, and whether a woman had a pregnancy before age 16.

[16] These results are similar to the findings in Fletcher et al. (2009).

cantly and negatively correlated with years of schooling in Stage 1, but there is no correlation in Stage 2. Smoking actually affects years of schooling at both stages though the effect is more precisely estimated for Stage 1 than Stage 2.

For other covariates examined in all three Model 1, Model 2, and Model 3, farther's education, black and Hispanic variables are shown to be significantly correlated with teen mother's attainment outcome in both stages. The results indicate that in the teen pregnancy sample, on average blacks and Hispanics are more likely to have higher years of schooling than whites. Father's education could have an impact on teen mothers' years of schooling. Besides, all ages in Stage 1 are significant while all ages in Stage 2 are not. This result is similar to the analysis using yearly data found in earlier studies.

### 2.5.2 Impact on Annual Earnings

The annual earnings of a teen mother represents her success in labor market, an important outcome for the analysis of the causal effects of teenage childbearing.[17] Summary results in Table 2.3 show that in Stage 1, all three models indicate that teenage childbearing negatively affects earnings. This suggests that in young adulthood teen mothers earn less than if they had delayed or avoided child birth. In particular, giving birth as a teen resulted in a loss of an annual estimated amount of $1,470 under Model 1. The results reported under Model 2 and Model 3 are negative and statistically insignificant. Note that both Model 2 and Model 3 are the IV models and the estimated coefficients are very similar.

In Stage 2, I find a negative impact of teen birth only in the OLS model. Specifically, the earnings of teen mother are $1,277 less than if she had not been a teen mother. Interestingly for both IV Models 2 and 3, although the estimated coefficients are positive, opposite to the result from OLS and the results in Stage 1 they are not

[17] As the number of respondents who reported zero annual earnings is small I include them in my analysis to increase the sample size.

statistically different from zero. These findings imply that there are no impact of teenage childbearing on annual earnings in Stage 2 under IV models.[18]

Intuitively, the above results suggest that teenage childbearing is a negative factor in the labor market because teen mother had to raise the kids and did not have the skills and experiences demanded from the employers. However, the effect goes from negative to zero. As they had to adapt to the hardship condition by working in the early adulthood, the experiences and skills developed in this stage might help the teen mothers to be well-accepted in the labor market in the second stage. This indicates that in long-run a teen mother could be able to overcome the disadvantage.

The detailed results in Table 2.6 and Table 2.7 for Stage 1 and Stage 2, respectively provide similar results in term of the effects across Age variable. In particular, all ages in Stage 1 are significantly correlated with annual earnings while all ages in Stage 2 are not. In addition

Results from both IV models are very similar in early ages (Stage 1). Note that the difference between two IV models is the incorporation in Model 3 of risk factor covariates such as smoking, drinking behaviors and the use of contraceptive methods. Similar to the findings for years of schooling, the use of contraceptive method as a risk factor does not affect earnings. However, smoking and drinking behaviors are significantly correlated with earnings in both the short run and long run. This is consistent with previous studies which find negative effects of smoking and drinking on wages and earnings (Hotz, McElroy and Sanders 2005).

### 2.5.3 Impact on Public Assistance Income

Teenage childbearing is considered a very costly social issue due to the potential public aid payment made to mother and the children born to her.[19] The previous

[18] This finding of a positive estimated coefficients of teen birth on annual earnings of teen mother in Stage 2 is consistent with the results from Hotz, McElroy, and Sander (2005).

[19] See Hoffman SD (2006) for further discussion.

section on annual earnings shows that having a child as a teen negatively affects the teen mother's income. It's therefore important to examine the link between teenage childbearing and the receipt of public assistance in short run and in the long run. In the NLSY79 data, respondents were asked about the receipt of public assistance from all resources such as Food Stamp, Aid to Families with Dependent Children (AFDC), Medicaid, and the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC). However, only the questions regarding Food Stamps and AFDC were answered consistently through all interviewed years by respondents. Therefore, I combine both the annual amount of Food Stamps and AFDC to be the amount received from public assistance.

As reported in Table 2.3, the impact on public assistance income in Stage 1 is large and statistically significant for all models. In particular, a teen mother in this stage received an additional amount of about \$1,007 according to Model 1, \$1,200 according to Model 2, and \$1,115 according to Model 3. In Stage 2, these numbers become very small and insignificant in all models (around \$125 in Model 1, \$260 in Model 2 and \$180 in Model 3). The results are consistent with the annual earning results as the teen mothers in the long run might equip themselves with better skills to help them more succeed in labor market. Also, their children at this stage are in school so that mothers can work more. Therefore, mothers earn more from working and receive less public assistance income. Note that estimated effects from both IV models are very close implying the miscarriage is a good instrument.

These findings are straightforward to interpret. Teen mothers receive more assistance from public programs when their children are small. In later years, they might receive less support from public as their children age. The detailed results in Table 2.8 and 2.9 are similar to the earlier findings.

### 2.5.4 Impact on Subsequent Fertility

My paper also presents empirical evidence on fertility of the teen mothers. I look at the number of children, not including the teen birth born at each stage. This measure helps us see whether having a first birth as a teen affects the possibility of having additional births in later. The impact of teen childbearing on the number of children born to teen mother in two separate stages is illustrated in the Table 2.10 and Table 2.11. The estimated coefficients of teen childbearing on the subsequent fertility is always less than 1 and statistically significant in all three models and in both stages.

In Stage 1, it is highest at 0.98 under OLS estimation (Model 1) and it falls to 0.62 and 0.53 under IV estimations (Model 2 and 3). Similarly, in Stage 2 the number falls from 0.72 to 0.46 and to 0.32. Note that in both stages, estimations reported in Model 3 take into account the risk factor covariates such as smoking, drinking and using contraceptive behavior. The estimated coefficient in Model 3 are slightly bigger than that of Model 2, i.e. 0.53 versus 0.62 in Stage 1 and 032 versus 0.46 in Stage 2. It implies that these risk factor covariates might contribute bad influences on subsequent fertility. Therefore, the statistical significance of coefficients at 1 percent might not be enough to suggest a strong correlation between teen birth and the number of children born to teen mother given the fact that all estimated coefficients are always less than 1. The small impact of teen birth on number of children born to teen mother could be caused by the relationship between having one child from a first birth as a teen and the addition number of children in the future. The detailed results in Table 2.10 and 2.11 show that except for OLS model, both IV models give estimated coefficients less than 1 at all ages as in Table 2.3.

## 2.6 Conclusion

This paper revisits the issue of teenage childbearing on socioeconomic outcomes. It adds further empirical evidence to the literature on this topic by separately examining the short and the long run impacts of teenage childbearing. I raise this research question because one might be interested in exploring whether the negative effects of teenage childbearing on teenage mothers are permanent or temporary and whether teen mothers develop any skills that support for them in their long-run. In the empirical analysis, I implement models that have been used in the literature. In addition, my Model 3 also incorporates three different risk factors in the analysis of miscarriage as an instrumental variable.

The results suggest the following: First, there are small negative impacts of childbearing on teen mother's education attainment in young adulthood (short run) but not in adulthood (long run). Second, the impact of teen birth on annual earnings goes from negative to zero when teen mother transitions from young adulthood to adulthood. Third, the causal effects of teen birth on public assistance income are quite clear and there are evidences that subsequent fertility of teen mothers are affected by teenage childbearing.

In sum, though teenage childbearing has negative impact on education, labor market outcome, and public assistance income in the short-run these effects disappear in the long-run. These findings imply that a teenage birth does not make a teen mother's socioeconomic outcomes very much worse, and it might be even no impact in the long-run. The disappearance of the disadvantage of giving birth early as a teen in the long-run might be explained by the hypothesis that teen mothers have advantages in labor market than young women who delayed their childbearing after teens in terms of work experiences and early participations. It also might be the hypothesis that the supports from public assistance programs in the short-run beneficed teen mothers whose socioeconomic status were low such as poor families and tough neighborhoods.

The findings on the effect of teenage childbearing in the short-run and the long-run have important social and public policy implications, especially welfare reform.

To make the life of teen mothers better, policy makers need to know where, when and how the policies might best intervene to their lives. Answers to these questions might depend on the dynamics of relationship between teenage childbearing and teen mothers' socioeconomic outcomes.

**Table 2.1: Background Characteristics of Teen Pregnancy and Non-Teen Pregnancy**

| Variable | Non-Teen Pregnancy | Teen Pregnancy | Total |
|---|---|---|---|
| Black | 0.246 | 0.424 | 0.297 |
| Hispanic | 0.181 | 0.171 | 0.178 |
| AFQT score | 90.233 | 73.185 | 85.571 |
| Family Income 1978 | $13,696.682 | $9,484.253 | $12,495.958 |
| Family in Welfare 1978 | 0.027 | 0.214 | 0.081 |
| In intact family at age 14 | 0.708 | 0.547 | 0.662 |
| In female-headed household at age 14 | 0.526 | 0.56 | 0.535 |
| Magazines in home at age 14 | 0.583 | 0.394 | 0.529 |
| Newspaper in home at age 14 | 0.747 | 0.654 | 0.726 |
| Library card in home at age 14 | 0.739 | 0.623 | 0.706 |
| Mother's Education | 10.768 | 9.966 | 10.539 |
| Father's Education | 10.834 | 10.101 | 10.625 |
| Missing Mother Education | 0.045 | 0.084 | 0.056 |
| Missing Father Education | 0.124 | 0.237 | 0.155 |
| Missing Family Income 78 | 0.189 | 0.21 | 0.195 |
| Number of observations | 2,443 | 974 | 3,417 |

Data Source: NLSY79; numbers is shown as mean of samples

**Table 2.2: Background Characteristics of Teen Pregnancy by Pregnancy Outcome**

| | Ended at Birth | Ended at Miscarriage | Ended at Abortion |
|---|---|---|---|
| Variables | Mean | Mean | Mean |
| Black | 0.476 | 0.397 | 0.250 |
| Hispanic | 0.184 | 0.162 | 0.130 |
| AFQT score | 67.229 | 80.691 | 91.655 |
| Family Income 1978 | \$8,129.931 | \$9,417.608 | \$14,287.668 |
| Family in Welfare 1978 | 0.258 | 0.221 | 0.055 |
| In intact family at age 14 | 0.528 | 0.559 | 0.610 |
| In female-headed household at age 14 | 0.531 | 0.485 | 0.685 |
| Magazines in home at age 14 | 0.331 | 0.471 | 0.590 |
| Newspaper in home at age 14 | 0.598 | 0.779 | 0.790 |
| Library card in home at age 14 | 0.588 | 0.618 | 0.750 |
| Mother's Education | 9.515 | 10.209 | 11.477 |
| Father's Education | 9.684 | 10.721 | 11.361 |
| Missing Mother Education | 0.093 | 0.088 | 0.050 |
| Missing Father Education | 0.261 | 0.309 | 0.130 |
| Missing Family Income 78 | 0.215 | 0.191 | 0.200 |
| Number of Observations | 706 | 68 | 200 |

Data Source: NLSY79

**Table 2.3: Effects of Teenage Childbearing on Socioeconomic Outcomes**

| Variable | Stage 1 (age 18 to age 24) | | | Stage 2 (age 25 to age 32) | | |
|---|---|---|---|---|---|---|
| Outcome | Model 1 | Model 2 | Model 3 | Model 1 | Model 2 | Model 3 |
| Years of Schooling | -1.79*** | -0.31 | -0.199 | -0.45* | -0.17 | -0.11 |
| | (0.22) | (0.46) | (0.49) | (0.184) | (0.371) | (0.39) |
| Annual Earnings | -1,470.40*** | -879.79 | -885.99 | -1,277.72*** | 775.38 | 1073.40 |
| | (250.63) | (537.37) | (566.34) | (413.51) | (833.60) | (869.28) |
| Public Assistance Income | 1,007.29*** | 1,212.90** | 1,115.93* | 125.35 | 260.10 | 179.50 |
| | (195.21) | (418.39) | (441.84) | (227.33) | (457.14) | (480.20) |
| Number of Children | 0.98*** | 0.62*** | 0.53** | 0.72*** | 0.45*** | 0.32*** |
| | (0.07) | (0.16) | (0.17) | (0.04) | (0.08) | (0.08) |
| Number of Observation | 5,736 | | | 4,928 | | |

Standard errors in parentheses, * p<.1, ** p<.05, *** p<.001

Annual Earning and Public Assistance are adjusted in 2000 dollar.

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

**Table 2.4: Impact of Teen Birth on Years of Schooling at Stage 1**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Teen Birth | -1.785*** | -0.307 | -0.199 |
| | (0.216) | (0.464) | (0.488) |
| Magazines in home at age 14 | 0.201 | 0.287 | 0.296 |
| | (0.209) | (0.211) | (0.210) |
| Newspaper in home at age 14 | 0.368 | 0.395 | 0.413 |
| | (0.212) | (0.213) | (0.212) |
| Library card in home at age 14 | 0.198 | 0.165 | 0.161 |
| | (0.201) | (0.202) | (0.201) |
| In female-headed household at age 14 | 0.090 | 0.143 | 0.038 |
| | (0.184) | (0.185) | (0.184) |
| In intact family at age 14 | -0.111 | -0.168 | -0.112 |
| | (0.198) | (0.198) | (0.198) |
| Mother's Education | 0.101* | 0.128** | 0.128** |
| | (0.039) | (0.040) | (0.040) |
| Father's Education | 0.122*** | 0.133*** | 0.124*** |
| | (0.034) | (0.034) | (0.034) |
| Family in Welfare 1978 | -0.354 | -0.525* | -0.405 |
| | (0.257) | (0.262) | (0.259) |
| Family Income 1978 | 0.000*** | 0.000*** | 0.000*** |
| | (0.000) | (0.000) | (0.000) |
| Black | 2.643*** | 2.560*** | 2.512*** |
| | (0.231) | (0.233) | (0.232) |
| Hispanic | 1.937*** | 2.003*** | 1.910*** |
| | (0.292) | (0.293) | (0.292) |
| AFQT score | 0.044*** | 0.047*** | 0.045*** |
| | (0.004) | (0.004) | (0.004) |
| Missing Mother Education | -0.300 | -0.358 | -0.269 |
| | (0.359) | (0.360) | (0.358) |
| Missing Father Education | -0.064 | -0.093 | -0.103 |
| | (0.240) | (0.241) | (0.241) |
| Missing Family Income 78 | -0.326 | -0.227 | -0.327 |
| | (0.259) | (0.260) | (0.259) |
| Observations | 5736 | 5736 | 5736 |

Standard errors in parentheses

* significant at 5 percent; ** significant at 10 percent; *** significant at 1 percent

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

**Table 2.4: (cont.)**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Born in 1963 | -3.984*** | -3.937*** | -4.159*** |
| | (0.420) | (0.421) | (0.419) |
| Born in 1962 | -3.631*** | -3.504*** | -3.727*** |
| | (0.425) | (0.427) | (0.425) |
| Born in 1961 | -3.626*** | -3.505*** | -3.527*** |
| | (0.416) | (0.418) | (0.415) |
| Born in 1960 | -2.697*** | -2.630*** | -2.784*** |
| | (0.416) | (0.417) | (0.415) |
| Born in 1959 | -2.266*** | -2.165*** | -2.239*** |
| | (0.426) | (0.428) | (0.425) |
| Born in 1958 | -0.249 | -0.288 | -0.441 |
| | (0.445) | (0.446) | (0.444) |
| Age 20 | -1.815*** | -5.968*** | -5.961*** |
| | (0.348) | (0.368) | (0.366) |
| Age 21 | -2.954*** | -7.106*** | -7.100*** |
| | (0.340) | (0.361) | (0.359) |
| Age 22 | -4.222*** | -8.379*** | -8.377*** |
| | (0.340) | (0.361) | (0.359) |
| Age 23 | -5.002*** | -9.159*** | -9.151*** |
| | (0.340) | (0.361) | (0.359) |
| Age 24 | -5.058*** | -9.219*** | -9.212*** |
| | (0.341) | (0.362) | (0.360) |
| Using contraception before first pregnancy | | | -0.070 |
| | | | (0.104) |
| Drinking alcohol before first pregnancy | | | -0.249* |
| | | | (0.114) |
| Smoking before first pregnancy | | | -0.405*** |
| | | | (0.121) |
| Constant | -1.810* | 0.507 | 0.814 |
| | (0.723) | (0.894) | (0.888) |
| Observations | 5736 | 5736 | 5736 |

Standard errors in parentheses

* significant at 5 percent; ** significant at 10 percent; *** significant at 1 percent

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

**Table 2.5: Impact of Teen Birth on Years of Schooling at Stage 2**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Teen Birth | -0.447* | -0.169 | -0.107 |
| | (0.184) | (0.371) | (0.390) |
| Magazines in home at age 14 | -0.087 | -0.076 | -0.048 |
| | (0.176) | (0.176) | (0.176) |
| Newspaper in home at age 14 | 0.093 | 0.102 | 0.100 |
| | (0.180) | (0.179) | (0.180) |
| Library card in home at age 14 | 0.034 | 0.025 | 0.023 |
| | (0.170) | (0.170) | (0.170) |
| In female-headed household at age 14 | -0.001 | -0.000 | -0.029 |
| | (0.154) | (0.154) | (0.154) |
| In intact family at age 14 | -0.529** | -0.540** | -0.536** |
| | (0.167) | (0.167) | (0.166) |
| Mother's Education | 0.056 | 0.061 | 0.061 |
| | (0.032) | (0.033) | (0.033) |
| Father's Education | -0.065* | -0.062* | -0.059* |
| | (0.029) | (0.029) | (0.029) |
| Family in Welfare 1978 | -0.284 | -0.317 | -0.277 |
| | (0.193) | (0.196) | (0.195) |
| Family Income 1978 | 0.000* | 0.000* | 0.000* |
| | (0.000) | (0.000) | (0.000) |
| Black | 0.891*** | 0.876*** | 0.847*** |
| | (0.194) | (0.195) | (0.195) |
| Hispanic | 0.778** | 0.791** | 0.726** |
| | (0.243) | (0.243) | (0.244) |
| AFQT score | 0.030*** | 0.030*** | 0.029*** |
| | (0.003) | (0.003) | (0.003) |
| Missing Mother Education | -0.384 | -0.398 | -0.347 |
| | (0.292) | (0.292) | (0.292) |
| Missing Father Education | 0.397* | 0.393* | 0.351 |
| | (0.199) | (0.199) | (0.200) |
| Missing Family Income 78 | 0.729*** | 0.745*** | 0.720*** |
| | (0.212) | (0.213) | (0.212) |
| Observations | 4928 | 4928 | 4928 |

Standard errors in parentheses

* significant at 5 percent; ** significant at 10 percent; *** significant at 1 percent

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

**Table 2.5: (cont.)**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Born in 1963 | 0.351 | -0.302 | -0.291 |
| | (0.384) | (0.332) | (0.332) |
| Born in 1962 | -0.311 | -0.954** | -0.975** |
| | (0.382) | (0.327) | (0.327) |
| Born in 1961 | 0.559 | -0.080 | -0.039 |
| | (0.375) | (0.305) | (0.305) |
| Born in 1960 | 0.465 | -0.185 | -0.206 |
| | (0.374) | (0.293) | (0.292) |
| Born in 1959 | -0.022 | -0.669* | -0.633* |
| | (0.378) | (0.289) | (0.289) |
| Born in 1957 | 0.135 | -0.527 | -0.485 |
| | (0.385) | (0.290) | (0.291) |
| Age 26 | 0.248 | -0.300 | 0.228 |
| | (0.289) | (0.311) | (0.288) |
| Age 27 | 0.140 | -0.407 | 0.117 |
| | (0.299) | (0.321) | (0.297) |
| Age 28 | 0.237 | -0.308 | 0.224 |
| | (0.292) | (0.313) | (0.291) |
| Age 29 | 0.156 | -0.390 | 0.154 |
| | (0.296) | (0.318) | (0.295) |
| Age 30 | 0.225 | -0.323 | 0.229 |
| | (0.305) | (0.325) | (0.304) |
| Used contraception before first pregnancy | | | -0.099 |
| | | | (0.094) |
| Drinking alcohol before first pregnancy | | | 0.033 |
| | | | (0.097) |
| Smoking before first pregnancy | | | -0.252* |
| | | | (0.106) |
| Constant | -6.342*** | -5.472*** | -5.897*** |
| | (0.589) | (0.697) | (0.698) |
| Observations | 4928 | 4928 | 4928 |

Standard errors in parentheses

* significant at 5 percent; ** significant at 10 percent; *** significant at 1 percent

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

**Table 2.6: Impact of Teen Birth on Annual Earnings at Stage 1**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Teen Birth | -1470.407*** | -879.791 | -885.989 |
| | (250.626) | (537.373) | (566.342) |
| Magazines in home at age 14 | -631.604** | -597.073* | -547.870* |
| | (243.144) | (244.214) | (243.611) |
| Newspaper in home at age 14 | 273.474 | 283.985 | 254.151 |
| | (246.449) | (246.069) | (245.708) |
| Library card in home at age 14 | 254.485 | 241.423 | 250.190 |
| | (233.479) | (233.219) | (232.457) |
| In female-headed household at age 14 | 1084.580*** | 1105.550*** | 1048.575*** |
| | (213.893) | (214.104) | (213.396) |
| In intact family at age 14 | -68.184 | -90.817 | -102.423 |
| | (229.457) | (229.691) | (229.080) |
| Mother's Education | 138.764** | 149.591** | 145.723** |
| | (45.852) | (46.578) | (46.459) |
| Father's Education | -10.308 | -5.960 | 1.211 |
| | (39.627) | (39.697) | (39.696) |
| Family in Welfare 1978 | -2196.112*** | -2264.426*** | -2191.114*** |
| | (298.523) | (302.923) | (300.783) |
| Family Income 1978 | 0.055*** | 0.058*** | 0.052*** |
| | (0.012) | (0.013) | (0.013) |
| Black | 174.718 | 141.362 | 58.387 |
| | (268.542) | (269.313) | (269.102) |
| Hispanic | 609.379 | 635.832 | 469.374 |
| | (338.969) | (338.916) | (339.104) |
| AFQT score | 57.452*** | 58.845*** | 55.547*** |
| | (4.426) | (4.556) | (4.557) |
| Missing Mother Education | -399.987 | -423.083 | -316.779 |
| | (416.895) | (416.422) | (415.121) |
| Missing Father Education | -378.201 | -389.572 | -508.195 |
| | (279.133) | (278.688) | (279.483) |
| Missing Family Income 78 | 406.934 | 446.254 | 358.096 |
| | (300.334) | (301.361) | (300.022) |
| Observations | 5736 | 5736 | 5736 |

Standard errors in parentheses

* significant at 5 percent; ** significant at 10 percent; *** significant at 1 percent

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

**Table 2.6: (cont.)**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Born in 1964 | 739.879 | 734.607 | 617.763 |
| | (519.810) | (518.719) | (518.027) |
| Born in 1963 | -202.459 | -183.445 | -267.618 |
| | (487.828) | (487.028) | (486.317) |
| Born in 1962 | -297.328 | -246.837 | -357.372 |
| | (493.155) | (493.780) | (492.382) |
| Born in 1961 | -606.803 | -558.464 | -546.692 |
| | (482.728) | (483.268) | (481.521) |
| Born in 1960 | 587.328 | 614.042 | 498.344 |
| | (482.982) | (482.431) | (480.830) |
| Born in 1959 | 379.284 | 419.400 | 406.512 |
| | (494.794) | (494.794) | (492.951) |
| Born in 1958 | 1376.057** | 1360.389** | 1262.309* |
| | (517.117) | (516.168) | (515.235) |
| Age 20 | 805.012* | 1852.190*** | 1860.078*** |
| | (403.821) | (426.028) | (424.534) |
| Age 21 | 1755.112*** | 2802.443*** | 2810.261*** |
| | (395.374) | (418.201) | (416.734) |
| Age 22 | 2298.589*** | 3344.072*** | 3346.682*** |
| | (395.217) | (418.024) | (416.554) |
| Age 23 | 3344.058*** | 4389.468*** | 4397.810*** |
| | (395.272) | (418.129) | (416.661) |
| Age 24 | 4231.554*** | 5275.343*** | 5282.366*** |
| | (396.537) | (419.323) | (417.848) |
| Used contraception before first pregnancy | | | -190.201 |
| | | | (120.400) |
| Drinking alcohol before first pregnancy | | | 259.035* |
| | | | (131.819) |
| Smoking before first pregnancy | | | -633.098*** |
| | | | (140.365) |
| Constant | -1913.061* | -3693.420*** | -3123.034** |
| | (839.541) | (1034.677) | (1029.175) |
| Observations | 5736 | 5736 | 5736 |

Standard errors in parentheses

* significant at 5 percent; ** significant at 10 percent; *** significant at 1 percent

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

**Table 2.7: Impact of Teen Birth on Annual Earnings at Stage 2**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Teen Birth | -1277.725** | 775.381 | 1073.409 |
| | (413.508) | (833.600) | (869.285) |
| Magazines in home at age 14 | -1022.502** | -942.538* | -824.990* |
| | (393.720) | (394.515) | (392.384) |
| Newspaper in home at age 14 | 891.442* | 957.673* | 908.249* |
| | (402.796) | (403.254) | (400.512) |
| Library card in home at age 14 | 454.266 | 389.072 | 403.751 |
| | (381.757) | (382.241) | (379.139) |
| In female-headed household at age 14 | 1547.327*** | 1549.639*** | 1465.002*** |
| | (346.236) | (346.049) | (343.970) |
| In intact family at age 14 | 101.918 | 20.928 | 33.072 |
| | (373.540) | (374.428) | (371.368) |
| Mother's Education | 222.775** | 258.790*** | 252.828*** |
| | (72.442) | (73.508) | (73.026) |
| Father's Education | -4.065 | 14.865 | 55.492 |
| | (64.850) | (65.158) | (64.914) |
| Family in Welfare 1978 | -2137.318*** | -2380.891*** | -2213.252*** |
| | (432.105) | (440.327) | (435.249) |
| Family Income 1978 | 0.164*** | 0.175*** | 0.166*** |
| | (0.021) | (0.021) | (0.021) |
| Black | 2457.605*** | 2345.705*** | 2106.275*** |
| | (435.732) | (437.279) | (434.710) |
| Hispanic | 2554.636*** | 2646.618*** | 2244.388*** |
| | (545.625) | (546.293) | (544.378) |
| AFQT score | 120.675*** | 125.490*** | 117.648*** |
| | (6.948) | (7.148) | (7.153) |
| Missing Mother Education | -1296.776* | -1399.531* | -1156.157 |
| | (655.714) | (656.359) | (651.134) |
| Missing Father Education | 927.030* | 899.555* | 583.717 |
| | (446.755) | (446.617) | (445.305) |
| Missing Family Income 78 | 1333.605** | 1454.704** | 1444.278** |
| | (476.087) | (477.741) | (473.719) |
| Observations | 4928 | 4928 | 4928 |

Standard errors in parentheses

* significant at 5 percent; ** significant at 10 percent; *** significant at 1 percent

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

**Table 2.7: (cont.)**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Born in 1963 | -263.304 | -5239.014*** | -5090.480*** |
| | (860.674) | (746.471) | (740.312) |
| Born in 1962 | 1305.463 | -3586.744*** | -3765.270*** |
| | (855.575) | (735.140) | (728.633) |
| Born in 1961 | 1063.567 | -3802.234*** | -3617.335*** |
| | (841.671) | (684.361) | (680.195) |
| Born in 1960 | 1633.176 | -3323.850*** | -3452.811*** |
| | (839.204) | (657.699) | (652.236) |
| Born in 1959 | 2363.532** | -2560.063*** | -2410.896*** |
| | (847.491) | (649.409) | (644.875) |
| Born in 1957 | 1717.171* | -3325.017*** | -3106.581*** |
| | (863.366) | (652.338) | (648.459) |
| Age 26 | -785.860 | -955.038 | -877.116 |
| | (647.292) | (698.523) | (641.470) |
| Age 27 | 230.096 | 60.932 | 119.571 |
| | (669.764) | (721.748) | (663.769) |
| Age 28 | -181.954 | -335.735 | -247.630 |
| | (655.673) | (703.920) | (649.718) |
| Age 29 | -135.146 | -293.844 | -147.347 |
| | (663.773) | (714.493) | (657.685) |
| Age 30 | -132.525 | -304.254 | -123.954 |
| | (684.842) | (729.725) | (678.565) |
| Using contraception before first pregnancy | | | 287.584 |
| | | | (209.851) |
| Drinking alcohol before first pregnancy | | | 776.157*** |
| | | | (216.898) |
| Smoking before first pregnancy | | | -2079.686*** |
| | | | (236.975) |
| Constant | -6036.871*** | -3339.509* | -2871.021 |
| | (1320.127) | (1567.110) | (1556.922) |
| Observations | 4928 | 4928 | 4928 |

Standard errors in parentheses

* significant at 5 percent; ** significant at 10 percent; *** significant at 1 percent

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

**Table 2.8: Impact of Teen Birth on Receipt of Public Assistance at Stage 1**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Teen Birth | 1007.294*** | 1212.903** | 1115.928* |
| | (195.207) | (418.385) | (441.842) |
| Magazines in home at age 14 | 66.907 | 78.929 | 65.116 |
| | (189.379) | (190.139) | (190.057) |
| Newspaper in home at age 14 | -235.317 | -231.657 | -224.330 |
| | (191.953) | (191.583) | (191.693) |
| Library card in home at age 14 | 267.348 | 262.801 | 263.499 |
| | (181.852) | (181.578) | (181.356) |
| In female-headed household at age 14 | -778.222*** | -770.921*** | -738.679*** |
| | (166.597) | (166.696) | (166.485) |
| In intact family at age 14 | -950.826*** | -958.706*** | -968.506*** |
| | (178.719) | (178.832) | (178.721) |
| Mother's Education | 5.098 | 8.868 | 9.782 |
| | (35.713) | (36.264) | (36.246) |
| Father's Education | 49.050 | 50.564 | 49.997 |
| | (30.864) | (30.907) | (30.969) |
| Family in Welfare 1978 | 3145.452*** | 3121.670*** | 3079.261*** |
| | (232.513) | (235.849) | (234.661) |
| Family Income 1978 | 0.005 | 0.006 | 0.008 |
| | (0.010) | (0.010) | (0.010) |
| Black | 1123.520*** | 1111.908*** | 1146.328*** |
| | (209.162) | (209.680) | (209.945) |
| Hispanic | 607.570* | 616.780* | 681.299* |
| | (264.016) | (263.872) | (264.558) |
| AFQT score | -16.228*** | -15.743*** | -14.555*** |
| | (3.447) | (3.548) | (3.555) |
| Missing Mother Education | 108.120 | 100.079 | 59.249 |
| | (324.711) | (324.215) | (323.864) |
| Missing Father Education | -375.100 | -379.059 | -347.120 |
| | (217.411) | (216.980) | (218.044) |
| Missing Family Income 78 | 263.251 | 276.939 | 315.319 |
| | (233.923) | (234.632) | (234.068) |
| Observations | 5736 | 5736 | 5736 |

Standard errors in parentheses

* significant at 5 percent; ** significant at 10 percent; *** significant at 1 percent

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

**Table 2.8: (cont.)**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Born in 1964 | 704.417 | 702.582 | 780.005 |
| | (404.869) | (403.861) | (404.148) |
| Born in 1963 | 1106.640** | 1113.259** | 1182.569** |
| | (379.958) | (379.188) | (379.409) |
| Born in 1962 | 678.460 | 696.037 | 783.600* |
| | (384.108) | (384.445) | (384.141) |
| Born in 1961 | 288.029 | 304.857 | 310.752 |
| | (375.986) | (376.260) | (375.667) |
| Born in 1960 | -17.572 | -8.272 | 55.892 |
| | (376.184) | (375.609) | (375.128) |
| Born in 1959 | -145.479 | -131.514 | -108.315 |
| | (385.384) | (385.234) | (384.584) |
| Born in 1958 | -876.044* | -881.499* | -817.559* |
| | (402.771) | (401.876) | (401.970) |
| Age 20 | 560.734 | 1265.425*** | 1262.904*** |
| | (314.527) | (331.695) | (331.207) |
| Age 21 | 735.203* | 1439.949*** | 1437.323*** |
| | (307.948) | (325.601) | (325.122) |
| Age 22 | 496.796 | 1200.898*** | 1199.831*** |
| | (307.826) | (325.463) | (324.982) |
| Age 23 | 866.223** | 1570.299*** | 1567.321*** |
| | (307.868) | (325.545) | (325.065) |
| Age 24 | 727.866* | 1431.378*** | 1428.789*** |
| | (308.854) | (326.475) | (325.992) |
| Used contraception before first pregnancy | | | -16.389 |
| | | | (93.932) |
| Drinking alcohol before first pregnancy | | | -18.845 |
| | | | (102.841) |
| Smoking before first pregnancy | | | 287.085** |
| | | | (109.508) |
| Constant | 1540.859* | 580.929 | 403.161 |
| | (653.900) | (805.574) | (802.929) |
| Observations | 5736 | 5736 | 5736 |

Standard errors in parentheses

* significant at 5 percent; ** significant at 10 percent; *** significant at 1 percent

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

**Table 2.9: Impact of Teen Birth on Receipt of Public Assistance at Stage 2**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Teen Birth | 125.348 | 260.107 | 179.508 |
| | (227.331) | (457.149) | (480.207) |
| Magazines in home at age 14 | 115.635 | 120.883 | 95.691 |
| | (216.452) | (216.354) | (216.759) |
| Newspaper in home at age 14 | -47.320 | -42.973 | -37.019 |
| | (221.442) | (221.146) | (221.249) |
| Library card in home at age 14 | 150.443 | 146.164 | 145.626 |
| | (209.875) | (209.622) | (209.442) |
| In female-headed household at age 14 | -552.510** | -552.358** | -526.958** |
| | (190.348) | (189.775) | (190.015) |
| In intact family at age 14 | -456.851* | -462.167* | -467.368* |
| | (205.358) | (205.338) | (205.150) |
| Mother's Education | -16.430 | -14.066 | -13.480 |
| | (39.826) | (40.312) | (40.341) |
| Father's Education | -64.968 | -63.726 | -70.193 |
| | (35.652) | (35.733) | (35.860) |
| Family in Welfare 1978 | 2396.988*** | 2381.001*** | 2339.359*** |
| | (237.555) | (241.477) | (240.438) |
| Family Income 1978 | 0.001 | 0.002 | 0.004 |
| | (0.012) | (0.012) | (0.012) |
| Black | 1058.215*** | 1050.870*** | 1094.357*** |
| | (239.549) | (239.805) | (240.141) |
| Hispanic | 537.563 | 543.601 | 622.190* |
| | (299.964) | (299.589) | (300.723) |
| AFQT score | -24.342*** | -24.026*** | -22.545*** |
| | (3.820) | (3.920) | (3.952) |
| Missing Mother Education | 572.605 | 565.860 | 512.661 |
| | (360.487) | (359.949) | (359.697) |
| Missing Father Education | 224.746 | 222.943 | 280.977 |
| | (245.609) | (244.926) | (245.993) |
| Missing Family Income 78 | 46.519 | 54.468 | 66.339 |
| | (261.735) | (261.995) | (261.690) |
| Observations | 4928 | 4928 | 4928 |

Standard errors in parentheses

* significant at 5 percent; ** significant at 10 percent; *** significant at 1 percent

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

**Table 2.9: (cont.)**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Born in 1963 | -608.877 | 1738.903*** | 1715.057*** |
| | (473.166) | (409.367) | (408.960) |
| Born in 1962 | -1454.205** | 899.056* | 934.982* |
| | (470.362) | (403.153) | (402.508) |
| Born in 1961 | -1518.846** | 836.148* | 792.351* |
| | (462.719) | (375.306) | (375.750) |
| Born in 1960 | -2135.996*** | 213.010 | 237.917 |
| | (461.362) | (360.684) | (360.305) |
| Born in 1959 | -2414.859*** | -63.658 | -99.265 |
| | (465.918) | (356.138) | (356.239) |
| Born in 1957 | -2518.640*** | -175.223 | -226.738 |
| | (474.646) | (357.744) | (358.219) |
| Age 26 | -643.179 | -1160.522** | -623.877 |
| | (355.857) | (383.072) | (354.358) |
| Age 27 | -698.261 | -1215.603** | -674.971 |
| | (368.211) | (395.809) | (366.676) |
| Age 28 | -679.963 | -1196.295** | -662.744 |
| | (360.464) | (386.032) | (358.915) |
| Age 29 | -419.011 | -935.667* | -415.587 |
| | (364.918) | (391.830) | (363.315) |
| Age 30 | 445.043 | -72.467 | 439.538 |
| | (376.500) | (400.184) | (374.850) |
| Using contraception before first pregnancy | | | -12.541 |
| | | | (115.925) |
| Drinking alcohol before first pregnancy | | | -103.111 |
| | | | (119.818) |
| Smoking before first pregnancy | | | 389.043** |
| | | | (130.909) |
| Constant | 6742.601*** | 4751.516*** | 4113.913*** |
| | (725.756) | (859.409) | (860.069) |
| Observations | 4928 | 4928 | 4928 |

Standard errors in parentheses

* significant at 5 percent; ** significant at 10 percent; *** significant at 1 percent

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

**Table 2.10: Impact of Teen Birth on Subsequent Fertility at Stage 1**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Teen Birth | 0.984*** | 0.617*** | 0.526** |
| | (0.072) | (0.155) | (0.163) |
| Magazines in home at age 14 | 0.153* | 0.132 | 0.136 |
| | (0.070) | (0.070) | (0.070) |
| Newspaper in home at age 14 | 0.034 | 0.027 | 0.013 |
| | (0.071) | (0.071) | (0.071) |
| Library card in home at age 14 | -0.051 | -0.043 | -0.042 |
| | (0.067) | (0.067) | (0.067) |
| In female-headed household at age 14 | -0.101 | -0.114 | -0.083 |
| | (0.061) | (0.062) | (0.061) |
| In intact family at age 14 | -0.017 | -0.002 | -0.023 |
| | (0.066) | (0.066) | (0.066) |
| Mother's Education | -0.018 | -0.025 | -0.026 |
| | (0.013) | (0.013) | (0.013) |
| Father's Education | -0.015 | -0.017 | -0.013 |
| | (0.011) | (0.011) | (0.011) |
| Family in Welfare 1978 | 0.259** | 0.302*** | 0.271** |
| | (0.086) | (0.087) | (0.087) |
| Family Income 1978 | -0.000*** | -0.000*** | -0.000*** |
| | (0.000) | (0.000) | (0.000) |
| Black | 0.131 | 0.151 | 0.152* |
| | (0.077) | (0.078) | (0.077) |
| Hispanic | 0.042 | 0.026 | 0.029 |
| | (0.097) | (0.098) | (0.098) |
| AFQT score | -0.003* | -0.003** | -0.003* |
| | (0.001) | (0.001) | (0.001) |
| Missing Mother Education | 0.284* | 0.298* | 0.280* |
| | (0.120) | (0.120) | (0.119) |
| Missing Father Education | -0.068 | -0.061 | -0.076 |
| | (0.080) | (0.080) | (0.080) |
| Missing Family Income 78 | -0.154 | -0.178* | -0.154 |
| | (0.086) | (0.087) | (0.086) |
| Observations | 5736 | 5736 | 5736 |

Standard errors in parentheses

* significant at 5 percent; ** significant at 10 percent; *** significant at 1 percent

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

**Table 2.10: (cont.)**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Born in 1964 | 3.890*** | 3.893*** | 3.943*** |
| | (0.149) | (0.149) | (0.149) |
| Born in 1963 | 3.963*** | 3.951*** | 4.014*** |
| | (0.140) | (0.140) | (0.140) |
| Born in 1962 | 3.289*** | 3.257*** | 3.307*** |
| | (0.142) | (0.142) | (0.142) |
| Born in 1961 | 2.651*** | 2.621*** | 2.627*** |
| | (0.139) | (0.139) | (0.139) |
| Born in 1960 | 2.166*** | 2.149*** | 2.182*** |
| | (0.139) | (0.139) | (0.138) |
| Born in 1959 | 1.727*** | 1.702*** | 1.722*** |
| | (0.142) | (0.142) | (0.142) |
| Born in 1958 | 0.803*** | 0.813*** | 0.845*** |
| | (0.149) | (0.149) | (0.148) |
| Age 20 | 0.759*** | 1.265*** | 1.263*** |
| | (0.116) | (0.123) | (0.122) |
| Age 21 | 1.427*** | 1.933*** | 1.931*** |
| | (0.114) | (0.120) | (0.120) |
| Age 22 | 1.638*** | 2.145*** | 2.145*** |
| | (0.114) | (0.120) | (0.120) |
| Age 23 | 2.390*** | 2.897*** | 2.895*** |
| | (0.114) | (0.120) | (0.120) |
| Age 24 | 3.270*** | 3.778*** | 3.776*** |
| | (0.114) | (0.121) | (0.120) |
| Used contraception before first pregnancy | | | 0.057 |
| | | | (0.035) |
| Drinking alcohol before first pregnancy | | | 0.134*** |
| | | | (0.038) |
| Smoking before first pregnancy | | | 0.006 |
| | | | (0.040) |
| Constant | -3.847*** | -3.897*** | -3.937*** |
| | (0.241) | (0.298) | (0.296) |
| Observations | 5736 | 5736 | 5736 |

Standard errors in parentheses

* significant at 5 percent; ** significant at 10 percent; *** significant at 1 percent

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

**Table 2.11: Impact of Teen Birth on Subsequent Fertility at Stage 2**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Teen Birth | 0.718*** | 0.457*** | 0.321*** |
| | (0.042) | (0.085) | (0.087) |
| Magazines in home at age 14 | 0.283*** | 0.272*** | 0.262*** |
| | (0.040) | (0.040) | (0.039) |
| Newspaper in home at age 14 | 0.011 | 0.003 | -0.008 |
| | (0.041) | (0.041) | (0.040) |
| Library card in home at age 14 | -0.087* | -0.079* | -0.072 |
| | (0.039) | (0.039) | (0.038) |
| In female-headed household at age 14 | -0.153*** | -0.153*** | -0.114*** |
| | (0.035) | (0.035) | (0.034) |
| In intact family at age 14 | -0.064 | -0.054 | -0.069 |
| | (0.038) | (0.038) | (0.037) |
| Mother's Education | -0.020** | -0.025*** | -0.027*** |
| | (0.007) | (0.007) | (0.007) |
| Father's Education | -0.001 | -0.003 | -0.003 |
| | (0.007) | (0.007) | (0.006) |
| Family in Welfare 1978 | 0.330*** | 0.361*** | 0.312*** |
| | (0.044) | (0.045) | (0.043) |
| Family Income 1978 | -0.000*** | -0.000*** | -0.000*** |
| | (0.000) | (0.000) | (0.000) |
| Black | 0.299*** | 0.313*** | 0.326*** |
| | (0.044) | (0.045) | (0.043) |
| Hispanic | 0.281*** | 0.269*** | 0.302*** |
| | (0.055) | (0.056) | (0.054) |
| AFQT score | -0.004*** | -0.005*** | -0.004*** |
| | (0.001) | (0.001) | (0.001) |
| Missing Mother Education | 0.325*** | 0.338*** | 0.300*** |
| | (0.067) | (0.067) | (0.065) |
| Missing Father Education | -0.175*** | -0.172*** | -0.157*** |
| | (0.045) | (0.045) | (0.044) |
| Missing Family Income 78 | 0.014 | -0.001 | 0.027 |
| | (0.048) | (0.049) | (0.047) |
| Observations | 4928 | 4928 | 4928 |

Standard errors in parentheses

* significant at 5 percent; ** significant at 10 percent; *** significant at 1 percent

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

**Table 2.11: (cont.)**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Born in 1963 | -0.207* | 0.464*** | 0.466*** |
| | (0.088) | (0.076) | (0.074) |
| Born in 1962 | -0.446*** | 0.215** | 0.245*** |
| | (0.087) | (0.075) | (0.073) |
| Born in 1961 | -0.549*** | 0.108 | 0.063 |
| | (0.086) | (0.070) | (0.068) |
| Born in 1960 | -0.505*** | 0.164* | 0.173** |
| | (0.085) | (0.067) | (0.065) |
| Born in 1959 | -0.377*** | 0.288*** | 0.253*** |
| | (0.086) | (0.066) | (0.064) |
| Born in 1957 | -0.465*** | 0.215** | 0.156* |
| | (0.088) | (0.066) | (0.065) |
| Age 26 | -0.528*** | -0.635*** | -0.508*** |
| | (0.066) | (0.071) | (0.064) |
| Age 27 | -0.418*** | -0.524*** | -0.394*** |
| | (0.068) | (0.073) | (0.066) |
| Age 28 | -0.299*** | -0.407*** | -0.283*** |
| | (0.067) | (0.072) | (0.065) |
| Age 29 | -0.174** | -0.282*** | -0.171** |
| | (0.067) | (0.073) | (0.066) |
| Age 30 | -0.095 | -0.201** | -0.098 |
| | (0.070) | (0.074) | (0.068) |
| Used contraception before first pregnancy | | | 0.043* |
| | | | (0.021) |
| Drinking alcohol before first pregnancy | | | 0.103*** |
| | | | (0.022) |
| Smoking before first pregnancy | | | 0.172*** |
| | | | (0.024) |
| Constant | 3.016*** | 2.762*** | 2.644*** |
| | (0.134) | (0.160) | (0.155) |
| Observations | 4928 | 4928 | 4928 |

Standard errors in parentheses

* significant at 5 percent; ** significant at 10 percent; *** significant at 1 percent

Model 1: OLS - Model 2: IV - Model 3: IV with all covariates

# Chapter 3

# An Examination of the Persistence of the Impact of Teenage Childbearing on Labor Market Outcomes Using the Add Health Data

## 3.1 Introduction

The consequences of teenage childbearing on women's future outcomes have received considerable attention from economists and other scientists. The results in this literature are quite mixed. Earlier studies found strong evidence of "short-run" causal effects of teenage childbearing on outcomes such as earnings, educational attainment, receipt of public assistance, and later fertility. For instance, it is more difficult for teen mothers to finish high school and to enter the labor force a couple of years after they gave birth. Teen mothers tend to receive low earnings and have to rely on government support. However, recent research papers looking at the long run effects of teenage childbearing have found that it either does not affect mothers' outcomes if it does or to only a very limited extent. This is particularly true for the labor market outcomes. In addition to the need to check the robustness of these estimated results using better data, it is important to understand and explain the difference in the short- and long-run effects.

Among very few papers that examine the long run effects, Hotz, McElroy and Sanders (2005) look at the changes of the effects from year to year. Although these authors do not formally examine the short run and long run effects, their results show that the effects of teenage childbearing on a woman's life 10 or 15 years after giving birth are small. Chapter 2 of this dissertation revisits this analysis and provides the first complete analysis of the short run versus long run effects of teenage childbearing using the National Longitudinal Survey of Youth cohort 1979 (NLSY79), the same data as in Hotz, McElroy and Sanders (2005).[20] Specifically, Chapter 2 aggregates teen mother's outcomes into two separate life stages, "young adulthood" between 18 to 24 years old (short run) and "adulthood" between 25 to 32 years old (long run). I find similar results to those of Hotz, McElroy and Sanders (2005) on various outcomes after controlling for different risk factors associated with miscarriage, the instrument used in estimation. Importantly, I find that teenage childbearing has a small but statistically significant impact on annual earnings of teen mothers in the short-run but no impact in the long-run. These findings support the hypothesis that the short-run disadvantage of teenage childbearing in terms of labor market outcomes diminish in the long-run. Teen mother might even gain advantages in the long run by gaining greater experiences and skills early in their work histories.

Building on the work of Hotz, McElroy and Sanders (2005) and Chapter 2, this chapter extends the analysis of teenage childbearing in two ways. First, this chapter utilizes the complete Wave 1, Wave 3 and Wave 4 of the restricted National Longitudinal Study of Adolescent Health (Add Health) data-set[21] to examine the short and long run impact of teenage childbearing on labor market outcomes. The Add Health data is very rich compared to the NLSY79 as it includes surveys of school, personnel, parents, and the respondents as well as providing contextual data such as

[20] My second chapter of this dissertation is based on Nguyen (2013).

[21] This research uses data from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Persons interested in obtaining Data Files from Add Health should contact Add Health, The University of North Carolina at Chapel Hill, Carolina Population Center, 206 W. Franklin Street, Chapel Hill, NC 27516-2524 (addhealth_contracts@unc.edu). No direct support was received from grant P01-HD31921 for this analysis.

relationships, families, social networks, neighborhoods, schools, and states. In addition, computer-assisted personal interview (CAPI) technology allows the respondents to answer sensitive questions by computer rather than by an open verbal conversation. Therefore, using the Add Health data help reduce the bias in self-reports of pregnancy outcomes and will likely yield more accurate information about miscarriage, the instrumental variable for teenage childbearing I used in the previous chapter.

The two earlier studies using this data-set, Fletcher and Wolfe (2009) and Lee (2010) only used the two waves (Wave 1 and Wave 3) which only allow them to examine only the outcomes in what I define as young adulthood. While Fletcher and Wolfe (2009) used instrumental variable methods Lee (2010) used different technique, i.e. propensity matching method. Second, I extend Chapter 2 as well as the existing literature by looking for reasons why the negative effects of teenage childbearing found in the short run diminish in the long run.

My results are consistent with those in the previous chapter in terms of labor market outcomes. The causal effects of having a birth as a teen on annual earnings are significantly negative within 6 years after giving birth (short run). However, the effects are very small or not significant when women are ages from 25 to 32 years old (long run). In addition, I examine the information value of a new variable, namely the age at which a woman starts working full-time, in diagnostically testing the hypothesis that the lack of a long-run impact of teenage childbearing on annual earnings is due to their earlier participation in the labor market. First, my findings show that the age at which a teen mother starts full-time work has a negative relationship with her annual earnings in adulthood stage. The test statistic for this coefficient of this variable is significantly different from 0 suggesting that this variable should be included in the model and has some informational value in the explanation of future annual income of teen mothers. Secondly, after controlling for this new variable, I find that the effects of teenage childbearing on annual earnings does not reappear. Moreover, the coefficient of teenage childbearing becomes even more positive. These findings point out that the inclusion of this new variable makes being a teen mother earning even more in the labor market. Therefore, entering the labor market might not be the only reason that explains why the impact of teenage childbearing disappears in the long run.

The rest of the chapter is organized as follows: Section 3.2 discusses data and summarizes empirical approach. Section 3.3 presents the empirical findings and Section 3.4 concludes.

## 3.2 Data

### 3.2.1 Add Health Data and Studies on Teenage Childbearing

This chapter is the first that employs all waves of the restricted-use Add Health data to research the causal effects of teenage childbearing. In its mandate, the Add Health data was designed to study adolescent health and is considered to be the largest and most comprehensive longitudinal survey of adolescents that has ever been made.[22] The data was initiated by surveying a nationally representative sample of students in grades 7 to 12 in their schools. Those students were then surveyed again in 1994-95, 1996, 2001-2002, and 2007-2008 at home. While Waves I and II of this data look at the factors associated with adolescents' health and risk behaviors, Wave III surveyed respondents when they were between 18 and 26 years old (young adulthood period) about the linkage between their experiences and behaviors and the "short-run" outcomes in their transition to adulthood.[23] The recent Wave IV conducted between January 2007 and February 2008 is similar to Wave III except

[22] See http://www.cpc.unc.edu/projects/addhealth/about

[23] Those factors in Waves I and II are such as personal traits, families, friendships, romantic relationships, peer groups, schools, neighborhoods, and communities

the questionnaires were completed when the teenagers were between the ages of 25 and 32 (adulthood or the "long run"). Wave IV provides the opportunity to examine the longer run trajectories of the relationship between a respondent's behaviors as a teen and her outcomes when she assumes mature adult roles and responsibilities. With respect to research topic of this chapter, the behavior of interest is teenage childbearing. Wave IV allows me to trace the complete trajectories of causal effects of teenage childbearing on the outcomes of women both in the short and long run.[24]

The only studies to date, Fletcher and Wolfe (2009) and Lee (2010) use the Add Health data to examine the consequences of teenage childbearing. However, they only look at the short-run outcomes of teen mothers utilizing the data up to Wave III. As pointed out by Hotz et al. (2005) and Fletcher and Wolfe (2009), the Add Health data has an advantage over other data-sets such as NLYS1979 because it helps reduce the bias in the self-reports of pregnancies. In particular, the Add Health survey used a computer-assisted personal interview (CAPI) technology allowing the respondents to answer sensitive questions by a laptop rather than by an open verbal conversation. The self-reported pregnancy outcomes in Add Health are consistent with the official Vital Statistics while other data-sets are not. For an example of the discrepancy between Add Health and NLYS1979 data regarding the self-reports, 25% of first pregnancies were ended in abortion and 16% were ended in miscarriage for Add Health, while those statistics were under reported at 18% and 7%, respectively for NLYS1979. The advantages of the restricted Add Health data over other data-sets

[24] Note that attrition is an important issue that needs to be addressed in the use of Add Health data as it could potentially lead to the attrition bias in the estimation in this chapter. This chapter does not account for this attrition issue as I build on the earlier findings of Chantala et al. (2004) and Brownstein et al. (2010). In particular, to address the attrition issue of the Add Health data previous studies have investigated the potential magnitude of nonresponse bias at Wave 3 (see Chantala et al. 2004) and Wave 4 (see Brownstein et al. 2010). The response rates for Wave 3 Survey are 75.6% and 80% for Wave 4. Regarding Wave 3, Chantala et al. (2004) show that the bias is very small (less than 1%) for the demographic characteristics, school experiences, health reports, attitudes and physical activities. Brownstein et al. (2010) show that bias and relative bias are small in magnitude for nearly all measures in Wave 4. Therefore, they conclude that Wave 3 and Wave 4 adequately represent the same population as Wave 1 does.

are therefore evident.

Both Lee (2010) and Fletcher and Wolfe (2009) found evidence of negative consequences of teenage childbearing on mother's socioeconomic outcomes including labor market outcomes and educational attainment in their young adulthood. While Lee (2010) used the propensity matching score method to stress the importance of correct estimation of the consequences, Fletcher and Wolfe (2009) used instrumental variable (IV) method with miscarriage as an instrument. While propensity score uses observable measures to construct a weight to match the treatment and control group, IV method relies on an instrumental variable. The advantage of IV method is that it is simple to use and a good instrument variable would yield precise estimates. In this literature of adolescents, it has been well-established that miscarriage is a good instrumental variable for teen birth. I follow this approach in this paper.[25]

Lee (2010) suggested that more research should be done to address the long-term effects of teenage childbearing for this young cohort. Building on these two papers and Chapter 2, this chapter looks at the effects transitioning to the adulthood period. In particular, I provide more complete analysis by incorporating the Wave IV data. Similar to Fletcher and Wolfe (2009) and Chaper 2, I use miscarriage as an instrument in the estimation of the effects of giving birth as a teen. I also follow Chapter 2 and consider the trajectory of these effects between the short period and the long run-period to test whether effects are persistence.

### 3.2.2 Variable Description

Many studies have found that teenage childbearing has a negative effect on labor market outcomes of teen mother several years after they gave birth. However, whether these effects persist in the long-run, in particular the age range between 25 and 32 years old remains uncertain. I use Wave I, Wave III, and Wave IV of the Add Health

[25] See chapter 2 of this dissertation for more discussion on miscarriage as an instrument.

data to address this question. In the following I describe the variables in my empirical examination of the data.

Firstly, I look at a sample of teenagers whose first pregnancy occurred prior to their $18^{th}$ birthday. This sample is referred to in this chapter as the teen pregnancy sample. In this sample, there are three sub-samples that are constructed on the basis of the outcome of that pregnancy: birth, abortion, miscarriage and stillbirth. Stillbirth is merged into miscarriage outcome because it is a rare event. After combining information on the date of the first pregnancy, the birthday of respondents, and the first pregnancy's outcomes I end up with a sample of 631 teen pregnancies including 372 women whose pregnancy ended at a live birth, 101 women whose pregnancy ended at miscarriage, and 158 women whose pregnancy ended at abortion.[26]

I use the annual earnings of women as measure of their labor market outcomes. The information on earnings taken from Wave III and Wave IV of data set. Wave III pertains to the young adulthood stage when respondents are between 18 and 24 years old, and Wave IV captures the adulthood stage when respondents were between 25 and 32 years old. In both waves, the earnings are measured by the responses to the following questions: "How much income did you receive from earnings - that is, wages or salaries, including tips, bonuses, and overtime pay, and income from self-employment ?" and "What is the best guess of the income you received from earnings?" if the respondent's answer is "Do not know" in the previous question.[27] Note that for each wave, respondents were only interviewed once so that for each stage woman only were observed once.[28]

[26] Wave I reported about 5,000 pregnancies when the respondents were on average 22 years old. By looking at the first pregnancies the sample size reduces to around 3,600 pregnancies. For teenage childbearing examination, focusing on pregnancies that ended before age 18 years and 9 months the final sample includes only 631 observations.

[27] As the number of respondents who reported zero annual earnings is small I include them in my analysis to increase the sample size.

[28] This approach is different from previous chapter which used the NLSY79. In previous chapter, respondents were interviewed on yearly basis so that women were observed more than once for each stage.

For explanatory variables, I use variables that are considered to be correlated with how teenage childbearing affects the teen mother's labor market outcomes. The first set is family background variables such as family income, family structure (two biological parents, biological mother only, biological father only),[29] race (hispanic, white, asian, black)[30], whether an individual was born in U.S., and individual's age at first interview.[31] The second set consists of variables that are linked with an individual's pregnancy history such as whether the individual used birth control before pregnancy, whether the individual wanted to have baby before getting pregnant. All variables were obtained directly from questions in the In-home Parent Survey or Wave I In-home Respondent Survey.

The last set of explanatory variables include an individual's characteristics such as individual's age at each stage[32], heath status at birth (excellent, very good, good, fair, and poor)[33], education (years of schooling), and age-standardized test scores (Add Health Picture Vocabulary Test - AHPVT).[34] The information on test scores is available for each wave of data set. The information on years of schooling in young adulthood and adulthood is taken from the question "What is the highest grade or year of regular school you have completed ?" (H2ED1 in Wave III and H4ED1 in

[29] In Section C of Parent Questionaires, the respondents are asked whether either biological mother or biological father lives in the household.

[30] The information about race is reported in Parent Questionaire, Part A.

[31] The information about family background is available in Parent Questionaires. Section C reports whether either biological mother or biological father lives in the household. Section A reports race information and other characteristics.

[32] I include both individual's age at the first interview (Wave 1) and individual's age at each stage in my analysis although these variables might be correlated with each other. The individual's age at the first interview is considered as one component of individual's background. Meanwhile, the individual's age at each stage might contritute to individual's current outcomes. The inclusion of both variables might help to separate the effect of age at different time of period on the individual's outcomes.

[33] The information is taken from the question about specific child's health and health behaviors in PC18-Parent Questionaire Survey "How would you rate {NAME}'s general physcial health?"

[34] Add Health Picture Vocabulary Test (AHPVT) is a computerized, abridged version of the Peabody Picture Vocabulary Test. In this test, the interviewer reads a word aloud and the respondent selects the illustration that best fits its meaning. Each word has four simple, black-and-white illustrations arranged in a multiple-choice format. For example, the word "furry" has illustrations of a parrot, dolphin, frog, and cat from which to choose. There are 87 items on the AHPVT, and raw scores have been standardized by age. The score range is between 24 and 130.

Wave IV). Table 3.1 reports the means and standard deviations of the variables used in the analysis.

## 3.3 Empirical Approaches and Findings

### 3.3.1 Empirical Approaches

My empirical strategy is to examine the effect of giving birth as a teen on women's annual earnings in young adulthood and adulthood. The empirical model can be summarized as follows:

Let $Y_i$ be the annual earnings of woman $i$ and $B_i$ be a dummy variable that indicates whether the pregnancy of this woman ends, with her giving a live birth ($B_i$ = 1). I begin with linear models in order to estimate the the causal effect of $B_i$ on $Y_i$,[35] i.e.

$$Y_i = \alpha + \beta B_i + \delta F_i + \theta X_i + \varepsilon_i$$

where $\alpha$, $\beta$, $\delta$, $\theta$ are model parameters, $F_i$ is a vector of family background's covariates , $X_i$ is a vector of teen mother's characteristics, and $\varepsilon_i$ is an error term. In this equation, $\beta$ measures the marginal impact of teenage childbearing on a mother's annual earnings.

As pointed out in the previous chapter, $E(Y_i^0|X, B_i = 1)$ is not always observed. One solution is to use an instrumental variable (IV) for $B_i$. However, the choice and validity of any instrumental variable are crucial for an unbiased estimation.

I follow the methodology in previous chapter by using miscarriage ($M_i$) to be an instrument for $B_i$ where $M_i = 1$ if a woman experienced a miscarriage and $M_i = 0$ otherwise. The instrumental variable method will produce an IV estimate for $\beta$.[36] As

[35] For example, Fletcher, Jason and Barbara (2009) use this frame work for estimation.
[36] See Hotz, McElroy, and Sander (2005)

a robustness check, I also implement a nonlinear IV model in which miscarriage, the desires to have a baby, and use of birth control are instrumental variables.

I implement three models. Model 1 is the ordinary least square (OLS) model in which I only control for teen birth by using a dummy variable specifying whether a woman gives birth as a teen and including other covariates associated with backgrounds of teen mother such as family structure (two-biological parents, biological mother only, biological father only, step-family), parent's education (less than high school, high school degree and/or GED, some colleges, college degree)[37], race, cognitive ability, and family income. I also control for the respondent's characteristics such as whether she was born in U.S., her PVT score, her health status, and a list of variables related to the respondent's fertility history including whether she used birth control before the 1st pregnancy, whether she wanted a child before pregnancy, and whether abortion is funded publicly.

Model 2 is the two-stage least square model using miscarriage as an instrumental variable for giving birth as a teenager. And finally, in Model 3, I use the covariates in Model 2 and use three variables as instruments: miscarriage, whether the respondent desired to have baby before pregnancy, and whether she used birth control before pregnancy.

Following the empirical strategy in Chapter 2, another important step is to define young adulthood and adulthood stages so as I could distinguish the short and long run effects of teenage childbearing. Specifically, I analyze the effects of teen births on teen mothers in two separate stages: Stage 1 from 18 to 24 years old and Stage 2 from 25 to 32 years old. In stage 1, the young adulthood stage, a woman could complete her education and start participating in labor market. In stage 2, the adulthood stage, a woman has more responsibilities from her job, marriage, and parenthood.[38] I

[37] The parent's education is defined as the maximum of parents' education.

[38] Another reason for this classification is that it is appropriate for Wave III and Wave IV in the Add Health data.

implement all three models for each stage and make comparisons both by model and by stage.

Note that in the examination of the persistence of the consequences of teenage childbearing, one hypothesis discussed in Chapter 2 of this dissertation is that giving birth as a teenager exposed a teen mother to more work experiences leading to greater compensation for her in the long-run in the labor market. This could explain why teenage childbearing does not cause negative affect mothers' labor market in the long run. Using the Add Health Data, I could provide a diagnostic test of this hypothesis. In particular in my empirical strategy, I add one more variable, namely the age at which a woman has the first full-time job as a proxy measure for work experience. I incorporate this variable in all models to see if this variable has information value for checking this hypothesis.

### 3.3.2 Empirical Findings

As my interest in this paper is to understand better the trajectory of the effect of teenage childbearing I focus my discussion on the labor market outcomes. Another consideration could be the educational attainment and marital status. However, previous studies have shown that teenage childbearing has negatively effects in the short run but no impact in the long run for these outcomes. I report the effects on these outcomes in Table 3.5 of this paper. Most of the findings are similar to the finding in Chapter 2. I therefore focus my empirical discussion on labor market outcomes only.

#### 3.3.2.1 The Effect of Teenage Childbearing on Annual Earnings in the Short-run

The short-run effects of teenage childbearing on annual earnings are reported in Table 3.2. As mentioned in the previous chapter, the short-run or young adulthood stage is defined as the ages between 18 and 24 years old. Column 1 presents the results from OLS estimation, Column 2 and 3 show the results from IV estimations. It is evidence that having a live birth as a teenager has negative effect on annual earnings in the short run. In particular, teenage childbearing reduces annual earnings from $2400 up to $2,600 depending on the model used. The OLS results show that having a baby as a teen will lower annual income by $2,600 and the IV models produce negative effects of $2,465 to $2,555. These results are consistent with results in other Add Health Data papers that giving birth as a teen became a disadvantage for women in labor market when they are 18 and 24 years old (Fletcher and Wolfe 2009). They also confirm the well-established findings of the earlier studies that teenage childbearing has negative effects not only on educational attainment but also on labor market outcomes. Similar to the results in Chapter 2, the other determinants that significantly affect annual earnings of teen mothers statistically at this stage are years of schooling, family income, and age of teen mother at the time of analysis. All estimated coefficients of these variables are positive and statistically significant. In addition, a teen mother who was born in United States has on average higher annual earnings than one born outside the United States.

Note that the estimated effects from the IV estimation are smaller in magnitude than that estimated in OLS model. The OLS model overestimates the effect of teenage childbearing on annual earnings. This finding is in line with my previous results in Chapter 2 of this dissertation even though I used a different data set - NLSY79. Later in this chapter I will implement the same models but add one more variable - the age at which the respondent has a first full-time job to see whether this variable correlates with her annual earnings after controlling for both her cognitive skills and educational attainment.

#### 3.3.2.2 The Effect of Teenage Childbearing on Annual Earnings in the Long-run

Table 3.3 reports the estimation results for the long run. At this stage the subjects were between 25 and 32 years old and therefore somewhat removed from their teenage years. Again, I find in the OLS model that having a live birth as a teen has a negative effect on annual earnings in long-run. In particular, giving a birth as a teen lower annual earnings by \$1,856 annually. This amount is smaller than the effect estimated for the short run by about \$800. The estimated coefficients of teenage childbearing from the IV models (Column 2 and Column 3) are both positive but statistically insignificant. This finding implies that teenage childbearing might has no impact in long-run. Compared to the results in the short-run the sign of estimated coefficients of teenage childbearing change from negative to positive sign although they are not statistically significant. This is an interesting result. The disadvantage of having a birth as a teen seems to disappear by the time a woman is between 25 to 32 years old. The variables with statistically significant effects on earnings are years of schooling, PVT test score, and family structure (two biological parents, only biological mother).

In sum, by comparing the two stages (short-run and long-run), I find the effect of teenage childbearing on annual earnings of teen mother is negative in both short-run and long-run using OLS estimation. However, OLS estimation are biased and overestimates these effects. In the IV models, I find no long run impacts on the annual earnings of teen mothers. On average, giving birth as a teen does not lead to less annual earnings. Though not significant, the positive point estimate between teenage childbearing and the labor market outcomes is intriguing.

#### 3.3.2.3 The Effect of Teenage Childbearing on Annual Earnings After Controlling for the Age at which Teen Mother Has the First Full-Time Job

As found in previous sections, OLS models indicate that negative effect of teenage childbearing on teen mother's annual earnings in the short-run is less than that in long-run. In the IV models, the negative effect in short-run does not persist in to the long-run. It is important to try to explain the findings. One proposed argument in the literature is that mothers might overcome their disadvantage of having a child early because their early participation in the labor market might help them gain more experiences and skills compared to other women at their age. I diagnostically check this hypothesis by incorporating a new variable that has never been used in this literature, namely the age at which a teen mother has her first full-time job. As teen mothers are more likely to enter the labor market sooner, they could learn more skills and have more experiences which become positive factors in the long run. I choose this variable as it could be a proxy measure for a woman's work experience and skills. I then could examine the informational value of this new variable in checking whether teen mothers got compensated for the short-run disadvantages by entering the labor market early (lower level of education).

The first two columns of Table 3.4 report the results for the short-run when teen mothers were in young adulthood stage. Estimates in Column 1 are obtained using OLS estimation and the estimates in column 2 are estimated using IV estimation. Note that, in the short-run after adding the new variable - the age at which the subject has the first full-time job, the estimated coefficients of teenage childbearing become statistically insignificant. This result contrasts the findings in Table 3.2. In addition, the estimated coefficient of the age at first full-time job also is statistically different from zero. These findings imply that an earlier participation of a teen mother in the labor market does affect her annual earnings in the short run.

The last two columns in Table 3.4 present the similar analysis but show the results in the long-run. Both OLS and IV models are implemented. As seen in this table, the estimated effects of teenage childbearing are statistically insignificant and positive in all models. Compared to the findings in Table 3.3, these results suggests that in the OLS model the negative effects of teenage childbearing on annual earnings disappear when teen mother transitions to the long-run and there is still no negative impact of teenage childbearing in the IV model. Although years of schooling at this stage are still positively correlated with annual earnings, individual age in long-run no longer affects this outcome. Interestingly, the new covariate - the age at which a teen mother had the first full-time job becomes negatively correlated with annual earnings in both the short- and long-run stage. However, the negative effect of teenage childbearing does not reappear after controlling for the age at which a woman had her first full-time job in both the short- and long-run. The test statistic for the coefficient of this age variable is significantly different from 0 suggesting that this variable should be included in the model and has some informational value in the explanation of future annual income of teen mothers. Moreover, the coefficient of teenage childbearing becomes even more positive. These findings point out that the inclusion of this new variable makes being a teen mother earning even more in the labor market. Therefore, entering the labor market might not be the only reason that explains why the impact of teenage childbearing disappears in the long run. This finding implies a rejection of the hypothesis that the negative impact of teenage childbearing might be improved in the long-run by more working experience and/or participating in labor market early.

## 3.4 Conclusion

In this chapter, I extend the analysis of teenage childbearing by employing the Wave I, Wave III, and Wave IV of the Add Health data to examine the short- and long-run impact of teenage childbearing on the labor market outcomes. My results show that there are significant short-run consequences of teenage childbearing on labor market earning. However, these effects disappear in the long run. Though earlier studies have pointed out these mixed effects, there is no paper which actually formally aggregates the short- and long-run analysis using the complete Add Health Data. In addition, I also contribute to the literature by proposing a way to explain the diminishing effects of teenage childbearing in the long run. I test the hypothesis whether teen mother overcomes their disadvantage in the long-run due to her earlier participating in labor market. I do it by incorporating a new variable in the analysis - the age at which a woman had the first full-time job. The results show that this variable is negatively correlated with annual earnings in the long run. In addition, no impact of teenage childbearing is found after controlling for this variable. This finding provides information on the hypothesis is that although an earlier participation in labor market positively affects annual earnings in the long run it might not be a reason for teen mothers to overcome their disadvantage in the long run. A formal test on this hypothesis is for future research.

Note that in previous chapter, the effect of teenage childbearing on teen mothers' outcomes is analyzed by using the NLSY79 in which the teenage births were occurred between 1971 and 1982. In the context that the demographic characteristics of teenage mothers and the economic environment have changed so quickly and dramatically the examination of longer run effect of teenage childbearing requires using newer data on women who were teenage mothers. Therefore, using the Add Health data - newest cohort of teen mothers has important implications in investigating the changes of teenage childbearing across time periods.

**Table 3.1: Summary Statistics of Teen Pregnancy**

| Variables | Wave | Mean | SD | Min. | Max. |
|---|---|---|---|---|---|
| Two Biological Parents | 1 | 0.334 | 0.472 | 0 | 1 |
| Biological Mother only | 1 | 0.379 | 0.485 | 0 | 1 |
| Biological Father only | 1 | 0.046 | 0.210 | 0 | 1 |
| Other Family Structure | 1 | 0.263 | 0.441 | 0 | 1 |
| Parent with Less than High School | 1 | 0.057 | 0.232 | 0 | 1 |
| Parent with High School Diploma or GED | 1 | 0.162 | 0.368 | 0 | 1 |
| Parent with Some College | 1 | 0.197 | 0.398 | 0 | 1 |
| Parent with College Degree | 1 | 0.136 | 0.343 | 0 | 1 |
| Hispanic | 1 | 0.190 | 0.393 | 0 | 1 |
| Black | 1 | 0.311 | 0.463 | 0 | 1 |
| Asian | 1 | 0.044 | 0.206 | 0 | 1 |
| White | 1 | 0.472 | 0.500 | 0 | 1 |
| Family Income in 1994 | 1 | $32,433 | 48,445 | 0 | 426000 |
| Born in US | 1 | 0.796 | 0.404 | 0 | 1 |
| Excellent Health | 1 | 0.182 | 0.386 | 0 | 1 |
| Very Good Health | 1 | 0.361 | 0.481 | 0 | 1 |
| Good Health | 1 | 0.372 | 0.484 | 0 | 1 |
| Fair Health | 1 | 0.081 | 0.273 | 0 | 1 |
| Poor Health | 1 | 0.003 | 0.056 | 0 | 1 |
| Age at Wave 1 | 1 | 15.385 | 1.755 | 11 | 20 |
| PVT Score in Wave 1 | 1 | 96.970 | 12.893 | 56 | 130 |
| Age at the First Full-time Job | 1 | 19.041 | 2.519 | 10 | 31 |
| Used birth control before pregnancy | 1 | 0.336 | 0.473 | 0 | 1 |
| Wanted Child before Pregnancy | 1 | 0.214 | 0.410 | 0 | 1 |
| Public Abortion Funding | 1 | 0.330 | 0.470 | 0 | 1 |
| First Contraception by Age 15 | 1 | 0.463 | 0.499 | 0 | 1 |
| Age at Stage 1 | 3 | 21.803 | 1.633 | 18 | 24 |
| Age at Stage 2 | 4 | 28.266 | 1.636 | 25 | 32 |
| Years of Schooling at Stage 1 | 3 | 12.36 | 1.89 | 7 | 20 |
| Years of Schooling at Stage 2 | 4 | 15.21 | 3.25 | 9 | 20 |
| Annual Earnings at Stage 1 in 2001 | 3 | $12,309 | 16,290 | 0 | 300000 |
| Annual Earnings at Stage 2 in 2007 | 4 | $30,325 | 21,317 | 0 | 150000 |
| Number of Observations | | 631 | | | |

**Table 3.2: The Effect of Teenage Childbearing on Annual Earnings in the Short-run**

| Variables | OLS | 2SLS(1) | 2SLS(2) |
|---|---|---|---|
| Live Birth | -2632.5*** | -2465.2** | -2555.4** |
| | (-5.87) | (-2.03) | (-2.09) |
| Two Biological Parents | -1614.5 | -1889.7 | -1846.7 |
| | (-0.26) | (-0.38) | (-0.37) |
| Biological Mother only | -635.3 | -895.1 | -817.1 |
| | (-0.11) | (-0.23) | (-0.21) |
| Biological Father only | -2259.5 | -2376.1 | -2313.1 |
| | (-0.50) | (-0.77) | (-0.75) |
| Other Family Structure | -98.59 | -322.1 | -289.8 |
| | (-0.02) | (-0.08) | (-0.07) |
| Parent with High School Diploma or GED | -1260.9 | 2398.7 | 2536.4 |
| | (-0.40) | (1.46) | (1.55) |
| Parent with Some College | -5311.7* | -1832.6 | -1875.9 |
| | (-1.67) | (-1.21) | (-1.23) |
| Parent with College Degree | -2854.5 | 676.0 | 552.1 |
| | (-0.85) | (0.43) | (0.35) |
| Family Income | 130.1*** | 127.3 | 127.6 |
| | (5.39) | (1.34) | (1.53) |
| Hispanic | -2631.0 | -2744.3 | -2671.0 |
| | (-0.51) | (-0.98) | (-0.94) |
| Black | -7109.8 | -7215.5* | -7281.4* |
| | (-1.36) | (-1.82) | (-1.80) |
| Asian | -1979.2 | -2337.6 | -2356.8 |
| | (-0.38) | (-0.54) | (-0.53) |
| White | -6657.6 | -6816.2 | -6920.0 |
| | (-1.23) | (-1.53) | (-1.53) |
| Born in US | 2561.6 | 2483.2* | 2413.9* |
| | (1.58) | (2.06) | (2.01) |
| Excellent Health | -2626.7 | -2606.3 | -3339.4 |
| | (-0.23) | (-0.57) | (-0.71) |
| Very Good Health | -2637.8 | -2615.2 | -3494.0 |
| | (-0.23) | (-0.61) | (-0.79) |
| Good Health | -3446.6 | -3515.0 | -4272.6 |
| | (-0.30) | (-0.89) | (-1.03) |
| Fair Health | -3961.5 | -4091.3 | -4774.8 |
| | (-0.34) | (-0.99) | (-1.11) |
| Age at Wave 1 | -781.5 | -756.2 | -743.7 |
| | (-0.98) | (-1.06) | (-1.05) |
| PVT Score at Stage 1 | -15.15 | -23.15 | -17.79 |
| | (-0.28) | (-0.26) | (-0.20) |
| Years of Schooling | 5366.1** | 5386.3*** | 5326.9*** |
| | (2.04) | (2.64) | (2.61) |
| Age at Stage 1 | 1935.1** | 1941.0*** | 1940.6*** |
| | (2.24) | (3.14) | (3.14) |
| Used birth control before Pregnancy | -712.2 | -537.5 | |
| | (-0.52) | (-0.42) | |
| Wanted Child before Pregnancy | 2205.6 | 1965.0 | |
| | (1.36) | (1.21) | |
| Public Abortion Funding | -102.3 | -133.6 | -92.65 |
| | (-0.07) | (-0.13) | (-0.09) |
| First Contraception by Age 15 | 609.1 | 698.9 | 769.5 |
| | (0.47) | (0.48) | (0.52) |
| Constant | -13837.0 | -16012.7 | -16012.4 |
| | (-0.78) | (-1.15) | (-1.14) |
| Observations | 631 | 631 | 631 |

t statistics in parentheses
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$
OLS: Ordinary Least Square; 2SLS (1): 2SLS with miscarriage as IV
2SLS (2): 2SLS with miscarriage, desires to have baby, and use of birth control as IVs

**Table 3.3: The Effect of Teenage Childbearing on Annual Earnings in the Long-run**

| Variables | OLS | 2SLS(1) | 2SLS(2) |
|---|---|---|---|
| Live Birth | -1856.3** | 2944.0 | 2707.0 |
| | (-2.08) | (1.20) | (0.89) |
| Two Biological Parents | 14028.9* | 14528.2** | 14529.8** |
| | (1.69) | (2.21) | (2.30) |
| Biological Mother only | 11358.6 | 11634.9* | 11641.6** |
| | (1.45) | (1.93) | (2.03) |
| Biological Father only | -1762.7 | -1665.0 | -1906.0 |
| | (-0.30) | (-0.38) | (-0.46) |
| Other Family Structure | 9450.0 | 9135.6 | 9139.5 |
| | (1.15) | (1.43) | (1.49) |
| Parent with High School Diploma or GED | 2687.4 | -1860.4 | -1895.0 |
| | (0.65) | (-0.70) | (-0.71) |
| Parent with Some College | 519.9 | -3052.5 | -3243.5 |
| | (0.13) | (-1.22) | (-1.30) |
| Parent with College Degree | 4451.7 | 1235.8 | 1062.7 |
| | (1.01) | (0.42) | (0.36) |
| Hispanic | 1296.2 | 2474.3 | 3015.4 |
| | (0.19) | (0.43) | (0.53) |
| Black | -1818.0 | -466.3 | -676.3 |
| | (-0.26) | (-0.08) | (-0.11) |
| Asian | 1527.0 | 4272.9 | 3732.6 |
| | (0.22) | (0.66) | (0.57) |
| White | -5104.2 | -3606.9 | -3791.2 |
| | (-0.72) | (-0.58) | (-0.61) |
| Family Income | 57.07* | 67.65* | 64.79* |
| | (1.76) | (1.76) | (1.72) |
| Born in US | 1456.3 | 1428.2 | 1376.5 |
| | (0.69) | (0.67) | (0.64) |
| Excellent Health | -183.3 | -1589.3 | -0.964 |
| | (-0.01) | (-0.08) | (-0.00) |
| Very Good Health | -560.6 | -2064.3 | -299.8 |
| | (-0.04) | (-0.11) | (-0.02) |
| Good Health | -3701.4 | -5296.9 | -3749.3 |
| | (-0.25) | (-0.28) | (-0.21) |
| Fair Health | -733.0 | -2136.7 | -652.6 |
| | (-0.05) | (-0.11) | (-0.04) |
| Age at Wave 1 | 561.5 | 668.4 | 780.6 |
| | (0.62) | (0.77) | (0.89) |
| Years of Schooling | 5898.7** | 6005.8** | 5875.5** |
| | (2.12) | (2.04) | (2.01) |
| PVT Score at Stage 2 | 241.9*** | 294.1*** | 288.4*** |
| | (3.27) | (3.86) | (3.83) |
| Age at Stage 2 | 36.63 | -237.9 | -252.3 |
| | (0.04) | (-0.25) | (-0.26) |
| Used Birth Control before Pregnancy | -1701.3 | -2000.2 | |
| | (-0.95) | (-1.22) | |
| Wanted Child before Pregnancy | -787.2 | -2392.0 | |
| | (-0.37) | (-1.05) | |
| Public Abortion Funding | 2961.8 | 3632.0* | 3349.6* |
| | (1.45) | (1.92) | (1.77) |
| Constant | -18733.1 | -19142.4 | -22682.6 |
| | (-0.71) | (-0.68) | (-0.83) |
| Observations | 631 | 631 | 631 |

t statistics in parentheses
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$
OLS: Ordinary Least Square; 2SLS (1): 2SLS with miscarriage as IV
2SLS (2): 2SLS with miscarriage, desires to have baby, and use of birth control as IVs

**Table 3.4: The Effect on Annual Earnings with Controlling for the Age at First Full-time Job**

| Variables | OLS-SR | 2SLS-SR | OLS-LR | 2SLS-LR |
|---|---|---|---|---|
| Live Birth | -3785.9 | -3166.6 | 4039.7 | 2999.7 |
| | (-0.83) | (-0.69) | (1.21) | (0.98) |
| Two Biological Parents | -2172.9 | -2116.0 | 15168.2** | 15147.9** |
| | (-0.42) | (-0.41) | (2.29) | (2.31) |
| Biological Mother only | -1202.2 | -1113.3 | 12355.8** | 12447.2** |
| | (-0.29) | (-0.27) | (2.04) | (2.07) |
| Biological Father only | -2305.4 | -2245.0 | -1747.4 | -1852.0 |
| | (-0.72) | (-0.70) | (-0.40) | (-0.42) |
| Other Family Structure | -847.2 | -795.9 | 10258.7 | 10384.1 |
| | (-0.20) | (-0.19) | (1.58) | (1.60) |
| Parent with High School Diploma or GED | 2064.0 | 2218.6 | -1228.0 | -1198.8 |
| | (1.29) | (1.40) | (-0.47) | (-0.46) |
| Parent with Some College | -2006.8 | -2039.0 | -2750.4 | -2653.0 |
| | (-1.28) | (-1.30) | (-1.12) | (-1.08) |
| Parent with College Degree | 892.3 | 786.6 | 790.5 | 823.4 |
| | (0.56) | (0.49) | (0.27) | (0.28) |
| Family Income | 130.5 | 130.5 | 61.28 | 58.63 |
| | (1.34) | (1.35) | (1.52) | (1.47) |
| Hispanic | -2174.8 | -2117.2 | 1319.2 | 1431.2 |
| | (-0.80) | (-0.76) | (0.23) | (0.25) |
| Black | -6306.8* | -6397.8* | -2310.1 | -2403.5 |
| | (-1.71) | (-1.70) | (-0.40) | (-0.41) |
| Asian | -1576.8 | -1638.6 | 2769.0 | 2327.0 |
| | (-0.38) | (-0.39) | (0.43) | (0.36) |
| White | -6133.5 | -6266.2 | -5018.3 | -5118.2 |
| | (-1.46) | (-1.46) | (-0.82) | (-0.83) |
| Born in US | 2414.1** | 2350.0** | 1647.9 | 1714.0 |
| | (2.04) | (2.00) | (0.76) | (0.79) |
| Excellent Health | -2393.6 | -3155.8 | -2417.0 | -2187.5 |
| | (-0.53) | (-0.68) | (-0.13) | (-0.12) |
| Very Good Health | -2685.7 | -3573.6 | -2364.6 | -1924.3 |
| | (-0.64) | (-0.83) | (-0.13) | (-0.11) |
| Good Health | -3691.7 | -4471.4 | -5289.2 | -5138.8 |
| | (-0.96) | (-1.10) | (-0.28) | (-0.28) |
| Fair Health | -4246.0 | -4938.7 | -2191.1 | -1813.4 |
| | (-1.05) | (-1.17) | (-0.11) | (-0.10) |
| Age at Wave 1 | -794.1 | -783.2 | 856.3 | 897.9 |
| | (-1.11) | (-1.09) | (0.98) | (1.01) |
| Age at Stage 1 | 2000.3*** | 2002.1*** | -494.2 | -556.8 |
| | (3.19) | (3.18) | (-0.51) | (-0.58) |
| Used birth control before pregnancy | -631.7 | | -1800.8 | |
| | (-0.49) | | (-1.11) | |
| Wanted Child before Pregnancy | 1970.8 | | -2403.1 | |
| | (1.22) | | (-1.06) | |
| Years of Schooling | 5425.0*** | 5373.3*** | 273.2*** | 267.2*** |
| | (2.61) | (2.58) | (3.73) | (3.70) |
| Public Abortion Funding | 22.63 | 70.29 | 3360.3 | 3191.1 |
| | (0.02) | (0.07) | (1.76) | (1.68) |
| First Contraception by Age 15 | 524.2 | 592.9 | -2302.6 | -2375.6 |
| | (0.37) | (0.42) | (-1.37) | (-1.43) |
| Age at the First Full-time Job | -468.5* | -455.5* | -947.7** | -942.8** |
| | (-1.71) | (-1.68) | (-2.34) | (-2.32) |
| Constant | -9532.7 | -9733.9 | -29891.5 | -28746.2 |
| | (-0.57) | (-0.58) | (-1.01) | (-0.99) |
| Observations | 631 | 631 | 631 | 631 |

t statistics in parentheses
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$
OLS: Ordinary Least Square; 2SLS: 2SLS with miscarriage as IV
SR: Short-run; LR: Long-run

# Chapter 4

# Birth Spacing and Outcomes in Adolescence, Young Adulthood and Adulthood

## 4.1 Introduction

Over the last several decades, researchers have demonstrated the importance of family structure in human investment. In particular, empirical evidence on the strong relationship between family structure and children's outcomes have been established. As examples of research in this area, Astone and McLanahan (1991) find that children who live with single parents or step-parents during adolescence receive less school work support than children who live with both their biological parents. Hanuskek (1992) points out the trade-off between the number of children in a family and children's scholastic performance. Jenkins and Astington (1996) find that the presence of siblings can compensate for slower language development in developing false belief understanding. Black, Devereux, and Salvanes (2005) use a rich data set on the entire population of Norway and report that higher birth order has a significant and large negative effect on children's education. Thomson and McLanahan (2012) confirm lower human capital investment in children from non-traditional families such as those headed by single-mothers or containing step-parents. In general, it is commonly accepted that family structure components have strong impacts on children's outcomes.

As a component of family structure, the spacing of births - defined as the difference in age between siblings - has received considerable attention from researchers, notably demographers and sociologists. Demographers examine birth spacing as evidence of fertility behavior to predict the average time between generations and the rate of population growth. Sociologists are interested in birth spacing as it could affect household behaviors including emotional connections of parents and children as well as the interactions between family members. For instance, Stewart et. al. (1987) model the family adjustment process and the impact on the first born child of the birth of a subsequent child. They find that mothers dramatically decrease their interactions with the first born child over time but that father's frequency of interactions remains stable. McCall (1984) shows that the IQ performance of children who experience the birth of a younger sibling drop compared to single children and last-born children from families of same size. Similarly, Baydar, Greek, and Brooks-Gunn (1997) show that the birth of a sibling, especially in the case of short birth intervals, significantly affects the older child due to the diminishing interactions with mother.

Though there is well-established research on birth spacing in demography and sociology, a limited body of economic research has looked at the impact of birth spacing. In a closely related area of research involving fertility decisions, economists have recognized the important relationships between a child's socioeconomic outcomes and birth order, number of siblings, and family size. In the very few papers that address the birth spacing, it is modeled as an outcome of parents' decisions to maximize their family's utility in term of fertility, health, education, investment, consumption, and labor outcomes. Theoretical models often predict a pattern of optimal spacing in both static and dynamic framework (Becker (1960); Becker, Landes, and Michael (1977); Rosenzweig and Wolpin (1986); Barro and Becker (1989)). Empirically, birth spacing could be measured as a discrete or continuous outcome with which parents choose the time to have a birth (Newman 1983).

Outcomes of children early in life, in particular personality and health outcomes such as cognitive development, premature birth, and mortality status have been the central area of research on birth spacing. Still, a handful of studies estimate the impact of birth spacing on later-life outcomes. In these studies, educational outcomes are measured along one dimension such as the decision to drop out of high school, elementary school test scores, and post-secondary school attendance. Broman, Nichols and Kennedy (1975) examine the correlation between birth spacing and children's measured intelligence at pre-school age and argue that younger siblings born after long intervals have higher intelligence scores than those born after short intervals. Buckles and Munnich (2012) use the National Longitudinal Survey of Youth, cohort 1979 (NLSY79) and show that closely-spaced births of less than two years decrease test scores of the older child at ages 5 and 7 years old. They found, however, no causal impact of birth spacing on test scores for the younger siblings. Powell and Steelman (1993) used data from the High School and Beyond Survey to investigate the effect of birth spacing on high school attrition and post-secondary school attendance. They found that a low proportion of the closely spaced births within the family decreases the likelihood a respondent dropped out of high school and increases the odds she attended post-secondary school.[39] Galbraith (1982) looked at a sample of college students and showed no relationship between sibling spacing and intellectual development. Each of these studies, however, only investigates the relationship between birth spacing and the educational attainment of siblings at one particular stage of life cycle such as pre-school, elementary school, high school, or young adulthood. These studies are not able to test whether the effects they estimate persist as siblings age.

Building on the previous research, I examine the effects of birth spacing on siblings' outcomes using the sample of sibling-pairs from the restricted National Longitudinal Study of Adolescent Health (Add Health) data set.[40] I test the persistence of birth

[39] The proportion of close spacing is based on the respondent's reports of the number of siblings within two years of the respondent's age and the number of siblings more than two years older.

[40] This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from

spacing effects during the stage when siblings transition from adolescence to young adulthood. I contribute to the literature on birth spacing in the following ways. First, I investigate the relationship between the spacing of births and later-life outcomes of siblings along various dimensions including educational achievement (percentile rank on test scores, years of schooling, college attendance, and the completion of college degree), labor market outcomes (annual earnings) and engagement in risky behaviors (cigarette smoking). To the best of my knowledge, there are no previous papers that investigate the association between birth spacing and labor market outcomes or the relationship between birth spacing and engagement in risky behaviors. Second, the outcomes in this paper are measured at much later stages in the lives of siblings, considered as long-run effects, starting from adolescence (from 12 to 18 years old) to young adulthood (from 19 to 24 years old) and adulthood (from 25 to 32 years old). In addition to looking at how the birth spacing effects differ between younger and older siblings, I also examine how these effects vary with the sex composition of sibling pairs. Lastly, I use the latest Add Health data and implement family fixed-effect model to tackle the potential endogeneity of birth spacing that has been pointed out by earlier literature. See Buckles and Munnich (2012) for example.

My results show that in the long-run, birth spacing does not have an causal impact on siblings' educational attainment in terms of test scores and years of schooling in adolescence and young adulthood. This finding confirms the conclusion from the previous literature (Galbraith 1982) that shows no relationship between sibling spacing and intellectual development among college students. I do find that wider birth spacing has a positive effect on the possibility of enrolling in college in young adulthood and the likelihood of obtaining in a college degree. Therefore, the effect of birth spacing does persist when siblings transition to adulthood. I find positive effects of

23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth).

wider birth spacing on the likelihood of enrolling in college are larger for the younger siblings than for the older sibling. I find no effect of birth spacing on annual earnings after controlling for education. Also, the probability of smoking cigarette is unrelated to birth spacing. In sum, all results suggest that the effects of birth spacing on siblings' outcomes are different in the short- and long- run. My findings also support the claim of social scientists that the main effects of family structure usually occur prior to adulthood because social interactions between siblings and parental investment in terms of time and finances are more crucial to children's development during these life stages. One hypothesis is that as individuals age, the outcomes might be driven more by environmental factors other than family structure. Unfortunately, the Add Health data does not include information to allow me to identify specific environmental factors.

The rest of the paper is organized as follows: Section 4.2 discusses the literature on the spacing of births. Section 4.3 describes the data and outlines the models for estimation, and Section 4.4 presents results. Extensions, limitations and directions for future research are discussed in Section 4.5 and Section 4.6 concludes.

## 4.2 Literature Review

Researchers have shown that birth spacing is an important determinant of human capital investment in children. First, as shown by medical professionals, closely-spaced births are related to poor health outcomes for the younger child. The younger child is more likely to be born pre-term or have a low birth-weight, and tends to be small for gestational age (Rawlings, S., Rawlings, B., and Read 1995). The younger child also has a higher risk of being diagnosed as autistic (Cheslack-Postava, Liu and Bearman 2011), or developing childhood type 1 diabetes (Cardwell et al. 2012).

Secondly, sociologists have pointed out that a short birth interval might limit parental investment in children (Baydar, Hyle, and Brooks-Gunn 1997). The limited parental investment takes the form of lessened time and financial allocations, leading to negative impacts on children' future outcomes. Also, as family size increases, the amount of resources being allocated to each child falls which results in negative effects, especially as children age. Powell and Steelman (1995) find that the spacing between children's births has a strong positive relationship with the amount of expenditure spent on each child in previous year.[41] The parents' reduced parental time spent reading, helping with school work could directly affect the future outcomes of children (Price 2008).

However, closer spacing of births could be associated with the future socioeconomic outcomes of siblings in a positive way. For example, expenditures on toys, cloths, baby furniture, and activities can be shared amongst closely spaced siblings. Moreover, interactions and rivalries between closely spaced siblings might have positive effects on their future outcomes. The older sibling might benefit from teaching the younger sibling, and in turn, the younger sibling also might benefit from accepting directions from or observing good behaviors of the older sibling.

Finally, birth spacing and future outcomes for the children are possibly connected through changes in parents' lives such as relationship status (divorce, remarriage, or in new relationships) and employment status (layoff, maternal leave, or unemployment). Those changes in parents' status will affect the amount of time and financial resources allocated to the children. Therefore, there are clear channels that birth spacing could be linked with the future outcomes of siblings.

The estimation of the effect of birth spacing on siblings' outcomes is challenging as birth spacing is an endogenous variable. Research on household fertility behavior suggests birth spacing can be influenced by numerous complex factors within and across families. In addition to observable determinants of birth spacing (mother's

[41] Measurements of birth spacing in Powell and Steelman (1995) include the number of siblings and a proportion of siblings closely spaced in family.

age at first birth, previous birth intervals, parents' education, parents' wages and income, sex composition of births, and birth order), unobservable factors (child-specific endowments, family endowments, and parents' characteristics) are correlated with birth spacing and affect siblings outcomes. Rosenzweig and Wolpin (1988) indicate that parents might choose a shorter birth interval between their children if the older sibling was born healthy, i.e. good endowment. The sibling's endowment is therefore positively correlated with both a closely-spaced birth and his/her future outcomes. With regard to parents' characteristics, parents who are considered careful planners might choose longer birth intervals. Such characteristics might have negative relationship with a closely-spaced birth and a positive relationship with children's future outcomes. The unobserved determinants of birth spacing have been referred to as the unobserved heterogeneity "across and within" families that could lead to a bias in the estimation of birth spacing effects. Of important note, the sign of bias depends on the sign of correlation between the unobservable variables and birth spacing.

Methodologies have been proposed to tackle this endogeneity issue. Rosenzweig and Schultz (1983) examine the linkage between fertility (including birth spacing) and child mortality in the United States by applying a two-stage demand/production estimation procedure. The fitted values of the demand function for children which controls for local prices and household income are used to estimate the mortality production function. The logic behind this strategy is that parent's decisions regarding fertility and investments in infant health are jointly determined by the environment of the household. Bhalotra and Soest (2008) analyze the relationship between birth spacing, neonatal mortality, and fertility by jointly estimating three equations. They allow for unobserved heterogeneity in each equation and correlations between error terms.[42] They use maximum likelihood methods, and account for all correlations and

[42] Although Add Health data has information on determinants of birth spacing such as how long the respondent was breast fed, mother's age at first marriage, mother's age at first child, etc. I still could not use joint this information for joint estimation strategy because the small sample size drops dramatically if these covariates are controlled for.

for censoring in the birth spacing equations. They find that the relationship between neonatal mortality of younger children in a family is limitedly explained by birth intervals.

Several papers have used instrumental variables (IV) to deal with the endogeneity of birth spacing. The validity of instrumental variables is a major concern. Rosenzweig and Wolpin (1986) examine the effects of parental fertility choices on child health outcomes. The authors take into account the inter and intra-family endowment heterogeneity in the estimation by using the lagged characteristics of parents as instrumental variables for birth spacing. They argue that lagged parent characteristics such as education and income might be good instruments for fertility choices (including birth spacing) as the lagged characteristics are correlated with preference of birth spacing. However, the lagged characteristics of parents would be invalid instruments if they contain time-invariant characteristics of parents such as parents' tastes or parents's abilities. Such characteristics are persistent and might affect all siblings in the family.

Buckles and Munnich (2012) use miscarriage between two live births as instrumental variable for birth spacing. The logic of using miscarriage as an instrumental variable for birth spacing is supported by the evidence that miscarriage is strongly correlated with birth spacing. In particular, spacing is increased as if the mother experienced a miscarriage in-between live births. However, the exogeneity of miscarriage has been questioned. Lang and Ashcraft (2006) point out that miscarriage would be random only in the absence of abortions. Fletcher and Wolfe (2009) raise a concern about the correlation between miscarriage and engagement in risky behaviors (such as smoking, alcohol use, and drug use) prior to or during pregnancy. Buckles and Munnich (2012) implement an IV method that utilizes variation in spacing driven by miscarriages.[43] However, Buckles and Munich (2012) use the National Longitudi-

[43] Bucklers and Munnich (2012) do not include women who had experienced an abortion after the first live birth in the sample. They also do robustness check with controlling for risky behaviors

nal Youth of Survey (1979) in which miscarriage is thought to be under reported by the respondents raising the issue of measurement error. This may explain why the standard errors in Buckles and Munich (2012) are large and why only the small effect of birth spacing on test scores could not be detected.

My paper extends the literature on birth spacing in several ways. Firstly, the effects of birth spacing on sibling outcomes in different dimensions have been analyzed separately in previous studies. Previous authors study either educational outcomes or health outcomes. I investigate the linkage between the spacing of births along a number of dimensions including various measures of educational achievement (age-standardized test score, years of schooling, college attendance, and college degree), labor market outcomes (annual earnings), and the probability of engaging to risky behaviors (cigarette smoking). To the best of my knowledge, there are no previous studies that investigate the association between birth spacing and labor market outcomes or studies the engagement in risky behaviors. By using the same data set to analyze the birth spacing effect on these outcomes the estimated results will give a more complete picture of how important birth spacing is for the future of children.

Secondly, previous studies only look at the outcomes in early stages of life. For example, Buckles and Munnich (2012) only consider test scores of children who are between 5 and 7 years old. Also, Powell and Steelman (1993) only focus on siblings who are high-school age, i.e. aged between 15 and 18 years old, and measure their educational outcomes by high school graduation and post-secondary school attendance. In this paper, I examine the effect of birth spacing on outcomes overtime, starting at adolescence (between 13 and 18 years old) to young adulthood (between 19 and 24 years old) and adulthood (between 25 and 32 years old). The reason for following the sibling over their lifetimes is to test whether birth spacing effects persist over time. If a longer birth interval is expected to benefit the older sibling in term of higher test

such as alcohol use, drug use, and smoking in their estimation.

scores (see Buckles and Munnich 2012), we might expect it to continue to benefit this child in the future in terms of further educational attainment (years of schooling, college attendance, college degree) and better labor market outcomes (higher annual earnings). However if the birth spacing effect does not persist, it means that it is overcome with age individual outcomes are driven by other external or internal factors. External factors might include peer and neighborhood effects. Internal factors might be the change in family resources allocated to and across siblings according to variation in ability or outcomes across siblings.

Finally, I use family fixed-effect to control for the endogeneity of birth spacing. As discussed in the previous section, the endogeneity of birth spacing occurs because birth spacing is correlated with unobserved variables within and across families. In this paper, I focus on controlling unobserved variables within family that affect all siblings in household at the same time as these time-invariant factors should help to identify the effect of birth spacing on siblings' outcomes within the same period of time. Parent's characteristics and family resources in terms of time and finance spent on siblings are examples for time-invariant family characteristics. Fortunately, the family fixed effect method can address the within-family endogeneity although it could not help to solve the across-family endogeneity.

## 4.3 Data

### 4.3.1 Add Health Sampling Design and Weights for Sibling Sample

I use the National Longitudinal Study of Adolescent Health (Add Health), a longitudinal study of a nationally representative sample of adolescents in grades 7-12 in 1994-1995 academic year in the United States. The Add Health cohort have been surveyed (in-home interviews) in four different waves with the most recent in 2008

(Wave IV) when the sample was aged from 25 to 32.[44]

In the sampling design, at first, 80 high schools were selected non-randomly from the Quality Education Database (QED) consisted of 26,666 U. S. High Schools. The main selection criteria is the school size. For the 7 grader selection, 52 feeder schools (junior high & middle) that regularly sent their graduates to the above high schools were added. After the total 132 schools were determined, all students in grades 7 through 12 attending at those school were asked to complete the In-school Questionnaire.

In the next step of choosing students for in-home interview, 27,559 students were finally selected. They consist of four sub-samples: (1) the core sample of 16,044 students drawn equally from 12 student-level strata (sex- and grade-based); (2) the PAIRS sample which includes all students at two high schools; (3) non-genetic supplemental sample drawn based on their responses about ethnicity (high education Blacks, Cubans, Puerto Rican, Chinese) and disability to the In-school Questionnaire; (4) the genetic supplemental sample selecting various types of sibling pairs (twins, unrelated siblings, half siblings, and full siblings) based on students' response to the In-school Questionnaire.

The birth spacing issue examined in this paper involves the genetic (siblings) supplemental sample. Twins are not considered in my empirical analysis.[45] One limitation of the sibling sample involves the selection design; especially, the sampling of sibling data is not random. In particular, all students who are identified as twins or unrelated siblings such as adopted kids (based on their responses to the In-school

[44]The estimation results of this paper do not take into account the attrition bias due to the different Waves in the data. However, note that, regarding the attrition issue for the genetic sample in the Add Health data, the response rates, the bias remaining, and relative bias were reported in Chantala et al. (2004) and Brownstein et al. (2010). In particular, the response rate for genetic sample in Wave 3 was 79.62% and this rate for Wave 4 was between 77.6% and 86% for different types of sibling pairs. Both studies show that non-response bias is negligible and Wave 3 and Wave 4 adequately represent the same population surveyed at Wave I.

[45]The nature of the interaction between twins is fundamentally different from that of the interaction between siblings with an age gap. I therefore choose to drop twins from my analysis.

Questionnaire) were included with certainty in the genetic supplemental sample. Half- and full-siblings were selected systematically and the necessary condition is that both siblings were in the 7th through 12th grade. Note that the selection of full siblings came from the Core, Pairs, or non-genetic sample and the rest of the In-school Questionnaire survey sample (samples of 132 schools). The question sample also includes siblings who are not in the 132 schools sampled. This creates a major difficulty in the construction of the sampling weights for the sibling sample. Except for the full sibling sample, half siblings, unrelated siblings, and twins in the genetic supplements were sampled independently from the other samples (Core, PAIRS, or Non-genetic sample).

The sibling sample oversamples certain groups of adolescents and therefore, does not represent the national population of adolescents. If there are no appropriate adjustments for oversampling, the estimates and empirical implications of research using this data set could be misleading for the population. Add Health provides a list of general sampling weights to produce nationally representative description statistics. For an illustration of weight implementation, I report in Table 4.2 the descriptive statistics computed using the cross section weight "GSWGT1". These weights adjust data in the sample design in which the adolescents were chosen with a known probability of being selected from 1994-1995 enrollment rosters of US schools. Contrasting the mean and standard error, the summary statistics calculated without weights with those calculated with weights, the only significant differences could be seen in race variables as shown in this table. Regarding the oversampling of the twin sample in the sibling data, I provide the summary statistics in Table 4.3 calculated for a twin-excluded sample and a twin-included sample. There are few differences between the two groups of statistics. This reflects the fact that though twins are chosen with certainty from 132 schools, their selection was independently sampled from the Core, Pairs, and non-genetic samples, and therefore excluding twins from the sample is unlikely to affect the analysis of other sibling pairs within the 132 schools.

In addition, I perform a mean test between the sibling sample I use in my analysis and the sample weighted. The results are reported in Table 4.4 in which the last column shows the p-value of mean test. Out of 18 variables, only 6 variables show that the mean test could not be rejected including male, parent's education, parent's marital status, annual earning at adulthood, years of schooling in young adulthood, and child's age at young adulthood. This result implies that the difference between two samples are large at most variables. Notably, number of sibling in household and race have large differences in means between two samples. It means that the sibling sample represents extremely large family and black, Hispanic population.

Regarding to the sibling sample used in this paper, as pointed out by Chantala (2001), the major issue with the sibling sample design is that sampling weights could not be computed for adolescents not included in the sample collected for use in making national estimates (Probability Sample). As discussed above, the sample of genetic pair of siblings is not nested in the probability sample. In fact, there is no weight information for the pairs in which additional siblings did not attend any of the 132 schools (They were interviewed to increase the sample size for genetic analysis). Chantala (2001) argued that because 35.83% of the pairs in the genetic sample do not have weight information for both members, one must carry out the analysis without any sampling weight adjustment. To the best of my knowledge, most papers using genetic sample in the Add Health data do not use sampling weights (See, for instance, Fletcher (2008), Fletcher and Wolfe (2014), Thompson (2014)). In this paper, I follow the above suggestion and do not use weights for sibling sample as this is the limitation of this data set. My estimates, therefore only represent the sample of adolescents who were interviewed in the Add Health data. My estimation do not represent the national population of siblings and this consideration should be kept in mind when interpreting the findings of this paper.

### 4.3.2 Sibling Pair and Variable Description

I mainly focus on the "all siblings sample" in Add Health that includes 3,139 pairs of siblings.[46] In the original set-up, sibling pairs can be full-siblings (1,251 pairs), half-siblings (442 pairs), twins (741 pairs), and unrelated/unidentified siblings (705 pairs of step-children, adopted children, and cousins). Each pair of siblings within a particular family consists of any one older sibling and any one younger sibling. For example, if the family has two siblings, it has one pair. If there are three siblings there are three pairs. Similarly, four siblings give rise to 6 pairs and five siblings create ten pairs. As my interest is to examine how the spacing of births is associated with the future outcomes of siblings, I restrict my sample to only pairs of siblings in which two siblings are consecutive. In particular, only the birth spacing of the first and the second child, of the second and the third child, and so on are taken into account of my analysis. I exclude the sample of sibling pairs in which birth spacing between two siblings equals to zero. The reason is that it is impossible to define which sibling is older for later analysis. This restriction also excludes all twins case as "special" case because in the theory of family structure the interaction between twins is different from other types of sibling pairs. Restricting the sample to those individuals without missing data on outcome and control variables (including imputed parents' education, family income, and mothers' age at birth allows a final sample of 834 pairs (528 pairs of full-siblings, 137 pairs of half-siblings, and 169 pairs of non-related siblings) or 1,668 individuals.[47] If the sample is divided by age group there are a sample of 834

[46] As Add Health uses school-based design so that respondents in Add Health survey might not be representative of all adolescents. Therefore, I use weighted information that is available in Add Health data to see whether there is any big difference between weighted sample and unweighted sample. My analysis indicates that there is minor difference in summary statistics between two sample. For brevity, I do not show the results in this chapter and they are availalbe upon request.

[47] A limitation of siblings sample in Add Health data is that siblings pairs are chosen non-randomly among school-based adolescents (i.e. the proportion of twin pairs is large). However, due to the

older siblings and a sample of 834 younger siblings.

In addition to the "all sibling sample", I also carry out analysis limited to the biological siblings sample in which siblings within a family share either a biological mother or biological father. In particular, only full-siblings and half-siblings are used for analysis. The differences between the estimated effects of birth spacing on different type of siblings might show which factors (genetic factors or allocated resources) are more important in human capital investment.

I examine the linkage between the spacing of births and the future outcomes of siblings in various aspects, i.e. educational attainment, labor market outcomes, and engagement in risky behaviors. The information about these outcomes is taken from three Waves of data (Wave 1, Wave 3 and Wave 4).[48] Wave 1 is associated with adolescence stage when siblings are between 13 and 18 years old, Wave 3 is related to young adulthood stage when siblings are between 19 and 24 years old, and Wave 4 is associated with adulthood stage when siblings are between 25 and 32 years old.

For educational outcomes, I use percentile ranks on test scores, years of schooling, a binary variable for attending one or more years in post-secondary school (college attendance), and a binary variable for completing any 4-year program at college study (college degree). I follow-up the siblings' outcomes in two subsequent stages for testing the persistence of the birth spacing effect on educational outcomes. In particular, percentile rank on test scores is examined in both adolescence and young adulthood; years of schooling are investigated in both young adulthood and adulthood, college attendance is looked at in young adulthood and college degree is analyzed in adulthood.

The analyzed test score is the Add Health Picture Vocabulary Test (AHPVT).[49]

---

unavaibility of the weighted information for siblings sample in Add Health Data, my sample could not be weight-adjusted in the empirical implemetation of this chapter.

[48] I do not use Wave 2 in the analysis as the respondents are only interviewed one year after Wave 1.

[49] AHPVT is a computerized, abridged version of the Peabody Picture Vocabulary Test—Revised. The Peabody Picture Vocabulary provides a quick estimate of verbal ability and scholastic aptitude.

All information on raw test scores, age-standardized test scores, and percentile ranks of test scores are available in the data. However, I use the percentile ranks on scores for the purpose of comparison between two Waves of data.[50]

The information on years of schooling in adolescence and young adulthood is taken from the question "What is the highest grade or year of regular school you have completed ?" (H2ED1) in Wave 3 and H4ED1 in Wave 4. College attendance is analyzed as a binary variable for whether the respondent had one or more years in post-secondary school by Wave 3 and by Wave 4. College degree is also a binary variable for whether the respondent has earned a 4-year college degree by Wave 4.[51]

For labor market outcomes, I use annual earnings of the respondent at Wave 4 when the respondents were on average 29 years of age . At Wave 4, the respondents are asked "In 2006/2007/2008, how much income did you receive from earnings - that is, wages or salaries, including tips, bonuses, and overtime pay, and income from self-employment?" and "what is the best guess of the income you received from earnings?" (H4EC2 and H4EC3). I choose annual earnings at Wave 4 to analyze the effect of birth spacing on labor market outcomes as it is the latest information on labor market. Moreover, the respondent is 29 years old on average and at this age, he/she has most likely completed schooling.

For the probability of engaging in risky behaviors, I investigate how birth spacing affects the probability of smoking cigarettes. I only examine this effect in adolescence since at this stage siblings are greatly affected by each other as well as by the other members of the family. Smoking behavior is measured as a dichotomous variable that equals "1" if the respondent reports cigarette smoking and equals "0" otherwise. The respondent is asked the following question "During the past 30 days, on how many

[50] I also perform a robustness check with age-standardized test scores. The results using age-standardized test scores are similar to those using percentile rank on test score.

[51] A sample of respondents who have high school diploma or certificate on general education development (GED) is used as a robustness check for the analysis of college attendance and college degree outcomes.

days did you smoke cigarettes?" (H1TO5) at Wave 1. The respondent is considered to be engaged in smoking if the answer is different from zero.

"Birth spacing" measures the difference in age between the index child who is called the older sibling and the subsequent child who is called the younger sibling. The age of each sibling is calculated by subtracting the date of birth (H1GI1M, H1GI1Y) from the current date (IMONTH, IYEAR). "All sibling sample" consist of either full sibling, half-sibling, and non-related sibling pairs. I measure the spacing of births in two different ways. In the first way, the spacing of births is considered as continuous variable. It is measured as an absolute value of the difference of age in months between two siblings. In the second way, birth spacing is a dummy variable that equals one if the difference in months between two siblings is greater than 24 and equals zero otherwise.

For other explanatory variables, I use variables that are considered to be correlated with how birth spacing affects siblings' outcomes. They are health status of child at birth, birth order, the gender and race of the child, the number of siblings in a family, family income, and characteristics of parents (age, marital status, and education).[52] Most variables were obtained directly from questions in the In-home Parent Survey or Wave 1 In-home Respondent Survey. The birth order of siblings in the "all siblings sample" (full-siblings, half-siblings, and unrelated siblings) is calculated from ranking siblings by their ages.

Table 4.1 reports the means and standard deviations of the variables used in Add Health for the sample of siblings (1,688 individuals), a sample of older siblings (834 individuals), and a sample of younger siblings (834 individuals). The average age of

---

[52] As the "all sibling sample" contains both biological and non-related siblings, in the empirical analysis I use characteristics of a parent (mother and farther) which could be either biological parent or residential parent. The rational for considering non-related siblings (which leads to the consideration of residential parent) is that I could be able to expand the sample size in the analysis of the sibling pairs. In measuring these characteristics, parent's education is the maximum of mother's and father's education. Parent's age at child birth is the minimum of father's and mother's age at child birth. Parent's marriage status is a dummy which indicates whether the mother or father in the household is married.

older siblings is 17 years old in adolescence, 23 years old in young adulthood, and 31 years old in adulthood. These figures for the younger siblings are 14, 21, and 28, respectively.

Regarding the measured outcomes, the older siblings have better educational attainment and annual earnings on average compared to the younger siblings. The older siblings are in a higher percentile rank on average on the AHPVT than the younger siblings in both adolescence and young adulthood (48.66 versus 46.73 and 51.84 versus 45.98). As seen in Table 4.1 the difference in percentile ranks between the older siblings and the younger siblings is much bigger when the siblings are in young adulthood. Regarding years of schooling, the older siblings only have slightly more years of schooling on average than the younger siblings in both young adulthood and adulthood (13.40 versus 12.73, 14.27 versus 14.03, respectively). The older siblings are more likely to attend college and complete college study than the younger siblings do (53% versus 47% for college attendance and 32% versus 29% for college degree). This explains why the older sibling tends to earn more than the younger sibling (annual incomes on average $36,703 versus $30,496). The older sibling is also less likely to engage in cigarette smoking compared to the younger sibling on average (62% versus 63%). There are more older siblings who are males than young siblings. The other characteristics are similar between the older siblings and the younger siblings. The older siblings are in a lower birth order (3.90 on average) in the family than the younger siblings (4.91 on average).[53]

[53] I construct the birth order variable by ranking the age of all siblings in the household. Under this construction, the older siblings are in lower birth order than the younger siblings. As shown in the descriptive statistics, on average the number of siblings in these households is about 4. The average birth order is 4 for older siblings and 5 for younger siblings, respectively. These descriptive statistics of birth order can be explained by the fact that this paper looks at sibling pairs and excludes all twin sample from the empirical analysis. In addition, the average age of parents at birth for the sibling pairs is about 25. This means that these parents have 4 children by the age of 25 which could include either biological siblings or non-related siblings (adopted-child, cousin, or step-child).

## 4.4 Empirical Models

As pointed out earlier, heterogeneity is a challenge to this research. In this context, I take into account the heterogeneity across families by implementing two empirical approaches. In the first approach, the baseline empirical models, I use ordinary least square estimation (OLS) with extended covariates for family characteristics. In the second approach, I estimate family fixed-effect models. In addition, I add an interaction term for interaction between birth spacing and an indicator for the older sibling. The addition of the interaction term helps examine the differences in the birth spacing effect between older and younger siblings.[54] In both approaches, the standard errors are clustered by family as there could be more than one pair of siblings in one family.

For baseline empirical specifications, I estimate:

$$Y_{isf} = \alpha + \beta_1 Gap_s + \beta_2 Older_i + \gamma Gap_s * Older_i + \theta X_i + \pi Family_f + \varepsilon_{isf}$$

where $s, i$ and $f$ are indexes for siblings pair, the individual, and family, respectively. $Y_{isf}$ is an outcome of child $i$ in siblings pair $s$ of family $f$. In the empirical implementation, I look at the various outcomes of siblings described above. I measure birth spacing in two ways. First, $Gap_s$ is measured as the number of months between the two sibling pairs. Second, it is measured as an indicator that takes 1 if the space is greater than 24 months or 0 otherwise. $Older_i$ is an indicator for whether individual $i$ is the older sibling. $Gap_s * Older_i$ is an interaction term that measures the interaction between birth spacing and being the older sibling, $X_i$ is a vector of characteristics of child $i$ including gender, race, health status at birth, birth order, and age of mother at child's birth. $Family_f$ is a vector of all commonly observable family characteris-

[54] Buckles and Munnich (2012) run separate samples for the younger sibling and for the siblings in their analysis.

tics including number of siblings, marital status of the family's head, education of the family's head, and family income. $\varepsilon_{isf}$ is standard error that captures all unobservable factors that affect siblings outcomes and is clustered by family .

Of important note, $Gap_s * Older_i$ indicates whether there exists any differences in the effect of birth spacing on siblings' outcomes across two groups: younger siblings and older siblings. The advantage of this approach is that we can interpret the results by looking at the coefficients for these two different groups.[55] An alternative approach is to run separate regressions for older siblings and younger siblings (Buckles and Munnich 2012).[56] I also implement this approach in my empirical investigation.[57]

For binary variables such as college attendance, college degree, and smoking behavior, I use logistic regression models to estimate the effect of spacing on the siblings' outcomes. For both OLS and logistic models, the standard errors are clustered by family.

The OLS regression (baseline models) described above controls for some family characteristics by including the variable $Family_f$. There might be other unobserved factors in the family are associated with the individual's outcomes as discussed in previous section, i.e. parents' characteristics, tastes, and abilities. In the second empirical approach, I estimate the following fixed effect models that could control for unobserved heterogeneity:

$$Y_{isf} = \alpha + \beta_1 Gap_s + \beta_2 Older_i + \gamma Gap_s * Older_i + \theta X_i + \pi_f + \varepsilon_i$$

where $Gap_s$, $Older_i$, and $X_i$ are the same as in the OLS. Note that I exclude the

[55] Possibilities include (a) only the intercepts differ across groups, (b) intercepts and some subset of the slope coefficients differ across groups, or (c) all of the coefficients, both intercepts and slope coefficients, differ across groups.

[56] Buckles and Munnich (2012) show that birth spacing has a significant effect in the older sibling and an insignificant effect in the younger sibling. However, the sample size for the older sibling is larger than the sample size for the younger sibling. It would, therefore be very misleading to say that birth spacing was important for the older sibling but not for the younger sibling.

[57] However, running separate models for each group can be quite unwieldy, estimating many more coefficients than may be necessary.

$Family_f$ variable and include the fixed effect variable $\pi_f$ in the fixed effect models. As discussed in previous section, the family fixed-effect model only can address the endogenous issue of birth spacing within family.

To guarantee that fixed-effects or random effects are preferred in the specification of the model, I implement the Hausman test (Hausman 1978) in which the alternative hypothesis of the fixed effect models is tested against the null hypothesis of the random effects models. In all of my results, the statistics are found to be statistically significant. This supports the use of fixed-effect models.

The OLS models are not nested in the fixed effects models. $\pi_f$ captures all common family characteristics shared among siblings including both observable and unobservable variables.

In addition, since this analysis involves "double counting", i.e. an individual is the "older" sibling in the one pair but the "younger" sibling in another pair, I carry out robustness check for running separate models for the younger sibling and the older sibling. For the OLS and family fixed-effect models, the robust standard errors are clustered by family.

## 4.5 Results

I test for the effects of educational outcomes (percentile rank on test score, years of schooling, college attendance, and completion of college degree), labor market outcomes (annual earnings), and engagement in risky behaviors (cigarette smoking). In each table, I present results from the OLS and family fixed-effect models. Also, I estimate models with two measures of birth spacing are shown in the results. The first two columns in each table show the results of models where birth spacing is measured as the difference in months of two sibling births and the last two columns report the results for model including a dummy variable for whether two subsequent siblings were born at least two years apart.

As discussed in the Section 4.2 on the Add Health data, this research does not take into account the weight adjustments on the siblings sample as they are not available at the time this analysis was done. The empirical findings of this paper are therefore restricted to the Add Health sample and may not represent the entire adolescence population of the United States. In view of this, there are two important notes. First, this research builds on other papers that provide empirical evidences on siblings using the Add Health data such as Fletcher and Wolfe (2009), Fletcher (2014), and Thomson (2014). Their estimates are not weight adjusted for the same reason. Secondly, as pointed out in the discussion of the data, the use of weights is impractical as 35.83% of the pairs in the genetic sample do not have weight information for both members. Even with the availability of the computation techniques for weights (provided by the Add Health), some students are sampled not from the list of original 132 schools and therefore the weights at the first selection stage are not available for this group. Therefore, the results in below sessions are only interpreted as the results drawn from a sample of adolescents who were interviewed in the Add Health data and they were not representative for the whole adolescent population in the U.S.

### 4.5.1 Effect of Birth Spacing on Percentile Ranks on Test Scores

The analyses for the effect of birth spacing on percentile ranks of test scores are showed in Tables 4.5 and Table 4.6. Table 4.5 contains results from OLS models and Table 4.6 shows results from family-fixed effect models. In each table, the effect of birth spacing is examined in both the adolescent stage (the respondents are between 13 and 18 years old) and the young adulthood stage (the respondents are between 19 and 24 years old).[58]

Columns 2 and 4 in Table 4.5 reports estimated coefficients of variables from the OLS estimation when birth spacing is measured as the difference in months between two siblings in one pair and the respondents are in adolescence. The effect of birth spacing and the interaction term of birth spacing and an indicator for being the older sibling in a pair are not statistically different from zero. The results from family fixed-effect model shown in columns 2 and 4 of Table 4.6 are similar. The findings of both models suggest that birth spacing does not affect test scores in adolescence. Columns 3 and 5 of each Table show the results from OLS and family-fixed effect models when birth spacing is measured as a dummy variable. Similarly, the estimated effects of birth spacing variable and the interaction term are not statistically different from zero. The results suggest that there is no effect of birth spacing on test scores in adolescence and young adulthood. The variables that have positive effects on percentile ranks of test scores of siblings in OLS models (Table 4.5) are being white, parent's education, parent's age at child birth. The only negative effect on test scores in both stage is numbers of siblings in household. It is interesting to note that the positive effect of family income found in adolescence disappears in young adulthood.

The results in Buckles and Munnich (2012) showed that a longer birth spacing interval increases the test scores of older siblings at ages between 5 to 7 years old. My finding implies that the effect does not persist and disappear when the siblings transition to adolescence and young adulthood. In the long-run, test performance is affected by a complex set of environmental factors such as school and neighborhood characteristics in addition to family resources. This might explain why the effect of birth spacing on siblings on test scores in early life stage does not persist as the siblings age. More evidence to support this theory is that the positive effect of family income on test scores in adolescence becomes statistically insignificant in young adulthood

[58] I also implemented an alternative analysis with test score outcome instead of percentile ranks on tests score. Results were qualitatively similar to those reported here, and are available upon request

(Table 4.5 versus Table 4.6).

### 4.5.2 Effect of Birth Spacing on Years of Schooling

I examine years of schooling in young adulthood and in adulthood. This is the number of years of schooling the respondent completed by Wave 3 (young adulthood) and by Wave 4 (adulthood). I do not use years of schooling as an educational outcome at the adolescent stage as there are no differences in years of schooling between siblings in adolescence. Most individuals have completed 12 years of schooling by Wave 3 (young adulthood stage). However, years of schooling at young adulthood is 13.40 on average for older siblings and 12.73 for younger siblings as seen in Table 4.1. These numbers at adulthood are 14.27 and 14.03, respectively. In comparison to the young adulthood stage, we can see that the difference in years of schooling between younger siblings and older sibling are smaller in adulthood. Younger siblings seem to catch up with older siblings in term of years of schooling when they transition to adulthood.

Tables 4.7 and 4.8 show the results of birth spacing on years of schooling when siblings are 23 years old (young adulthood) and 29 years old (adulthood) on average. Table 4.7 represents estimated coefficients from the OLS. Although the estimated effects of birth spacing on years of schooling in both young adulthood and adulthood are statistically insignificant, the positive effects of other factors are recognized such as child's age, parent's education, parent's marriage status, parent's at child's birth, and family income. The other factors which contribute to negative effects on years of schooling include being male and numbers of siblings in household.

Table 4.8 reports the estimated coefficients in fixed effect models. No impact of birth spacing on years of schooling is found and these results are similar to these in Table 4.7. Interestingly, although birth spacing does not affect years of schooling birth order is negatively correlated with years of schooling. In particular, being the older sibling in household is more likely to have higher years of schooling than being the younger sibling. This finding is consistent with the literature on birth order. This literature supports for the evidence that older sibling received more family resources than younger sibling to get higher education.

I also implemented the same analysis using the biological sibling sample (unrelated siblings are excluded). The estimated results shown in Table 4.13. They indicate that birth interval does not affect years of schooling in adulthood stage. However, there is evidence of the negative effect of birth order on years of schooling.[59] The results indicate that on average an older sibling has more time of schooling than younger sibling. In conjunction with the statistically insignificant effect of birth spacing, this finding implies that among dimensions of family structure, birth order might be more important than birth spacing in determining the degree to which parents invest in their children. This finding is consistent with earlier findings in the literature that the shares of the resources devoted by parents to their children's education are decreasing with birth order (Booth and Kee 2005).

One other notable finding in Tables 4.7 and 4.8 is that being male is negatively correlated with years of schooling. Controlling for family characteristics, males have less years of schooling than females do in young adulthood and adulthood. Later in this paper, I will report the results of an alternative analysis in which the relationship between siblings in terms of ages replaced by the relationship in terms of sex composition in sibling pairs.

### 4.5.3 Effect of Birth Spacing on College Attendance

I continue to investigate how birth spacing affects educational attainment by looking at the probability of attending college at young adulthood when siblings are 22 years

[59] The respondents in Wave 1 survey of the Add Health data were asked: "How many children have your biological parents had together?" and "Which child are you—the first, the second, or what?" for birth order information. Therefore, in biological siblings sample the older siblings are in lower birth order than the younger siblings.

old on average. The college attendance is coded as "1" if the sibling has one or more years in post-secondary school and "0" otherwise. Therefore, I use logit and fixed effect logit models in the analysis. Columns 2 and 3 in Tables 4.9 report the marginal effects of covariates on the likelihood of enrolling in college from the OLS models. Results from the family-fixed models are shown in Columns 2 and 3 in Table 4.10.

Columns 2 in each table show the results for the estimations in which birth spacing is measured as the difference in months between two siblings. Birth spacing does not have statistically significant effects. Columns 3 in each table show the results for the models in which birth spacing is measured as a dummy variable. Here, the marginal effects of birth spacing are statistically different from zero. The positive coefficients of birth spacing in columns 3 show that spacing siblings at least 2 years apart increases the likelihood of enrolling in college by 12.17 percentage points. The result for the family fixed effect is similar but much bigger in magnitude. In particular, being at least 2 years apart the sibling is more likely to attend college at approximately 78 percentage points.[60]

Another statistically significant result is shown in interaction term. Recall that the interaction term is measured by multiplying birth spacing and an indicator for whether the individual is the older sibling in a sibling pair. The coefficient of interaction term reveals how the effect of birth spacing differs between the older sibling and the younger sibling. Therefore, a negative coefficient on the interaction term in columns 3 implies that birth spacing has a larger effect on the younger sibling than it does on the older sibling in terms of the possibility of enrolling in college. In particular, the results in column 3 of Table 4.9 indicate that being at least 2 years apart between siblings increases the likelihood of enrolling in college by 7.56 percentage points for the younger sibling compared to the older sibling.[61] The results in the family-fixed

[60] The reason for a significantly large effect of birth spacing in fixed effect model might be a small sample used in analysis. This is one of limitation of this analysis.

[61] From Table 4.9, column 2, I derive -0.0756 using the formula -0.0756=-0.1217+0.0459

effect (Column 3 in Table 4.10) show a similar effect. Spacing of more than two years apart will increase the younger sibling's possibility of enrolling in college by 30.67 percentage points compared to the older sibling.[62][63]

This finding suggests that a closely-spaced birth reduces the possibility of attending college for siblings, especially for the younger sibling. This effect might be caused by limited family resources allocated to siblings for educational investment. This conclusion is supported by the finding in the previous part in which percentile ranks on test scores of siblings are not affected by birth spacing (seen in Table 4.5 and 4.6) and by a strongly positive correlation between years of schooling and family income in young adulthood (see columns 2 and 3 in Tables 4.7 and 4.8).[64] The finding suggests that family resources still play a very important role in investment in children's education, especially in pursuing a higher education.[65]

Also, the correlation of being male and college attendance is statistically negative. It suggests that males are less likely to attend one or more years at post-secondary school than females. Another negative correlation is between numbers of siblings in household and college attendance. Other covariates that have positive effects on the possibility of enrolling in college include parent's education, parent's marriage status, parent's age at child birth, and family income.[66]

---

[62] From Table 4.10 column 3, I derive 0.3607 using the formula -0.3607=-0.731-0.4243

[63] I do robustness check by controlling for test scores or percentitle ranks on test scores of adolescents in the models as a measurment of individual's ability. The results still show statistically significant effect of birth spacing on the likelihoof of attending college in young adulthood.

[64] I also do robustness check by using an interaction between birth spacing and family income instead of the interaction term between birth spacing and an indicator for the older sibling. The results show estimated coefficients of interaction term are positive and statistically different from zero in OLS models.

[65] As a robustness check I include an interaction term which is the multiplication of birth spacing and family income in the estimation. Interestingly, although birth spacing does not affect years of schooling birth order is negatively correlated with educational outcome in fixed-effect models. This finding supports for the evidence that older sibling received more family resources than younger sibling to get higher education.

[66] I also implement an alternative analysis in biological sibling sample. The results were qualitatively similar to those reported here.

### 4.5.4 Effect of Birth Spacing on College Degree

To have a complete view of the effect of birth spacing on educational outcomes, I examine whether an individual has a 4-year college degree by Wave 4 - adulthood. By this wave, the siblings are approximately 29 years old on average. the dependent variable is a binary variable which equals "1" if a sibling had a college degree by Wave 4 and equals "0" otherwise. Columns 4 and 5 in Tables 4.9 and 4.10 summarizes the results predicting college degree for both logit and family fixed effect logit models. Again, the results show that the estimated effects of birth spacing and interaction term are statistically different from zero in case birth spacing is measured as a dummy variable (see Columns 5 in Tables 4.9 and 4.10). The positively estimated effects of birth spacing indicates that greater birth spacing leads to more likelihood of completing 4-year college degree in adulthood. In particular, spacing siblings at least 2 years apart increases the likelihood sibling gets college degree by about 9 percentage points in the logit model and 71 percentage points in the fixed effect model. However, the coefficients of interaction terms are statistically insignificant. These findings suggest no differential effect of birth spacing between the younger sibling and the older sibling on obtaining college degree. This result is different from the previous session in which birth spacing is found to have more effect on the younger sibling than the older sibling in likelihood of attending college. Again, the results show that males are less likely to obtain college degree in adulthood than females. Additionally, family income is positively correlated with the likelihood of having college degree in adulthood even though this effect is small. One explanation for this finding is that in the adulthood stage siblings have other financial resources such as labor market earnings or loans approval for college education.

Combined with the findings in the previous sections, these findings indicate that birth spacing continues to have effects on post-secondary education when siblings transition to adulthood stage. Moreover, the family income is positively correlated with the possibility of enrolling in college and obtaining a college degree. This suggests that family resources are important for siblings in pursuing and completing their post-secondary education. However, the different effect of birth spacing on higher education between the younger and older siblings does not persist from young adulthood to adulthood. It could be explained by the fact that whenever siblings are already in colleges they can find other resources rather than family finance, and therefore, birth spacing and birth order do not make any differences between the younger one and the older one.

### 4.5.5 Effect of Birth Spacing on Annual Earnings

This paper is the first paper to look at the relationship between birth spacing and the labor income of siblings. Labor income is measured in terms of annual earnings in adulthood when siblings are 29 years old on average. I only investigate the effect of birth spacing on annual earnings in adulthood so that I can control for completed educational attainment. Table 4.11 shows the results for the association between birth spacing and annual earnings, adding a variable for educational attainment (years of schooling) to the models. Both OLS and family fixed-effect estimation reveal no statistically significant effect of birth spacing on earning. I also run a robustness check by excluding years of schooling from the analysis and the results still show that birth spacing does not affect annual earnings in adulthood. This finding suggests that birth spacing does not have impact on labor market outcomes in adulthood. Also, there is no relationship between birth order and annual earnings. In addition, the effect of family income is even statistically insignificant. Therefore, the finding of no association of birth spacing and labor market outcome is not surprising. Overall, these results suggest that after controlling for educations, family structure does not affect labor market outcomes in the long-run of siblings' life time. Another notable finding in this part is males earn more than females in adulthood.

### 4.5.6 Effect of Birth Spacing on Engagement in Risky Behaviors

This paper also is the first paper to examine the correlation between birth spacing and engagement in risky behaviors. I choose cigarette smoking in adolescence for my analysis. A respondent is considered to engage in risky behaviors if he or she has ever smoked a cigarette during adolescence. At this stage, the individual is 16 years old on average. I choose adolescence to analyze the effect of birth spacing on cigarette smoking behavior because at this stage the sibling influence is stronger and the interaction between siblings are more frequent than at other stages. Adolescents may imitate each other and they can learn both bad and good behaviors quickly from other members of family. Therefore, in this analysis I also consider an indicator for whether that mother smokes cigarettes.

As smoking behavior is a binary variable I again use logits and fixed effect logits to estimate the effect of birth spacing. Table 4.12 presents the marginal effects of these models. These marginal effects of birth spacing are never statistically different from zero. The results suggest that birth spacing does not affect the likelihood of engaging in cigarette smoking. Also, the results for birth order are the same. Only mother's cigarette smoking is strongly correlated with respondent's likelihood to smoke.

### 4.5.7 Extensions, Limitations, and Directions of Future Research

In this section I provide empirical results for two important extensions of empirical analysis. Finally, I emphasize the limitations of this paper and discuss some directions to improve it's empirical findings.

#### 4.5.7.1 First Extension - Structure of Sibling Pairs

As extensions, I estimate separate models for three samples of sibling pairs: two brothers pairs, two sisters pairs, and brother-sister pairs in biological sibling sample. The numbers of observations for each sample are 209, 449, and 176, respectively.

For percentile ranking on test scores, the results from running separate models for each sample are similar with the results from the previous models. In particular, birth spacing does not affect siblings' percentile ranking on test scores in young adulthood and adulthood. In addition, there is no effect of birth spacing on the likelihood of college completion, annual earnings and cigarette smoking. [67]

Tables 4.14 - 4.17 show summaries of results for key variables including birth spacing, an indicator for the older sibling, and the interaction of these variables from models of years of schooling. Table 4.14 shows the results for brothers sample. Columns 1 - 4 present the effect of birth spacing on brothers' years of schooling in young adulthood and columns 5 - 8 indicate this effect in adulthood. Note that the estimated coefficients of interaction terms in columns 1, 5, and 6 are positive and statistically different from zero. The results indicate that the older brother benefits more from larger birth spacing than does the younger brother does in terms of years of schooling. However, the sign of birth spacing effect becomes negative. In particular, a closely-spaced birth between brothers increases the years of schooling for younger and older brothers. These results appear in both young adulthood and adulthood. This finding suggests that parents might find it difficult to differentiate between two sons in terms of educational investment if these sons are spaced closely. In addition, two closely-spaced brothers might be more competitive with each other.

Table 4.15 shows the results from the same analyses with the sample of sisters pairs. The positive and statistically significant estimated coefficients of birth spacing

[67] I only show the tables for the effect of birth spacing on years of schooling and the likelihood of attending college.

(as a dummy variable) in columns 2 and 3 imply that birth spacing has positive effect on sisters' years of schooling in young adulthood. In particular, spacing two more years apart between sisters increase the years of schooling by 0.6 for the younger sister and by 0.24 for the older sister under OLS model. These increases are 0.3 and 1.2, respectively under family fixed-effect model. However, this effect does not persist when sisters transition to adulthood.

Table 4.16 presents the analysis for the sample of brother-sister pairs. Column 2 shows that estimated coefficient of interaction term is positive and joint statistically different from zero. The results suggest that birth spacing has positive effect on the years of schooling for the older sibling but no effect is found for the younger sibling in a mixed pair. The finding is consistent with the results in Table 4.11 (column 2) and Table 4.12 when the analysis is carried out for the sample of all siblings pairs.

Table 4.17 indicates the results for examining the linkage between birth spacing and the likelihood of enrolling a 4-year college, on the sample of sisters-pairs. Columns 3 and 4 show a positive and statistical significance of estimated marginal effect. It means that birth spacing has a positive effect on the probability of attending college for both younger and older sisters in young adulthood. However, this effect disappears in adulthood. I do the same analysis on the sample of brothers pairs and the sample of brother-sister pairs and I find no effect of birth spacing.

#### 4.5.7.2 Second Extension - The Interactions Between Siblings in Engagement in Risky Behaviors

In previous section, I find that birth spacing does not affect the smoking decisions of biological siblings. In this section I extend further this analysis to examine the determinants of their smoking decisions. Regarding the determinants, I am interested in the interactions of members within a family. Table 4.18 presents smoking behavior of the older sibling after controlling for whether younger sibling smokes cigarettes and versus. In this analysis, I use logit models for the sample of 126 older sibling.[68] In Model 1, I only control for birth spacing in months and an indicator for whether the younger sibling smokes cigarettes. In Model 2, I add an indicator for whether mother smokes cigarettes. The following models (Models 3-6) add more variables on family's and child's characteristics. The findings report no effect of birth spacing on the older sibling's smoking behavior. The only variables that determine the smoking decision of the older sibling are the mother and the younger sibling smoking. Table 4.19 shows results from the same analysis for 126 younger siblings. The results are stable and no effect of birth spacing on the younger sibling's smoking behavior. A notable finding is that mother's smoking behavior becomes statistically insignificant in all models. The only causal effect on the younger sibling's smoking behavior is the older sibling's cigarette smoking. This finding suggests that older sibling's smoking behavior plays the most important role in influencing whether the younger sibling smokes.

#### 4.5.7.3 Limitations and Directions of Future Research

There are two limitations of this paper that need to be emphasized. First, this paper, as well as other research papers using the genetic sample in the Add Health data share the same limitation that the sampling weights are unavailable for sibling sample. The weight construction is important so that the results could be nationally representative. The empirical findings in this paper are restricted to the Add Health data set. I will highlight the weight construction as a direction for future research next.

Secondly, a part of the empirical results comes from the estimation of family fixed effect models. However, the sample size of sibling pairs used for the fixed effect model is small. The larger sample size would help confirm the precision in the estimation of the paper. To improve this problem, large national data sets involving sibling pairs are needed. Add Health is expanding the genetic sample. In addition, Thompson

[68] See the next section for the discussion on these limitations of using small sibling sample for the fixed effect models.

(2014) pointed out that the Health and Retirement Survey (HRS) and the National Health and Nutrition Examination Survey (NHANES) started collecting data on the genetic markers of respondents.

In improving the empirical approach of this paper so that the findings could be nationally representative, the sampling weights for the sibling sample need be properly approximated (due to the unavailability of the weights). One solution is to follow the suggestion by Chantala (2001) to compute the pair weights. I highlight the formula as follows:

If the pair of the Pair of Adolescents (i and j) were sampled from the same feeder or high school (k), the formula is

$$PAIRWT_{i,j} = \frac{WEIGHT_i * WEIGHT_j}{SCHOOLWT_k}$$

If the The Pair of Adolescents (i and j) were sampled from a high school and the associated feeder school, the formula is

$$PAIRWT_{i,j} = \frac{WEIGHT_i * WEIGHT_j}{SCHOOLWT_{High\_School}}$$

The only information needed for these formulas is the sampling weight of each adolescent ($WEIGHT_i$ and $WEIGHT_j$) and the sampling weight for the school from which they were sampled ($SCHOOLWT$). Note that the $SCHOOLWT$ is not available in the Add Health data set but is available upon request.

## 4.6 Conclusion

I investigate the linkage between birth spacing and various outcomes of siblings using the sub-sample of sibling-pairs from the restricted-use National Longitudinal Study of Adolescent Health (Add Health) data-set.

My empirical results show that birth spacing does not affect siblings' percentile ranks of test scores and years of schooling in adolescence and young adulthood. The findings suggest that although wider birth spacing might benefit siblings' test scores at very young ages, this effect does not persist when the siblings transition to adolescence and young adulthood. In the long-run, the test performance might be affected by a complex set of environmental factors such as school and neighborhood characteristics in addition to parental investment in their children. For example, an individual's test performance might be influenced by his/her classmates, teachers, or friends in neighborhood.

However, I do find that wider birth spacing has a positive effect on the likelihood of enrolling in college in young adulthood and possibility of obtaining a college degree in adulthood for both the younger and older siblings. This finding implies that the effect of birth spacing persists when siblings transition to adulthood. I also find that birth spacing this effect is different on the older sibling and the younger sibling in young adulthood. A greater birth spacing has more impact on college enrollment for the younger siblings than for the older siblings. These findings can be useful in social, public, and economic policies in terms of support for post-secondary education. For example, policies on the tax value of the personal exemption or college aid rules should address the effect of birth spacing on post-secondary education to provide financial aid resources for college students who have closely-spaced siblings. How tax policies and financial policies reflect these implications might be potential research questions for my future research.

I find no effect of birth spacing on annual earnings in adulthood after controlling for educational attainment or the probability of engaging in cigarettes smoking in adolescence. Other notable findings are that brothers have fewer years of schooling and less likelihood of attending college than sisters do and that birth order is strong correlated with later-life educational outcomes after controlling for family characteristics. The findings suggest that the allocation of family resources to and across family members changes overtime to adapt to family's and siblings' abilities. Moreover, although birth spacing is strongly related to post-secondary education it is less powerful than birth order in determining parental investment in children.

Though the Add Health data set is one of the largest data set on adolescents, its limitation on the unavailability of sampling weights of sibling samples make the empirical findings in this paper non-representative for the U.S. I also discuss this limitations and highlight the directions for future research.

**Table 4.1: Summary Statistics - All Siblings Sample**

| Variables | Wave | All Sibling Sample | | Older Siblings | | Younger Siblings | |
|---|---|---|---|---|---|---|---|
| | | Mean | Std. | Mean | Std. | Mean | Std. |
| Birth Spacing (Months) | 1 | 25.98 | 13.96 | 25.98 | 13.96 | 25.98 | 13.96 |
| Birth Spacing (Dummy) | 1 | 0.51 | 0.50 | 0.51 | 0.50 | 0.51 | 0.50 |
| Test Score in Adolescence | 1 | 47.70 | 28.37 | 48.66 | 29.55 | 46.73 | 27.12 |
| Test Score in Young Adulthood | 3 | 48.91 | 29.01 | 51.84 | 29.68 | 45.98 | 28.02 |
| Years of Schooling in Young Adulthood | 3 | 13.07 | 1.97 | 13.40 | 2.14 | 12.73 | 1.72 |
| College Attendance in Young Adulthood | 3 | 0.50 | 0.50 | 0.53 | 0.50 | 0.47 | 0.50 |
| Years of Schooling in Adulthood | 4 | 14.15 | 2.27 | 14.27 | 2.31 | 14.03 | 2.22 |
| College Degree in Adulthood | 4 | 0.30 | 0.46 | 0.32 | 0.46 | 0.29 | 0.45 |
| Annual Earnings in Adulthood (thousands) | 4 | 33.60 | 35.20 | 36.70 | 42.06 | 30.50 | 26.30 |
| Smoking Behavior | 4 | 0.62 | 0.48 | 0.62 | 0.49 | 0.63 | 0.48 |
| Child's Age in Adolescence | 4 | 15.59 | 1.73 | 16.66 | 1.32 | 14.51 | 1.39 |
| Child's Age in Young Adulthood | 3 | 21.88 | 1.78 | 22.98 | 1.38 | 20.78 | 1.42 |
| Child's Age in Adulthood | 4 | 29.34 | 1.80 | 30.44 | 1.41 | 28.24 | 1.45 |
| Older Sibling | 1 | 0.50 | 0.50 | 1.00 | 0.00 | 0.00 | 0.00 |
| Male | 1 | 0.47 | 0.50 | 0.48 | 0.50 | 0.45 | 0.50 |
| Hispanic | 1 | 0.15 | 0.35 | 0.15 | 0.35 | 0.14 | 0.35 |
| White | 1 | 0.57 | 0.50 | 0.57 | 0.50 | 0.58 | 0.49 |
| Black | 1 | 0.21 | 0.41 | 0.21 | 0.41 | 0.21 | 0.41 |
| Number of Siblings in HH | 1 | 4.47 | 2.41 | 4.55 | 2.49 | 4.39 | 2.32 |
| Birth Order | 1 | 4.41 | 0.59 | 3.90 | 0.32 | 4.91 | 0.29 |
| Parent's Education | 1 | 12.90 | 2.36 | 12.90 | 2.36 | 12.90 | 2.36 |
| Parent's Marriage Status | 1 | 0.73 | 0.42 | 0.73 | 0.42 | 0.73 | 0.42 |
| Parent's Age at Child's Birth | 1 | 24.95 | 5.77 | 24.04 | 5.89 | 25.85 | 5.51 |
| Family Income in 1994 (thousands) | 1 | 46.05 | 54.52 | 46.05 | 54.54 | 46.05 | 54.54 |
| Observations | | 1668 | | 834 | | 834 | |

Parent is either mother or father who lives in the household

**Table 4.2: Summary Statistics of Sibling's Characteristics Computed With and Without General Sampling Weights**

| Variable | All Sample Unweighted | | All Sample Weighted | |
|---|---|---|---|---|
| | Mean | Std. | Mean | Std. |
| Test Scores at Adolescence | 50.14 | 28.95 | 53.02 | 28.43 |
| Child's Age at Adolescence | 15.58 | 1.73 | 15.37 | 1.79 |
| Male | 0.47 | 0.50 | 0.49 | 0.50 |
| Hispanic | 0.16 | 0.36 | 0.11 | 0.31 |
| White | 0.57 | 0.50 | 0.71 | 0.45 |
| Black | 0.20 | 0.40 | 0.14 | 0.34 |
| Asian | 0.06 | 0.23 | 0.03 | 0.17 |
| Number of Biological Siblings | 2.66 | 1.40 | 2.58 | 1.31 |
| Parent's Education | 13.05 | 2.39 | 13.02 | 2.33 |
| Parent's Marital Status | 0.73 | 0.42 | 0.75 | 0.41 |
| Parent's Age at Child's Birth | 26.00 | 6.01 | 25.73 | 5.78 |
| Years of Schooling at Adulthood | 14.43 | 2.31 | 14.29 | 2.31 |
| College Degree at Adulthood | 0.34 | 0.47 | 0.33 | 0.47 |
| Annual Earnings at Adulthood (thousands) | 37.55 | 45.89 | 35.15 | 37.97 |
| Child's Age at Adulthood | 29.35 | 1.78 | 28.84 | 1.66 |
| Number of Siblings in Household | 4.02 | 2.27 | 2.81 | 2.23 |
| Test Scores at Young Adulthood | 51.03 | 29.25 | 52.26 | 28.61 |
| Years of Schooling at Young Adulthood | 13.28 | 1.96 | 13.07 | 1.89 |
| College Attendance at Young Adulthood | 0.56 | 0.50 | 0.54 | 0.50 |
| Child's Age at Young Adulthood | 21.90 | 1.76 | 21.37 | 1.65 |
| Observations | 10454 | | 10454 | |

**Table 4.3: Descriptive Statistics of Individual Sibling Sample**

| Variables | All Sibling Sample (Including twins) | | Sibling Sample (Excluding twins) | |
|---|---|---|---|---|
| | Mean | Std. | Mean | Std. |
| Test Score at Adolescence | 46.39 | 28.58 | 46.45 | 28.47 |
| Child's Age at Adolescence | 15.56 | 1.72 | 15.55 | 1.76 |
| Male | 0.48 | 0.50 | 0.48 | 0.50 |
| Hispanic | 0.15 | 0.35 | 0.15 | 0.35 |
| White | 0.56 | 0.50 | 0.56 | 0.50 |
| Black | 0.23 | 0.42 | 0.22 | 0.42 |
| Asian | 0.05 | 0.22 | 0.06 | 0.23 |
| Number of Biological Siblings | 2.90 | 1.53 | 2.76 | 1.51 |
| Parents' Education | 14.81 | 32.27 | 19.06 | 36.41 |
| Head of Household's Education | 12.88 | 2.41 | 12.82 | 2.37 |
| Mother's Marital Status | 0.70 | 0.43 | 0.70 | 0.43 |
| Mother's Age at Child's Birth | 25.56 | 6.03 | 25.18 | 6.10 |
| Years of Schooling at Adulthood | 14.21 | 2.33 | 14.11 | 2.29 |
| College Degree at Adulthood | 0.31 | 0.46 | 0.30 | 0.46 |
| Annual Earnings at Adulthood (thousands) | 34.99 | 39.60 | 34.64 | 39.36 |
| Child's Age at Adulthood | 29.32 | 1.78 | 29.31 | 1.82 |
| Number of Siblings in Household | 4.50 | 2.46 | 4.51 | 2.45 |
| Test Score at Young Adulthood | 47.77 | 29.23 | 47.61 | 29.18 |
| Years of Schooling at Young Adulthood | 13.13 | 1.97 | 13.05 | 1.96 |
| College Attendance at Young Adulthood | 0.51 | 0.50 | 0.49 | 0.50 |
| Child's Age at Young Adulthood | 21.86 | 1.76 | 21.85 | 1.80 |
| Observations | 3156 | | 2353 | |

Parent is either mother or father who lives in the household

**Table 4.4: Mean Test between Sibling Sample and Weighted Sample**

| Variables | Wave | Sibling Sample | | Weighted Sample | | Mean Test* |
|---|---|---|---|---|---|---|
| | | Mean | Std. | Mean | Std. | p-value |
| Test Scores at Adolescence | 1 | 47.70 | 28.37 | 53.02 | 28.43 | 0.0000 |
| Child's Age at Adolescence | 1 | 15.59 | 1.73 | 15.37 | 1.79 | 0.0000 |
| Male | 1 | 0.47 | 0.50 | 0.49 | 0.50 | 0.1293 |
| Hispanic | 1 | 0.15 | 0.35 | 0.11 | 0.31 | 0.0000 |
| White | 1 | 0.57 | 0.50 | 0.71 | 0.45 | 0.0000 |
| Black | 1 | 0.21 | 0.41 | 0.14 | 0.34 | 0.0000 |
| Parent's Education | 1 | 12.90 | 2.36 | 13.02 | 2.33 | 0.0512 |
| Parent's Marital Status | 1 | 0.73 | 0.42 | 0.75 | 0.41 | 0.0652 |
| Parent's Age at Child's Birth | 1 | 24.95 | 5.77 | 25.73 | 5.78 | 0.0000 |
| Years of Schooling in Adulthood | 4 | 14.15 | 2.27 | 14.29 | 2.31 | 0.0212 |
| College Degree in Adulthood | 4 | 0.30 | 0.46 | 0.33 | 0.47 | 0.0152 |
| Annual Earnings in Adulthood (thousands) | 4 | 33.60 | 43.20 | 35.15 | 37.97 | 0.1291 |
| Child's Age in Adulthood | 4 | 29.34 | 1.80 | 28.84 | 1.66 | 0.0000 |
| Number of Siblings in Household | 4 | 4.47 | 2.41 | 2.81 | 2.23 | 0.0000 |
| Test Scores in Young Adulthood | 3 | 48.91 | 29.01 | 52.26 | 28.61 | 0.0000 |
| Years of Schooling in Young Adulthood | 3 | 13.07 | 1.97 | 13.07 | 1.89 | 1.0000 |
| College Attendance in Young Adulthood | 3 | 0.50 | 0.50 | 0.54 | 0.50 | 0.0024 |
| Child's Age in Young Adulthood | 3 | 21.88 | 1.78 | 21.37 | 1.65 | 1.0000 |
| Observations | | 1,668 | | 10,454 | | |

*Two samples, 2-sided t test at 95% confidential interval

Weighted Sample includes respondents who were interviewed in all waves of In-home Survey

Parent is either mother or father who lives in the household

**Table 4.5: Percentile Ranks on Test Scores - All Siblings Sample - OLS Models**

| Variables | Adolescence | | Young Adulthood | |
|---|---|---|---|---|
| | Coef./t | Coef./t | Coef./t | Coef./t |
| Birth Spacing (Months) | -0.053 | | 0.073 | |
| | (-0.694) | | (0.988) | |
| Birth Spacing*Older[1] | 0.066 | | 0.060 | |
| | (0.607) | | (0.577) | |
| Birth Spacing (Dummy) | | 0.803 | | 2.647 |
| | | (0.416) | | (1.298) |
| Dummy*Older[2] | | -1.305 | | -0.471 |
| | | (-0.524) | | (-0.188) |
| Older Sibling | 2.944 | 4.792 | 0.483 | 2.393 |
| | (0.773) | (1.567) | (0.129) | (0.737) |
| Child's Age in Adolescence | -0.613 | -0.389 | | |
| | (-0.919) | (-0.618) | | |
| Male | 2.718* | 2.751* | 1.269 | 1.296 |
| | (1.741) | (1.759) | (0.822) | (0.837) |
| Hispanic | 3.905 | 3.809 | 0.080 | 0.141 |
| | (0.793) | (0.773) | (0.015) | (0.026) |
| White | 14.679*** | 14.622*** | 9.712** | 9.795** |
| | (3.351) | (3.334) | (2.024) | (2.010) |
| Black | -5.782 | -5.821 | -10.970** | -10.882** |
| | (-1.239) | (-1.243) | (-2.103) | (-2.059) |
| Number of Siblings in HH | -0.756** | -0.748** | -0.928** | -0.915** |
| | (-2.008) | (-1.985) | (-2.281) | (-2.246) |
| Birth Order | 0.204 | 0.151 | -2.806 | -2.468 |
| | (0.076) | (0.056) | (-0.951) | (-0.835) |
| Parent's Education | 2.962*** | 2.967*** | 2.859*** | 2.884*** |
| | (7.390) | (7.428) | (6.698) | (6.774) |
| Parent's Marriage Status | 2.099 | 2.087 | 1.087 | 1.120 |
| | (1.056) | (1.047) | (0.463) | (0.476) |
| Parent's Age at Child's Birth | 0.317** | 0.302** | 0.509*** | 0.503*** |
| | (2.324) | (2.219) | (3.564) | (3.547) |
| Family Income in 1994 (thousands) | 0.053*** | 0.053*** | 0.022 | 0.020 |
| | (4.637) | (4.615) | (1.545) | (1.426) |
| Child's Age in Young Adulthood | | | 1.130* | 1.236* |
| | | | (1.663) | (1.936) |
| Observations | 1668 | 1668 | 1668 | 1668 |
| R-sqr | 0.256 | 0.256 | 0.237 | 0.236 |

t statistics in parentheses;
* p<.10, ** p<.05, *** p<.01
[1]: Interaction term between birth spacing (months) and being the older sibling
[2]: Interaction term between birth spacing (dummy) and being the older sibling
Parent is either mother or father who lives in the household

**Table 4.6: Percentile Ranks on Test Scores - All Siblings Sample - Fixed Effect Models**

| Variables | Adolescence | | Young Adulthood | |
|---|---|---|---|---|
| | Coef./t | Coef./t | Coef./t | Coef./t |
| Birth Spacing (Months) | -0.041 | | 0.250 | |
| | (-0.230) | | (0.948) | |
| Birth Spacing*Older[1] | 0.101 | | -0.057 | |
| | (0.482) | | (-0.333) | |
| Birth Spacing (Dummy) | | -0.427 | | 0.926 |
| | | (-0.171) | | (1.035) |
| Dummy*Older[2] | | -3.476 | | -3.242 |
| | | (-1.099) | | (-1.124) |
| Older Sibling | -5.468 | -2.784 | -1.188 | -1.399 |
| | (-1.047) | (-0.789) | (-0.300) | (-0.507) |
| Child's Age in Adolescence | -1.447 | 0.557 | | |
| | (-0.612) | (0.376) | | |
| Male | 2.803 | 2.891 | 1.126 | 1.226 |
| | (1.459) | (1.500) | (0.601) | (0.654) |
| Hispanic | 12.910 | 12.233 | 7.013 | 6.908 |
| | (1.323) | (1.247) | (0.722) | (0.708) |
| White | 9.129 | 9.067 | 4.806 | 4.726 |
| | (1.134) | (1.123) | (0.574) | (0.569) |
| Black | -2.519 | -2.568 | -0.060 | -0.109 |
| | (-0.300) | (-0.307) | (-0.007) | (-0.012) |
| Number of Siblings in HH | -0.259 | -0.266 | -0.959* | -0.986* |
| | (-0.439) | (-0.454) | (-1.734) | (-1.790) |
| Birth Order | -8.256* | -5.740 | -4.156 | -3.863 |
| | (-1.835) | (-1.550) | (-1.147) | (-1.241) |
| Child's Age in Young Adulthood | | | 2.065 | 2.370* |
| | | | (1.062) | (1.896) |
| Observations | 1668 | 1668 | 1668 | 1668 |
| R-sqr | 0.022 | 0.024 | 0.055 | 0.057 |

t statistics in parentheses; Hausman Test: Prob>chi2 = 0.0312 & 0.0131

* p<.10, ** p<.05, *** p<.01

[1]: Interaction term between birth spacing (months) and being the older sibling

[2]: Interaction term between birth spacing (dummy) and being the older sibling

Parent is either mother or father who lives in the household

**Table 4.7: Years of Schooling- All Siblings Sample - OLS Models**

| Variables | Young Adulthood | | Adulthood | |
|---|---|---|---|---|
| | Coef./t | Coef./t | Coef./t | Coef./t |
| Birth Spacing (Months) | 0.007 | | 0.009 | |
| | (1.258) | | (1.430) | |
| Birth Spacing*Older[1] | -0.009 | | -0.010 | |
| | (-1.213) | | (-1.156) | |
| Birth Spacing (Dummy) | | 0.350 | | 0.259 |
| | | (1.556) | | (1.510) |
| Dummy*Older[2] | | -0.298 | | -0.103 |
| | | (-1.582) | | (-0.475) |
| Older Sibling | 0.321 | 0.236 | 0.255 | 0.089 |
| | (1.268) | (1.074) | (0.810) | (0.332) |
| Child's Age in Young Adulthood | 0.294*** | 0.289*** | | |
| | (6.318) | (6.736) | | |
| Male | -0.276** | -0.266** | -0.581*** | -0.575*** |
| | (-2.520) | (-2.445) | (-4.779) | (-4.731) |
| Hispanic | -0.542 | -0.563* | -0.106 | -0.119 |
| | (-1.607) | (-1.665) | (-0.307) | (-0.348) |
| White | -0.364 | -0.392 | 0.074 | 0.050 |
| | (-1.191) | (-1.281) | (0.255) | (0.171) |
| Black | -0.136 | -0.160 | 0.222 | 0.203 |
| | (-0.412) | (-0.482) | (0.672) | (0.617) |
| Number of Siblings in HH | -0.091*** | -0.090*** | -0.095*** | -0.094*** |
| | (-3.641) | (-3.623) | (-3.250) | (-3.241) |
| Birth Order | -0.053 | -0.068 | -0.128 | -0.137 |
| | (-0.287) | (-0.372) | (-0.542) | (-0.586) |
| Parent's Education | 0.179*** | 0.177*** | 0.278*** | 0.276*** |
| | (6.222) | (6.209) | (8.132) | (8.100) |
| Parent's Marriage Status | 0.556*** | 0.556*** | 0.569** | 0.571** |
| | (3.692) | (3.703) | (3.290) | (3.298) |
| Parent's Age at Child's Birth | 0.031*** | 0.030*** | 0.040*** | 0.041*** |
| | (2.718) | (2.690) | (3.553) | (3.571) |
| Family Income in 1994 (thousands) | 0.003*** | 0.003*** | 0.004*** | 0.004*** |
| | (3.691) | (3.576) | (2.734) | (2.629) |
| Child's Age in Adulthood | | | 0.132** | 0.114** |
| | | | (2.542) | (2.377) |
| Observations | 1668 | 1668 | 1668 | 1668 |
| R-sqr | 0.201 | 0.204 | 0.209 | 0.210 |

t statistics in parentheses;
* p<.10, ** p<.05, *** p<.01
[1]: Interaction term between birth spacing (months) and being the older sibling
[2]: Interaction term between birth spacing (dummy) and being the older sibling
Parent is either mother or father who lives in the household

**Table 4.8: Years of Schooling- All Siblings Sample - Fixed Effect Models**

| Variables | Young Adulthood | | Adulthood | |
|---|---|---|---|---|
| | Coef./t | Coef./t | Coef./t | Coef./t |
| Birth Spacing (Months) | -0.007 | | 0.002 | |
| | (-0.417) | | (0.097) | |
| Birth Spacing*Older[1] | 0.013 | | 0.008 | |
| | (1.099) | | (0.637) | |
| Birth Spacing (Dummy) | | -0.307 | | -0.634 |
| | | (-1.186) | | (-1.429) |
| Dummy*Older[2] | | -0.150 | | 0.184 |
| | | (-0.667) | | (0.734) |
| Older Sibling | -0.214 | 0.052 | -0.133 | 0.001 |
| | (-0.805) | (0.270) | (-0.462) | (0.003) |
| Child's Age in Young Adulthood | -0.028 | 0.134 | | |
| | (-0.204) | (1.495) | | |
| Male | -0.403*** | -0.405*** | -0.667*** | -0.676*** |
| | (-3.071) | (-3.085) | (-4.470) | (-4.517) |
| Hispanic | -0.215 | -0.259 | -0.117 | -0.117 |
| | (-0.429) | (-0.524) | (-0.174) | (-0.176) |
| White | 0.048 | 0.039 | -0.066 | -0.062 |
| | (0.092) | (0.076) | (-0.105) | (-0.100) |
| Black | -0.559 | -0.542 | 0.230 | 0.237 |
| | (-0.869) | (-0.833) | (0.303) | (0.314) |
| Number of Siblings in HH | -0.026 | -0.031 | 0.016 | 0.016 |
| | (-0.649) | (-0.776) | (0.302) | (0.294) |
| Birth Order | -0.670*** | -0.472** | -0.574** | -0.497** |
| | (-2.718) | (-2.253) | (-2.247) | (-2.010) |
| Child's Age in Adulthood | | | -0.146 | -0.116 |
| | | | (-1.095) | (-1.329) |
| Observations | 1668 | 1668 | 1668 | 1668 |
| R-sqr | 0.132 | 0.132 | 0.053 | 0.055 |

t statistics in parentheses; Hausman Test: Prob>chi2 = 0.0000 & 0.0008

* p<.10, ** p<.05, *** p<.01

[1]: Interaction term between birth spacing (months) and being the older sibling

[2]: Interaction term between birth spacing (dummy) and being the older sibling

Parent is either mother or father who lives in the household

**Table 4.9: College Attendence and College Degree- All Siblings Sample - Logit Models**

| Variables | Young Adulthood | | Adulthood | |
|---|---|---|---|---|
| Birth Spacing (Months) | .0018 | | .0015 | |
| | (1.076) | | (1.039) | |
| Birth Spacing*Older[1] | -.0025 | | -.0012 | |
| | (-1.185) | | (-.5948) | |
| Birth Spacing (Dummy) (d) | | .1217*** | | .0898** |
| | | (2.754) | | (2.321) |
| Dummy*Older (d)[2] | | -.1215** | | -.0656 |
| | | (-2.229) | | (-1.432) |
| Older Sibling (d) | .0567 | .0459 | .0389 | .0432 |
| | (.7334) | (.6822) | (.5357) | (.6846) |
| Child's Age in Young Adulthood | .0346** | .0355*** | | |
| | (2.381) | (2.598) | | |
| Male (d) | -.0816** | -.0794** | -.0663** | -.0658** |
| | (-2.536) | (-2.468) | (-2.381) | (-2.364) |
| Hispanic (d) | -.0821 | -.0901 | .0014 | -.0035 |
| | (-.8975) | (-.98) | (.0174) | (-.0428) |
| White (d) | -.0308 | -.0394 | .0463 | .0402 |
| | (-.3706) | (-.4705) | (.6718) | (.5829) |
| Black (d) | .0341 | .0281 | .0089 | .0045 |
| | (.3657) | (.299) | (.112) | (.0566) |
| Number of Siblings in HH | -.0294*** | -.0293*** | -.034*** | -.0338*** |
| | (-3.066) | (-3.047) | (-3.68) | (-3.65) |
| Birth Order | -.0345 | -.0394 | -.0358 | -.037 |
| | (-.5892) | (-.6667) | (-.6754) | (-.6976) |
| Parent's Education | .0498*** | .0498*** | .0468*** | .0467*** |
| | (5.929) | (5.93) | (6.186) | (6.187) |
| Parent's Marriage Status (d) | .1712*** | .1719*** | .1222*** | .1214*** |
| | (3.857) | (3.866) | (3.544) | (3.519) |
| Parent's Age at Child's Birth | .0061* | .0059* | .0108*** | .0106*** |
| | (1.903) | (1.816) | (3.948) | (3.838) |
| Family Income in 1994 (thousands) | 9.3e-04* | 9.2e-04 | 9.5e-04*** | 9.3e-04** |
| | (1.685) | (1.637) | (2.599) | (2.497) |
| Child's Age in Adulthood | | | .0118 | .0121 |
| | | | (.9892) | (1.078) |
| Observations | 1668 | 1668 | 1668 | 1668 |

Marginal effects; $t$ statistics in parentheses
(d) for discrete change of dummy variable from 0 to 1
* $p<.10$, ** $p<.05$, *** $p<.01$
[1]: Interaction term between birth spacing (months) and being the older sibling
[2]: Interaction term between birth spacing (dummy) and being the older sibling
Parent is either mother or father who lives in the household

**Table 4.10: College Attendence and College Degree- All Siblings Sample - Fixed Effect Models**

| | Young Adulthood | | Adulthood | |
|---|---|---|---|---|
| Birth Spacing (Months) | .0125 | | .0116 | |
| | (1.21) | | (1.068) | |
| Birth Spacing*Older[1] | -.0136 | | -.0103 | |
| | (-1.059) | | (-.7119) | |
| Birth Spacing (Dummy) (d) | | .7816*** | | .7066** |
| | | (2.847) | | (2.424) |
| Dummy*Older (d)[2] | | -.731** | | -.5705 |
| | | (-2.204) | | (-1.601) |
| Older Sibling (d) | .4505 | .4243 | .391 | .3835 |
| | (1.074) | (1.216) | (.7758) | (.8938) |
| Child's Age in Young Adulthood | .1446 | .1637* | | |
| | (1.596) | (1.958) | | |
| Male (d) | -.4697** | -.4525** | -.48** | -.466** |
| | (-2.519) | (-2.433) | (-2.379) | (-2.304) |
| Hispanic (d) | -1.247** | -1.282** | -.873 | -.8949 |
| | (-2.43) | (-2.476) | (-1.542) | (-1.577) |
| White (d) | -.5457 | -.5878 | -.2031 | -.2392 |
| | (-1.172) | (-1.247) | (-.395) | (-.4631) |
| Black (d) | -.8255 | -.8546* | -1.036* | -1.054* |
| | (-1.614) | (-1.653) | (-1.846) | (-1.869) |
| Number of Siblings in HH | -.2296*** | -.2276*** | -.2794*** | -.2768*** |
| | (-4.064) | (-4.067) | (-3.615) | (-3.601) |
| Birth Order | -.1268 | -.1301 | -.1859 | -.1982 |
| | (-.4249) | (-.4349) | (-.5305) | (-.5592) |
| Child's Age in Adulthood | | | .0171 | .0296 |
| | | | (.1911) | (.3596) |
| Observations | 1668 | 1668 | 1668 | 1668 |

Marginal effects; $t$ statistics in parentheses
(d) for discrete change of dummy variable from 0 to 1
* p<.10, ** p<.05, *** p<.01
[1]: Interaction term between birth spacing (months) and being the older sibling
[2]: Interaction term between birth spacing (dummy) and being the older sibling
Parent is either mother or father who lives in the household

**Table 4.11: Annual Earnings in Adulthood - All Siblings Sample**

| Variables | OLS | | FE | |
|---|---|---|---|---|
| | Coef./t | Coef./t | Coef./t | Coef./t |
| Birth Spacing (Months) | -0.004 | | -0.003 | |
| | (-1.348) | | (-0.437) | |
| Birth Spacing*Older[1] | 0.005 | | 0.002 | |
| | (1.128) | | (0.402) | |
| Birth Spacing (Dummy) | | -0.010 | | -0.186 |
| | | (-0.131) | | (-0.878) |
| Dummy*Older[2] | | 0.048 | | -0.051 |
| | | (0.444) | | (-0.466) |
| Older Sibling | -0.152 | -0.067 | -0.107 | -0.043 |
| | (-1.148) | (-0.604) | (-0.681) | (-0.356) |
| Child's Age in Adulthood | 0.055** | 0.062*** | 0.079 | 0.114** |
| | (2.345) | (2.806) | (1.142) | (2.232) |
| Years of Schooling in Adulthood | 0.113*** | 0.112*** | 0.109*** | 0.109*** |
| | (8.340) | (8.234) | (4.780) | (4.772) |
| Male | 0.361*** | 0.361*** | 0.374*** | 0.374*** |
| | (6.401) | (6.387) | (4.319) | (4.306) |
| Hispanic | 0.084 | 0.077 | 0.434 | 0.430 |
| | (0.523) | (0.475) | (1.348) | (1.340) |
| White | 0.035 | 0.029 | 0.413 | 0.421 |
| | (0.240) | (0.197) | (1.281) | (1.305) |
| Black | -0.177 | -0.183 | 0.354 | 0.359 |
| | (-1.092) | (-1.126) | (0.985) | (1.003) |
| Number of Siblings in HH | -0.020 | -0.020 | -0.001 | -0.003 |
| | (-1.618) | (-1.586) | (-0.053) | (-0.102) |
| Birth Order | -0.040 | -0.045 | -0.001 | 0.050 |
| | (-0.367) | (-0.407) | (-0.004) | (0.329) |
| Parent's Education | 0.010 | 0.010 | 0.189 | 0.183 |
| | (0.773) | (0.782) | (1.273) | (1.216) |
| Parent's Marriage Status | 0.055 | 0.054 | | |
| | (0.789) | (0.775) | | |
| Parent's Age at Child's Birth | 0.005 | 0.005 | | |
| | (1.197) | (1.066) | | |
| Family Income in 1994 (thousands) | 0.000 | 0.000 | | |
| | (0.113) | (0.132) | | |
| Observations | 1668 | 1668 | 1668 | 1668 |
| R-sqr | 0.146 | 0.145 | 0.102 | 0.102 |

t statistics in parentheses; Hausman Test: Prob>chi2 = 0.0421

* p<.10, ** p<.05, *** p<.01

[1]: Interaction term between birth spacing (months) and being the older sibling

[2]: Interaction term between birth spacing (dummy) and being the older sibling

Parent is either mother or father who lives in the household

**Table 4.12: Smoking Behaviors in Adolescence - All Siblings Sample**

| Variables | Logit1 | FE1 | Logit2 | FE2 |
|---|---|---|---|---|
| | ME/t | ME/t | ME/t | ME/t |
| Birth Spacing (Months) | .002 | .0102 | | |
| | (.7815) | (.867) | | |
| Birth Spacing*Older[1] | -.004 | -.0129 | | |
| | (-1.161) | (-.8156) | | |
| Birth Spacing (Dummy) (d) | | | .0255 | .2732 |
| | | | (.3946) | (.9242) |
| Dummy*Older (d)[2] | | | -.0697 | -.2876 |
| | | | (-.776) | (-.7266) |
| Older Sibling (d) | -.0126 | -.0095 | -.0714 | -.1793 |
| | (-.1114) | (-.0157) | (-.7434) | (-.3458) |
| Child's Age in Adolescence | .0295 | .1388 | .0249 | .1262 |
| | (1.465) | (1.491) | (1.275) | (1.426) |
| Parent's Smoking Behavior (d) | .1161** | .6149*** | .1151** | .6087*** |
| | (2.442) | (2.838) | (2.42) | (2.816) |
| Male (d) | .0187 | .1431 | .0185 | .1438 |
| | (.3851) | (.72) | (.3842) | (.7241) |
| Hispanic (d) | -.0301 | -.3824 | -.0312 | -.389 |
| | (-.2301) | (-.7658) | (-.2407) | (-.7797) |
| White (d) | .1782 | .416 | .1736 | .4035 |
| | (1.508) | (.9274) | (1.484) | (.9044) |
| Black (d) | .0243 | -.3507 | .0212 | -.3685 |
| | (.194) | (-.668) | (.1694) | (-.7007) |
| Number of Siblings in HH | 8.8e-05 | -.0024 | 8.3e-05 | -5.2e-04 |
| | (.0078) | (-.0491) | (.0073) | (-.0109) |
| Birth Order | -.0803 | .1312 | -.0793 | .1323 |
| | (-.8865) | (.2923) | (-.8803) | (.2951) |
| Parent's Education | -.0057 | | -.0065 | |
| | (-.5399) | | (-.613) | |
| Parent's Marriage Status (d) | -.0458 | | -.0469 | |
| | (-.7948) | | (-.8124) | |
| Parent's Age at Child's Birth | -7.7e-04 | | -6.8e-04 | |
| | (-.194) | | (-.1721) | |
| Family Income in 1994 (thousands) | 1.9e-04 | | 2.1e-04 | |
| | (.6145) | | (.657) | |
| Observations | 535 | 535 | 535 | 535 |

ME/t: Maginal effects; $t$ statistics in parentheses
(d) for discrete change of dummy variable from 0 to 1
* p<.10, ** p<.05, *** p<.01
[1]: Interaction term between birth spacing (months) and being the older sibling
[2]: Interaction term between birth spacing (dummy) and being the older sibling
Parent is either mother or father who lives in the household

**Table 4.13: Years of Schooling - Biological Siblings Sample**

| Variables | OLS1 Coef./t | FE1 Coef./t | OLS2 Coef./t | FE2 Coef./t |
|---|---|---|---|---|
| Birth Spacing (Months) | 0.003 | -0.009 | | |
| | (0.280) | (-0.389) | | |
| Month*Older[1] | 0.009 | 0.019 | | |
| | (0.611) | (1.114) | | |
| Birth Spacing (Dummy) | | | 0.216 | -0.539 |
| | | | (1.073) | (-0.753) |
| Dummy*Older[2] | | | 0.251 | 0.312 |
| | | | (0.833) | (1.071) |
| Older Sibling | -0.152 | -0.213 | -0.037 | -0.007 |
| | (-0.375) | (-0.648) | (-0.161) | (-0.026) |
| Child's Age in Adulthood | 0.125** | -0.068 | 0.121** | -0.003 |
| | (2.104) | (-0.411) | (2.051) | (-0.028) |
| Male | -0.460** | -0.596** | -0.461** | -0.605** |
| | (-3.208) | (-3.172) | (-3.217) | (-3.175) |
| Hispanic | -0.280 | -1.092 | -0.293 | -1.039 |
| | (-0.711) | (-1.554) | (-0.745) | (-1.502) |
| White | -0.300 | -1.119 | -0.308 | -1.060 |
| | (-0.915) | (-1.633) | (-0.943) | (-1.608) |
| Black | 0.076 | -0.087 | 0.082 | -0.054 |
| | (0.198) | (-0.151) | (0.216) | (-0.096) |
| Number of Biological Siblings | 0.106 | 0.190 | 0.110 | 0.172 |
| | (1.501) | (0.781) | (1.580) | (0.767) |
| Birth Order | -0.220** | -0.269* | -0.220** | -0.253 |
| | (-2.225) | (-1.666) | (-2.247) | (-1.597) |
| Biological Parent's Education | 0.273*** | | 0.270*** | |
| | (7.024) | | (6.966) | |
| Biological Parent's Marriage Status | 0.697*** | | 0.704*** | |
| | (3.507) | | (3.561) | |
| Biologial Parent's Age at Child's Birth | 0.084*** | | 0.083*** | |
| | (4.381) | | (4.321) | |
| Family Income in 1994 (thousands) | 0.006*** | | 0.006*** | |
| | (6.470) | | (6.274) | |
| Observations | 848 | 848 | 848 | 848 |
| R-sqr | 0.235 | 0.063 | 0.239 | 0.064 |

t statistics in parentheses
* p<.10, ** p<.05, *** p<.01
[1]: Interaction term between birth spacing (months) and an indicator for the older sibling
[2]: Interaction term between birth spacing (dummy) and an indicator for the older sibling
Parent is either biological mother or biological father

**Table 4.14: Years of Schooling - Biological Brothers Sample**

| Variables | Young Adulthood | | | | Adulthood | | | |
|---|---|---|---|---|---|---|---|---|
| | OLS1 | FE1 | OLS2 | FE2 | OLS3 | FE3 | OLS4 | FE4 |
| | Coef./t | Coef./t | Coef./t | Coef./t | Coef./t | Coef./t | Coef./t | Coef./t |
| Birth Spacing (Months) | -0.017* | -0.006 | | | -0.019** | -0.008 | | |
| | (-1.866) | (-0.277) | | | (-2.007) | (-0.579) | | |
| Interaction Term (Months*Older) | 0.043** | 0.047 | | | 0.043** | 0.034** | | |
| | (2.532) | (1.558) | | | (2.510) | (2.006) | | |
| Birth Spacing (Dummy) | | | -0.250 | 0.638 | | | -0.294 | 0.087 |
| | | | (-0.818) | (1.566) | | | (-0.999) | (0.259) |
| Interaction Term (Dummy*Older) | | | 0.737 | 0.420 | | | 0.527 | 0.403 |
| | | | (1.439) | (0.732) | | | (1.113) | (1.077) |
| Older Sibling | -0.181 | -0.735 | 0.477 | -0.323 | -0.345 | -0.139 | 0.388 | |
| | (-0.380) | (-1.480) | (1.299) | (-0.873) | (-0.789) | (-0.324) | (1.203) | |
| Observations | 418 | 418 | 418 | 418 | 418 | 418 | 418 | 418 |
| R-sqr | 0.242 | 0.241 | 0.232 | 0.232 | 0.293 | 0.293 | 0.293 | 0.293 |

* $p<.10$, ** $p<.05$, *** $p<.01$

Parent is either biological mother or biological father who lives in the household

**Table 4.15: Years of Schooling - Biological Sisters Sample**

| Variables | Young Adulthood | | | | Adulthood | | | |
|---|---|---|---|---|---|---|---|---|
| | OLS1 | FE1 | OLS2 | FE2 | OLS3 | FE3 | OLS4 | FE4 |
| | Coef./t | Coef./t | Coef./t | Coef./t | Coef./t | Coef./t | Coef./t | Coef./t |
| Birth Spacing (Months) | 0.013 | -0.053* | | | 0.015 | 0.015 | | |
| | (1.359) | (-1.894) | | | (0.933) | (1.005) | | |
| Interaction Term (Months*Older) | -0.016 | 0.041** | | | -0.011 | -0.006 | | |
| | (-0.835) | (2.039) | | | (-0.487) | (-0.351) | | |
| Birth Spacing (Dummy) | | | 0.601** | 0.301*** | | | 0.238 | 0.143 |
| | | | (2.398) | (6.963) | | | (0.637) | (0.386) |
| Interaction Term (Dummy*Older) | | | -0.372 | 0.476 | | | -0.082 | -0.008 |
| | | | (-0.757) | (1.115) | | | (-0.150) | (-0.022) |
| Older Sibling | 0.197 | 0.077 | 0.012 | 0.456 | 0.064 | 0.038 | -0.170 | |
| | (0.364) | (0.167) | (0.032) | (1.158) | (0.095) | (0.074) | (-0.366) | |
| Observations | 898 | 898 | 898 | 898 | 898 | 898 | 898 | 898 |
| R-sqr | 0.320 | 0.290 | 0.329 | 0.290 | 0.239 | 0.239 | 0.237 | 0.239 |

* p<.10, ** p<.05, *** p<.01
Parent is either biological mother or biological father who lives in the household

**Table 4.16: Years of Schooling - Bioglogical Brother & Sister Sample**

| Variables | Young Adulthood | | | | Adulthood | | | |
|---|---|---|---|---|---|---|---|---|
| | OLS1 | FE1 | OLS2 | FE2 | OLS3 | FE3 | OLS4 | FE4 |
| | Coef./t | Coef./t | Coef./t | Coef./t | Coef./t | Coef./t | Coef./t | Coef./t |
| Birth Spacing (Months) | -0.001 | -0.031 | | | -0.001 | 0.005 | | |
| | (-0.101) | (-1.523) | | | (-0.111) | (0.464) | | |
| Interaction Term (Months*Older) | 0.013 | 0.045** | | | 0.014 | 0.012 | | |
| | (1.043) | (2.531) | | | (0.985) | (0.975) | | |
| Birth Spacing (Dummy) | | | 0.220 | -0.527 | | | 0.034 | 0.148 |
| | | | (1.123) | (-0.554) | | | (0.143) | (0.586) |
| Interaction Term( Dummy*Older) | | | 0.141 | 0.435 | | | 0.170 | 0.134 |
| | | | (0.402) | (1.283) | | | (0.478) | (0.516) |
| Older Sibling | -0.083 | -0.012 | 0.177 | -0.041 | -0.198 | -0.076 | 0.073 | |
| | (-0.233) | (-1.183) | (0.679) | (-0.146) | (-0.505) | (-0.231) | (0.272) | |
| Observations | 352 | 352 | 352 | 352 | 352 | 352 | 352 | 352 |
| R-sqr | 0.251 | 0.244 | 0.253 | 0.232 | 0.224 | | 0.223 | |

* p<.10, ** p<.05, *** p<.01

Parent is either biological mother or biological father who lives in the household

**Table 4.17: College Attendence - Biological Sisters Sample**

| Variables | Young Adulthood | | | | Adulthood | | | |
|---|---|---|---|---|---|---|---|---|
| | Logit1 | FE1 | Logit2 | FE2 | Logit3 | FE3 | Logit4 | FE4 |
| main | | | | | | | | |
| Birth Spacing (Months) | .0048 | .0287 | | | .001 | .0221 | | |
| | (1.09) | (1.155) | | | (.297) | (.7823) | | |
| Interaction Term (Months*Older) | -8.2e-05 | -.004 | | | .002 | 5.5e-04 | | -2.8e-04 |
| | (-.0129) | (-.1179) | | | (.3912) | (.0156) | | (-.0088) |
| Birth Spacing (Dummy) (d) | | | .2454*** | 1.301** | | | .1391 | .8558 |
| | | | (2.666) | (2.07) | | | (1.564) | (1.27) |
| Interaction Term (Dummy*Older) (d) | | | -.006 | .1137 | | | -.0683 | |
| | | | (-.0413) | (.1503) | | | (-.5279) | |
| Older Sibling (d) | -.0221 | .3856 | -.0116 | .2897 | -.1366 | -.5181 | -.0443 | -.4795 |
| | (-.119) | (.4108) | (-.0944) | (.4326) | (-.8891) | (-.5082) | (-.3726) | (-.4921) |
| Observations | 898 | 898 | 898 | 898 | 898 | 898 | 898 | 898 |

Marginal effects; $t$ statistics in parentheses
(d) for discrete change of dummy variable from 0 to 1
* $p<.10$, ** $p<.05$, *** $p<.01$
Parent is either biological mother or biological father who lives in the household

**Table 4.18: Robustness Check - Biological Older Sibling's Smoking Behavior**

| | Logit 1 | Logit 2 | Logit 3 | Logit 4 | Logit 5 | Logit 6 |
|---|---|---|---|---|---|---|
| Birth Spacing (Months) | -.0011 | -.0018 | -.0017 | -.0023 | -.0043 | -.0045 |
| | (-.2597) | (-.4261) | (-.3922) | (-.533) | (-.8869) | (-.9341) |
| Younger Sibling's Smoking Behavior (d) | .1905** | .1827** | .1792** | .1797** | .1672* | .1708* |
| | (2.169) | (2.038) | (1.987) | (1.993) | (1.83) | (1.865) |
| Mother's Smoking Behavior (d) | | .1673* | .158* | .1781* | .1779* | .1762* |
| | | (1.937) | (1.699) | (1.878) | (1.869) | (1.844) |
| Mother's Education | | | -.0102 | -.0027 | .0011 | -1.5e-05 |
| | | | (-.5121) | (-.1276) | (.0516) | (-6.8e-04) |
| Family Income in 1994 (thousands) | | | 3.4e-04 | 4.7e-04 | 4.4e-04 | 4.7e-04 |
| | | | (.5485) | (.7026) | (.6387) | (.6602) |
| Older Sibling's Education (Test Score) | | | | -.0045 | -.0049 | -.0049 |
| | | | | (-1.076) | (-1.157) | (-1.157) |
| Child's Age in Adolescence | | | | | .0395 | .0394 |
| | | | | | (.9309) | (.9275) |
| Black (d) | | | | | | .144 |
| | | | | | | (.7592) |
| Observations | 126 | 126 | 126 | 126 | 126 | 126 |

Marginal effects; $t$ statistics in parentheses
(d) for discrete change of dummy variable from 0 to 1
* $p<.10$, ** $p<.05$, *** $p<.01$
Parent is either biological mother or biological father who lives in the household

**Table 4.19: Robustness Check - Biological Younger Sibling's Smoking Behavior**

| | Logit 1 | Logit 2 | Logit 3 | Logit 4 | Logit 5 | Logit 6 |
|---|---|---|---|---|---|---|
| Birth Spacing (Months) | -8.5e-04 | -3.7e-04 | -4.7e-04 | -4.2e-04 | .0019 | .0022 |
| | (-.1985) | (-.0831) | (-.1054) | (-.094) | (.3996) | (.4687) |
| Smoking Behavior of Older Sibling (d) | .1943** | .191** | .1883** | .1911** | .186* | .1889* |
| | (2.172) | (2.043) | (2.009) | (2.013) | (1.933) | (1.957) |
| Mother's Smoking Behavior (d) | | .0985 | .1217 | .1197 | .1186 | .1146 |
| | | (1.099) | (1.267) | (1.237) | (1.21) | (1.161) |
| Mother's Education | | | .0045 | .0031 | .0088 | .0094 |
| | | | (.221) | (.1412) | (.3917) | (.418) |
| Family Income in 1994 (thousands) | | | 7.3e-04 | 6.9e-04 | 6.3e-04 | 5.8e-04 |
| | | | (.746) | (.7047) | (.6225) | (.59) |
| Younger Sibling's Education (Test Score) | | | | 8.0e-04 | 2.1e-04 | 2.5e-04 |
| | | | | (.1888) | (.0483) | (.0593) |
| Child's Age in Adolescence | | | | | .0734* | .0703 |
| | | | | | (1.721) | (1.642) |
| Black (d) | | | | | | -.2363 |
| | | | | | | (-.7724) |
| Observations | 126 | 126 | 126 | 126 | 126 | 126 |

Marginal effects; $t$ statistics in parentheses
(d) for discrete change of dummy variable from 0 to 1
* $p<.10$, ** $p<.05$, *** $p<.01$
Parent is either biological mother or biological father who lives in the household

# References

An, C.-B., Haveman, R. and Wolfe, B. (1993), Teen out-of-wedlock births and welfare receipt: The role of childhood events and economic circumstances. *Review of Economics and Statistics*, 75, 195-208.

Ashcraft, A. and Lang, K. (2006). The Consequences of Teenage Childbearing. Working Paper

Augood, C., Duckitt, K., & Templeton, A. A. (1998). Smoking and female infertility: a systematic review and meta-analysis. *Human Reproduction*, 13(6), 1532-1539.

Astone and McLanahan. (1991). Family Structure, Parental Practices and High School Completion. *American Sociological Review*, Vol. 56, No. 3, pp. 309-320

Alison Aughinbaugh, Alison. (2004). The Impact of Attrition on the Children of the NLSY79. *Journal of Human Resources*, Vol. 39, No. 2, pp. 536-563

Barro, R. and Becker, G. (1989). Fertility Choice in a Model of Economic Growth. *Econometrica*, Vol. 57, No. 2 (Mar., 1989), pp. 481-501

Baydar, N., Greek, A., and Brooks-Gunns, J. (1997). A Longitudinal Study of the Effects of the Birth of a Sibling during the First 6 Years of Life. *Journal of Marriage and Family*, Vol. 59, No. 4 (Nov., 1997), pp. 939-956

Baydar, N., Hyle, P., and Brooks-Gunn, J. (1997). A Longitudinal Study of the Effects of the Birth of a Sibling During Preschool and Early Grade School Years. *Journal of Marriage and the Family*, Vol. 59, No. 4 (Nov., 1997), pp. 957-965.

Becker, G. (1960). An Economic Analysis of Fertility. Demographic and Economic Change in Developed Countries, Columbia University Press (p. 209 - 240)

Becke, G., Landes, E. and Michael, R. (1977). An Economic Analysis of Marital Instability", *Journal of Political Economy*, Vol. 85, No. 6 (Dec., 1977), pp. 1141-1187

Behrman, J. and Taubman, P. (1986). Birth Order, Schooling, and Earnings. *Journal of Labor Economics*, Vol. 4, No. 3, Part 2: The Family and the Distribution of Economic Rewards (Jul., 1986), pp. S121-S145

Bane, M. J., & Ellwood, D. T. (1986), Slipping into and out of Poverty: the Dynamics of Spells in, *Journal of Human Ressources*

Bhalotra, S. and Soes, A. (2008). Birth Spacing, Fertility and Neonatal, Mortality in India: Dynamics, Frailty and Fecundity. *Journal of Econometrics*, April 2008, v. 143, iss. 2, pp. 274-90

Bhargava, A. (2003). Family planning, gender differences and infant mortality: evidence from Uttar Pradesh, *India. Analysis of Data on Health*, Volume 112, Issue 1, January 2003, Pages 225–240

Black, S., Devereux, P., and Salvanes, K. (2005). The More the Merrier? The Effect of Family Size and Birth Order on Children's Education. *The Quarterly Journal of Economics* , Vol. 120, No. 2 (May, 2005), pp. 669-700

Broman, S., Nichols P., and Kennedy, W. (1975). Preschool IQ: Prenatal and Early Developmental Correlates. Hillsdale, N.J.: L. Erlbaum Associates.

Brownstein, N., Kalsbeek, W. D., Tabor, J., Entzel, P., Daza, E., & Mullan Harris, K. (2010). Non-response in Wave IV of the National Longitudinal Study of Adolescent Health. http://www.cpc.unc.edu/projects/addhealth/data/guides/W4_nonresponse.pdf

Bronars, Stephen G. and Jeffrey Grogger (December 1994): "The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment," American Economic Review, 84: 1141-56.

Buckles, K. and Munnich, E. (2012). Birth Spacing and Sibling Outcomes. *Journal of Human Resources*, Vol. 47, No. 3, pp. 613-642

Card, J.J. and L. Wise (1978), Teenage Mothers and Teenage Fathers: The Impact of Early Childbearing on Parents' Personal and Professional Lives, *Family Planning Perspectives*, 10:100-205.

Cardwell, C., Svensson, J., Waldhoer, T., Ludvigsson, J., Sadauskaitė-Kuehne, V., Roberts, C., Parslow, R., Wadsworth, E., Brigis, G., Urbonaitė, B., Schober, E., Devoti, G., Ionescu-Tirgoviste, C., Beaufort, C., Soltesz, G., Patterson, C. (2012). Interbirth Interval Is Associated With Childhood Type 1 Diabetes Risk. *Diabetes*, 61(3): 702–707

Chantala, Kim. 2001. Constructing Weights to Use in Analyzing Pairs of Individuals from Add Health Data, working paper.

Chantala K, Kalsbeek WD, Andraca E. Nonresponse in wave III of the Add Health Study, 2004. http://www.cpc.unc.edu/projects/addhealth/files/W3nonres.pdf. Accessed June 17, 2005.

Cheslack-Postava, Liu, and Bearman. (2011). Closely Spaced Pregnancies Are Associated With Increased Odds of Autism in California Sibling. *Pediatrics*, Vol. 127 No. 2 February 1, 2011, pp. 246 -253 (doi: 10.1542/peds.2010-2371)

Cleland, J. and Sathar, Z. (1984). The Effect of birth spacing on childhood mortality in Pakistan. Population Studies. *Journal of Demography*, Volume 38, Issue 3, 1984

Curtis, K. M., Savitz, D. A., & Arbuckle, T. E. (1997). Effects of cigarette smoking, caffeine consumption, and alcohol intake on fecundability. American journal of epidemiology, 146(1), 32-41.

DaVanzo, J., Lauren, H., Abdur, R., and Mizanur, R. (2008). The Effects of Pregnancy Spacing on Infant and Child Mortality in Matlab, Bangladesh: How They Vary by the Type of Pregnancy Outcome That Began the Interval. *Population Studies*, July 2008, v. 62, pp. 131-54

Fletcher, J. (2008). Adolescent Depression: Diagnosis, Treatment, and Educational Attainment. *Health Economics*, 17, 1215-1235.

Fletcher, Jason B. and Barbara L. Wolfe (2009), Education and Labor Market Consequences of Teenage Childbearing: Evidence Using the Timing of Outcomes and Community Fixed Effects. *Journal of Human Resources*, 44: 303-25.

Fletcher, Jason (2014), The Effects of Childhood ADHD on Adult Labor Market Outcomes. *Health Economics*, 23: 159-181

Ford, J. H., & MacCormac, L. (1995). Pregnancy and lifestyle study: the long-term use of the contraceptive pill and the risk of age-related miscarriage. Human Reproduction, 10(6), 1397-1402.

Geronimus Arline and Sanders Korenman (November 1992), The Socioeconomic Consequences of Teen Childbearing Reconsidered. *Quarterly Journal of Economics*, 107: 1187-1214.

Geronimus, Arline and Sanders Korenman. (1993), The socioeconomic costs of teenage childbearing: Evidence and interpretation, *Demography,* 30.2 (1993): 281-290.

Geronimus, A. T., Korenman, S., & Hillemeier, M. M. (1994), Does young maternal age adversely affect child development? Evidence from cousin comparisons in the United States, *Population and Development Review*, 585-609.

Galbraith, R. (1982). Sibling spacing and intellectual development: A closer look at the confluence models. *Developmental Psychology*, Vol 18(2), Mar 1982, 151-173. doi: 10.1037/0012-1649.18.2.151

Gribble, J. (1993). Birth Intervals, Gestational Age, and Low Birth Weight: Are the Relationships Confounded?. *Population Studies*, March 1993, v. 47, iss. 1, pp. 133-46

Grogger, J. and S. Bronars (1993), The Socioeconomic Consequences of Teenage Childbearing: Results from a Natural Experiment, Unpublished manuscript, University of California- Santa Barbara.

Hao, L., V. J. Hotz, and G. Z. Jin (1999), Games Daughters and Parents Play: Teenage Childbearing, Parental Reputation, and Strategic Transfers, Unpublished manuscript, UCLA, July 1999.

Hausman, J. (1978). Specification Tests in Econometrics. *Econometrica*, Vol. 46, No. 6 (Nov., 1978), pp.1251-1271

Hanushek, E. (1992). The Trade-off between Child Quantity and Quality. *Journal of Political Economy.* Vol. 100, No. 1 (Feb., 1992), pp. 84-117.

Hassan M. and Killick S. (2004), Negative lifestyle is associated with a significant reduction in fecundity. *Fertility and Sterility*, 81 (2), pp. 384-392.

Haveman, R. and Wolfe, B. (1995). The Determinants of Children's Attainments: A Review of Methods and Findings. *Journal of Economic Literature*, Vol. 33, No. 4 (Dec., 1995), pp. 1829-1878

Hayes, C. (1987), Risking the Future: Adolescent Sexuality, Pregnancy, and Childbearing. Vol. I, Washington, DC: National Academy Press.

Heckman, J. and Walker, J. (1990). The Relationship Between Wages and Income and the Timing and Spacing of Births: Evidence from Swedish Longitudinal Data. *Econometrica*, Vol. 58, No. 6 (Nov., 1990), pp. 1411-1441

Hoffman SD (2006), By the Numbers: The Public Costs of Adolescent Childbearing, Washington, DC: National Campaign to Prevent Teen Pregnancy.

Hoffman, S. (2008), Updated Estimates of the Consequences of Teen Childbearing for Mother, Kids Having Kids, The Urban Institute Press, Second Edition.

Hoffman, S., M. Foster, and F. Furstenberg (1993), Reevaluating the Costs of Teenage Childbearing. *Demography*, 30:1-13.

Holmlund, Helena (Summer 2005), Estimating Long-Term Consequences of Teenage Childbearing. *Journal of Human Resources*, 40: 716-43.

Hotz, V. Joseph, Charles H. Mullin, Seth G. Sanders (1997), Bounding Causal Effects Using Data from a Contaminated Natural Experiment: Analyzing the Effects of Teenage Childbearing. *Review of Economic Studies*, 64: 575-603.

Hotz, V. Joseph, Susan Williams McElroy, and Seth G. Sanders (2005), Teenage Childbearing and Its Life Cycle Consequences. *Journal of Human Resources*, 40: 683-715.

Imbens, G. (2004), Semiparametric Estimation of Average Treatment Effects Under Exogeneity: A Review. *Review of Economics and Statistics*, 86, 4-29.

Jenkins, J. and Astington, J. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology*, 32, 70–78.

Kahn, J. R., & Anderson, K. E. (1992). Intergenerational patterns of teenage fertility, *Demography*, 29(1), 39-57.

Klepinger, D., Lundberg, S., & Plotnick, R. (1999). How does adolescent fertility affect the human capital and wages of young women?, *Journal of Human Resources*, 421-448.

Kost K, Henshaw S., and Carlin L (2010), U.S. Teenage Pregnancies, Births and Abortions: National and State Trends and Trends by Race and Ethnicity.

Kost K. and Henshaw S. (2012). U.S. Teenage Pregnancies, Births and Abortions, 2008: National Trends by Age, Race and Ethnicity, New York: Guttmacher Institute.

Lee, D. (2010), The early socioeconomic effects of teenage childbearing: A propensity score matching approach, *Demography Research*, 23, 697–736.

Marini, M. (1984), Women's Educational Attainment and the Timing of Entry into Parenthood, *American Sociological Review*, 49: 491-511.

McCall, R. (1984). Developmental Changes in Mental Performance: The Effect of the Birth of a Sibling, *Child Development*, Vol. 55, No. 4, pp. 1317-1321

McCauley, A. P., and C. Salter. (1995). October 1995. Meeting the needs of young adults, *Population Reports*.

Newman, J. (1983). Economic Analyses of the Spacing of Births. *The American Economic Review*, Vol. 73, No. 2, Papers and Proceedings of the Ninety-Fifth Annual Meeting of the American Economic Association (May, 1983), pp. 33-37

Nguyen, N. (2013). Teenage Childbearing and Socioeconomic Outcomes of Teen Mothers in Young Adulthoold and Adulthood using NLSY79 Data, working paper

Price, J. (2008). Parent-Child Quality Time: Does Birth Order Matter? *Human Resources*, Vol. 43, No. 1, pp. 240-265

Powell, B. and Steelman, L. (1993). The Educational Benefits of Being Spaced Out: Sibship Density and Educational Progress. *American Sociological Review*, Vol. 58, No. 3 (Jun., 1993), pp. 367-381
Powell, B. and Steelman, L. (1995). Feeling the Pinch: Child Spacing and Constraints on Parental Economic Investments in Children. *Social Forces*, Vol. 73, No. 4 (Jun., 1995), pp. 1465-1486

Rawlings, S., Rawlings, B., and Read. (1995). "Prevalence of Low Birth Weight and Preterm Delivery in Relation to the Interval between Pregnancies among White and Black Women", *Journal of Medicine*, 332:69-74

Ribar, D. C. (1994), Teenage fertility and high school completion, *The Review of Economics and Statistics*, 413-424.

Rosenbaum, P. and D. Rubin (1983), The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70, 41-55.

Rosenzweig, M. and Schultz, P. (1983). Consumer Demand and Household Production: The Relationship Between Fertility and Child Mortality. *The American Economic Review*, Vol. 73, No. 2, Papers and Proceedings of the Ninety-Fifth Annual Meeting of the American Economic Association (May, 1983), pp. 38-42

Rosenzweig, M. and Wolpin, K. (1986). Evaluating the Effects of Optimally Distributed Public Programs: Child Health and Family Planning Interventions. *The American Economic Review*, Vol. 76, No. 3, pp. 470-482

Rosenzweig, M. and Wolpin, K. (1988). Heterogeneity, Intrafamily Distribution, and Child Health. *The Journal of Human Resources*, Vol. 23, No. 4 (Autumn, 1988), pp. 437-461
Seth Sanders, Jeffrey Smith and Ye Zhang (2007). Teenage Childbearing and Maternal Schooling Outcomes: Evidence from Matching. Working paper.

Stewart, R., Mobley, L., Van Tuyl, S., and Salvador, M. (1987). The Firstborn's Adjustment to the Birth of a Sibling: A Longitudinal Assessment. *Child Development*, Vol. 58, No. 2, pp. 341-355.

Sullivan, Jeremiah M., Shea Oscar Rutstein, and George T. Bicego. 1994. Infant and child mortality. Demographic and Health Surveys, Comparative Studies, No. 15. Calverton, Maryland: Macro International.

Thomson, E. and McLanahan, S. (2012). Reflections on Family Structure and Child Well-Being: Economic Resources vs. Parental Socialization. *Social Forces* , Vol. 91, No. 1, pp. 45-53

Thompson, Owen. (2014). Economic Background and Educational Attainment: The Role of Gene-Environment Interactions (2014). *Journal of Human Resources*. 49(2), 263-294.

Upchurch, D. and J. McCarthy (1990), The Timing of a First Birth and High School Completion, *American Sociological Review*, 55, 224-234.

Waite, L. and K. Moore (1978), The Impact of an Early First Birth on Young Women's Educational Attainment, *Social Forces*, 56, 845-865

Zajonc, R. (1976). Family configuration and intelligence: Variations in scholastic aptitude scores parallel trends in family size and the spacing of children. *Science*, Vol 192 (4236), Apr 1976, 227-236