# An Empirical Economic Study on

# Human Capital Investment Behaviours in Uganda

A Dissertation

Submitted to the National Graduate Institute for Policy Studies (GRIPS)

in Partial Fulfillment of the Requirements for the Degree of

Ph.D. in Development Economics

by

Nagashima Masaru

March 2020

# Abstract

In this dissertation, I attempt to understand factors that affect, as well as are affected by, human capital investment behaviours in developing countries.

As a consequence of human capital investment, I analyse the relationship between female education and brideprice payment practice using data from Uganda. Education is arguably one of the important components of human capital, which can increase productivity and welfare. Brideprice is a wealth transfer at marriage sent from the groom to the parents of the bride. For the analysis, I employ a novel approach referred to as the Regression Kink Design and use the introduction of free primary education as the source of exogenous variation. I find that an increase in female education, induced by a reduction in the cost of education, led to a decline in the share of females whose marriage involved brideprice payment. With results from additional analyses and extensive exploration of related literature, I hypothesize the change in cultural practice as stemming from a potential trade-off faced by altruistic parents of the bride between the instantaneous utility from immediate transfer of brideprice at marriage and the future utility from a better marital life of their daughter. This finding contributes to the growing literature on culture and institutions, and has a large implication to policies pertaining to education and its consequences.

As a cause of human capital investment, I analyse the determinants of birth spacing behaviours that improve various aspects of human capital accumulation of mothers and children such as health and education, and eventually long-run outcomes such as marriage quality. In countries with high fertility rates, birth spacing tends to be short, which may be one of the factors that suppress human capital stock in developing countries. Using two data sets from Uganda, I show that birth intervals are shorter than recommended by the World Health Organization, and also that the intervals are further shortened if females experience a pregnancy loss due to miscarriage and stillbirth. The estimated impacts are comparable to the negative effect of infant mortality on the birth spacing for the pregnancy that immediately follows the infant death, reported by other studies in economic literature. My additional results suggest that an actual experience of pregnancy loss leads females to revise their perceived probability of pregnancy loss, based on which they adapt their subsequent reproductive schedule. This study contributes a new finding to the economics literature on reproductive behaviours, as well as suggestive evidence of the mechanism for the behavioural change that has not been comprehensively investigated before.

Based on these studies, I discuss the possible implications for further studies in economics and related policies in developing countries.

# Acknowledgements

First and foremost, I would like to express my sincere gratitude to Professor Chikako Yamauchi at the National Graduate Institute for Policy Studies (GRIPS), who has always encouraged me to continue my study ever since I first made an appointment back in April of the year of 2015. Space limitation does not allow me to list up when I was saved by her support and how, such as when I developed the first research idea, when I made a questionnaire to collect data from the field, and the like. She has provided me with such extensive support in every single aspect of my research that I otherwise could not have come all the way to the point of writing and finally submitting this dissertation.

I am also grateful to Professors Alistair Munro (GRIPS), Stephan Litschig (GRIPS), and Yoshito Takasaki (The University of Tokyo) who accepted to join my doctoral committee without a moment's hesitation. In particular, Professor Munro has once taken my supervisorship for a year and constantly supported my research whenever I was in need of help. Professor Litschig has commented constructively on my studies which has helped me substantially improve the analyses. I have benefited so much from my interactions with them and their instructive comments.

My appreciation extends to Professors Yosuke Yasuda, Makoto Hasegawa, Tetsushi Sonobe, Minchung Hsu, Julen Esteban-Pretel, Shinsuke Ikeda, Yoichi Arai, Roberto Leon-

Gonzalez, Tomoya Matsumoto, Keijiro Otsuka, Jonna P. Estudillo, and Hitoshi Ieda, who taught me the courses at GRIPS and formed the academic basis for my Ph.D. work.

I am also indebted to Faith Masekesa, Le Thu Hang, Le Thu Trang, Opoku Kwadwo, Francisco M. P. Mugizi, Opoku Kwadwo, Ekapirak Tanyasorn, Doyananda Debnath, Nguyen Kim Lan, and Stefan Angrick all whom I studied together with in my batch, encouraging each other at all times. Some of them have already acquired Ph.D., others have left school, and the others are still trying their best in pursuit of the degree. If I had missed any one of them in my Ph.D. life, I would not have survived to date. I also thank staff, and ex-staff, at GRIPS who has always helped me with any academic, as well as administrative, problems. Their just being there meant so much to me every single moment of my life at school.

I also acknowledge the financial support by Japan Student Service Organization and Japan Society for Promotion of Science (JSPS). In particular, studies in this dissertation were supported by JSPS Research Fellowship Grant No. 18J11688.

Last but not least, my special thanks reach out to my parents, Osamu and Keiko Nagashima, as well as my sister, Megumi Nagashima, for their kindness and generosity.

# Dedication

To my parents, Osamu and Keiko Nagashima.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# INTRODUCTION

## 1.1 Human Capital Investment Behaviours in Developing Countries

Importance of human capital in developing countries cannot be overstated: at the macro level, an increase in human capital can *ceteris paribus* boost production and accelerate economic growth; at the micro level, it implies higher productivity of an economic agent, which can have positive welfare consequences. The returns to human capital investment are large particularly in the developing world, for females, in private sector, with lower absolute level of education (Montenegro and Patrinos, 2014; Psacharopoulos and Patrinos, 2018). Human capital can not only benefit the one who invest and accumulate it but also others, which is known as externality (Moretti, 2004; Wantchekon et al., 2015). These thoughts have formed the basis of economic research on human capital investment and policy interventions, and by now have created a large body of literature that is still far from seeing the end of expansion.

Since human capital is accumulated through investment, it is crucial to understand what affects the decision to invest in human capital, let alone whether or not the decision responds to a change in economic incentives. At the same time, it is worth questioning what other variables co-vary when such investment decisions change the stock of human capital. In other words, more comprehensive understanding of economics of human capital accumulation necessitates investigation into the causes and consequences of the investment behaviour in many more life aspects.

This dissertation is an attempt to understand human capital investment behaviours in the developing world. More specifically, the ultimate objective of this dissertation is to showcase causes and consequences of human capital investment behaviours. For this purpose, I conduct two case studies: an analysis of the impact of female education on brideprice culture for the consequence part, and an analysis of the impact of miscarriage and stillbirth on birth spacing behaviours. For analysis, I consider two major components of human capital, education and health.

Chapter 2 presents a study of a consequence of human capital investment, where I analyse the relationship between female education and brideprice practice. Brideprice refers to a wealth transfer at marriage from the groom to the parents of the bride. In the past, it used to be observed in many parts of the world, including the contemporaneously developed countries. In Sub-Saharan Africa, it is still widely practiced though gradually diminishing. It is not well understood how and why the cultural practice has declined (Anderson, 2007). My analysis, therefore, is an investigation whether the cultural change is a consequence of human capital accumulation, which is one of the variables that dramatically change over the course of economic development. By so doing, it contributes a new case study to the growing literature on culture and institution (Alesina and Giuliano, 2015). In other words, it is an attempt to figure out factors that have made, and are making, brideprice practices disappear in the past developed, and contemporary developing, worlds, and thereby provide a new perspective on how culture evolves in relation to human capital investment during economic development.

Chapter 3 presents a study on a cause of human capital investment, where I examine determinants of birth spacing intervals. As appropriate birth spacing has been shown to benefit the health of mothers and new-borns, it can be considered health investment. In addition, short spacing intervals may be one of the reasons why the stock of human capital is low in Sub-Sahara Africa, since birth intervals are shown to increase birth weight of new-born babies, which help accumulate education and health capital at older ages. Despite its importance, however, little has been studied as to what factors affect the decisions of parents on birth spacing behaviours, which is the research gap that my analysis attempts to fill. One study by Bhalotra and van Soest (2008) shows that infant mortality (death of a child within a month of birth) negatively affects birth spacing for the next birth, but it excludes from its analysis miscarriage and stillbirth, which may also affect birth spacing behaviours. Inspired by this paper, I study whether the experience of miscarriage and stillbirth affects birth spacing behaviours as does infant mortality. I also seek to understand the underlying mechanism that governs the birth spacing decisions which seems to be missing from the previous study, since in order to infer implications for public policies, the knowledge on the mechanism for the behavioural change is indispensable: otherwise, resources may not be efficiently utilised due to improper targeting or inappropriate programme design.

## 1.2 Organization of This Dissertation

This dissertation is composed of two chapters that present detailed discussion on two themes. In Chapter 2, the relationship between female education and brideprice practice is analysed using data from Uganda. I summarise the existing literature and clarify the research gap that the analyses attempts to fill. Then I describe the data and the econometric model, and discuss the results and the preferred interpretation of the results.

In Chapter 3, the focus is put on birth spacing behaviours in relation to pregnancy loss due to miscarriage and stillbirth. I review the related medical and economic literature, and show the responses to such experience in terms of birth spacing using data from Uganda. Then I discuss the potential mechanism pertaining to belief updating, and present suggestive evidence on it.

In Chapter 4, I conclude with discussion on the implications from two chapters to the economic literature on human capital investment in developing countries.

# Chapter 2

# Female Education and Brideprice

## 2.1 Introduction

Brideprice is a marital wealth transfer sent from the groom to the parents of the bride.[1] Brideprice used to be practiced in many parts of the world, including contemporaneously high-income countries, where the amount of brideprice payment is reported to have been so significant that it sometimes represented a large burden, particularly for poor households (Anderson, 2007). Although the practice has been largely discontinued in such countries, it is still prevalent in many parts of Africa. Anecdotes indicate that brideprice serves as a token of gratitude to the bride's parents for their efforts in raising their daughter. Observationally, brideprice is practiced more frequently in areaswith high virilocality—the bride leaves her natal family and joins the groom's family upon marriage— polygyny—a

[1] Anecdotes suggest that there are cases in which the groom's parents help him with this wealth transfer. There are also cases where this wealth transfer is only agreed upon at the establishment of the marriage, and the actual payment is made later. In this particular study, we use data from Uganda, where it is usually only the groom who pays and the transfer can be made later. We discuss these particular contexts in more detail in the subsequent sections.

man marries more than one woman—and high female engagement in household agriculture (Anderson, 2007; Goody, 2011).

The economic literature on brideprice suggests that it compensates the bride's family for her labour income that she would earn and contribute to her family if she were not married (Becker, 1991; Anderson, 2007). In this framework, brideprice positively reflects female education to the extent that human capital matters for her foregone economic activities.[2] Recent studies have also considered brideprice practice as an indicator of greater female agricultural engagement after marriage (e.g., Jacoby 1995). Partly due to this economic contribution, females from ethnicities that have historically practiced brideprice are found to be less likely to suffer from domestic violence (Alesina et al., 2016). When the costs of primary education are lowered, only parents from the ethnic groups practicing brideprice are shown to invest in female education (Ashraf et al., 2019). In addition to the human capital hypothesis, upon which these studies are based, assortative matching in the marriage market also implies a positive association between brideprice and female education. That is, educated women are likely to marry educated men who are likely from wealthy families.

This stream of the literature takes brideprice culture for granted and, within a given cultural framework, analyses the relationship between the human capital accumulation of females and brideprice payment. The positive association between these two variables implies that brideprice practice may intensify over the course of economic development in a country, which is typically accompanied by the enhancement of female education.

[2] The positive association between brideprice and female human capital is also posited in social sciences other than economics (e.g., Bell 1998; Goody 2011).

However, history suggests the opposite: brideprice culture has *disappeared* in the course of economic development. Little is known about how brideprice culture has diminished and how it is related to the process of economic development, a point that is clearly raised by Anderson (2007).

Other studies have examined the strategic behaviours of the husband, the bride, and the bride's parents bargaining over the brideprice payment. For example, the existence of a brideprice payment at marriage is found to be correlated with an increased probability of divorce (Platteau and Gaspart, 2007). Gaspart and Platteau (2010) models and empirically tests the possibility that the parents of the bride may receive a lower brideprice if they worry about the chance of their daughter's divorce and the resulting loss of altruistic utility. Although these analyses consider bargaining over the *amount of brideprice*, it is possible that these strategic motives lead to the daughter's parents not receiving brideprice at all. We thus intend to empirically analyse when such a cultural decline occurs and whether it is related to an increase in female education.

Specifically, we estimate the effect of female education on the cultural practice of brideprice payment, exploiting the introduction of universal primary education (UPE) in Uganda. This reform abolished school fees for all pupils who were enrolled in primary schools in and after 1997. The reform was announced in December 1996, initiated in January 1997, and thus almost surely unpredictable by Ugandan families. Since it affected everyone as long as they were in school at the time of the introduction, the benefits of the reform were lower for older cohorts and greater for younger cohorts. As such, the increased years of education shows a clear kink when shown in a graph on the year of birth.

This exogenous source of variation is used in our fuzzy regression kink design estimation to investigate the relationship between female education and brideprice culture.[3]

While it is difficult to collect historical data from countries that have already experienced the decline in brideprice practices, countries that are going through relevant cultural changes provide an ideal case for studying the factors contributing to the disappearance of brideprice culture. Uganda is a particularly suitable case, as almost 100% of its population had practiced brideprice until approximately the 1980s. However, recently, this practice has been diminishing (Anderson, 2007). Together with this decline in brideprice payments has been a significant increase in female educational attainment due primarily to the UPE, which was not anticipated.

We first find that being born one year later increases educational attainment for females in the treatment group by approximately .2 to .4 years in addition to the general trend observed for females in the control group. Exploiting this as an exogenous variation, we then find that a one-year increase in female education reduces brideprice practice by approximately 9.9 to 12.3 percentage points. At the same time, we find no evidence that more years of education lead to an increase in the non-agricultural job status of women. We do not find any change in assortative matching in terms of male and female education nor in marital characteristics such as polygyny and love marriage. These results indicate that the decline in brideprice payment is not due to a change in the valuation of female productivity or the matching pattern in the marriage market. In other words, the

[3] We find that the reform increased educational attainment among females but not among males, whose educational attainment had already been above the primary level prior to the UPE. This finding is consistent with that of Keats (2018) and implies that the effect of education is not realized by males but rather by females.

results suggest that the decline in brideprice culture may not be consistently explained by conventional theory, such as assortative matching and human capital compensation.

From our estimation results, we present a hypothesis that is comprehensive enough to explain the decline in brideprice practice. That is, better educated females and their parents may become more aware of the potential downsides of brideprice, such as differential practices of sexual infidelity between spouses (Bishai and Grossbard, 2010), domestic violence (Kaye et al., 2005), and divorce (Platteau and Gaspart, 2007; Gaspart and Platteau, 2010), which are likely to reduce the utility of married females.[4] When the parents of better educated females make a decision about whether to receive brideprice payment, those with altruism toward their daughters face the tradeoff between the immediate payment of brideprice and the future sound marital life of their daughters. If their altruistic utility increases with female education, then they may choose not to receive brideprice payment. This interpretation is similar to that of Gaspart and Platteau (2010), who predict the possible increase in the bride's influence on her parents' behaviour due to her educational attainments.

We contribute to the brideprice literature by providing a new piece of evidence that the cultural practice can change and decline when female education increases. Most studies in brideprice literature have examined the *intensive margin*: within a given cultural framework, they analyse the relationship between the amount of brideprice and socio-economic variables including human capital investment. Our study considers the *extensive*

[4] From local media reports, it seems that brideprice practice is still common in Uganda, where a few highly educated persons express concerns over these potential downsides of brideprice (New Vision, 2016a,b; Daily Monitor, 2016a,b, 2017, 2018a,b,c, 2019).

*margin*: we investigate when and why the brideprice culture declines. Our finding suggests that female education is likely one of the factors that facilitate the cultural decline and help explain the disappearance of brideprice practice in contemporary high-income countries and its current decline in Africa. More broadly, our study is related to the growing literature on culture and institutions (Alesina and Giuliano, 2015). In particular, it shows evidence of the way in which an institution that is aimed at accelerating human capital accumulation affects local culture in the course of economic development.

The rest of this paper is structured as follows. The next section provides a brief overview of the related economic literature. After a summary on primary education reform in Uganda, we describe our dataset and identification strategy. Then, we present the estimation results and discuss our interpretation. The last section concludes with implications for Uganda and for future studies.

## 2.2 Literature Review

### 2.2.1 Marriage and brideprice in Sub-Saharan Africa

Wealth transfers upon marriage are termed differently depending upon who sends and receives them. They are called brideprice if the groom and his family send them to the bride's parents and dower if they send them to the newly married couple, particularly the bride. Similarly, transfers from the bride's side are called dowry if the groom receives them and groomprice if the groom's family receives them (Papps et al., 1983). Among these four channels, we focus on brideprice, which is predominant in Sub-Saharan Africa

(Fafchamps and Quisumbing, 2007).

The economic research related to brideprice can be traced back to at least the seminal work by Becker (1991) on the marriage market, in which women and men search for their marrying partner. His model suggests that assortative matching and compensation for human capital can affect brideprice.[5] First, assortative matching—men and women with similar traits are more likely to marry than are those with dissimilar traits—suggests that a more educated woman is likely to marry a more educated man who is also likely to come from a wealthier family. Then, a more educated woman may be paid a larger brideprice for than a less educated woman. In Sub-Saharan Africa, assortative matching is found to have a strong influence on marital formation and wealth transfer in Ethiopia (Fafchamps and Quisumbing, 2005a,b).

Second, the human capital compensation hypothesis suggests that brideprice is larger for women with larger human capital because it compensates their parents for letting go of a member of family that could provide labour (Becker, 1991). This hypothesis can be particularly important in Sub-Saharan Africa, where women's contribution to household agricultural production is significant and virilocality is common (Anderson, 2007). However, empirical evidence on the human capital compensation hypothesis is relatively scarce in the literature; the available studies include Platteau and Gaspart (2007) and Ashraf et al. (2019). Available studies include Platteau and Gaspart (2007) and Ashraf

[5] Marriage squeeze, another potential factor that can affect brideprice, is a situation in which the numbers of men and women are unbalanced in the marriage market, at which point marital payment arises for the purpose of clearing the market. An excess supply of men in the marriage market incentivizes them and their parents to pay a higher brideprice (Becker, 1991). This situation, however, seems less relevant in Sub-Saharan Africa, where son preference is not found (Rossi and Rouanet, 2015). In contrast, marriage squeezes have been investigated intensively in South and East Asia, where son preference is observed (e.g., Rao 1993; Francis 2011).

et al. (2019).

Platteau and Gaspart (2007) show that brideprice is significantly higher for educated women compared to uneducated women in Senegal based on small cross-sectional data. Ashraf et al. (2019) employs a quasi-experimental method to examine the impact of female education on brideprice. Exploiting the regional variation in school openings and the fact that only a subset of ethnic groups practice brideprice payment, their triple difference estimation shows that school construction booms in Indonesia and Zambia increased the bride's levels of education only among brideprice-practicing ethnic groups, which then led to a larger brideprice payments. The above authors claim that brideprice culture incentivizes parents to invest in girls' education, which indeed leads to higher brideprice payments.

One seemingly missing piece in these studies is the viewpoint of whether the cultural practice of brideprice payment has declined and whether it has anything to do with female education. Instead, within the given cultural framework, these studies consider how education and brideprice interact. That is, the human capital compensation hypothesis claims that brideprice increases if females have greater human capital since it can make her family better off through larger household income and more efficient home production. The assortative matching hypothesis also states that brideprice is larger for highly educated women than for less educated women because highly educated women are likely to be matched with a highly educated and wealthy man. These hypotheses are concerned with the intensive, and not the extensive, margin of cultural practice, which seems to be a serious theoretical limitation, as societies experiencing major changes such as an educational reform can exhibit changes in the extensive margin; *i.e.*, individuals can stop practicing brideprice. This paper sheds light on this issue and attempts to empirically investigate when cultural practice declines.

Another strand of the literature has analysed the strategic behaviours of the groom, the bride, and her parents to further deepen the economic understanding of the causes and consequences of brideprice practice. Bishai and Grossbard (2010) consider the Ugandan case where brideprice is refundable and postulate and show that brideprice payment is associated with a lower probability of the wife's adultery but not with that of the husband. The proposed logic behind this finding is that the payment of brideprice that the husband can recover upon marital disruption due to the wife's sexual infidelity increases his intrahousehold bargaining power, suppressing his wife's extramarital relations but not his own. Gaspart and Platteau (2010) analyse the decisions of the husband and the bride, in addition to her parents, who are altruistic towards the bride, where the agents face a bargaining situation over the brideprice. To the extent that the bride's parents care about the probability of her marital dissolution modelled as a function of brideprice or that the parental altruistic utility depends on the bride's persuasion, which may be affected by her educational attainment, her parents may decrease the amount of brideprice for the groom, for which the authors show empirical evidence.[6] These studies imply that an increase in female education, which may lead to a rise in her bargaining against her husband or her parents, may lead to an overall decline in brideprice practice. Our study attempts to

[6] As Gaspart and Platteau (2010) note, these analyses do not entirely replace the approach based on marriage market interactions (Becker, 1991); rather, they complement the behavioural analyses by bringing individual strategic motives into the analytical framework.

provide evidence of this possibility.

### 2.2.2 Literature on culture and institutions

From a broader perspective of the economic literature, our present paper can be contextualized as an empirical study on the relationship between culture and institutions. According to a recent survey (Alesina and Giuliano, 2015), one approach is to build a model in which culture and institution affect each other and analyse their co-evolution (e.g., Alesina et al. 2015 and Fernández 2013). Studies using this approach, however, inevitably face empirical difficulty in identifying the two-way influencing system (Alesina and Giuliano, 2015).

The other approach is to analyse the one-way influence of culture on institutions or vice versa. In this stream of the literature, it seems that many studies have looked at the effects of culture on economic variables: examples include the effect of individualistic norms on agricultural production (Olsson and Paik, 2016) and the effect of witchcraft beliefs on social capital formation (Gershman, 2016). A relatively limited number of studies have examined the effects of institutions on cultural variables, such as the land titling effect on individualism and materialism (Di Tella et al., 2007) and the impact of commercial legislation on church going and religious donations (Gruber and Hungerman, 2008). Our present paper adds a new piece of evidence to this strand of the literature, as it investigates the effect of a formal institution—the primary education policy—on a cultural behaviour—brideprice practice.

The most closely related study to ours is Ashraf et al. (2019), which examines the effect of female education on brideprice. While the above authors consider the intensive margin of the brideprice culture—whether female education increases brideprice—our current study focuses on the extensive margin—whether or not female education renders a decline in brideprice culture. By providing new empirical evidence on the effect of females on brideprice culture, our study contributes to the emerging literature on the relationship between culture and institutions.

## 2.3 Universal Primary Education in Uganda

Uganda has an educational system that consists of 7 years of primary education, 6 years of secondary education, and 3 to 5 years of tertiary education. Children are supposed to commence schooling at the age of 6. The net enrolment rate was low before the start of universal primary education (UPE). Among children aged 6 to 12 years, 62.1% attended primary school in 1992 (Deininger, 2003). There were also large disparities across income levels and geographic locations. For example, the attendance rate was higher in urban areas (78%) than in rural areas (66%) (Demographic and Health Survey, 2004).

A major impediment was then said to be the costs of schooling, both direct and indirect, borne by parents and family. Tuition paid by households covered more than 80% of the total finances of schools (Nishimura et al., 2008). Households also paid other costs of education, including costs for uniforms, textbooks, and contributions to the Parents and Teachers Association (PTA). These costs were likely a particularly heavy burden for poor households. In fact, 45.7% of children from low-income households attended school, while the equivalent figure was 81.7% for those from high-income households (Demographic and Health Survey, 2004).

To tackle the financial problem of education, a reform called 'universal primary education' (UPE) was initiated in January 1997, which eliminated school fees (Uganda Bureau of Statistics, 2003). Specifically, the UPE scheme provided each school with enough funding to cover private education costs such as tuition and PTA contributions (Ministry of Education and Sports, 1999). Uganda had, by then, put into practice a variety of educational reforms likely to bring about qualitative improvements of education, such as curriculum changes, teacher training, and primary completion examination criteria. However, it was not until December 1996 that the abolition of school fees was announced by President Museveni, who was elected in May 1996 (Grogan, 2008). The announcement was followed by a governmental advertising campaign, which informed nearly all parents and guardians of school-age children of the reform (Demographic and Health Survey, 2004).

As a result, the UPE brought about monumental changes in Uganda. The number of enroled children aged 6 to 8 increased from 2.7 million in 1996 to 5.3 million in 1997 and further to 7.3 million in 2002 (Uganda Bureau of Statistics, 2003). With the gross enrolment rate rising from 74.3% in 1995 to 135.8% in 2000, the reform was said to have achieved universal access to primary education (Riddell, 2003). The effect of UPE was found to be larger for girls and poor households than for boys and richer households (Deininger, 2003; Nishimura et al., 2008). Additionally, the reform reduced delayed enrolment and increased the probability of completing higher grades (Nishimura et al., 2008). Although there are still quite a few people who never attended school or dropped out of school (Demographic and Health Survey, 2001), it would not be an exaggeration to say that Uganda's UPE has been successful in increasing educational attainment. We take advantage of this sudden increase in education brought about by the UPE to investigate the impact of education on brideprice practice. The exogeneity of this policy change is discussed in more detail in Section 2.5.2.

## 2.4 Data

### 2.4.1 Survey design

This study uses the data obtained from the fifth wave of the Research on Poverty, Environment, and Agricultural Technologies (RePEAT) survey in Uganda, undertaken from September through December 2015. The RePEAT survey comprises a panel dataset from the first wave in 2003 to the fifth wave in 2015. In 2003, 10 randomly chosen households were interviewed from each of 94 randomly chosen rural villages from the eastern, western and central districts (Yamano et al., 2004). In 2015, the survey was extended to cover five more randomly selected households in each of the formerly surveyed villages, and 23 additional villages (15 households each) were added from two northern districts. In total, the 2015 RePEAT survey constitutes a dataset of 1,755 households from 117 villages. The RePEAT survey is designed to collect data on the agricultural activities of rural households. In the fifth wave, a specially made questionnaire was added to ask about individual characteristics and behaviours at the time of marriage, including brideprice payment.

The questions about brideprice payments were asked only for each female's first marriage for the following reasons. First, the brideprice agreed upon at the time of the first marriage is likely to correctly measure the value of parental investment in the bride's human capital to the extent that her parents do not anticipate the divorce and remarriage of their daughter when she marries for the first time (Arunachalam and Logan, 2016). Second, the decision-making mechanism for brideprice in a remarriage may be different than that in the first marriage, and these differences may depend on many factors such as the gender of the person who remarries;[7] whether the remarriage follows divorce, separation, or widowhood; and unobservable factors including feelings towards remarriage. This study avoids analytical complications arising from these complexities by focusing only on women's first marriage.

This study sample consists of females who had ever been married at the time of the survey and were between 24 and 49 years of age.[8,9] We imposed this age restriction expecting to obtain a representative sample of females in this age group in rural Uganda because the median age at first marriage was 17.9, and the share of females who had never married in the age group between 25 and 29 years was 5.6%, with a sharp decline from the 23.9% in the age group between 20 and 24 years (Demographic and Health Survey, 2012).

[7] For example, Fafchamps and Quisumbing (2005b) find that men and women remarry with significantly different probabilities in Ethiopia, the country right to the north of Uganda.

[8] Attempts were made to ask questions to the females themselves, but a proxy interview was allowed if the female was unavailable and the alternative respondents knew a great deal about her first marriage.

[9] Another attempt was made to collect information from male respondents since if a male interviewee had remarried and his first partner was no longer available in the RePEAT survey, asking him about his first marriage would increase the sample size of the study. However, the use of such male-queried data was abandoned because for such female partners, other critical information (collected in the education and demography sections that were used for existing household members only) is missing, and thus, regression analyses cannot be performed.

### 2.4.2 Major variables

Females aged 24 to 49 years who have ever been married were asked whether their first marriage partner agreed to pay anything as a form of brideprice.[10] The survey also collected information about the characteristics of the first marriage and the women in their first marriage, such as the year of the marriage and whether the marriage was based on love or arrangement.[11] Furthermore, the survey asked the religion and location of residence before and after the first marriage to identify any changes due to the marriage.

To construct a variable measuring educational background, this study uses the survey responses on the highest grade of education achieved by each individual. Comparing these responses to the education system in Uganda, the minimum years of schooling required to achieve the person's highest grade of education is calculated. This variable is used as the years of schooling throughout the paper.

[10] If the answer to this question was 'yes', then females were further asked about the amount of brideprice agreed to be paid in cash, cattle, or other forms. The value of cattle and other transfers was queried in real terms (respondents were asked to report 'how much it would cost to buy the same amount of cattle (or other) now'), while the cash payment was questioned in the nominal amount. This approach was intended to suppress recall bias; however, the inflation adjustment turned out to be so problematic that the cash amount was barely comparable to cattle and other brideprice amounts due to rampant inflation rates in the late 1980s and early 1990s, when there was internal conflict in Uganda. Therefore, the data on the number of brideprice payments are not used in this study.

[11] The survey attempted to collect information about the land holdings of the female's natal family. However, the variables contain too many missing values and thus cannot be used for analysis. Moreover, the survey did not collect information about gift exchange and reciprocity.

## 2.5 Estimation Strategy

### 2.5.1 Regression kink design estimation

In evaluating the impact of education on brideprice, it is important to note that the UPE introduced free primary education to all those who were enroled in primary school in and after 1997, regardless of their year of birth.[12] Since older females were less likely to remain enroled in primary schools in any given year, they were less likely to be exposed to the reform. This aspect is clearly illustrated in Figure 2.1, where the share of females who were enroled in primary schools in 1997 is zero for cohorts born in or before 1978[13], whereas it is increasing for younger cohorts born in or after 1979. This aspect is likely to have resulted in a kinked increase in their years of education, which is likely to show accelerated growth for younger females who were increasingly exposed to free primary education for a longer period of time. We confirm this finding in Figure 2.2, where educational attainments are stable for females of non-UPE-exposed cohorts (born in or before 1978), while they exhibit a steady increase for females of UPE-exposed cohorts (born in or after 1979).[14] Our aim is to exploit this kinked increase in the years of education

[12] In the original planning scheme, up to four children per household were supposed to be the target of free primary education. However, everyone was eventually provided with free primary education as lon as they were enroled in primary school (Grogan, 2008).

[13] The measurement of enrolment status in 1997 is based on the survey question that asked about the year in which each person left primary education. In the survey, we asked about the attendance of a UPE-funded school and found quite a few females who reported having attended UPE-funded schools and reported having already left their primary schools by 1997. This fact may be due to misreporting, but it is more likely that our question was simply misunderstood: we asked whether they attended a UPE-funded school and did not clarify whether the school was free while they were enroled there. In this case, since the UPE abolished tuition fees at all primary schools, females may have been confused about this question, particularly if they acquired some posterior knowledge about the nature of the UPE to answer our survey question.

[14] The vertical line indicates the year 1979, our cutoff that divides the control and treatment cohorts in our estimation described below. For details, see Section 2.5.2.

to causally estimate the effect of female education on brideprice transactions in Uganda. The nature of the reform implies that the pattern of exposure is better characterized by regression kink design (RKD) estimation than by regression discontinuity design (RDD) estimation, the latter of which has been used by previous studies (Keats, 2018; Behrman, 2015). While the existence of a jump in the share of females exposed to the reform, or in the years of education, is a necessary condition for the RDD estimation, the figures also fail to exhibit such a jump at the cutoffs proposed by previous studies (1983 in Keats (2018) and 1984 in Behrman 2015).

[Figures 2.1 and 2.2 about here]

Based on these considerations, we employ fuzzy RKD (FRKD) estimation, which takes the years of education as the treatment variable and the year of birth as the running variable.[15] Let $y$ be a generic outcome, $s$ be the years of education, $z$ be the year of birth, and $c$ be the cutoff. Then, the treatment-on-the-treated (TOT) parameter in the FRKD, $\tau$, is expressed (Card et al., 2015) as follows:

$$\tau = \frac{\lim_{z_0 \to +c} \left. \frac{d\,\mathrm{E}(y|z)}{dz} \right|_{z=z_0} - \lim_{z_0 \to -c} \left. \frac{d\,\mathrm{E}(y|z)}{dz} \right|_{z=z_0}}{\lim_{z_0 \to +c} \left. \frac{d\,\mathrm{E}(s|z)}{dz} \right|_{z=z_0} - \lim_{z_0 \to -c} \left. \frac{d\,\mathrm{E}(s|z)}{dz} \right|_{z=z_0}} \tag{2.1}$$

where the change in the first-order derivative of the conditional expectation of the outcome

[15] It would be ideal if we had more precise measurements, such as date of birth. However, birth year is the most precise measure in our data.

at the cutoff is evaluated by the change in the first-order derivative of the conditional expectation of the years of education.

To obtain the estimate of the TOT parameter, we follow the literature (Lundqvist et al., 2014; Manoli and Turner, 2014; Tirgil et al., 2018) and estimate the following model by two-stage least squares regression, specified as follows:

$$y_i = \alpha_0 + \alpha(z_i - c) + \beta\hat{s}_i + W_i\phi + u_i \tag{2.2}$$

$$s_i = \gamma_0 + \gamma(z_i - c) + \delta\mathbf{I}\{z_i \geq c\}(z_i - c) + W_i\psi + v_i \tag{2.3}$$

where $\mathbf{I}\{A\}$ denotes an indicator function that takes the value of one if condition A in the following bracket holds and zero otherwise, $\hat{s}_i$ denotes the predicted years of education from equation (2.3), and $W$ denotes premarital controls. This two-equation model is estimated for females of $h$ cohorts below and above the cutoff, $c$. The parameter $\beta$ represents the parametric analogue of the TOT parameter, $\tau$, based on the assumption that the conditional expectation functions of $y$ and $s$ are linearly specified.[16] For statistical inference, we compute the heteroscedasticity-robust standard errors and perform the usual two-sided significance test.[17] For statistical inference, we compute the heteroscedasticity-

[16] As in Ashraf et al. (2019), we considered using an additional exogenous variation to compare females from ethnicities with and without the historical practice of brideprice. However, we abandoned the use of such information after merging the historical data of the Ethnographic Atlas (Murdock, 1967) and its updates (Gray, 1999) since all the ethnic groups in our data (except for non-Uganda nationals) have had a substantial amount of brideprice payments in the past, as shown in Appendix Table 2.B.1.

[17] According to the extant literature, everyone enrolled in primary school was the beneficiary of the UPE, which suggests that exposure to the UPE is equivalent to leaving primary school in or after 1997. We do not use this information explicitly in estimation, primarily because it changes the research question, from the effect of a one-year increase in female schooling on the cultural practice of brideprice to that of UPE exposure, at the expense of richer information on treatment intensity.

robust standard errors and perform the usual two-sided test for significance.[18]

### 2.5.2 Identification assumptions and the choice of cutoff

The identification of $\beta$ hinges upon the exogeneity of the introduction of the UPE, which in this case is that Ugandan females could not precisely manipulate whether or not they attended the UPE-funded school (Lee and Lemieux, 2010). This condition is likely to be met in our setting for the following two reasons.

First, the results of the presidential election preceding the introduction of the UPE were *ex ante* uncertain. In 1996, along with the incumbent Yoweri Museveni, two other candidates were running for presidential office: Paul Ssemogerere and Kibirige Mayanja. Although Museveni won over three times more votes than any other candidates in the end, he lost in quite a few districts in the northern region and some in the central and eastern regions to the second-place candidate, Ssemogerere (Uganda Electoral Commission, 1996). Moreover, Museveni's then slogan of anti-multiparty politics was said to be unpopular (The Independent, 1996). Given the limited information network and coverage in Uganda in 1996, it is unlikely that even those voters who supported Museveni in his winning constituencies were able to predict his popularity in other places, let alone his overall victory in the race. In addition, there was another election in June 1996 for the members of Parliament.[19] These two elections within a year are likely to have created

[18] It is found from the literature that everyone enrolled in primary school was the beneficiary of the UPE. This suggests that exposure to the UPE is equivalent to leaving primary school in or after 1997. We do not use this information explicitly in estimation, primarily because it changes the research question, from the effect of a one-year increase in female schooling on the cultural practice of brideprice to that of UPE exposure, at the expense of richer information about the intensity of treatment.

[19] Large national projects such as the UPE usually require approval by Parliament in regard to the budget and implementation plan. Thus, Museveni's victory in the presidential election would have been insufficient

large uncertainty over the politics of Uganda thereafter.

Second, President Museveni was said to be reluctant to implement the UPE. He arguably placed a larger emphasis on infrastructure development in his economic development planning. Furthermore, the government as a whole, and not just the newly elected president himself, showed little to no interest in pursuing the removal of primary school tuition, despite the call for it by international society (Stasavage, 2005). All these facts support that the reform was indeed suddenly introduced.[20]

To strengthen the validity of this identifying assumption, we choose the year of birth of 1979 as the cutoff, which is the threshold below which the share of females exposed to the UPE is zero and above which it is increasing (Figure 2.1). This cutoff choice is likely to further increase the credibility of our identification since with this cutoff, the average years of education show a clear slope change with no level change (Figure 2.2), which satisfies the continuity assumption of the treatment variable in the RKD estimation (Card et al., 2015).[21] Additionally, compared to later years chosen by previous studies (Behrman, 2015; Keats, 2018), our cutoff raises the opportunity cost of schooling, which intuitively is higher at an older age, when females can otherwise work as family labourers or for wages outside of the household. The larger opportunity cost makes it more difficult

to put the UPE into practice.

[20] His manifesto (Museveni, 1996a) states that he was planning to initiate a reform to allow parents to send four children per household to school for free in 1997. However, 1997 was noted only in the written manifesto: it was never addressed in his oral speech (Museveni, 1996b). That is, information on the timing of the UPE introduction was available only to those who were literate and able to obtain a copy of his manifesto or those who were in touch with someone who could read the manifesto. The proportion of the politically literate in our rural sample data is likely to be small, judging from the fact that the government launched a massive political campaign to publicize the UPE reform after its announcement in December 1996 (Grogan, 2008): if the reform had already been well known to the public, then the government would not have needed such a massive campaign.

[21] Specifically, Assumption 3a of Card et al. (2015) rules out the case in which the treatment variable (years of education in our setting) has no level change at the cutoff.

for females of older cohorts to manipulatively remain enroled in primary schools. In sum, the sudden introduction of the UPE and our choice of cutoff are likely to validate our identification.

Our identification may fail if the effect of the UPE on brideprice practices is brought about by any factors other than a change in female education in a kinked manner. One such factor is male education, and the direction of the potential bias in the estimate of $\beta$ arising from the effect of UPE on male education is ambiguous. For instance, the bias may be positive if males with larger educational attainments are more likely to offer brideprices or pay a larger brideprice, as their natal family is likely to be wealthier or they earn more in the labour market. Conversely, the bias may be negative if such males have better negotiation skills and persuade the bride's parents into a marriage with no brideprice or a lower brideprice. However, as we discuss in Sections 2.6.4 and 2.6.5, the UPE did not significantly affect male education.

Our identification may also fail if this educational reform triggered reforms of other policies and legal systems related to brideprice practices. One possibility is the amendment of the penal code that re-defined the defilement of youths under the age of eighteen, which is the same as the legal marital age in Uganda. Specifically, this amendment extended the definition of defilement from having sexual intercourse with a girl aged 18 or younger to performing any sexual acts with any person younger than 18 years of age (Doya, 2017). However, the amendment of the penal code came into effect in 2007, when females in our oldest treatment cohort were 28 years old and thus not subject to the reform. Therefore, it is unlikely that our estimation exploiting the kink for cohorts born in 1979 or later will fail due to this concern.

Another potential confounding reform is the marriage law, for which many bills have been drafted and discussed, since the Constitutional Court declared in 2004 that the laws governing marriage and divorce were inconsistent with the Constitution adopted in 1995 (Okello, G. M. & Ors., v. Attorney General, 2004). Uganda Law Reform Commission (2010) created the basis for a series of bills, including those in 2009, 2013, and 2017. However, none of these bills has been put into effect to date, and thus, no major reform has been made with regard to the marriage act. Therefore, these bills have had barely any meaningful impact on the marriages in our data.[22]

## 2.6 Results

### 2.6.1 Descriptive analyses

Table 2.1 presents the summary statistics of the major variables from our data for the older control group (born in or before 1978 ) and the younger treatment group (born in or after 1979). Panel A shows that the proportion of females from the Langi, an ethnic group mostly in the northern region of Uganda, is larger in the treatment group than in the control group. This finding may be related to the internal conflict in the 1980s and 1990s, but we do not have a clear explanation for these significant differences. Therefore, dummies for the location of residence at age 7 and ethnicities are always included in our

[22] Similarly, policies pertaining to HIV and AIDS, such as condom provision and anti-stigma campaigns (Tumushabe, 2006), are unlikely to have created a kinked relationship for cohorts born in our cutoff year and contaminate our findings.

regression analyses.

Panel B shows that the treatment females are indeed more educated than the control females. The years of education increased by approximately two years on average, and the treatment females enroled in primary school at an earlier age and left school at an older age than did the control females. The treatment females were more likely to repeat at least one grade in primary school and to proceed to secondary and tertiary education.

[Table 2.1 about here]

Panel C presents a summary of marital characteristics. While the probability of having ever married is lower for the younger treatment females, this is most likely because they are younger than those in the control group.[23] When we focused on those who have ever married (which we refer to as the *married sample*), the proportion of love marriages slightly decreased and that of local mating increased. The share of females in a polygynous union shows a relatively large and significant change between the treatment and control groups. Finally, we observe a large change in brideprice practice. In particular, the proportion of females who had a brideprice paid for their first marriage was significantly lower for treatment females than for control females.

Our key outcome variable, whether females had a brideprice paid for their marriage, demonstrates a kink at the cutoff. Figure 2.3 plots the proportion of females who received

[23] One potential concern is that brideprice payment is observed only for females who have ever married. If the composition of females who have ever married is different for the older and younger cohorts, then the treatment effect estimates of the observable variables for these females may be biased. However, we find that marital age does not affect our regression results. Specifically, we limit the sample to those women who married at age 24 or younger since we can observe marriages that occurred below 24 years of age for all the cohorts in our data. This approach may be an endogenous sample selection, but even so, it did not change our results and conclusions. For more details, see Section 2.6.4.

a brideprice for their first marriage and shows that the slope of the linear fit for the control cohorts is almost flat, while a downward trend is found for the treatment cohorts, thereby making a kink at the cutoff. This kink indicates that there may have been a behavioural change in the cultural practice of brideprice. We investigate this more rigorously through regression analyses.

[Figure 2.3 about here]

### 2.6.2 Tests for identification assumptions

The identification assumptions of the female education effect on marriage and brideprice payment behaviours imply that there exists no kink in the predetermined covariates at the cutoff. Figures 2.4a to 2.4n show the trend of several premarital covariates, including parental years of education, premarital religion, the region of residence when females were seven years old, and major ethnicities. For most of these variables, we fail to find a clear slope change at the cutoff. These observations are further confirmed in Appendix Table 2.B.2, where the slope change, expressed by the coefficient estimate of the interaction term, $\mathbf{I}\{\text{Year of birth} \geq 1979\} \cdot (\text{Year of birth} - 1979)$, is insignificant across these pretreatment covariates. The exceptions include a decrease in the share of females from the central region and an increase in that from the northern region.[24] Although these significant changes are small in magnitude and not robust across bandwidths, we always control for past residence and ethnicity in the following regression analyses.

[24] Correspondingly, we observe a small decrease in the share of Baganda females (Figure 2.4i), most of whom are from the central region, and an increase in the shares of Langi and Acholi females (Figures 2.4l and 2.4m, respectively), most of whom are from the northern region.

[Figure 2.4 about here]

Another important testable implication is whether the density of year of birth is smooth around the cutoff of 1979. For our first check, we use the histogram shown in Figure 2.5. Due to the small sample size of each cohort, the distribution is relatively noisy.[25] However, no noticeable sorting of density around the cutoff appears to be present. The non-existence of sorting is further checked by the statistical test proposed by Frandsen (2017), which detects sorting around the cutoff when the running variable is discrete.[26] The results shown in Table 2.2 reveal that the running variable indeed has a smooth density around our cutoff. The absence of skewed density is not surprising since the cutoff is not used for any administrative purpose for UPE exposure in practice.

[Figure 2.5 about here]

[Table 2.2 about here]

### 2.6.3 Main results

To more rigorously examine the impact of female education on brideprice, we now turn to the regression analysis. Table 2.3 shows the estimated kink coefficient in the first-stage equation (2.3). The interaction term, $\mathbf{I}\{z \geq 1979\} \cdot (z - 1979)$, generally has a positive and significant coefficient estimate. For bandwidths of 8 years or longer, the F statistic moves close to or larger than the usual rule-of-thumb value of 10. The positive coefficient

[25] In Section 2.6.4, we show that our results are robust to this noise.

[26] The standard density test proposed by McCrary (2008) is inappropriate in our setting since this method requires the running variable to be strictly continuous, whilst we use year of birth, which takes integer values and is thus discrete (Frandsen, 2017).

is likely to reflect the kink in the years of education at the cutoff, which increased only for the treatment cohorts (Figure 2.2), implying that the relationship between the year of birth and years of education became positively steeper for the cohorts exposed to UPE.

[Table 2.3 about here]

Table 2.4 presents the estimated effect of female education on whether a brideprice was paid at the first marriage. It shows that female education had a negative and significant effect on the probability of having a brideprice paid for marriage. These estimates are consistent with Figures 2.2 and 2.3, suggesting that an increase in female education for cohorts born in or after 1979 has led to a decrease in the brideprice receipt status for the same cohorts. The statistically significant estimates from Tables 2.3 and 2.4 imply that being born one year later increased female education by .19 to .26 years, which in turn decreased the share of females with a brideprice payment for marriage by 1.9 to 2.8 percentage points, respectively. For bandwidths shorter than 10 years, the coefficient estimates for the treatment effect are insignificant but still negative. These results suggest that an increase in female education led to a decline in receiving a brideprice for marriage.

[Table 2.4 about here]

One may be concerned that the increased years of education may have changed the marital characteristics that caused the decline in brideprice practice. Table 2.5 presents the estimates of the treatment effect on marriage characteristics. We find in Panel A that females with more years of education married at an older age, a finding consistent with those of prior studies using data from Uganda (e.g., Masuda and Yamauchi 2018 and Keats

2018). This result raises the concern that age at marriage may bias our main estimates through a potentially differential marriage probability. We therefore conduct a robustness check by limiting the sample of females to only those who married at younger ages and confirm the robustness of our main findings. For details, see Section 2.6.4.

[Table 2.5 about here]

In contrast, Panels B and C show that the levels of love marriage and local mating did not change much due to female education. The significant estimate for love marriage with a bandwidth of 11 years is likely to be just by chance, as Figure 2.6 shows no substantial kink in its trend at the cutoff. The share of polygynous unions shows a decrease over years according to the summary statistics (Table 2.1), but Panel D shows null effect estimates. Figure 2.7 shows a consistent downward trend of polygynous marriage with no kink at the cutoff. The long-term decline in the presence of polygyny is consistent with previous studies reporting the negative linkage between education and polygyny (Tertilt (2005), with an equilibrium model, and Fenske (2015), with empirical evidence).[27] In Panel E, the estimates for the non-divorcee indicator are all small in magnitude and insignificant.[28] These findings suggest that the change in brideprice receipt for females with more years of education is unlikely due to a change in marital characteristics.

[Figures 2.6 and 2.7 about here]

[27] The dummy for polygamy concerns the current marital union and, thus, may not necessarily indicate that the first marriage was polygynous, which is a limitation of our data.

[28] The divorce status may be right-censored, which is another limitation of our data.

### 2.6.4 Robustness checks and additional analyses

#### Robustness checks of the main results

Our analysis is intended to reveal the effect of female education on brideprice payment practices, exploiting free primary education in Uganda as the source of exogenous variation. However, it is possible that males, not just females, benefited from the UPE and that highly educated males accumulated more wealth to pay to their brides' families for marriage. Alternatively, it is also possible that such males develop better negotiation skills and persuade the brides' parents to agree to a lower brideprice. These possibilities imply that our estimation will identify the mixed effect of both male and female education instead of the effect of only female education. To examine this implication, we repeat the first-stage regression for males with our RePEAT data. The results in Appendix Table 2.B.5 show that the UPE did not affect the educational attainments of males: in particular, the point estimates are generally small and insignificant throughout, which is consistent with Keats (2018), who found that male education did not respond to the Ugandan UPE, and with Deininger (2003) and Nishimura et al. (2008), who reported lesser impacts of the UPE on male education. This finding supports our identification exploiting the UPE to estimate the effects of female education on brideprice practices and marital behaviours.

Additionally, our specification of estimation equations may introduce bias into the estimated effect of female education. In particular, our identification may fail if there exists a discontinuous jump at the cutoff[29] or non-linearity in the expectation function of

[29] Card et al. (2015) shows that, if there is a discontinuity in the conditional expectation function, the identification of the TOT parameter in the RKD framework fails.

brideprice payment practice.[30] Appendix Table 2.B.6 shows that our main findings are robust to allow for a discontinuous jump at the cutoff.[31] In addition, Appendix Table 2.B.7 shows that the regression equation is better specified as a linear, rather than quadratic, function based on the Akaike information criterion (AIC).[32] Therefore, it is unlikely that our main results are severely biased due to the linear specification.

Alternatively, one may be concerned that our choice of cutoff is invalid and that those used by previous studies, such as 1983 (Keats, 2018) or 1984 (Behrman, 2015), are preferable. Figures 2.1 and 2.2 show that the kinks in the probability of exposure to the UPE, as well as in the years of education, started in 1979 in our dataset. If we use later years as the cutoff, then we expect that the first-stage kink will be estimated to be smaller and potentially insignificant, which could make the treatment effect estimate more imprecise and unstable. The estimated results using 1983 as the cutoff in Appendix Table 2.B.8 are in line with this view, showing insignificant coefficient estimates for the first stage and the treatment effect. The point estimates are also inflated, which is likely due to the weak first-stage estimates. These results indicate that it is important to choose the year from which the increasing exposure to free education started.

[30] The linear specification in RKD estimation may spuriously produce a significant treatment effect when the underlying true model is a smooth function that continuously changes its slope around the cutoff.

[31] To allow for a potential jump at the cutoff, we estimate the following equations:

$$y_i = \alpha_0 + \alpha_1(z_i - c) + \beta\hat{s}_i + W_i\phi + u_i$$
$$s_i = \gamma_0 + \gamma_1(z_i - c) + \mathbf{I}\{z_i \geq c\}[\delta_0 + \delta_1(z_i - c)] + W_i\psi + v_i.$$

[32] Specifically, we compare the AIC from the estimation of the following two reduced-form equations:

$$y_i = \gamma_0 + \gamma_1(z_i - c) + \delta_1\mathbf{I}\{z_i \geq c\}(z_i - c) + W_i\phi + r_i, \text{ and}$$
$$y_i = \gamma_0 + \gamma_1(z_i - c) + \gamma_2(z_i - c)^2 + \mathbf{I}\{z_i \geq c\}[\delta_1(z_i - c) + \delta_2(z_i - c)^2] + W_i\psi + r_i.$$

Furthermore, Figure 2.5 shows that the density of the year of birth is somewhat noisy, which may arise from misreporting and possibly bias the treatment effect estimate. The drop in the density for the year of birth of 1981 may be particularly worrisome.[33] To address this concern, we re-estimate the model by excluding each of the birth cohorts within the bandwidth. The results, partly reported in Appendix Figures 2.A.2a to 2.A.2d, show that the estimated effects and their confidence intervals are robust to such sample restriction.[34] Similarly, the change in the composition of females between the treatment and control cohorts in terms of ethnicity and region of residence at age seven may affect our main estimates if controlling for them is inadequate. To address this issue, we re-estimate the effects by excluding the Langi females and those who lived in the northern region at age seven. The results are reported in Appendix Table 2.B.9, showing that our main estimates are robust, which suggests that the noise in the density of the year of birth, as well as the predetermined covariates, is unlikely to drive our estimates.

The potential confounding effect of the legal trial against brideprice practice in Uganda is worth noting. That is, some in Uganda have criticized brideprice practice, claiming that such a culture, combined with virilocality, may dehumanize females, referring to the payment of brideprice as treating females as a '*commodity*' (Wendo, 2004; MIFUMI Project, 2009). If such a social debate confounds the estimation of the effect of female education, then it should affect the marriage of only females in the treatment cohorts in an intensifying manner. However, this result is unlikely because such a debate cannot affect

[33] The density test also suggests that the year of birth of 1981 may raise the concern of potential sorting, although the policy or survey design does not provide any incentive for manipulative reporting.

[34] The results for other bandwidths are not shown for brevity but are available upon request.

only the marriage of females born in or after the cutoff year of 1979. Furthermore, even when we allow for a differential trend of brideprice payment for females who married in or after 2007,[35] which is the year in which the trial against brideprice practice was initiated, the estimated negative effect of female education is essentially unchanged (Appendix Table 2.B.10).

In addition, we consider the potential selection in terms of marriage and censoring by marital age. First, we examine whether the choice to marry is a result of increased female education. Table 2.B.3 shows the regression results for an indicator for having ever married using all the females in our data, regardless of marital history.[36] In Panel A, female education is found to have a significant negative effect on marital probability. However, Panel B shows that the estimates from a quadratic specification are much smaller and insignificant when all bandwidths are available.[37] Here, we present the results from the quadratic model since the AIC for the reduced-form regressions (Panel C) suggests that the linear model is not always preferable to the quadratic model. Therefore, our data are inconclusive regarding the possibility of any selective marital behaviours.

Second, given this inconclusiveness, we examine whether the probability of brideprice payment is a function of females' marital age.[38] In our data, the marital information is

[35] Specifically, we estimate the following equations:

$$\begin{aligned} y_i &= \alpha_0 + \alpha_1(z_i - c) + \beta\hat{s}_i + \theta_0 m_i + \mathbf{I}\{m_i \geq 2007\}(\theta_1 + \theta_2 m_i) + W_i\phi + u_i \\ s_i &= \gamma_0 + \gamma_1(z_i - c) + \delta\mathbf{I}\{z_i \geq c\}(z_i - c) + \rho_0 m_i + \mathbf{I}\{m_i \geq 2007\}(\rho_1 + \rho_2 m_i) + W_i\psi + v_i \end{aligned}$$

where $m_i$ denotes the year in which female $i$ married.

[36] Correspondingly, Figure 2.A.1 shows the share of married females in each cohort for all the females in our data, regardless of marital history.

[37] The estimates from the quadratic model are less precise, as expected (Card et al., 2016; Gelman and Imbens, 2019). However, the insignificance does not solely stem from the larger standard errors since the point estimates are also considerably smaller.

[38] For example, parents of younger females may be more likely to receive a brideprice if marital age is

censored at age 24 for the youngest cohort compared to age 49 for the oldest cohort (when using the largest bandwidth of 13 years). When we replicate the regressions, limiting the sample to those who married at the age of 24 or younger, the estimated effects are strikingly similar to the main results, despite the potentially endogenous sample restriction (Appendix Table 2.B.4).[39] This finding suggests that at least for females who are already married, the effects of female education are not systematically different across ages at marriage. Moreover, regardless of age at marriage, Table 2.5 shows that marital characteristics are not systematically different for the younger treatment females and older control females. Therefore, our findings on the effect of female education on brideprice practice are unlikely to be substantially biased due to a potentially endogenous choice to marry.

### 2.6.5 Mechanism through which female education reduces brideprice practices

We have shown that an increase in female education reduces the probability of receiving a brideprice for a first marriage. We have also shown that this cultural change does not seem to have been caused by a change in marital behaviours. Based on our above findings, we now examine the mechanism of the cultural change brought about by educational reform.

First, as we discuss in Section 2.2.1, the human capital compensation hypothesis suggests that a brideprice is paid to compensate for a bride's human capital. For example,

linked to virginity, which may be valued in the marriage market (Ambrus et al., 2010), or the husband may feel more obliged to pay a brideprice for a younger bride, for whom a larger compensation may need to be made to her natal family.

[39] The number of observations is close to the results in Table 2.4, suggesting that most females marry at fairly young ages and are not actually subject to censoring.

an educated female is likely to hold a non-agricultural job that generates additional income. For this hypothesis to explain the decline in brideprice practice, then more educated females are less compensated when they marry, which contradicts the speculation by Becker (1991).[40] Therefore, it is unlikely that the human capital compensation hypothesis provides an adequate framework for the decline in brideprice practice.

Second, the assortative matching hypothesis suggests that educated women are more likely to marry educated men, who can afford to pay a higher brideprice. For this theory to account for the decline in brideprice practice, the correlation between females and their partners' education needs to have been weakened. However, the simple correlation coefficient of women's own and their partners' years of education became larger (0.408) for the younger treated females than for the older control females (0.312). More rigorously, we regress partners' years of education on the indicator for being born in or after 1979, female education, and their interaction term. The results in Appendix Table 2.B.12 show that on the one hand, female education is significantly correlated with women's partners' education, indicating the existence of assortative matching. On the other hand, the coefficient estimate for the interaction term is positive yet insignificant, suggesting that the assortative nature of marital mating becomes neither stronger nor weaker for the younger treated cohorts. These results show that the assortative matching hypothesis cannot explain the decline in brideprice practice.

[40] In addition, we test one of the predictions of human capital compensation theory by estimating the effect of female education on the probability of females having non-agricultural jobs. The estimated effects are small and insignificant, as shown in Table 2.B.11. Strictly speaking, the effect of female education on female employment should arise from the husband's *ex ante* expectation at marriage, while the analysis here is concerned with the *ex post* self-report by females at the time of the survey. The measurement may thus be inaccurate, which is a limitation of our data.

We now discuss an alternative hypothesis that can explain our findings. That is, more education may have changed women's perception of brideprice and led them to refuse to receive a brideprice for marriage. Several studies have shown that women in Uganda are worried about domestic violence and the sexual infidelity of their husbands who have paid a brideprice (Wendo, 2004; Kaye et al., 2005; MIFUMI Project, 2009; Bishai and Grossbard, 2010). Another study reports a higher probability of divorce for couples with brideprice payments in Senegal (Platteau and Gaspart, 2007).[41] In the media, articles on violence related to brideprice have been reported even after the Supreme Court decision in 2015.[42] The expectation of domestic violence, differential extramarital affairs, and divorce induced by the payment of a brideprice are likely to decrease the future utility of women. Then, their altruistic parents might become increasingly averse to brideprice, as their daughters become more educated and more aware of these possible drawbacks. In other words, the parents of educated females may choose their daughters' sound marital life in the future at the expense of a wealth transfer in the form of a brideprice. This interpretation is consistent with the findings of Gaspart and Platteau (2010), who argue that an increase in female education may help daughters strengthen their influence on their altruistic parents and persuade them to consider their daughters' potential loss of utility contingent on a large brideprice payment.

This mechanism may sound unique to Uganda, where there has been a social debate on whether brideprice practice should be banned (Wendo, 2004; MIFUMI Project, 2009).

[41] While divorce, per se, may not be a bad outcome, noting the differential treatment of men and women after divorce (Fafchamps and Quisumbing, 2005b), it may be the last resort for women.

[42] Examples include Daily Monitor (2016a,b, 2017, 2018a,b,c, 2019) and New Vision (2016a,b).

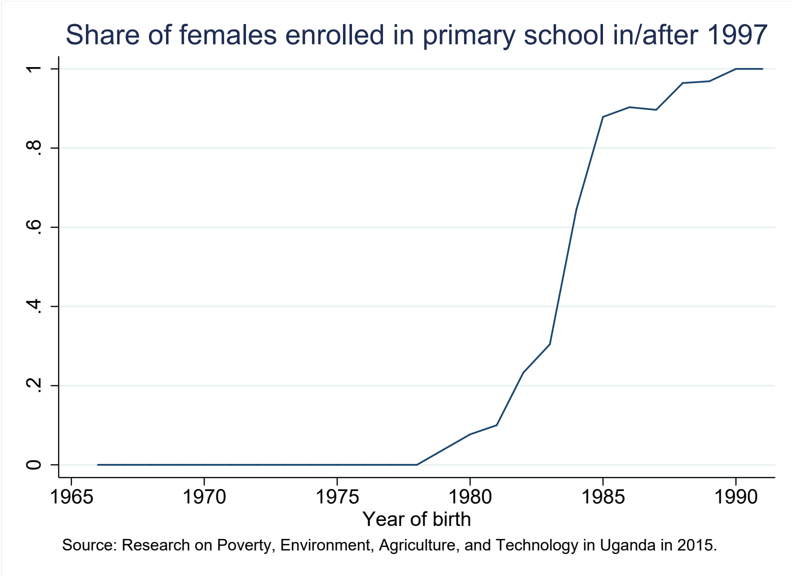
However, as discussed by Anderson (2007), the practice has declined, or is declining, where such a societal discourse does not seem intense or when the contemporary concept of human rights has not yet been fully developed. Moreover, we have found that even in Uganda, the presence of the concern of the legitimacy of brideprice practice does not confound our estimate of the negative effect of female education. Therefore, although entirely speculative, our results on the decline of cultural practice are likely to apply to other settings as well.

## 2.7 Conclusion

In this study, we have investigated the relationship between female education and brideprice payment practice in Uganda. We find that female education decreases the probability of receiving a brideprice for marriage. This probability had been high and stable for the older cohorts but started to decline for the cohorts exposed to the UPE, which resulted in the negative slope change in the probability of brideprice payment conditional on the year of birth. Our estimation results show that a year increase in female education reduces the probability of brideprice payment by approximately 10 to 12 percentage points. To the best of our knowledge, this finding is the first evidence of its kind in the economic literature on brideprice. These results imply that a change in the education system can alter cultural practices, which cannot be explained by conventional theories, namely, the human capital compensation and assortative matching hypotheses. We have, therefore, discussed an alternative conjecture: females with more years of education may become more averse to brideprice due to its potential detrimental consequences, such as the differential probability of extramarital affairs between spouses and domestic violence caused by the husband.

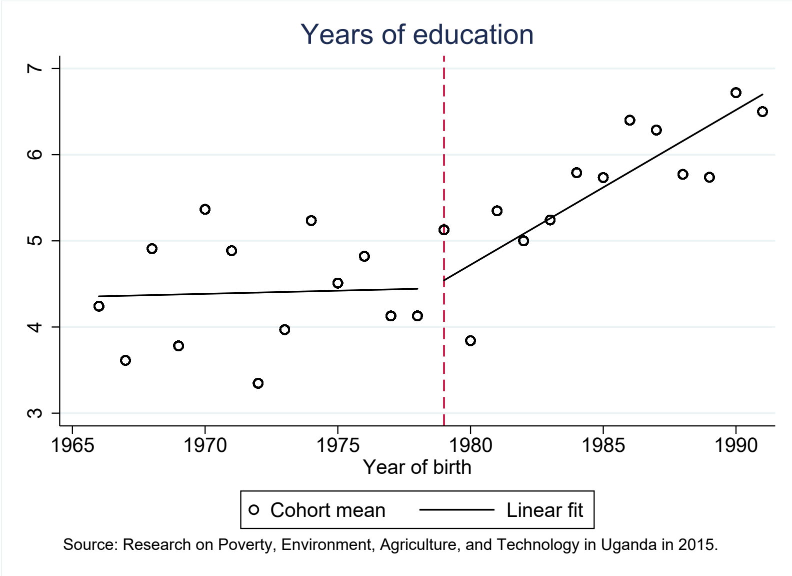
Our study is not free of limitations. First, our findings and conjecture on the negative effect of female education on brideprice practice need to be tested for external validity. An increase in the human capital of females is a common phenomenon during economic development and is thus likely to have contributed to the decline in brideprice practice in many countries. It would thus be beneficial to investigate whether similar changes have occurred in other countries that have experienced major educational reforms. Second, our study does not deny the existence of other mechanisms that may also play a role in the decline in brideprice culture. It is possible that the cultural decline reported the world over (Anderson, 2007) may also have been facilitated for reasons other than an increase in female education. Studies exploring these other mechanisms would thus enrich the understanding of the transformation of brideprice practice. Third, our estimation results are relevant only for the cohorts born in the small neighbourhood of the cutoff. The findings would be different if the cultural behaviour further away from the cutoff cohort is considered. It would be fruitful for future research to overcome these challenges and explore the impact of institutional reforms on cultural changes, particularly brideprice practice.

## 2.8 Figures.



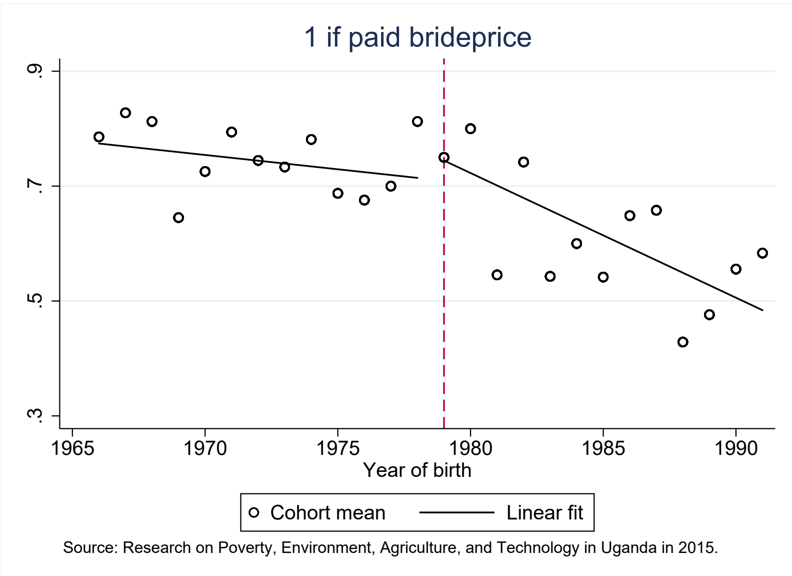

*Notes*. This figure plots the share of females in each birth cohort of the females aged 24 to 49 who have ever married, who were enroled in primary school in and after 1997. The enrolment status is based on the question about the years in which each female enrolled and left primary school, rather than the self-reported measure of their UPE receipt.

Figure 2.1: Share of females enroled in primary school in or after 1997.



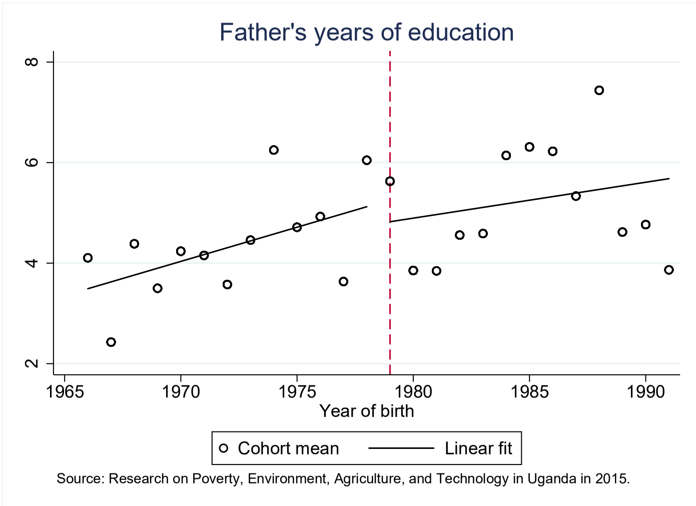

*Source*. Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015.
*Notes*. This figure plots the average years of education of females born in each year and its linear fit for females aged 24 to 49 who have ever married. Years of education is defined as the minimum years of schooling required to achieve the highest grade of education reported by the respondent. The dashed vertical line represents the year 1979, the cutoff of our analysis that is explained in detail in Section 2.5.2.

Figure 2.2: Years of education of females.

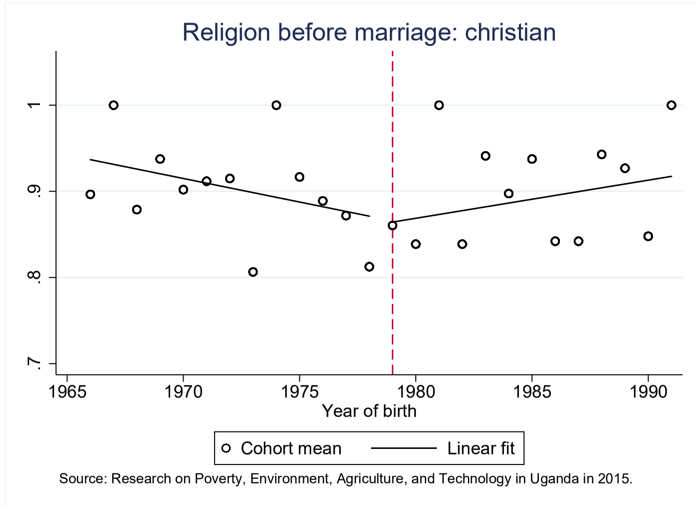



*Notes*. This figure shows the share of females who had agreed to a brideprice payment for their first marriage for each birth cohort and its linear fit. The dashed vertical line represents the year 1979, and the cutoff of our analysis that is explained in detail in Section 2.5.2.

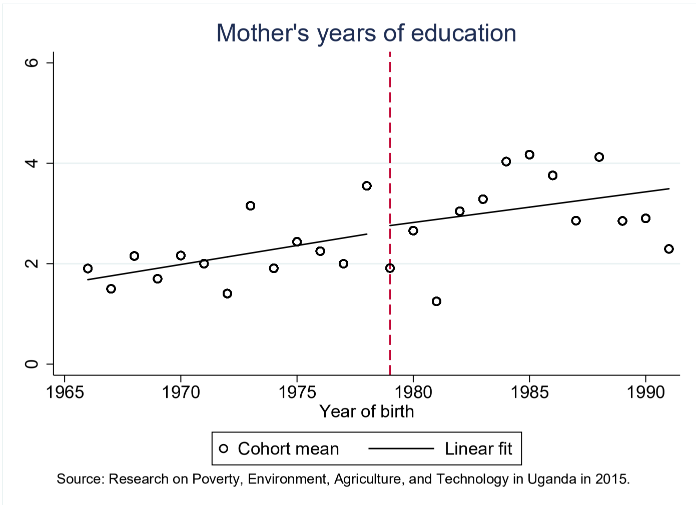
Figure 2.3: Share of females who had a brideprice for marriage.



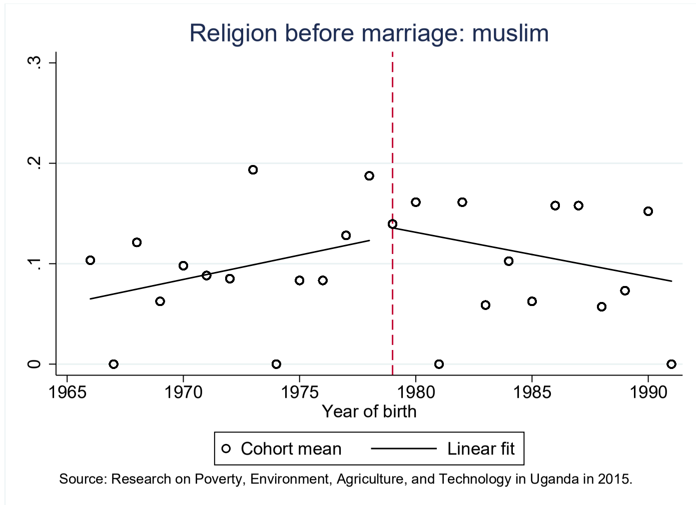

(a) Father's years of education.




(c) Share of females who were Christian prior to marriage.

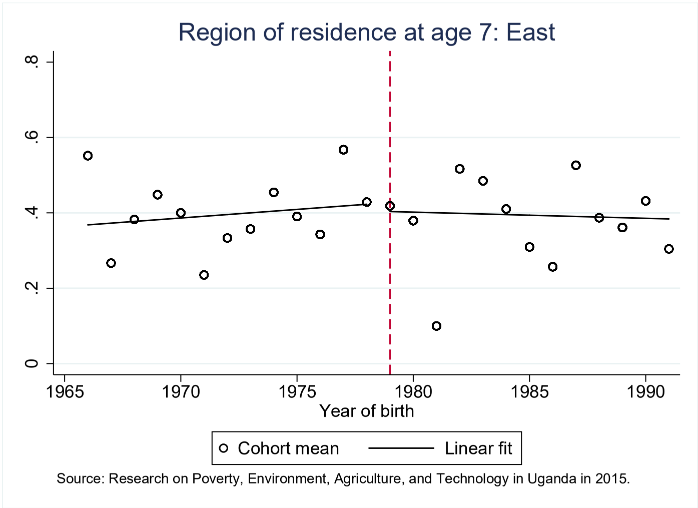



(b) Mother's years of education.

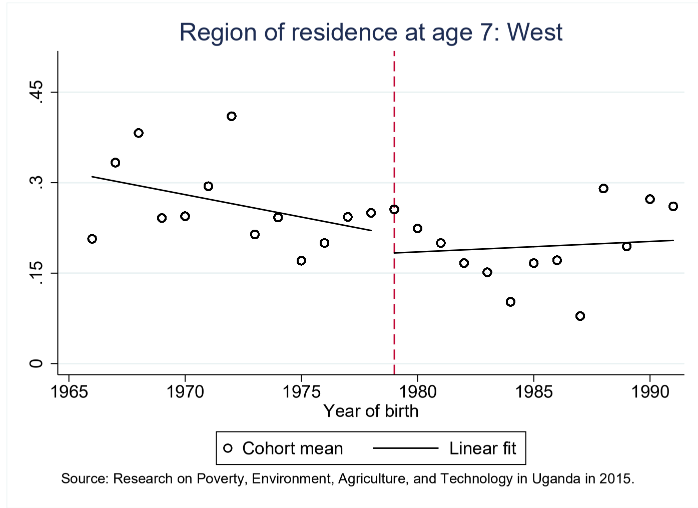



(d) Share of females who were Muslim prior to marriage.

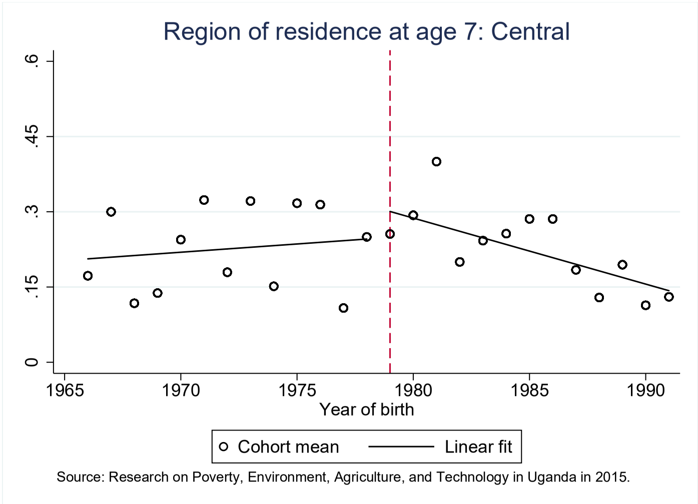
Figure 2.4: Trends in premarital covariates for females born in years around the cutoff.



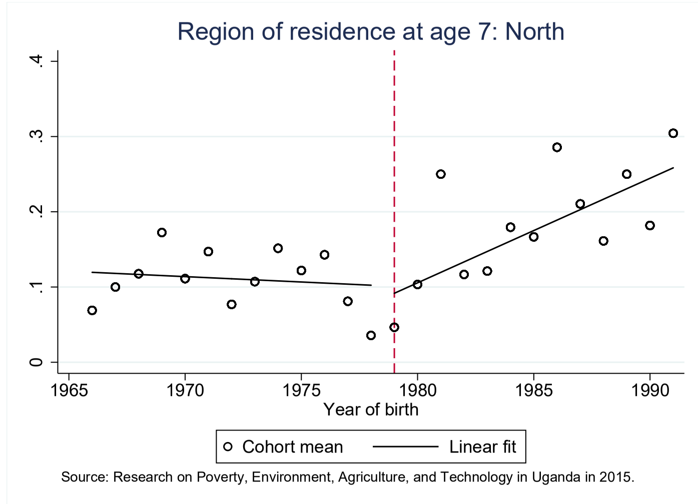

(e) Share of females who lived in the eastern region at age seven.




(g) Share of females who lived in the western region at age seven.

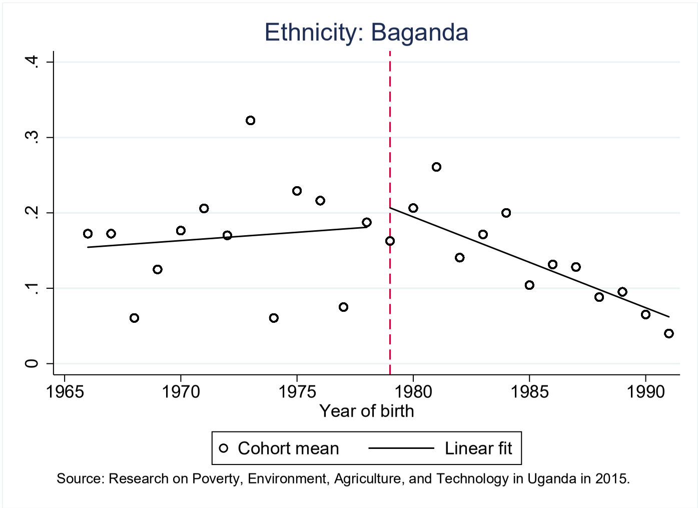



(f) Share of females who lived in the central region at age seven.

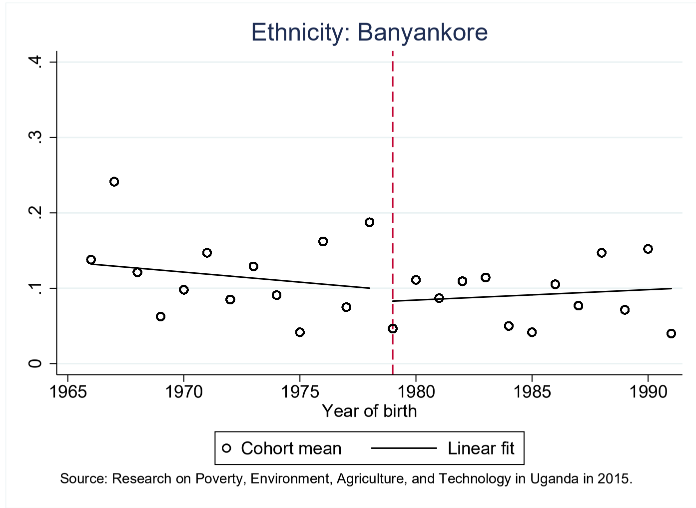



(h) Share of females who lived in the northern region at age seven.

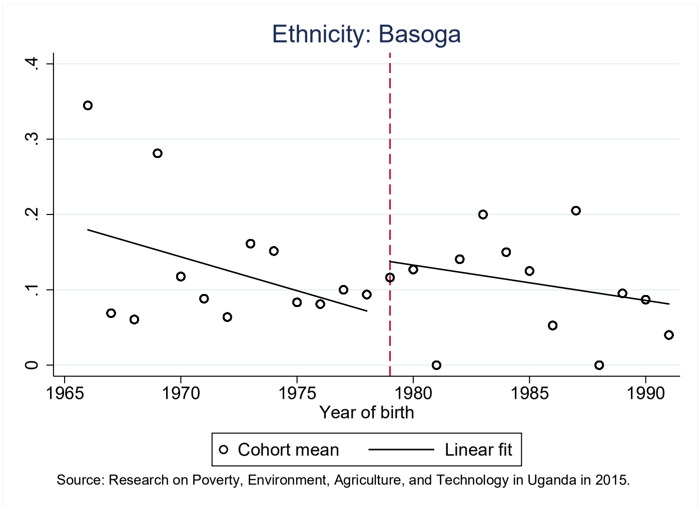
Figure 2.4: Trends in premarital covariates for females born in years around the cutoff (continued).



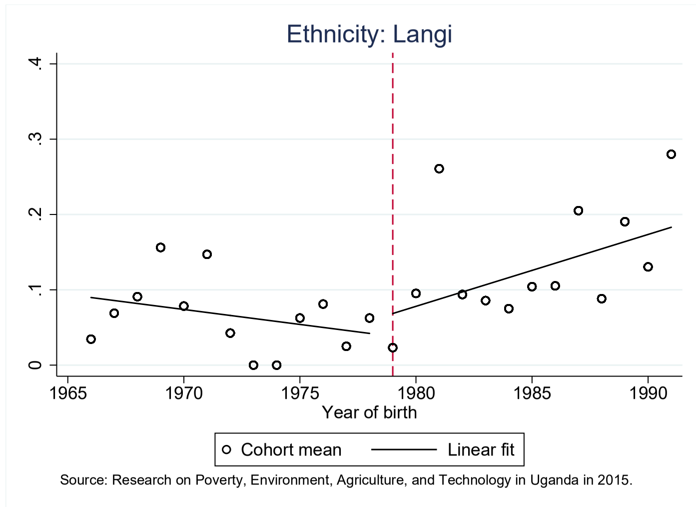

(i) Share of females whose ethnicity is Baganda.




(k) Share of females whose ethnicity is Banyankore.

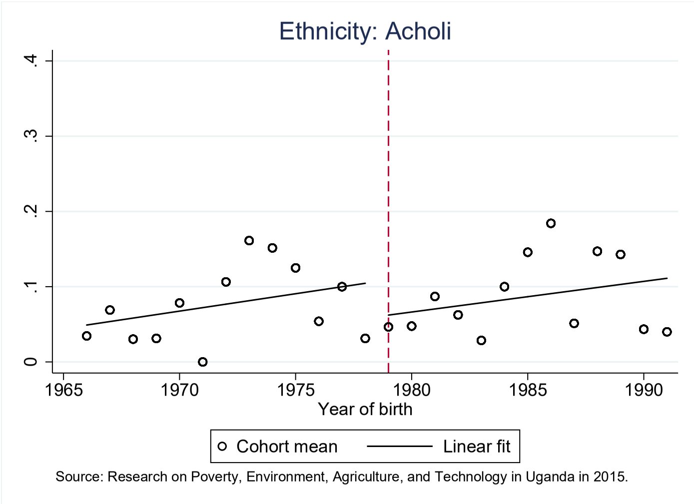



(j) Share of females whose ethnicity is Basoga.

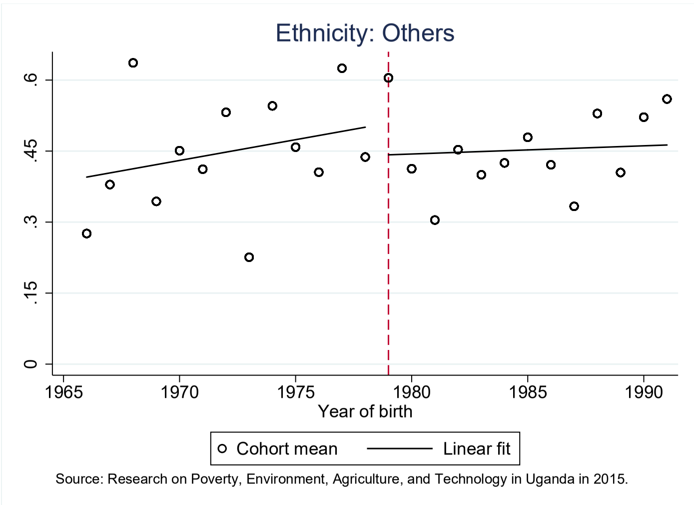



(l) Share of females whose ethnicity is Langi.

Figure 2.4: Trends in premarital covariates for females born in years around the cutoff (continued).

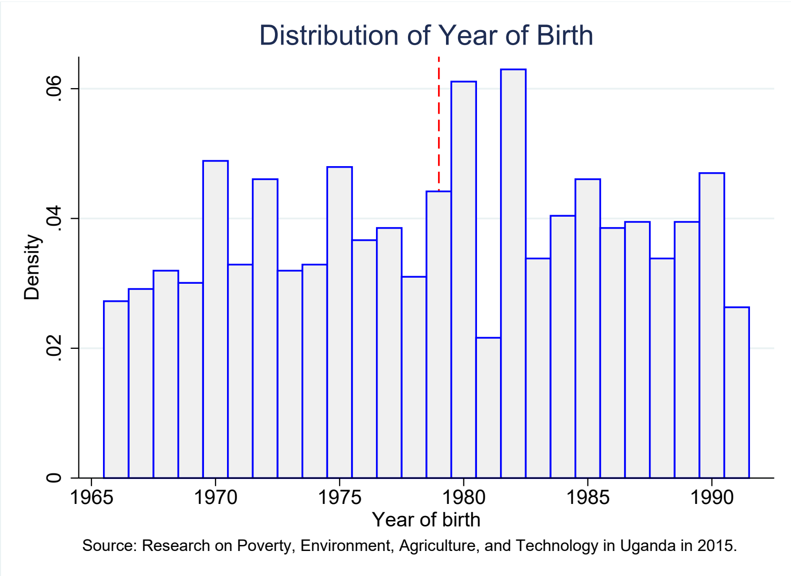



(m) Share of females whose ethnicity is Acholi.



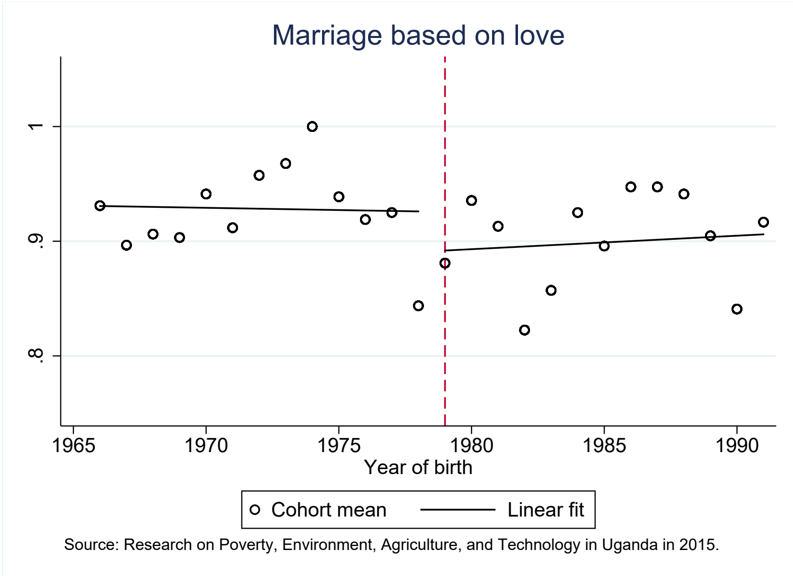

(n) Share of females whose ethnicity is any other.

Figure 2.4: Trends in premarital covariates for females born in years around the cutoff (continued).



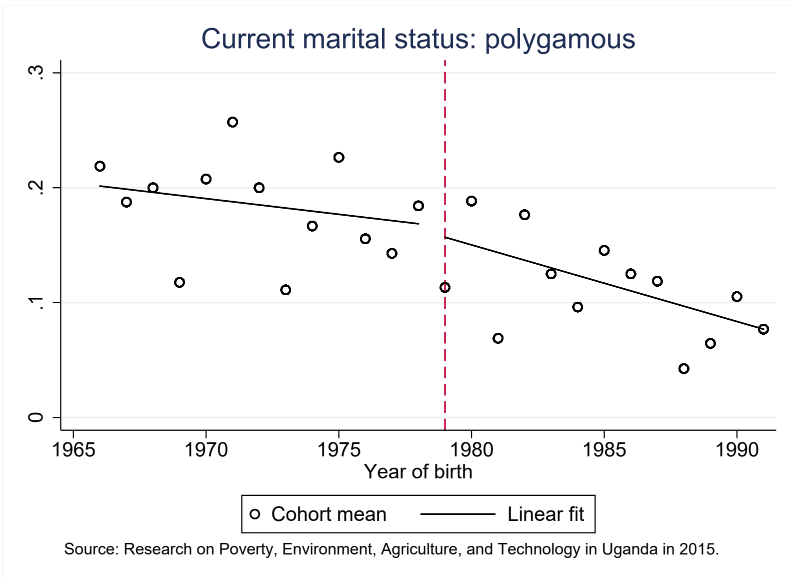

*Notes*. This figure shows the histogram of year of birth for the females aged 24 to 49 who have ever married. The dashed vertical line represents the year 1979, the cutoff of our analysis that is explained in detail in Section 2.5.2.

Figure 2.5: Histogram of year of birth.




*Notes*. This figure shows the share of females whose first marriage was based on love and its linear fit for females aged 24 to 49 who have ever married.

Figure 2.6: Share of females whose first marriage was based on love.




*Notes*. This figure shows the share of females whose current marital union is polygynous and its linear fit for females aged 24 to 49 who have ever married. Polygyny is a type of marital union in which one male marries more than one female.

Figure 2.7: Share of females whose current marital union is polygynous.

## 2.9 Tables.

# Tables.

Table 2.1: Summary Statistics of the Major Variables.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| Sample | Females born in 1966 - 1978. | | | | Females born in 1979 - 1991. | | | | (2) = (6) |
| Variables | N | Mean | Median | Std. Dev. | N | Mean | Median | Std. Dev. | *t* statistic |
| | Panel A. Demographic characteristics. | | | | | | | | |
| Age | 523 | 42.71 | 43 | 3.62 | 733 | 29.71 | 30 | 3.84 | -60.57*** |
| Region at age 7: Eastern | 468 | 0.387 | 0 | 0.488 | 662 | 0.361 | 0 | 0.481 | -0.88 |
| Region at age 7: Central | 468 | 0.235 | 0 | 0.424 | 662 | 0.248 | 0 | 0.432 | 0.49 |
| Region at age 7: Western | 468 | 0.259 | 0 | 0.438 | 662 | 0.210 | 0 | 0.408 | -1.91* |
| Region at age 7: Northern | 468 | 0.118 | 0 | 0.322 | 662 | 0.169 | 0 | 0.375 | 2.41** |
| Own ethnicity: Baganda | 504 | 0.177 | 0 | 0.382 | 702 | 0.168 | 0 | 0.374 | -0.39 |
| Own ethnicity: Basoga | 504 | 0.129 | 0 | 0.335 | 702 | 0.117 | 0 | 0.321 | -0.64 |
| Own ethnicity: Banyankore | 504 | 0.117 | 0 | 0.322 | 702 | 0.101 | 0 | 0.302 | -0.88 |
| Own ethnicity: Langi | 504 | 0.069 | 0 | 0.254 | 702 | 0.110 | 0 | 0.313 | 2.38** |
| Own ethnicity: Acholi | 504 | 0.077 | 0 | 0.267 | 702 | 0.088 | 0 | 0.284 | 0.68 |
| Own ethnicity: Any other | 504 | 0.431 | 0 | 0.496 | 702 | 0.416 | 0 | 0.493 | -0.51 |
| | Panel B. Education variables. | | | | | | | | |
| Years of education | 514 | 4.449 | 5 | 3.305 | 719 | 6.439 | 6 | 4.003 | 9.24*** |
| Partner's years of education | 433 | 6.367 | 6 | 3.603 | 477 | 6.964 | 7 | 3.549 | -2.52** |
| Primary: 1 if attended in any grade | 514 | 0.790 | 1 | 0.408 | 719 | 0.894 | 1 | 0.308 | 5.12*** |
| Primary: Age of enrolment | 341 | 7.639 | 7 | 1.500 | 587 | 7.305 | 7 | 1.567 | -3.18*** |
| Primary: Age of leaving school | 321 | 9.330 | 12 | 8.302 | 517 | 11.53 | 13 | 6.440 | 4.30*** |
| Secondary: 1 if attended in any grade | 514 | 0.123 | 0 | 0.328 | 719 | 0.300 | 0 | 0.459 | 7.52*** |
| Tertiary: 1 if attended in any grade | 514 | 0.008 | 0 | 0.088 | 719 | 0.061 | 0 | 0.240 | 4.82*** |

(Continues to the next page)

Table 2.1: Continued.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| | Females born in 1966 - 1978. | | | | Females born in 1979 - 1991. | | | | (2) = (6) |
| Variables | N | Mean | Median | Std. Dev. | N | Mean | Median | Std. Dev. | *t* statistic |
| | Panel C. First marriage variables. | | | | | | | | |
| 1 if ever married | 521 | 0.946 | 1 | 0.226 | 731 | 0.776 | 1 | 0.417 | -8.49*** |
| Age at first marriage | 465 | 17.66 | 18 | 4.746 | 535 | 17.90 | 18 | 4.064 | 0.87 |
| 1 if love marriage | 474 | 0.928 | 1 | 0.258 | 532 | 0.898 | 1 | 0.302 | -1.67* |
| Own pre-marital residence: Within LC1 | 475 | 0.236 | 0 | 0.425 | 535 | 0.290 | 0 | 0.454 | 1.94* |
| Own pre-marital residence: Within Subcounty | 475 | 0.251 | 0 | 0.434 | 535 | 0.245 | 0 | 0.430 | -0.21 |
| Own pre-marital residence: Within District | 475 | 0.208 | 0 | 0.407 | 535 | 0.200 | 0 | 0.400 | -0.33 |
| Own pre-marital residence: Within Uganda | 475 | 0.303 | 0 | 0.460 | 535 | 0.252 | 0 | 0.435 | -1.80* |
| Own pre-marital religion: Christian | 474 | 0.903 | 1 | 0.296 | 533 | 0.889 | 1 | 0.314 | -0.71 |
| Own pre-marital religion: Muslim | 474 | 0.095 | 0 | 0.293 | 533 | 0.111 | 0 | 0.314 | 0.82 |
| 1 if current = first marriage | 399 | 0.802 | 1 | 0.399 | 463 | 0.840 | 1 | 0.367 | 1.46 |
| 1 if in polygynous union | 521 | 0.184 | 0 | 0.388 | 731 | 0.115 | 0 | 0.319 | -3.46*** |
| 1 if having brideprice paid | 471 | 0.743 | 1 | 0.437 | 528 | 0.621 | 1 | 0.486 | -4.15*** |

*Source*. Research on Poverty, Environment, Agriculture, and Technology survey in Uganda in 2015. *Notes*. This table shows the summary statistics (number of observations (N), mean, median, and standard deviation) of the major variables for the sample women who were born from 1966 to 1991. The control group consists of females born from 1966 to 1978, while the treatment group consists of females born from 1979 to 1991. The amount of brideprice paid in cash, cattle, and other means is defined for only those who reported having agreed to brideprice payment in the previous question. The age of leaving primary school was asked of those who were born in 1972 or after and had completed at least some primary education.

Table 2.2: Density Test of Year of Birth.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Smoothing parameter (k) | | | |
| Year of birth | 0.00 | 0.01 | 0.05 | 0.10 |
| 1967 | 0.961 | 1.000 | 1.000 | 1.000 |
| 1968 | 0.634 | 0.830 | 0.836 | 0.777 |
| 1969 | 0.179 | 0.172 | 0.188 | 0.234 |
| 1970 | 0.014 | 0.025 | 0.030 | 0.047 |
| 1971 | 0.070 | 0.069 | 0.079 | 0.113 |
| 1972 | 0.050 | 0.079 | 0.088 | 0.121 |
| 1973 | 0.346 | 0.331 | 0.349 | 0.450 |
| 1974 | 0.385 | 0.384 | 0.400 | 0.447 |
| 1975 | 0.065 | 0.089 | 0.103 | 0.138 |
| 1976 | 0.443 | 0.460 | 0.478 | 0.523 |
| 1977 | 0.442 | 0.551 | 0.565 | 0.600 |
| 1978 | 0.185 | 0.178 | 0.194 | 0.241 |
| 1979 | 0.891 | 0.861 | 0.934 | 0.941 |
| 1980 | 0.000 | 0.000 | 0.001 | 0.002 |
| 1981 | 0.000 | 0.000 | 0.000 | 0.000 |
| 1982 | 0.000 | 0.000 | 0.000 | 0.000 |
| 1983 | 0.029 | 0.028 | 0.035 | 0.058 |
| 1984 | 0.868 | 1.000 | 1.000 | 1.000 |
| 1985 | 0.341 | 0.464 | 0.483 | 0.483 |
| 1986 | 0.650 | 0.645 | 0.657 | 0.694 |
| 1987 | 0.577 | 0.698 | 0.709 | 0.735 |
| 1988 | 0.502 | 0.499 | 0.512 | 0.556 |
| 1989 | 0.983 | 1.000 | 1.000 | 1.000 |
| 1990 | 0.045 | 0.066 | 0.076 | 0.104 |

*Source*. Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015. *Notes*. This table shows the results of density test proposed by Frandsen (2017) for the year of birth of females in our dataset. A smaller parameter value of $k \in [0, 1]$ makes the test stricter, where the null hypothesis is that there is no manipulative sorting of the running variable at the cutoff. The computation uses females aged 24 to 49 who have ever married.

Table 2.3: Estimated results of first-stage regression.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Years of education | | | | | | | | | | |
| Bandwidth | 13 years | 12 years | 11 years | 10 years | 9 years | 8 years | 7 years | 6 years | 5 years | 4 years | 3 years |
| **I**{Year of birth≥ 1979} | 0.187*** | 0.198*** | 0.253*** | 0.256*** | 0.375*** | 0.369*** | 0.244 | 0.542*** | 0.577** | 0.597 | 0.495 |
| ×(Year of birth−1979) | (0.059) | (0.067) | (0.075) | (0.091) | (0.105) | (0.124) | (0.152) | (0.201) | (0.250) | (0.365) | (0.672) |
| Year of birth - 1979 | 0.032 | 0.024 | -0.009 | -0.003 | -0.066 | -0.045 | 0.020 | -0.134 | -0.209 | -0.125 | -0.100 |
| | (0.031) | (0.036) | (0.042) | (0.048) | (0.054) | (0.065) | (0.077) | (0.104) | (0.130) | (0.183) | (0.280) |
| Observations | 894 | 847 | 776 | 708 | 651 | 572 | 504 | 426 | 364 | 301 | 206 |
| R-squared | 0.261 | 0.260 | 0.258 | 0.259 | 0.272 | 0.272 | 0.288 | 0.327 | 0.348 | 0.365 | 0.508 |
| F statistics | 9.895 | 8.853 | 11.41 | 7.962 | 12.82 | 8.887 | 2.565 | 7.276 | 5.314 | 2.682 | 0.54 |

*Source*. Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015. *Notes*. This table shows the coefficient estimate of the interaction between the year of birth minus the cutoff of 1979 and an indicator for it being equal to or larger than the cutoff from the regression of male years of education on a constant, the year of birth minus the cutoff, and the interaction. Standard errors, robust for heteroscedasticity, are reported in parentheses. Statistical significance is denoted by *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$. F statistic for the significance of the coefficient for the interaction term is reported in each panel. All regressions include the dummies for ethnicity and region of residence at age 7 as covariates. The regressions use females born in years within the indicated bandwidth of the cutoff. Regressions are run for females aged 24 to 49 who have ever married and born in years within the indicated bandwidth of the cutoff.

Table 2.4: The Estimated Impact on Brideprice Receipt Probability with All the Available Bandwidths.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidth. | 13 years | 12 years | 11 years | 10 years | 9 years | 8 years | 7 years | 6 years | 5 years | 4 years | 3 years |
| 1 if having brideprice paid | -0.099** | -0.123** | -0.110** | -0.109** | -0.031 | -0.040 | -0.081 | -0.031 | -0.039 | -0.082 | -0.393 |
| | (0.051) | (0.055) | (0.048) | (0.056) | (0.035) | (0.046) | (0.089) | (0.044) | (0.056) | (0.090) | (0.588) |
| Observations | 878 | 832 | 762 | 695 | 640 | 562 | 495 | 417 | 356 | 294 | 201 |

*Source*. Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015. *Notes*. This table shows the treatment effect estimate for the probability of receiving brideprice for all the available bandwidths in our data. Reported in parentheses are standard errors robust to heteroscedasticity. Statistical significance is denoted by *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$. All regressions include dummies for ethnicity and for the region of residence at age seven as covariates. The regressions are run for the females aged 24 to 49 who have ever married and who were born in years within the indicated bandwidth of the cutoff.

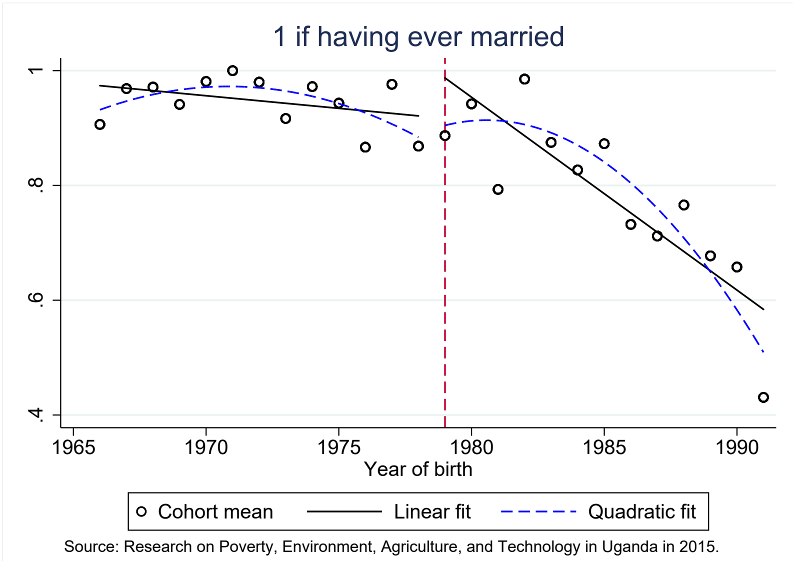
Table 2.5: The Estimated Impact of Female Education on Marital Characteristics.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidth | 13 years | 12 years | 11 years | 10 years | 9 years | 8 years | 7 years | 6 years | 5 years | 4 years | 3 years |
| | Panel A. Age at marriage | | | | | | | | | | |
| Treatment effect | 0.761** | 0.684** | 0.456 | 0.438 | 0.254 | -0.347 | -1.643 | -0.213 | -0.426 | -0.556 | 0.130 |
| | (0.362) | (0.389) | (0.358) | (0.433) | (0.369) | (0.468) | (1.456) | (0.524) | (0.621) | (0.896) | (1.455) |
| Observations | 879 | 834 | 764 | 698 | 644 | 567 | 499 | 422 | 361 | 298 | 204 |
| | Panel B. 1 if love marriage | | | | | | | | | | |
| Treatment effect | 0.027 | 0.027 | 0.043** | 0.033 | 0.018 | 0.009 | -0.013 | -0.023 | -0.056* | -0.092 | 0.020 |
| | (0.025) | (0.026) | (0.024) | (0.026) | (0.021) | (0.023) | (0.043) | (0.028) | (0.040) | (0.072) | (0.094) |
| Observations | 886 | 839 | 770 | 703 | 648 | 570 | 502 | 424 | 362 | 300 | 206 |
| | Panel C. 1 if living in premarital LC1 | | | | | | | | | | |
| Treatment effect | 0.056 | 0.029 | 0.012 | 0.021 | 0.008 | -0.012 | -0.026 | -0.029 | 0.014 | 0.076 | -0.005 |
| | (0.045) | (0.043) | (0.038) | (0.042) | (0.033) | (0.039) | (0.071) | (0.043) | (0.054) | (0.086) | (0.175) |
| Observations | 889 | 842 | 771 | 705 | 648 | 569 | 501 | 424 | 362 | 299 | 205 |
| | Panel D. 1 if polygyny | | | | | | | | | | |
| Treatment effect | 0.009 | -0.010 | -0.026 | -0.007 | 0.008 | 0.013 | -0.022 | -0.019 | 0.003 | -0.005 | -0.053 |
| | (0.037) | (0.036) | (0.032) | (0.038) | (0.031) | (0.037) | (0.064) | (0.035) | (0.042) | (0.054) | (0.125) |
| Observations | 894 | 847 | 776 | 708 | 651 | 572 | 504 | 426 | 364 | 301 | 206 |
| | Panel E. 1 if not divorced | | | | | | | | | | |
| Treatment effect | -0.015 | -0.048 | -0.047 | -0.046 | -0.021 | -0.013 | -0.054 | -0.015 | -0.036 | -0.015 | -0.021 |
| | (0.040) | (0.046) | (0.044) | (0.049) | (0.030) | (0.035) | (0.056) | (0.037) | (0.046) | (0.041) | (0.067) |
| Observations | 755 | 715 | 655 | 602 | 559 | 493 | 432 | 371 | 319 | 264 | 182 |

*Source*. Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015. *Notes*. This table shows the estimated treatment effect of female education on marital characteristics. Reported in parentheses are standard errors robust to heteroscedasticity. Statistical significance is denoted by *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$. All regressions include dummies for ethnicity and for the region of residence at age 7 as covariates. The regressions are run for females aged 24 to 49 who have ever married and were born in years within the indicated bandwidth of the cutoff. Due to missing values, the number of observations differs across regressions.

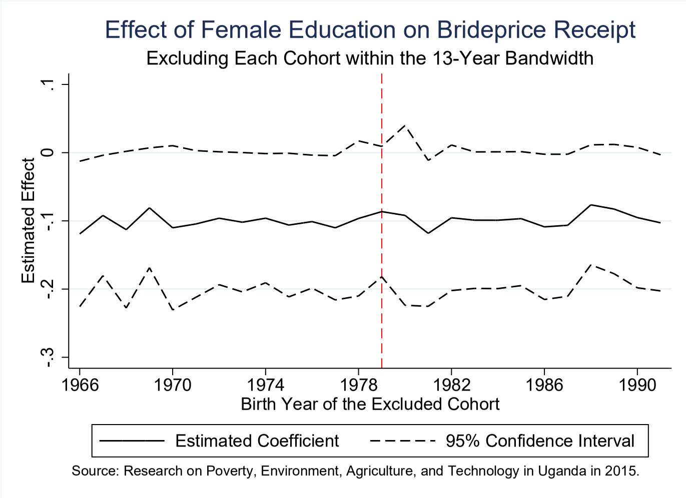
# Appendices for Chapter 2.

## Appendix 2.A Additional figures.

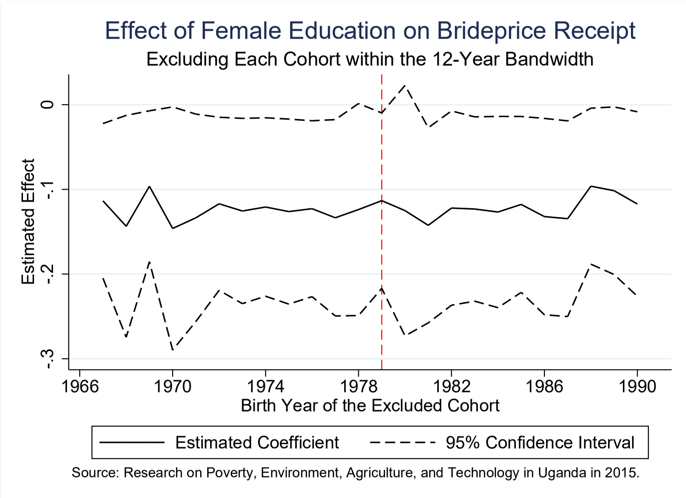



*Source*: Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015. *Notes*: This figure plots the share of females who have ever married for each birth cohort and its linear and quadratic fits. The dashed vertical line represents the year 1979, the cutoff of our analysis that is explained in detail in Section 2.5.2.

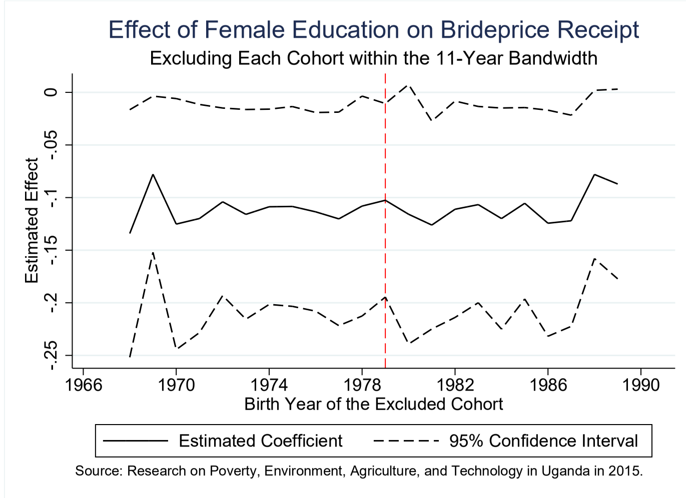
Figure 2.A.1: Share of females who have ever married.



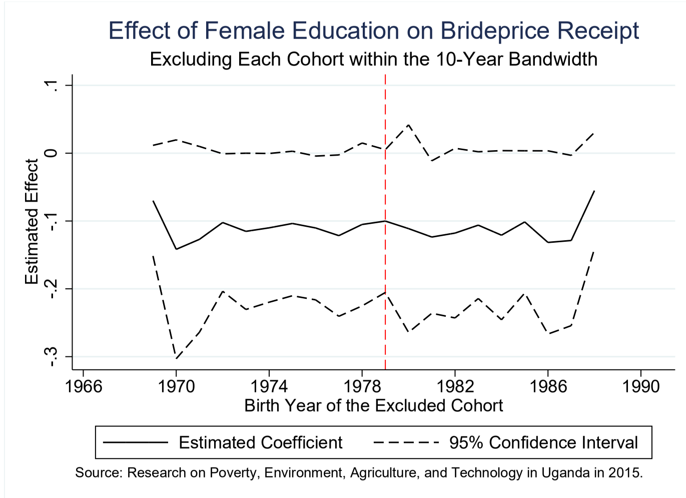

(a) Bandwidth = 13 years.




(b) Bandwidth = 12 years.




(c) Bandwidth = 11 years.




(d) Bandwidth = 10 years.

*Source*: Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015. *Notes*: These figures show the estimated effect of female education on brideprice receipt status and its 95% confidence intervals for each bandwidth by excluding each of the birth cohorts from the estimation sample. The dashed vertical line indicates the cutoff year of birth, 1979.

Figure 2.A.2: Robustness check by excluding each of the birth cohorts within the bandwidths.

## Appendix 2.B Additional tables.

Table 2.B.1: Matching Ethnicity Codes to Murdock (1967) and Gray (1999).

| (1) RePEAT Data | (2) Ethnographic Atlas | (3) Reference | (4) Brideprice practice |
|---|---|---|---|
| Acholi | Luo | O, S | Yes |
| Alur | Luo | S | Yes |
| Badama | Luo | J | Yes |
| Bafumbira | Rwanda-Rundi* | E, W | Yes |
| Baganda | Ganda | O, S | Yes |
| Bagisu | Gisu | O, S | Yes |
| Bagwere | Soga | E | Yes |
| Bahororo | Nyankole | J, E | Yes |
| Bakenyi | Soga | E | Yes |
| Bakiga | Nyankole | E | Yes |
| Bakonjo | Nyoro | E | Yes |
| Banyankore | Nyankole | O, S | Yes |
| Banyarwanda | Rwanda-Rundi* | W | Yes |
| Banyole | Gisu | O | Yes |
| Banyoro | Nyoro | O, S | Yes |
| Baruli | Soga | E | Yes |
| Barundi | Rwanda-Rundi* | W | Yes |
| Basoga | Soga | O, S | Yes |
| Batooro | Nyoro | O, S | Yes |
| Iteso | Teso | S | Yes |
| Jopadhola | Luo | S | Yes |
| Karimojong | Teso | O, S | Yes |
| Kuman | Luo | O | Yes |
| Langi | Lango | O, S | Yes |
| Samia | Gisu | W | Yes |
| Sebei | Kipsigis | S | Yes |
| Sabiny | Kipsigis | S, J | Yes |
| Not Ugandan | - | - | - |

*Notes*. This table relates the ethnic groups that appear in this study's data collected from the Research on Poverty, Environment, Agriculture, and Technology (RePEAT) in Uganda in 2015 to the data from *Ethnographic Atlas*, first written by Murdock (1967) and updated by Gray (1999). Column (3) indicates the source of information that is used to match the names of ethnic groups in the RePEAT and *Ethnographic Atlas*; O stands for Olson et al. (1996), S for Stokes (2009), J for the Joshua Project (Retrieved on the 5th of July, 2019 at https://joshuaproject.net/), E for Ethnologue (retrieved on the 5th of July, 2019 at https://www.ethnologue.com/), and W for Wikipedia. *Since the discussion still carries on as to where the ethnic groups in the area that is now contemporary Rwanda and Burundi come from, we assign a new code, 'Rwanda-Rundi'; this area traditionally employs brideprice practice, according to the Wikipedia.

Table 2.B.2: Kink Coefficient Estimates from Regressions of Predetermined Covariates.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidth | 13 years | 12 years | 11 years | 10 years | 9 years | 8 years | 7 years | 6 years | 5 years | 4 years | 3 years |
| | Panel A. Father's years of education | | | | | | | | | | |
| $\mathbf{I}\{$Year of birth$\geq$ 1979$\}$ | -0.058 | -0.028 | 0.095 | 0.166 | 0.122 | 0.189 | 0.189 | 0.254 | 0.055 | -0.315 | -1.467* |
| $\times$(Year of birth$-$1979) | (0.085) | (0.093) | (0.108) | (0.118) | (0.136) | (0.161) | (0.205) | (0.296) | (0.395) | (0.446) | (0.882) |
| Observations | 689 | 655 | 604 | 557 | 514 | 452 | 395 | 330 | 278 | 237 | 159 |
| R-squared | 0.020 | 0.024 | 0.022 | 0.031 | 0.016 | 0.017 | 0.010 | 0.003 | 0.006 | 0.003 | 0.021 |
| | Panel B. Mother's years of education | | | | | | | | | | |
| $\mathbf{I}\{$Year of birth$\geq$ 1979$\}$ | -0.016 | 0.012 | 0.073 | 0.133 | 0.145 | 0.214 | 0.270 | 0.408* | 0.155 | 0.164 | -0.624 |
| $\times$(Year of birth$-$1979) | (0.068) | (0.076) | (0.087) | (0.099) | (0.115) | (0.140) | (0.170) | (0.245) | (0.308) | (0.398) | (0.645) |
| Observations | 716 | 678 | 625 | 572 | 528 | 464 | 408 | 336 | 281 | 238 | 155 |
| R-squared | 0.025 | 0.028 | 0.030 | 0.037 | 0.029 | 0.037 | 0.034 | 0.016 | 0.007 | 0.002 | 0.006 |
| | Panel C. Pre-marital religion: Christian | | | | | | | | | | |
| $\mathbf{I}\{$Year of birth$\geq$ 1979$\}$ | 0.010* | 0.009 | 0.010 | 0.011 | 0.008 | 0.014 | 0.024* | 0.022 | 0.047** | 0.024 | 0.084* |
| $\times$(Year of birth$-$1979) | (0.005) | (0.006) | (0.007) | (0.008) | (0.010) | (0.011) | (0.013) | (0.018) | (0.021) | (0.031) | (0.048) |
| Observations | 1007 | 953 | 878 | 804 | 737 | 648 | 576 | 481 | 411 | 344 | 234 |
| R-squared | 0.004 | 0.004 | 0.003 | 0.003 | 0.002 | 0.003 | 0.006 | 0.003 | 0.014 | 0.003 | 0.010 |
| | Panel D. Pre-marital religion: Muslim | | | | | | | | | | |
| $\mathbf{I}\{$Year of birth$\geq$ 1979$\}$ | -0.009* | -0.008 | -0.009 | -0.011 | -0.007 | -0.014 | -0.023* | -0.022 | -0.047** | -0.026 | -0.095** |
| $\times$(Year of birth$-$1979) | (0.005) | (0.006) | (0.007) | (0.008) | (0.010) | (0.011) | (0.013) | (0.018) | (0.021) | (0.031) | (0.047) |
| Observations | 1007 | 953 | 878 | 804 | 737 | 648 | 576 | 481 | 411 | 344 | 234 |
| R-squared | 0.004 | 0.004 | 0.002 | 0.003 | 0.002 | 0.003 | 0.006 | 0.004 | 0.016 | 0.005 | 0.012 |

(Continues to the next page)

Table 2.B.2: Continued.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidth | 13 years | 12 years | 11 years | 10 years | 9 years | 8 years | 7 years | 6 years | 5 years | 4 years | 3 years |
| | Panel E. Premarital region of residence: Eastern | | | | | | | | | | |
| **I**{Year of birth≥ 1979} | -0.006 | -0.010 | -0.010 | -0.010 | -0.014 | -0.037** | -0.018 | 0.006 | 0.027 | 0.015 | -0.168** |
| ×(Year of birth−1979) | (0.009) | (0.009) | (0.011) | (0.013) | (0.014) | (0.017) | (0.021) | (0.027) | (0.035) | (0.046) | (0.076) |
| Observations | 944 | 892 | 818 | 748 | 688 | 605 | 536 | 455 | 388 | 322 | 221 |
| R-squared | 0.000 | 0.002 | 0.001 | 0.001 | 0.003 | 0.008 | 0.002 | 0.001 | 0.002 | 0.000 | 0.033 |
| | Panel F. Premarital region of residence: Central | | | | | | | | | | |
| **I**{Year of birth≥ 1979} | -0.018** | -0.016** | -0.017* | -0.013 | -0.002 | 0.005 | -0.005 | -0.002 | -0.017 | 0.003 | 0.084 |
| ×(Year of birth−1979) | (0.007) | (0.008) | (0.009) | (0.011) | (0.013) | (0.016) | (0.018) | (0.024) | (0.030) | (0.040) | (0.079) |
| Observations | 944 | 892 | 818 | 748 | 688 | 605 | 536 | 455 | 388 | 322 | 221 |
| R-squared | 0.007 | 0.005 | 0.004 | 0.002 | 0.000 | 0.000 | 0.001 | 0.000 | 0.001 | 0.000 | 0.013 |
| | Panel G. Premarital region of residence: Western | | | | | | | | | | |
| **I**{Year of birth≥ 1979} | 0.010 | 0.012 | 0.006 | 0.000 | -0.008 | 0.000 | -0.004 | -0.037* | -0.037 | -0.053 | -0.049 |
| ×(Year of birth−1979) | (0.008) | (0.008) | (0.009) | (0.011) | (0.012) | (0.015) | (0.018) | (0.022) | (0.029) | (0.038) | (0.072) |
| Observations | 944 | 892 | 818 | 748 | 688 | 605 | 536 | 455 | 388 | 322 | 221 |
| R-squared | 0.009 | 0.013 | 0.014 | 0.010 | 0.018 | 0.014 | 0.014 | 0.010 | 0.006 | 0.006 | 0.002 |
| | Panel H. Premarital region of residence: Northern | | | | | | | | | | |
| **I**{Year of birth≥ 1979} | 0.016*** | 0.016** | 0.021*** | 0.022*** | 0.025*** | 0.030*** | 0.021* | 0.033** | 0.038* | 0.048* | 0.137** |
| ×(Year of birth−1979) | (0.006) | (0.006) | (0.007) | (0.008) | (0.010) | (0.012) | (0.013) | (0.017) | (0.020) | (0.026) | (0.055) |
| Observations | 944 | 892 | 818 | 748 | 688 | 605 | 536 | 455 | 388 | 322 | 221 |
| R-squared | 0.019 | 0.014 | 0.017 | 0.014 | 0.017 | 0.017 | 0.008 | 0.009 | 0.007 | 0.009 | 0.034 |

(Continues to the next page)

Table 2.B.2: Continued.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidth | 13 years | 12 years | 11 years | 10 years | 9 years | 8 years | 7 years | 6 years | 5 years | 4 years | 3 years |
| | Panel I. Ethnicity: Baganda | | | | | | | | | | |
| **I**{Year of birth≥ 1979} | -0.015** | -0.015** | -0.015** | -0.010 | -0.006 | -0.004 | -0.004 | 0.012 | -0.014 | 0.007 | 0.0584 |
| ×(Year of birth−1979) | (0.006) | (0.006) | (0.007) | (0.009) | (0.010) | (0.013) | (0.015) | (0.021) | (0.025) | (0.034) | (0.067) |
| Observations | 1,016 | 962 | 887 | 812 | 746 | 656 | 584 | 489 | 418 | 350 | 238 |
| R-squared | 0.010 | 0.008 | 0.005 | 0.004 | 0.003 | 0.002 | 0.001 | 0.001 | 0.001 | 0.001 | 0.008 |
| | Panel J. Ethnicity: Basoga | | | | | | | | | | |
| **I**{Year of birth≥ 1979} | 0.003 | -0.002 | 0.000 | 0.005 | 0.005 | -0.002 | 0.007 | 0.023 | 0.022 | -0.003 | -0.061 |
| ×(Year of birth−1979) | (0.005) | (0.006) | (0.007) | (0.008) | (0.009) | (0.010) | (0.013) | (0.018) | (0.024) | (0.029) | (0.042) |
| Observations | 1,016 | 962 | 887 | 812 | 746 | 656 | 584 | 489 | 418 | 350 | 238 |
| R-squared | 0.003 | 0.000 | 0.000 | 0.001 | 0.002 | 0.001 | 0.003 | 0.005 | 0.005 | 0.002 | 0.007 |
| | Panel K. Ethnicity: Banyankore | | | | | | | | | | |
| **I**{Year of birth≥ 1979} | 0.004 | 0.006 | -0.001 | -0.001 | -0.004 | -0.003 | -0.012 | -0.005 | -0.001 | -0.003 | 0.035 |
| ×(Year of birth−1979) | (0.005) | (0.006) | (0.006) | (0.007) | (0.008) | (0.010) | (0.011) | (0.015) | (0.020) | (0.025) | (0.047) |
| Observations | 1,016 | 962 | 887 | 812 | 746 | 656 | 584 | 489 | 418 | 350 | 238 |
| R-squared | 0.002 | 0.002 | 0.000 | 0.000 | 0.001 | 0.001 | 0.002 | 0.000 | 0.001 | 0.000 | 0.005 |
| | Panel L. Ethnicity: Langi | | | | | | | | | | |
| **I**{Year of birth≥ 1979} | 0.013** | 0.012** | 0.015** | 0.013* | 0.013 | 0.005 | -0.003 | -0.007 | 0.005 | 0.028 | 0.127** |
| ×(Year of birth−1979) | (0.005) | (0.006) | (0.006) | (0.007) | (0.008) | (0.009) | (0.010) | (0.012) | (0.015) | (0.022) | (0.050) |
| Observations | 1,016 | 962 | 887 | 812 | 746 | 656 | 584 | 489 | 418 | 350 | 238 |
| R-squared | 0.018 | 0.011 | 0.012 | 0.008 | 0.012 | 0.005 | 0.012 | 0.013 | 0.012 | 0.012 | 0.052 |

(Continues to the next page)

Table 2.B.2: Continued.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidth | 13 years | 12 years | 11 years | 10 years | 9 years | 8 years | 7 years | 6 years | 5 years | 4 years | 3 years |
| | Panel M. Ethnicity: Acholi | | | | | | | | | | |
| **I**{Year of birth≥ 1979} | 0.000 | 0.003 | 0.008 | 0.010 | 0.012 | 0.022** | 0.028** | 0.029** | 0.022 | 0.027 | 0.028 |
| ×(Year of birth−1979) | (0.004) | (0.005) | (0.006) | (0.006) | (0.007) | (0.009) | (0.011) | (0.014) | (0.017) | (0.022) | (0.040) |
| Observations | 1,016 | 962 | 887 | 812 | 746 | 656 | 584 | 489 | 418 | 350 | 238 |
| R-squared | 0.002 | 0.002 | 0.006 | 0.004 | 0.003 | 0.008 | 0.012 | 0.017 | 0.014 | 0.008 | 0.002 |
| | Panel N. Ethnicity: Others | | | | | | | | | | |
| **I**{Year of birth≥ 1979} | -0.006 | -0.004 | -0.007 | -0.016 | -0.021 | -0.018 | -0.017 | -0.052** | -0.034 | -0.056 | -0.187** |
| ×(Year of birth−1979) | (0.009) | (0.009) | (0.011) | (0.012) | (0.014) | (0.017) | (0.020) | (0.026) | (0.034) | (0.045) | (0.080) |
| Observations | 1,016 | 962 | 887 | 812 | 746 | 656 | 584 | 489 | 418 | 350 | 238 |
| R-squared | 0.001 | 0.000 | 0.001 | 0.002 | 0.004 | 0.002 | 0.002 | 0.008 | 0.007 | 0.006 | 0.028 |

*Source*. Research on Poverty, Agriculture, Environment, and Technology in Uganda in 2015. *Notes*. This table shows the coefficient estimate of the interaction between the year of birth minus the cutoff of 1979 and an indicator for it being equal to or larger than the cutoff from the regression of selected predetermined covariates on a constant, the year of birth minus the cutoff, and the interaction. Reported in parentheses are standard errors robust for heteroscedasticity. Statistical significance is denoted by *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$. The regressions are run for females aged 24 to 49 who have ever married and were born in years within the indicated bandwidth of the cutoff. Due to missing values, the number of observations differs across regressions.

Table 2.B.3: Estimated Effects of Female Education on Marital Probability.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidth | 13 years | 12 years | 11 years | 10 years | 9 years | 8 years | 7 years | 6 years | 5 years | 4 years | 3 years |
| | Panel A. Treatment effect estimates with linear specification | | | | | | | | | | |
| Treatment effect | -0.111*** | -0.080*** | -0.062*** | -0.061*** | -0.058*** | -0.039** | -0.034 | -0.011 | 0.023 | 0.047 | -0.084 |
| | (0.022) | (0.023) | (0.020) | (0.026) | (0.023) | (0.021) | (0.035) | (0.023) | (0.030) | (0.039) | (0.099) |
| Observations | 1077 | 991 | 895 | 806 | 736 | 641 | 558 | 472 | 399 | 331 | 232 |
| | Panel B. Treatment effect estimates with quadratic specification | | | | | | | | | | |
| Treatment effect | -0.011 | -0.018 | -0.009 | -0.007 | 0.035 | -0.057 | 0.018 | -0.050 | 0.073 | 0.051 | -0.150 |
| | (0.030) | (0.029) | (0.045) | (0.030) | (0.056) | (0.047) | (0.029) | (0.054) | (0.255) | (0.076) | (0.134) |
| Observations | 1077 | 991 | 895 | 806 | 736 | 641 | 558 | 472 | 399 | 331 | 232 |
| | Panel C. Akaike information criterion (AIC) from the reduced-form regression of marital probability | | | | | | | | | | |
| Linear model | 669.2 | 572.4† | 475.6† | 397.9† | 309.8 | 222.1 | 176.0† | 155.7 | 81.7† | 79.9† | 105.5 |
| Quadratic model | 653.1† | 573.9 | 477.1 | 400.5 | 309.2† | 218.0† | 177.4 | 154.6† | 83.5 | 82.1 | 102.6† |

*Source*. Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015. *Notes*. Panels A and B show the treatment effect estimate of female years of education on the indicator for having ever married. Panel C shows the AIC values from the reduced-form regressions of the marital indicator on the first-stage regressors. Reported in parentheses are standard errors robust to heteroscedasticity. Statistical significance is denoted by *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$. † indicates that the AIC is smaller relative to the other model specification; *i.e.*, the model specification is more preferred to the other. All regressions include dummies for ethnicity and for the region of residence at age seven as covariates. The regressions are run for females aged 24 to 49 who married at the age of 24 or below and for females born in years within the indicated bandwidth of the cutoff.

Table 2.B.4: Robustness Check Using Only Females Who were 24 Years or Younger at Their First Marriage.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidth | 13 years | 12 years | 11 years | 10 years | 9 years | 8 years | 7 years | 6 years | 5 years | 4 years | 3 years |
| 1 if having brideprice paid | -0.084** | -0.121** | -0.105** | -0.116** | -0.024 | -0.030 | -0.084 | -0.024 | -0.035 | -0.075 | -0.715 |
| | (0.050) | (0.060) | (0.052) | (0.066) | (0.042) | (0.047) | (0.121) | (0.059) | (0.069) | (0.114) | (2.309) |
| Observations | 822 | 778 | 711 | 647 | 598 | 532 | 468 | 393 | 340 | 280 | 189 |
| Age at marriage | 0.761** | 0.684** | 0.456 | 0.438 | 0.254 | -0.347 | -1.643 | -0.213 | -0.426 | -0.556 | 0.130 |
| | (0.362) | (0.389) | (0.358) | (0.433) | (0.369) | (0.468) | (1.456) | (0.524) | (0.621) | (0.896) | (1.455) |
| Observations | 879 | 834 | 764 | 698 | 644 | 567 | 499 | 422 | 361 | 298 | 204 |
| 1 if love marriage | 0.027 | 0.027 | 0.043** | 0.033 | 0.018 | 0.009 | -0.013 | -0.023 | -0.056* | -0.092 | 0.020 |
| | (0.025) | (0.026) | (0.024) | (0.026) | (0.021) | (0.023) | (0.043) | (0.028) | (0.040) | (0.072) | (0.094) |
| Observations | 886 | 839 | 770 | 703 | 648 | 570 | 502 | 424 | 362 | 300 | 206 |
| 1 if living premarital LC1 | 0.056 | 0.029 | 0.012 | 0.021 | 0.008 | -0.012 | -0.026 | -0.029 | 0.014 | 0.076 | -0.005 |
| | (0.045) | (0.043) | (0.038) | (0.042) | (0.033) | (0.039) | (0.071) | (0.043) | (0.054) | (0.086) | (0.175) |
| Observations | 889 | 842 | 771 | 705 | 648 | 569 | 501 | 424 | 362 | 299 | 205 |
| 1 if polygynous union | 0.009 | -0.010 | -0.026 | -0.007 | 0.008 | 0.013 | -0.022 | -0.019 | 0.003 | -0.005 | -0.053 |
| | (0.037) | (0.036) | (0.032) | (0.038) | (0.031) | (0.037) | (0.064) | (0.035) | (0.042) | (0.054) | (0.125) |
| Observations | 894 | 847 | 776 | 708 | 651 | 572 | 504 | 426 | 364 | 301 | 206 |
| 1 if not divorced | -0.015 | -0.048 | -0.047 | -0.046 | -0.021 | -0.013 | -0.054 | -0.015 | -0.036 | -0.015 | -0.021 |
| | (0.040) | (0.046) | (0.044) | (0.049) | (0.030) | (0.035) | (0.056) | (0.037) | (0.046) | (0.041) | (0.067) |
| Observations | 755 | 715 | 655 | 602 | 559 | 493 | 432 | 371 | 319 | 264 | 182 |

*Source*. Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015. *Notes*. This table shows the treatment effect estimate for the probability of receiving brideprice for all the available bandwidths in our data. Reported in parentheses are standard errors robust to heteroscedasticity. Statistical significance is denoted by *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$. All regressions include dummies for ethnicity and for the region of residence at age seven as covariates. The regressions are run for females aged 24 to 49 who married at the age of 24 or below and for females born in years within the indicated bandwidth of the cutoff. Due to missing values, the number of observations differs across regressions.

Table 2.B.5: First-Stage Estimation Results of Partner's Years of Education.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Outcome | Male years of education | | | | | | | | | | |
| Bandwidth | 13 years | 12 years | 11 years | 10 years | 9 years | 8 years | 7 years | 6 years | 5 years | 4 years | 3 years |
| **I**{Year of birth≥1979} | -0.041 | -0.016 | -0.021 | 0.010 | 0.069 | 0.045 | 0.073 | -0.071 | -0.038 | -0.013 | 0.083 |
| ×(Year of birth−1979) | (0.068) | (0.077) | (0.087) | (0.101) | (0.119) | (0.145) | (0.182) | (0.218) | (0.285) | (0.409) | (0.942) |
| Observations | 747 | 705 | 660 | 612 | 563 | 506 | 452 | 369 | 315 | 249 | 179 |
| R-squared | 0.200 | 0.199 | 0.196 | 0.214 | 0.221 | 0.230 | 0.279 | 0.342 | 0.404 | 0.425 | 0.425 |
| F statistic | 0.370 | 0.043 | 0.059 | 0.009 | 0.339 | 0.097 | 0.162 | 0.106 | 0.018 | 0.001 | 0.008 |

*Source*. Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015. *Notes*. This table shows the coefficient estimate of the interaction between the year of birth minus the cutoff of 1979 and an indicator for it being equal to or larger than the cutoff from the regression of male years of education on a constant, the year of birth minus the cutoff, and the interaction. Reported in parentheses are standard errors robust for heteroscedasticity. Statistical significance is denoted by *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$. All regressions include dummies for the ethnicity and the district of residence at age seven as covariates. The regressions use males born in years within the indicated bandwidth of the cutoff.

Table 2.B.6: Robustness Check of Main Results Allowing for a Potential Jump at the Cutoff.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidth | 13 years | 12 years | 11 years | 10 years | 9 years | 8 years | 7 years | 6 years | 5 years | 4 years | 3 years |
| 1 if having brideprice paid | -0.107** | -0.126** | -0.111** | -0.111** | -0.033 | -0.047 | -0.086 | -0.031 | -0.026 | -0.080 | 0.010 |
| | (0.052) | (0.056) | (0.048) | (0.056) | (0.035) | (0.045) | (0.078) | (0.044) | (0.053) | (0.089) | (0.094) |
| Observations | 878 | 832 | 762 | 695 | 640 | 562 | 495 | 417 | 356 | 294 | 201 |

*Source*. Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015. *Notes*. This table shows the treatment effect estimate for the probability of receiving brideprice for all the available bandwidths in our data. Reported in parentheses are standard errors robust to heteroscedasticity. Statistical significance is denoted by *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$. All regressions include dummies for ethnicity and for the region of residence at age seven as covariates. The regressions are run for females aged 24 to 49 who have ever married and born in years within the indicated bandwidth of the cutoff.

Table 2.B.7: Akaike Information Criterion for the Reduced-Form Regression of Brideprice Receipt.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidth | 13 years | 12 years | 11 years | 10 years | 9 years | 8 years | 7 years | 6 years | 5 years | 4 years | 3 years |
| | Panel A. 1 if having brideprice paid | | | | | | | | | | |
| Linear model | 1118.4† | 1055.3† | 964.7† | 898.1† | 827.0† | 740.6† | 669.3† | 550.4† | 475.3† | 392.5† | 266.4 |
| Quadratic model | 1119.4 | 1057.8 | 966.9 | 901.6 | 829.7 | 743.3 | 672.2 | 552.1 | 476.4 | 395.2 | 265.6† |
| | Panel B. 1 if having brideprice paid (0 if unmarried) | | | | | | | | | | |
| Linear model | 1363.6† | 1275.8† | 1159.8† | 1074.2† | 987.3† | 870.3† | 777.9† | 647.1† | 549.0† | 460.5† | 318.4 |
| Quadratic model | 1364.9 | 1278.2 | 1161.6 | 1076.4 | 989.5 | 872.3 | 781.1 | 649.5 | 551.1 | 463.0 | 315.6† |
| | Panel C. 1 if having brideprice paid (1 if unmarried) | | | | | | | | | | |
| Linear model | 1324.5 | 1231.3† | 1122.2† | 1032.2† | 934.6† | 831.5† | 747.2† | 621.9† | 540.1† | 445.5† | 321.2† |
| Quadratic model | 1322.7† | 1234.8 | 1126.0 | 1035.7 | 937.1 | 833.4 | 750.0 | 623.5 | 542.7 | 449.2 | 322.5 |

*Source*. Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015. *Notes*. This table shows the values of Akaike Information Criterion for the reduced-form regression of brideprice receipt status. † indicates that the AIC is smaller relative to the other model specification, i.e., the model specification is more preferred to the other. Panel A use the females aged 24 to 49 who have ever married and were born in years within the indicated bandwidth of the cutoff. Panels B and C use all the females in our data, regardless of marital status, born in years within the indicated bandwidth of the cutoff, where the brideprice receipt status for never married females are all coded as one in Panel B and zero in Panel C. All the regressions include dummies for ethnicity and for the region of residence at age seven as covariates.

Table 2.B.8: Robustness Check of the Main Results with the Alternative Cutoff of the Year 1983.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Bandwidth | 9 years | 8 years | 7 years | 6 years | 5 years | 4 years | 3 years |
| | Panel A. First-stage estimation results | | | | | | |
| **I**{Year of birth$\geq$ 1983} | 0.175 | 0.085 | 0.014 | -0.152 | -0.064 | 0.102 | -0.738 |
| | (0.106) | (0.121) | (0.137) | (0.192) | (0.265) | (0.359) | (0.585) |
| Observations | 635 | 585 | 503 | 434 | 372 | 311 | 238 |
| R-squared | 0.250 | 0.260 | 0.282 | 0.307 | 0.333 | 0.395 | 0.464 |
| F statistics | 2.697 | 0.492 | 0.010 | 0.625 | 0.059 | 0.080 | 1.595 |
| | Panel B. Treatment effect estimates | | | | | | |
| 1 if having brideprice paid | -0.158* | -0.451 | -1.255 | 0.180 | -0.364 | -0.040 | 0.061 |
| | (0.110) | (0.651) | (6.237) | (0.320) | (1.743) | (0.470) | (0.111) |
| Observations | 623 | 574 | 494 | 425 | 364 | 304 | 235 |
| Age at marriage | 0.600 | 1.200 | 3.124 | -2.821 | 5.525 | 0.760 | 1.093* |
| | (0.588) | (1.310) | (6.879) | (7.095) | (25.619) | (2.975) | (0.769) |
| Observations | 628 | 578 | 498 | 429 | 368 | 310 | 237 |
| 1 if love marriage | 0.048 | 0.050 | 0.927 | -0.197 | -1.477 | 0.467 | -0.112* |
| | (0.046) | (0.100) | (4.271) | (0.293) | (11.020) | (1.192) | (0.086) |
| Observations | 629 | 580 | 500 | 431 | 370 | 310 | 237 |
| 1 if living in pre-marital LC1 | 0.017 | 0.097 | 2.407 | -0.183 | -0.238 | 0.045 | -0.003 |
| | (0.080) | (0.258) | (32.338) | (0.288) | (1.544) | (0.323) | (0.093) |
| Observations | 631 | 581 | 499 | 431 | 369 | 308 | 235 |
| 1 if in polygynous union | 0.000 | -0.020 | -0.863 | 0.044 | 0.044 | 0.235 | -0.059 |
| | (0.065) | (0.149) | (7.853) | (0.136) | (0.442) | (0.806) | (0.076) |
| Observations | 635 | 585 | 503 | 434 | 372 | 311 | 238 |
| 1 if not divorced | -0.002 | -0.061 | -0.435 | 0.020 | -0.010 | -0.068 | -0.024 |
| | (0.075) | (0.132) | (3.340) | (0.147) | (0.123) | (0.283) | (0.095) |
| Observations | 543 | 502 | 429 | 374 | 320 | 269 | 207 |

*Source*. Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015. *Notes*. This table shows the treatment effect estimate for the probability of receiving brideprice for all the available bandwidths in our data, using the alternative year of birth of 1983 as the cutoff. Reported in parentheses are standard errors robust to heteroscedasticity. Statistical significance is denoted by *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$. All regressions include dummies for ethnicity and for the region of residence at age seven as covariates. The regressions are run for females aged 24 to 49 who have ever married and born in years within the indicated bandwidth of the cutoff.

Table 2.B.9: Robustness Check of Main Results Excluding the Langi and Those from the Northern Region.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidth | 13 years | 12 years | 11 years | 10 years | 9 years | 8 years | 7 years | 6 years | 5 years | 4 years | 3 years |
| | Panel A. Excluding the Langi | | | | | | | | | | |
| 1 if having brideprice paid | -0.126** | -0.150** | -0.136** | -0.147** | -0.043 | -0.053 | -0.104 | -0.037 | -0.048 | -0.084 | -0.280 |
| | (0.064) | (0.070) | (0.063) | (0.080) | (0.045) | (0.054) | (0.108) | (0.045) | (0.058) | (0.082) | (0.280) |
| Observations | 813 | 773 | 709 | 651 | 602 | 532 | 473 | 397 | 337 | 278 | 191 |
| | Panel B. Excluding the northern region | | | | | | | | | | |
| 1 if having brideprice paid | -0.142** | -0.175** | -0.154** | -0.173** | -0.041 | -0.050 | -0.097 | -0.048 | -0.060 | -0.095 | -0.239 |
| | (0.072) | (0.082) | (0.072) | (0.098) | (0.051) | (0.067) | (0.132) | (0.050) | (0.066) | (0.105) | (0.209) |
| Observations | 758 | 720 | 660 | 606 | 561 | 495 | 441 | 371 | 317 | 264 | 182 |

*Source*. Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015. *Notes*. This table shows the treatment effect estimate for the probability of receiving brideprice for all the available bandwidths in our data, using the alternative year of birth of 1983 as the cutoff. Reported in parentheses are standard errors robust to heteroscedasticity. Statistical significance is denoted by *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$. All regressions include dummies for ethnicity and for the region of residence at age seven as covariates. The regressions are run for females aged 24 to 49 who have ever married and born in years within the indicated bandwidth of the cutoff.

Table 2.B.10: Robustness Check of Main Results Allowing for a Differential Trend in and after 2007.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidth | 13 years | 12 years | 11 years | 10 years | 9 years | 8 years | 7 years | 6 years | 5 years | 4 years | 3 years |
| 1 if having brideprice paid | -0.126* | -0.138** | -0.104** | -0.112* | -0.021 | -0.034 | -0.070 | -0.032 | -0.038 | -0.084 | -3.72 |
| | (0.079) | (0.076) | (0.060) | (0.073) | (0.044) | (0.053) | (0.080) | (0.048) | (0.054) | (0.094) | (54.72) |
| Observations | 867 | 823 | 754 | 688 | 635 | 558 | 491 | 414 | 354 | 292 | 199 |

*Source*. Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015. *Notes*. This table shows the treatment effect estimate for the probability of receiving brideprice for all the available bandwidths in our data, allowing for a differential trend for females whose first marriage took place before or after the year of 2007. Reported in parentheses are standard errors robust to heteroscedasticity. Statistical significance is denoted by *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$. All regressions include dummies for ethnicity and for the region of residence at age seven as covariates. The regressions are run for females aged 24 to 49 who have ever married and born in years within the indicated bandwidth of the cutoff.

Table 2.B.11: The Estimated Impact of Female Education on Whether Women have a Non-agricultural Job.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidth | 13 years | 12 years | 11 years | 10 years | 9 years | 8 years | 7 years | 6 years | 5 years | 4 years | 3 years |
| 1 if having non-agricultural job | -0.002 | 0.024 | -0.018 | -0.039 | -0.016 | 0.002 | -0.047 | 0.046 | 0.028 | 0.069 | 0.061 |
| | (0.043) | (0.045) | (0.040) | (0.047) | (0.036) | (0.042) | (0.083) | (0.045) | (0.052) | (0.066) | (0.150) |
| Observations | 891 | 844 | 773 | 705 | 648 | 570 | 502 | 424 | 362 | 300 | 205 |

*Source*. Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015. *Notes*. This table shows the estimated treatment effect of female education on whether females have a non-agricultural job for selected bandwidths. Reported in parentheses are standard errors robust to heteroscedasticity. Statistical significance is denoted by *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$. All regressions include dummies for ethnicity and for the region of residence at age 7 as covariates. The regressions are run for females aged 24 to 49 who have ever married and were born in years within the indicated bandwidth of the cutoff.

Table 2.B.12: Simple Regression for Assortative Matching.

| | (1) | (2) | (3) |
|---|---|---|---|
| Outcome | Partner's years of education | | |
| Female years of education | 0.466*** | 0.467*** | 0.432*** |
| | (0.089) | (0.091) | (0.103) |
| $\mathbf{I}\{\text{Year of birth} \geq 1979\}$ | 0.075 | -0.063 | 0.110 |
| | (0.652) | (0.712) | (0.971) |
| $\mathbf{I}\{\text{Year of birth} \geq 1979\}$ | 0.033 | 0.054 | 0.102 |
| $\times$(Years of education) | (0.112) | (0.120) | (0.139) |
| Observations | 397 | 354 | 353 |
| R-squared | 0.218 | 0.452 | 0.529 |
| Premarital controls | N | Y | Y |
| Year of marriage | N | N | Y |

*Source*. Research on Poverty, Environment, Agriculture, and Technology in Uganda in 2015. *Notes*. This table shows the selected coefficient estimate of female years of education from the regression of their partners' years of education. Reported in parentheses are standard errors robust for heteroscedasticity. Statistical significance is denoted by *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$. Premarital controls include dummies for ethnicity for the region of residence at age 7. The regressions are run for females aged 24 to 49 who have ever married and were born in years within the indicated bandwidth of the cutoff. Due to missing values, the number of observations differs across regressions.

# Chapter 3

# Pregnant in Haste? Evidence of Reproductive Behaviours in Uganda

## 3.1 Introduction

In many low income countries, and particularly in Sub-Saharan Africa (SSA), fertility rates as well as maternal and infant mortality rates remain very high: the total fertility rate was 4.9; 541 mothers died per 100,000 live births; and 56.4 infants died per 1,000 live births.[1] One of the factors contributing this situation, which however has not been well-studied in economics, is the short birth spacing. Appropriate spacing of births can reduce the total fertility rate by decreasing the number of children a woman can have during her reproductive period. It can decrease the burden on fatigued uterus, which reduces

[1] World Development Indicators by World Bank, retrieved on the 30th of August, 2017, at http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators.

nutrition and protection for a fetus[2] It can also help households to increase human capital investment for children by allowing them to increase the income levels (e.g., Buckles and Munnich (2012)). Given the influence of birth outcomes on socio-economic welfare throughout one's life (Fletcher, 2011; Black et al., 2007; Bharadwaj et al., 2018; Behrman and Rosenzweig, 2004), it is of great importance to investigate the causes of birth spacing. The results are likely to shed light on a healthy start of lives, which in turn leads to long-term welfare.

This study investigates whether pregnancy loss due to miscarriage and stillbirth affects birth spacing for the subsequent pregnancies. While it is known that variables such as household income and female education matter in explaining birth spacing (Heckman and Walker, 1990; Bhalotra, 2010; Kim, 2010), the effects of past pregnancy loss experience have not yet been fully examined. Related studies have found that infant mortality shortens subsequent birth spacing (Whitworth and Stephenson, 2002; Bhalotra and van Soest, 2008; Maitra and Pal, 2008; van Soest and Saha, 2018). However, since child mortality is observed only when the child is born alive, the analysis may produce a misleading conclusion if one only focuses on child mortality and ignores the impact of fetus loss.

Another related strand of literature examines the impact of birth spacing on socio-economic behaviours and outcomes of children and mothers (Buckles and Munnich, 2012; Karimi, 2014). These studies utilise the fact that miscarriage exogenously lengthens the spacing interval for the next birth since a woman loses time for the lost pregnancy and recovery (details are discussed later). In other words, it is well-established as the

[2] Economic studies also show that appropriate birth spacing increases birth weight (e.g., Rosenzweig and Wolpin (1988)).

first-stage results that pregnancy loss positively affects the birth interval. However, it is not well-known whether the intervals of the following births are affected.

We fill in this gap in the literature by more comprehensively studying the impact of miscarriage and stillbirth on the subsequent reproductive behaviors in the long run. For this purpose, we use data containing a large number of pregnancies per woman from a developing country, Uganda. This is important since the health risks of mothers and children are arguably larger in developing countries as shown in the mortality rates, and also seen in the lack of easy access to decent reproductive services and protective social institutions (Bhalotra, 2010). While it may be difficult to address this question in data from developed countries with lower fertility rates, the issue is of relevance to these countries as well. That is, as women choose to become pregnant at a later stage of their lives, the chance of pregnancy loss is biologically increased and so is the incidence of hasted conception, which can be detrimental to the maternal health.

Methodologically, the identification of the relationship between birth spacing behaviours and pregnancy loss due to miscarriage and stillbirth rests on whether pregnancy loss occurs at random. Medical studies suggest that it indeed does occur at random conditional on a few factors. The most common cause of miscarriage and stillbirth is chromosomal abnormality of the fetus, accounting for 50% to 80% of all cases (Simpson, 2007). The abnormalities are due to mal-separation of chromosomes during meiotic division, which is likely to be unpredictable conditional on a few factors that can be easily controlled for such as age and fixed effects (Simpson, 2007; Silver et al., 2007; Brown, 2008; Larsen et al., 2013).[3]

Our analysis using data from the Demographic and Health Survey (DHS) in Uganda in 2011, as well as those from the Research on Poverty, Environment, Agriculture, and Technology (RePEAT) in Uganda in 2015, brings quite a few pieces of new evidence. First, we find that women with a pregnancy loss lengthen the birth interval for the pregnancy immediately after the loss, but shorten the intervals for all the subsequent pregnancies. The longer interval for the first post-loss pregnancy is mechanical, as the child born just after a loss has the spacing interval not from the timing of the loss but from the birth of the child born just before the loss.[4] On the other hand, the intervals for all the subsequent pregnancies are significantly shorter by 4 to 8 months. We find the shortening effect gradually diminishes as women experience more successful live births, but it persists for the entire fertility after the pregnancy loss. To the best of our knowledge, this shortening effect has not been presented in the economic literature, perhaps because existing related studies examine fertility behaviours in richer countries where the fertility rate is lower than in Uganda. We also find that this shortening effect is heterogeneous across several dimensions: in particular, we find it stronger for women who were older when they lost a pregnancy, for women who had given fewer live births before the loss, and for women who were less educated.

We also attempt to explore the mechanism for why birth spacing intervals change. In

[3] Human reproduction is said to be very inefficient, in the sense that around 60% of all fertile eggs miscarry without being ever recognised, and 15% to 20% of the recognised pregnancies also end up miscarrying (Brown (2008), Larsen et al. (2013)).

[4] This lengthening effect can be considered random holding a planned number of children constant, and has thus been exploited as an exogenous variation by several studies (e.g., Buckles and Munnich (2012), Karimi (2014), Bratti et al. (2016)).

particular, we utilise an idea from a structural model by Mira (2007). She considers the process of updating the perceived probability of infant death, which can depart from its true probability, where the updating is modelled to occur every time females observe the survival status of the children they have already given birth to. While her study finds the change in birth spacing due to the belief updating is small, this hypothesis can explain, if any, a life-long impact of miscarriage and stillbirth on subsequent birth spacing behaviours. The belief updating mechanism of Mira (2007) is consistent with the behavioural theory which posits that the realization of an event with a small probability may change one's belief about its occurrence probability, and it may lead to a behavioural change thereafter (Hertwig et al., 2004). This theory suggests that a realisation of pregnancy loss may change mothers' belief on its probability, leading to a change in their birth spacing behaviours as long as they re-optimize their subsequent fertility schedule according to their updated belief.

We then analyse whether the behavioural change, the shortening of birth spacing intervals, stems from the updating of belief on the probability of pregnancy loss. Although beliefs are becoming increasingly important in development studies (Delavande, 2014), the surveys of our data did not ask about the subjective belief on pregnancy loss probability. Instead, we construct a measure for the realised probability of pregnancy loss using each woman's fertility history and loss experience. Our regression using this personal pregnancy loss probability shows that they adjust birth spacing in such a way that a larger probability of pregnancy loss leads to a shorter spacing interval. Moreover, our measure of the personal pregnancy loss probability explains the persistence of the birth spacing effect very well. These findings support the hypothesis that Ugandan females decide birth spacing according to the probability of pregnancy loss, adding new piece of evidence to the growing literature on belief updating and behavioural analyses.

The rest of the paper is organized as follows. Section 3.2 reviews the medical literature on miscarriage and stillbirth, and economic literature related to birth spacing. Section 3.3 describes the data set and major variables used in this study. Section 3.4 provides the details of our estimation strategy. Section 3.5 presents the results from summary statistics and regression analyses. Section 3.6 explores the mechanism for the findings on birth spacing behabiours. The last section concludes.

## 3.2 Related Literature

### 3.2.1 Medical studies on miscarriage and stillbirth

Miscarriage refers to the loss of pregnancy before 20 to 23 gestational weeks, and the loss which occurs later is termed stillbirth (Brown (2008), Larsen et al. (2013)). Pregnancy loss like these are also referred to as spontaneous abortion and sporadic pregnancy loss. Such a loss is not uncommon, accounting for 10% to 15% of all clinically recognized pregnancies (van den Berg et al., 2012). Since the cutoff length of gestation that divides miscarriage and stillbirth varies across studies even in the literature of obstetrics and gynecology,[5] and also since the data at hand (explained in more details later) do not distinguish the two, we refer to such terminated pregnancies as spontaneous pregnancy losses, or simply

[5] For instance, Brown (2008) uses 23 gestational weeks, whereas Larsen et al. (2013) seems to use 22 weeks. Silver et al. (2007) uses 20 weeks but makes clear the cutoff every time they cite other studies.

pregnancy loss, throughout the paper hereafter.

There are quite a few potential causes of pregnancy loss discussed in the literature, but by far the most important is the genetic reason (Brown (2008), Larsen et al. (2013), Ford and Schust (2009), Silver et al. (2007)). This type of pregnancy loss originates from the meiotic division, which may involve a malsegregation of the pairs of chromosomes[6] and result in chromosomal abnormality of the fetus, such as sex chromosomal polysomies. It is estimated that 50% (Brown, 2008) to 80% (Simpson, 2007) of all pregnancy losses are said to be associated with genetic problems. It seems a consensus that chromosomal abnormality is more likely observed for older mothers, but no other factors are reported to be consistently associated with the probability of such anomalies. Another genetic cause for pregnancy loss is said to be parental karyotype abnormality (Ford and Schust, 2009), although this is considered much less frequent (accounting for 3% to 6% of pregnancy loss cases, Larsen et al. (2013)).

Anatomic factors are also said to be associated with pregnancy loss. They include uterine malformation, which is observed in about 5% of all females but 15% of those who has an experience of pregnancy loss (Larsen et al., 2013), and uterine fibroid (Brown, 2008), although their causal influence has yet to be revealed (Brown, 2008). Some studies point to immunologic causes of pregnancy loss[7] and endocrine etiologies.[8] These factors are, however, found to be associated only through statistical correlations, and the causal

[6] Chromosomal abnormality, therefore, can be caused by both males and females. However, the extra chromosome, one of the most commonly reported causes of chromosomal abnormality, is mostly of maternal origin (Larsen et al., 2013). Therefore, although not strictly perfect, it may be enough to control for maternal characteristics in our particular context.

[7] Examples include antiphospholipid, antinuclear, and antithyroid antibodies.

[8] Examples include hypothyroidism.

mechanism has not been revealed yet (Ford and Schust, 2009). Thrombophilia is also suspected to be associated with miscarriage and stillbirth, though it is said to be caused by genetic attributes, and its causal impact on pregnancy loss has not been clarified (Tomer, 2014).

Infectious diseases are also said to increase the likelihood of pregnancy loss (Silver et al., 2007). Infection can be due to virus or bacteria, where the latter is said to be more common in developing countries. In developed countries where clinical data are available, around 10% to 25% of stillbirths are said to be caused by infection. Some infectious diseases have seasonality, and other infections may be related to sanitary conditions such as toilet.

Lifestyle factors, such as smoking and alcohol, are suspected to risk pregnancy losses, yet the evidence seems to be quite mixed and unclear. For instance, Brown (2008) cite several studies that found association between cigarette smoking and pregnancy loss, but Larsen et al. (2013) and Ford and Schust (2009) note that the past evidence is based on mere correlations. As for coffee intake, while coffee drinking is accepted in many countries (Larsen et al., 2013), Ford and Schust (2009) draws a study showing that caffeine has a dose-dependent relationship with pregnancy loss. Alcohol consumption seems an exception such that many studies agree upon its adverse effect (Brown, 2008; Silver et al., 2007; Larsen et al., 2013; Ford and Schust, 2009). Obesity is said to be associated with pregnancy losses only for the extreme with the body mass index larger than thirty (Larsen et al., 2013).

To summarize, by far the most important cause of pregnancy loss is genetic abnormality, which depends in part upon age but happens otherwise in an unpredictable manner during meiotic division. Anatomic factors are also large in proportion, and there could be other factors related to immunity, endocrine, and thrombosis. This suggest that, although the direct causes of these factors are not fully understood, these may be controlled for by a female's age and fixed effects to the extent that these factors persist and continue to affect females' reproductive activities. Infection may be time variant, but at least seasonality effects and sanitary conditions can be controlled for. Lastly, behavioural risk factors including smoking and drinking may also affect both pregnancy loss and spacing interval, but these factors are usually persistent and some can be checked in the data. Overall, most of the possible confounding factors in analysing the effect of pregnancy loss on birth spacing are likely to be controlled for using observable data.

### 3.2.2 Causes and consequences of birth spacing

While we study the impact of pregnancy loss on birth spacing, a large body of literature suggests that a longer interval improves the health of the child and the mother, underscoring the importance of understanding the determinants of birth interval. For example, a longer spacing interval is shown to increase birth weight (Rosenzweig and Wolpin, 1988) and reduce preterm delivery and pregnancy-related complications (Norton, 2005; Conde-Agudelo et al., 2006). Later socio-economic outcomes, such as children's schooling ((Buckles and Munnich, 2012; Sawada and Lokshin, 2009)) and marriage (Vogl, 2013), are also found to improve by spacing interval.

This literature faces the methodological difficulty of endogeneity in estimating the effect of birth spacing, which is a choice of individuals (Winikoff, 1983). Among the studies above, a few have already used miscarriage as an exogenous variation to address the endogeneity problem. One such study is Buckles and Munnich (2012), which estimates the effect of birth spacing on educational outcomes. They note the possible endogenity of birth spacing choice and thus instrument it by the presence of miscarriage that lengthens spacing interval for the next birth. They find that birth spacing increases test scores of children. Karimi (2014) considers the effect of spacing on mother's labour market outcomes, and find that the longer intervals increase mothers' labour force participation and push up their wage growth path. Bratti et al. (2016) estimates the effect of sibship size on the migration decision among children, finding that the raw correlation between family size and migration may not hold once the endogenous choice of family size is accounted for. In a similar vein, Hotz et al. (2005) and Miller (2011) use the first-pregnancy miscarriage that exogenously reduces teenage motherhood and find its positive impact on the females' subsequent careers.

Studies that consider spacing interval as an outcome are relatively scarce in economics. Among the few related studies, Heckman et al. (1985) stresses the importance of controlling for unobserved heterogeneity, without which a researcher may reach a misleading conclusion on the relationship between birth spacing and other socio-economic variables. Kim (2010) shows that in Indonesia, modern education for females increases their birth spacing intervals. Perhaps one of the most analysed topics may be whether an infant death shortens the birth spacing for the next child (Whitworth and Stephenson, 2002; Bhalotra and van Soest, 2008; Maitra and Pal, 2008; van Soest and Saha, 2018). In particular, Bhalotra and van Soest (2008) and van Soest and Saha (2018) consider the relationship between infant mortality and birth spacing where the primary interest lies in whether a neonatal death affects birth spacing for the next pregnancy, while noting the possibility that short birth spacing may increase infant mortality for subsequent pregnancies. They report that birth spacing of subsequent pregnancies is significantly shortened by neonetal death, even after accounting for the reverse effect of infant mortality on subsequent birth spacing.

Our study, related to the two strands of the literature, is considered unique in at least two ways. First, the spacing interval seen as an outcome has not been analysed in relation to spontaneous pregnancy loss. Studies on infant mortality such as Bhalotra and van Soest (2008) share a research question close to ours, but their focus is on past neonatal mortality. Second, existing studies on pregnancy loss may be insufficient to consider the fertility behaviours and policies in the developing world, since females get pregnant more times, and the analysis of longer-term reproductive behaviours is called for (Bhalotra, 2010). We provide evidence on the longer-term effects of pregnancy loss, and demonstrate that the lengthened interval for the pregnancy right after the loss is only one aspect of the entire impacts of pregnancy loss.

## 3.3 Data

### 3.3.1 Demographic and Health Survey in Uganda in 2011

In order to investigate the relationship between pregnancy loss and birth spacing, we primarily use the data of the Demographic and Health Surveys (DHS) in Uganda conducted in 2011. The data were collected by the Uganda Bureau of Statistics and ICF International Inc. in May through December in 2011. The survey covered a nationally representative sample of 10,086 households from which 9,247 females aged 15-49 years are found. These females were asked about household characteristics, their socio-economic activities, and reproductive behaviours. As to pregnancy-related variables, their pregnancy history was first queried. For each pregnancy that resulted in live birth, further information was recorded, such as the year and month of the birth and place of delivery.

One of the variables of interest in this study is the experience of miscarriage and stillbirth. The DHS first asks 'Have you ever had a pregnancy that miscarried, was [aborted] or ended in a stillbirth?'[9] If the respondent answers yes to this question, it further asks the year and month when the last such pregnancy ended. There are two issues about this measurement that are worth discussing at this moment. First, it is possible that a female experiences multiple and/or consecutive miscarriages or stillbirths, although the probability is small (Ford and Schust, 2009). The DHS asks about whether respondents had more than one pregnancy loss, but does not record when the prior losses occurred.

[9] The DHS Uganda 2011 questionnaire has this question: "Have you ever had a pregnancy that miscarried, was or ended in a stillbirth?" However, the recode manual (Demographic and Health Surveys, 2013) describes the variable as '[w]hether the respondent ever had a pregnancy that terminated in a miscarriage, abortion, or still birth, i.e., did not result in a live birth.' Therefore, we inserted the seemingly missing 'abortion' and enclosed it with square brackets.

Because of this measurement, the data at hand may underestimate the magnitude of the effect of miscarriage or stillbirth for females with multiple pregnancy losses. We discuss this in more details in subsection 3.3.2. Second, this measurement does not separate abortion from miscarriage or stillbirth. If abortion is selectively performed, our analysis using the DHS data may be biased in any arbitrary way. However, we use a different data set which measures miscarriage and stillbirth separately from abortion and show that results are qualitatively unchanged. For details, see Section 3.5.3 and Appendix 3.D.

Another important variable of ours is the birth spacing interval. In this study, it is defined as the interval, measured in months, between the end of the last pregnancy and the end of the current pregnancy.[10,11] The WHO defines birth spacing differently, as it measures the interval between the end of a pregnancy and the conception of the next one (World Health Organization, 2007). We use a different measurement, since our DHS data do not contain information about the length of gestation, which disallows us to compute a birth spacing measure consistent with the WHO definition. This difference is, however, unlikely to invalidate our econometric analysis, since it is absorbed by the intercept of an estimation equation.[12] These two variables, the pregnancy loss and birth spacing, are the main variables throughout the paper.

[10] Thus defined, birth spacing is missing for the first pregnancy of each female.

[11] The spacing interval for the birth that immediately follows a miscarriage or stillbirth can also be measured from the timing of the pregnancy loss. However, we define the interval this way, following other previous studies such as Buckles and Munnich (2012) and Karimi (2014) so that we can compare our results with theirs.

[12] One might be concerned of the possibility that the length of gestation for pregnancies conceived after pregnancy loss may systematically differ from that of pregnancies before loss. However, so far we have not found any convincing literature that suggests a systematic change in premature, or post-mature, births specific to post-loss pregnancies.

### 3.3.2 Descriptive analyses

Summary statistics of main child-level variables are presented in Table 3.1, while those of the mother-level variables are shown in Table 3.2. There are around 6,400 females with at least one pregnancy experience and 28,600 children born to them in our data.[13] Due to some missing data and unknown responses, the number of observations is not strictly the same across variables.

Of the primary interest in our study is the spacing interval between pregnancies. Measured as the period between the ends of one pregnancy and its previous one, the mean length is 31.9 months. This is shorter than is recommended by the WHO, which is at least 33 months, or 2 years plus 9 months. Consistently, the dummy for whether birth spacing is shorter than the WHO recommendation has the mean of 0.656, suggesting that more than half of the spacing intervals may be too short in the DHS data. The smaller number of observations for the spacing interval, by about 6,450, than for other easily observable measures such as the age at pregnancy termination presents the number of the first pregnancies, as the spacing interval cannot be defined for the first pregnancy of each female.

Out of all the pregnancies available in our data, approximately 10% were conceived after a pregnancy loss. When decomposed, 3.6% are the first post-loss pregnancy, 2.2% the second, and 1.5% the third. We will put together the fourth and subsequent post-

[13] The panel variables shown in Table 3.1 are summarised at the child level, where every child is counted as one observation regardless of the number of children born together. We note the possible selection of females who deliver multiple births (Bhalotra and Clarke, 2016). In our regressions, we primarily focus on pregnancies (multiple births are counted as one observation), while we confirm the robustness of our conclusion when we count each child as one observation.

loss pregnancies into one category hereafter, since the fourth and subsequent post-loss pregnancies account for less than one percent without the rounding of decimal numbers. Other variables in our data set that vary over the pregnancy history include female's age at the birth of the child, an indicator for whether the pregnancy ended in a single birth or multiple births, sex of the child,[14] and year and month of birth of the child (not shown for brevity).

Cross-section variables observed once for each female are summarised in Table 3.2. Panel A lists the number of children and the number of pregnancies of surveyed females in the DHS data. The two numbers differ for those who had one or more multiple births, although the difference is small. On average, females in our data had 3.25 pregnancies and 3.3 children by the time they were surveyed. When we focus on females who have ever given birth, the figures rise to 4.41 pregnancies and 4.48 children. These numbers are much larger when we limit the sample to older females: females older than 40 years had 7.04 pregnancies and 7.15 children.

Panel B of Table 3.2 shows the summary statistics for other cross-section variables. Around a quarter of females have ever experienced at least one pregnancy loss. It suggests that the reported experience of pregnancy loss is likely to be a good measure of the actual experience of pregnancy loss, since the variable does not seem to be picking up the variation of a limited number of females. The share of females who had more than one pregnancy loss in their lifetime is about 1.5%. Other cross-section variables include

[14] Sex cannot be defined for multiple births when the unit of analysis is a pregnancy rather than a child. In such a case, we define a third category for multiple births and use the three-category dummies in the regressions that follow.

females's age, education, marital status, religion, ethnicity, and region of residence.[15]

## 3.4 Empirical Strategy

### 3.4.1 Estimation model

This study is aimed at investigating the birth spacing response to pregnancy loss. Seeing the data structure as an unbalanced panel where the time dimension is expressed by the number of pregnancies, we estimate the following simple model by ordinary least squares (OLS):

$$y_{ij}^{\text{panel}} = \phi_j + \delta_1 D_{ij}^{\text{post}} + x_{ij}^{\text{panel}} \delta_2 + v_{ij} \tag{3.1}$$

where $\phi_j$ represents woman fixed effects (FE), and $x_{ij}^{\text{panel}}$ control covariates of pregnancy $i$ of woman $j$, including female's age at the termination of pregnancy $i$, the month and year dummies of pregnancy $i$, and the sexes of child $i$ and of the previous child. In equation (3.1), the parameter of interest is $\delta_1$, which measures the effect of a pregnancy loss on the spacing intervals of subsequent pregnancies, holding constant the other covariates. Because the major potential confounding factors identified in Section 3.2.1 such as woman FE and age are controlled for, we interpret $\delta_1$ as the causal average difference in spacing intervals between the pre- and post-loss pregnancies of the woman who experienced a pregnancy loss.

It is possible that the effect of a pregnancy loss changes in the course of pregnancy history. For example, the effect may be large when the memory of a pregnancy loss is still

[15] Statistics for religion, ethnicity, and region of residence are not shown for brevity.

fresh in the female's mind, but it may diminish as she experiences several successful live births thereafter. In order to examine this possibility, we modify equation (3.1) as

$$y_{ij}^{\text{panel}} = \phi'_j + \sum_{l \in L} \delta_1^l D_{ij}^{\text{post},l} + x_{ij}^{\text{panel}} \delta'_2 + v'_{ij} \tag{3.2}$$

where $D_{ij}^{\text{post},l}$ equals to one if pregnancy $i$ of woman $j$ is her $l$-th pregnancy since her last pregnancy loss episode and zero otherwise. For example, for woman $j$ who experienced a pregnancy loss after two live birth, $D_{4,j}^{\text{post},2} = 1$ since her fourth live birth is realised as her third live birth after her pregnancy loss experience, while $D_{5,j}^{\text{post},4+} = 0$ since her fifth live birth is the third live birth after her pregnancy loss, not the fourth or later pregnancy after her loss episode. The set $L$ is one of $\{1, 2+\}$, $\{1, 2, 3+\}$, and $\{1, 2, 3, 4+\}$, and denotes the decomposition of post-loss pregnancies, where $l+$ denotes the $l$-th and all the subsequent post-loss pregnancies.[16] In this equation, the parameters of interest are $\delta_1^l$, which represents the change in birth spacing at the $l$-th pregnancy after loss. If the effects differ over the post-loss pregnancies, we obtain differing estimates for $\delta_1^l$ which we interpret as the persistence of the effect of pregnancy loss. When we present estimation results, we write $L_1 \equiv \{1\}$, $L_2 \equiv \{1, 2+\}$, $L_3 \equiv \{1, 2, 3+\}$, and $L_4 \equiv \{1, 2, 3, 4+\}$.

[16] In our data, the maximum number of pregnancies conceived after the last loss is 13. However, 4+ is taken as the maximum number of post-loss pregnancies to be considered in this framework. This is because the proportion of pregnancies at the fourth or higher post-pregnancy loss is less than 1% in our data (Table 3.1), and thus the estimates become unstable and sensitive to only a few observations.

### 3.4.2 Trend of the birth spacing behaviour prior to pregnancy loss

In order to examine whether the estimation models in equations (3.1) and (3.2) identify the behavioural response to a pregnancy loss experience, we check the trend in birth spacing prior to a pregnancy loss. This corresponds to the standard test of the pre-treatment trend in the outcome in a difference-in-difference estimation,[17] relating to the parallel trend assumption that the two groups of subjects are indifferent were it not for the treatment. The idea behind the pre-trend check is to examine how similar the two groups of females—those who have ever had a pregnancy loss and those who have not—were in birth spacing practices before the pregnancy loss. The more similar they are prior to the pregnancy loss, the more similar the birth spacing behaviour would be had there not been the pregnancy loss, which adds credibility of our results as the causal estimate of the response to a pregnancy loss in birth spacing.

In order to examine this, we first regress birth spacing on parity, pregnancy loss status, and their interactions for pregnancies that should not be affected by pregnancy loss, i.e. all parities for females with no loss experience, and pre-loss parities for females with loss experience. The coefficients of the loss status and interaction terms are reported in Table 3.3. For a meaningful comparison, regressions use females who have attempted at least the number of pregnancies, unaffected by pregnancy loss, specified in the column title—for instance, column (1) uses females who have made at least three pregnancy attempts[18] and compare their spacing length at the second parity; likewise, column (2) uses females with

[17] Since equations (3.1) and (3.2) includes woman FE and the post-pregnancy loss dummy (dummies), our parameters of interest can be interpreted as a difference-in-difference estimator.

[18] Among females with no pregnancy loss, females with three live births are included, and among females with pregnancy loss, those with two live births and a pregnancy loss at her third parity are included.

at least four pregnancy attempts, in which case we can estimate the coefficients for both the dummy and trend potentially different between females with and without pregnancy loss. This sample restriction is not an arbitrary choice to make the two groups of females more similar for this pre-trend analysis, since it is based on the number of pregnancies as of the survey which can be controlled for by the fixed effects in our estimation equations (3.1) and (3.2).

The coefficient estimates in Table 3.3 fail to indicate a systematic difference in pre-loss birth spacing behaviours between females with and without pregnancy loss experience. We observe that the estimates for the loss status and differential trend are generally small and insignificant. Since even the significant estimates take different signs and are unstable, we conclude that the significance is just by chance. Therefore, the pre-trend is not as different between the two groups of females as to threaten the estimation of the effect of a pregnancy loss on birth spacing by equations (3.1) and (3.2).

Next, we graphically compare the birth spacing between females with and without pregnancy loss experience. Figure 3.1 shows that the pre-loss birth spacing intervals for females with pregnancy loss experience are close to the intervals for females with no such experience.[19] Average spacing intervals are noisier for females with pregnancy loss particularly at higher-order parities due to the small sample size shown at the bottom of the figure. However, the average intervals for the two groups are generally statistically indistinguishable from zero at each parity.[20] In addition, we add a dummy that takes the

[19] Spacing intervals for females with pregnancy loss disaggregated by the parity of the pregnancy loss are shown in Figure 3.A.1.

[20] See also the difference estimated at each parity presented in Table 3.B.1 and find that the differences are insignificantly estimated at any parity.

value of one if pregnancy $i$ of female $j$ is the last one that precedes her lost pregnancy to equations (3.1) and (3.2). The results are presented and discussed in Section 3.5.1 confirm our above findings that females before a pregnancy loss episode are similar to females with no such experience in terms of the birth spacing behaviour. These results suggest that the two groups of females are unlikely to be different in birth spacing intervals before the pregnancy loss.

### 3.4.3 Falsification test

Before moving onto the main part of the regression analysis, we discuss the results of a falsification test. The idea here is to test whether the pregnancy loss dummy does indeed have no correlation with pre-determined characteristics of women. While no correlation does not prove the exogeneity of pregnancy loss, it provides supportive evidence for it.

In our DHS data, pre-determined characteristics such as religion, ethnicity, and native languages should not be systematically correlated with pregnancy loss experience. Table 3.4 shows the results of the regressions of these variables on $D^{\text{ever}}$ in a cross-section setting. For religion (columns (1) and (2)) and native language (columns (6) to (8)), the coefficients are precisely estimated to be close to zero. For ethnicity, although the single largest ethnic group, the Baganda, and the top 10 ethnic groups have insignificant correlation (columns (3) and (5)), the top 5 ethnic groups are more likely to lose a pregnancy (column (4)). This may be due to the sampling error: the second largest ethnicity Banyankole (accounting for 8.88% of the sample women) and the third Basoga (8.18%) have the mean of $D^{\text{ever}}$ of 0.267 and 0.294, respectively, both higher than the overall mean of 0.243.

These results from the pre-trend check and the falsification test are supportive of our identification that utilises the pregnancy loss as an exogenous variation to examine changes in spacing intervals.

### 3.4.4 Selection into Higher-Order Parities

A potential threat to the identification of the effect of pregnancy loss on birth spacing behaviours that follow is that the pregnancy loss experience may affect the decision on whether to continue reproduction. If, for instance, females who lost a pregnancy are systematically more, or less, likely to continue reproduction after their loss experience at any post-loss parity, comparison of birth spacing behaviours between females with and without pregnancy loss may no longer permit causal interpretation, since the attributes correlated with the post-pregnancy loss (non-)attrition may be the driving factor for birth spacing difference and they may be unobservable in the data.

To see the severity of this potential selective (non-)attrition, we regress a set of birth indicators on observable characteristics and their full interaction with the pregnancy loss indicator, and test whether the interactions are significantly partially correlated with fertility progression. Specifically, using the sample females who have given at least $k$ live births without pregnancy loss or with loss at the $(k+1)$-st pregnancy attempt, we consider the equation of the form:

$$\Pr\{\text{give } (k+t)\text{-th successful birth}\}_j = \beta_0 + X_j\beta_1 + D_j^{\text{ever}}\left(\beta_2 + X_j\beta_3\right) + e_j, \qquad (3.3)$$

for $t = \{1, 2, 3, 4\}$, and test whether the estimates for $\beta_2$ and elements in vector $\beta_3$ are significantly different from null. That is, we examine whether the probability of giving the $(k + 1)$-st live birth differs between females with $k$ live birth and no loss and others with a loss after their first $k$ live births, as well as whether such differential probability, if any, is correlated with available observable covariates. We conduct this test for $k = 0, 1, ..., 7$ since the average number of children that females aged 45 and above have is 7.2 (Table 3.2). See Appendix 3.C for more details.

The results, presented and discussed in Appendix 3.C, fail to find any characteristics that systematically predict differential selective (non-)attrition between females with and without pregnancy loss, among the observable covariates included in the regressions. This suggests that, at least for the first four parities after pregnancy loss and among the variables considered, our analysis based on equations (3.1) and (3.2) may not suffer large selection bias related to pregnancy loss. Thus, we discuss the results from our main regression analyses as causal evidence in the next section.

## 3.5 Results

### 3.5.1 The effect of a pregnancy loss on spacing intervals

Our main results for the effect of pregnancy loss on spacing interval are presented in Table 3.5. Each column shows the results from separate regressions based on equations (3.1) and (3.2), with different sets of $D^{\text{post}}$ dummies (denoted by $L_1$ through $L_4$). Columns (1) through (4) use a pregnancy as the unit of observation where a multiple birth is counted as one observation, while columns (5) through (8) use a child as the unit of observation where one pregnancy is repeatedly used if it resulted in a multiple birth.

We find from Table 3.5 that the single dummy specification in column (1) does not reveal any significant change in spacing interval. However, once the first post-loss pregnancy is separated from the other subsequent pregnancies, we find that the first post-loss pregnancy has a longer spacing interval, whereas the subsequent pregnancies have significantly shorter spacing intervals. Point estimates in column (4) suggest that the first pregnancy after a loss have an approximately 4.5 months longer spacing interval, but the subsequent pregnancies have a shorter interval by 4.5 to 8.1 months. The positive effect for the first post-loss birth is mechanical, as the birth spacing in our analysis is measured as the interval between the end of the present pregnancy and that of its previous successful pregnancy. This positive effect has been used as an exogenous variation to examine the relationship between birth spacing and other socio-economic variables such as child education (Buckles and Munnich, 2012) and mother's career (Karimi, 2014). On the other hand, the negative effects for all but the first post-loss pregnancies have not been reported, to the best of our knowledge, in the related economic literature. The point estimates, −8.1 for the second, −5.7 for the third, and −4.5 for all the subsequent post-loss pregnancies, suggest that the negative effects diminish as females experience successful live births after the pregnancy loss, but appear to persist over the course of their pregnancy history.

Estimates in columns (5) to (8) of Table 3.5 show that the above results are virtually unchanged when all the children born in multiple births are used in the regressions. These results suggest the robustness of our findings that the first pregnancy after the loss has a longer spacing interval while the subsequent pregnancies have shorter spacing intervals.

Table 3.6 shows the estimated effects of pregnancy loss on different measures of birth spacing. In columns (1) to (4), we take the logarithm of spacing intervals, as the distribution of birth spacing is skewed to the right. In columns (5) to (8), we draw on the WHO recommendation of 33 months as the threshold to construct an indicator for whether the spacing interval is considered too short. Arguably, these two alternative measures are less sensitive to a few observations with extreme values.

The results in Table 3.6 suggest that our findings in Table 3.5 are robust to different measurements of the outcome. The sign of the estimates is flipped in columns (5) through (8), since the dichotomised outcome takes the value of one if the spacing interval is *shorter* than the WHO recommendation. The most stark difference from Table 3.5 is the decrease in the probability that the spacing is too short in the single dummy specification, shown in column (1). However, once the first post-loss pregnancy is separately handled from the other subsequent pregnancies, the estimated effects are all consistent with the estimates in Table 3.5.

### 3.5.2 Survival analysis for spacing intervals

One concern about the above estimates is that more pregnancies are observed for younger ages of mothers than for older ages. That is, for older females, pregnancy outcomes are observed for both their younger ages and older ages, while for younger females pregnancies are realised only for their younger ages, and their future outcomes have not yet to be seen. Thus, it may be suspected that the shorter spacing intervals after a pregnancy loss may apply only to older or younger mothers, driving the results in Tables 3.5 and 3.6. Moreover, the spacing interval data is generally censored unless females in the sample just deliver a child at the time of the survey interview. These econometric problems may create bias in our simple estimation results.

In order to address these concerns, we follow the past studies and estimate a similar model using the Cox proportional hazard estimation. Specifically, we consider the proportional hazard model of the form:

$$\lambda(t|x) = \lambda_0(t) \cdot \exp\left(\alpha_0 + \alpha_1 D_{ij}^{\text{post}} + x_{ij}^{\text{panel}}\alpha_2 + r_{ij}\right) \tag{3.4}$$

and

$$\lambda(t|x) = \lambda_0'(t) \cdot \exp\left(\alpha_0' + \sum_{l \in L} \alpha_1^l D_{ij}^{\text{post},l} + x_{ij}^{\text{panel}}\alpha_2' + r_{ij}'\right) \tag{3.5}$$

where $\lambda_0(t)$ and $\lambda_0'(t)$ are the baseline hazard functions left unspecified, and other variables are defined in the same way as in equation (3.1). Including in the covariate vector the age of female $j$ and its square at either the end of pregnancy $i$ or the survey interview, and year and month dummies, we estimate the parameters in the log relative hazard.

It should be noted that this estimation is similar to, but not the exact counterpart of, the panel estimation of the models (3.1) or (3.2). One reason is that the control covariates differ, since some of the panel covariates are undefined for censored duration.[21] Another reason is that the unobservable heterogeneity is not controlled for in this hazard estimation,

[21] Specifically, the child sex is undefined for censored durations, and thus omitted from the hazard estimation. This would not make much difference in the estimation, however, as the coefficient estimates for the child sex are always very small and insignificant throughout.

as compared to the fixed effects in the panel estimation.[22] Nevertheless, as the pooled OLS estimation omitting the female fixed effects produces the qualitatively similar results (see Appendix Table 3.B.4),[23] it is unlikely that the results would be strongly biased in this proportional hazard estimation.

We present the plot of survival probability as a function of months since the end of the previous pregnancy in Figure 3.2.[24] The figure shows that our main finding remains intact: the first post-loss pregnancy has a longer spacing interval, whereas all the subsequent pregnancies have shorter intervals. One major difference from the panel estimation is the insignificant coefficient estimate for the third post-loss pregnancy, which may stem from the relatively small sample size of females with this variable equal to one, or the difference in the model specification. However, since the overall finding is otherwise remarkably similar in both panel and hazard estimations, it can be said that the negative spacing effect of a pregnancy loss on the second and subsequent post-loss pregnancies is robust to the estimation model.

### 3.5.3 Measurement of the pregnancy loss

The above findings are obtained from the DHS data. However, the DHS data contain the timing of the most recent pregnancy loss only. Furthermore, they do not differentiate

[22] We attempted to include an unobserved heterogeneity term multiplicatively, which is assumed to be independent of the timing of the entry to reproductive activity and of all the variables in the model, but the practical estimation turned out to be so computationally demanding and unbearably time-consuming that we decided to abort the estimation programme.

[23] The pooled OLS estimation results with the same set of covariates appear to contain upward bias (columns (1) to (4) of Appendix Table 3.B.4). However, after conditioning on a few cross-section controls (age of females and its square, years of education, and age at the first sexual intercourse), estimated effects reported in columns (5) to (8) seem to be qualitatively similar to the main results.

[24] Coefficient estimates for the parameters in the log relative hazard are reported in Appendix Table 3.B.5.

miscarriage and stillbirth from induced abortion, either. We show that these differences in the measurement do not change the conclusion altogether, by using our secondary data collected from the RePEAT survey. One feature of the RePEAT data is that it attempts to separate the measurement of abortion from that of miscarriage and stillbirth, in a way that does not allow respondents to falsely report abortion experience as miscarriage, which is reported to be one of the sources of measurement errors in low income countries (Singh et al., 2018). Another is that it allows us to observe more than one pregnancy loss experience, with the timing information for each loss. While the details are found in Appendix 3.D, we highlight the main findings in this sub-section.

We first compare the main results for estimation equations (3.1) and (3.2) using the DHS and RePEAT data. The results presented in Table 3.D.4 show that the two data sets produce similar results. In particular, the main finding that the birth intervals are longer for the first post-loss pregnancy, and shorter for all the subsequent pregnancies in a persistent and diminishing manner. The estimated effects are less precise in the RePEAT data, which we attribute to the smaller sample size.

We then perform two estimation exercises to further consolidate our main findings. In one exercise, we exclude the reported induced abortions from the measurement of pregnancy loss. The results in Table 3.D.5 confirms that, although the estimates are imprecise, the pattern of the changes in birth spacing remains the same. Another exercise we conduct is to vary the way to count up the post-pregnancy loss parities. In particular, we (i) use the first, rather than the last, experience of pregnancy loss experience, and (ii) reset the post-loss counter every time a female encounters a pregnancy loss. The results in Table 3.D.6 show that the main findings are unchanged at least qualitatively. These results suggest that our main findings are robust to the inclusion of abortion, as well as the definition of post-pregnancy loss for Uganda.[25]

### 3.5.4 Robustness checks and additional analyses

We first examine whether our main findings are driven by older females for whom we observe more pregnancies, on average, than younger females whose pregnancy data may be right censored. To see this, we limit our sample to females aged 30, or 40, years at the time of data collection, and re-estimate equations (3.1) and (3.2). Estimation results in Appendix Table 3.B.6 show that the change in spacing intervals are strikingly similar to the main results when the estimation sample is limited to females of age 30 and older, or 40 and older. This suggests that our main findings are unlikely to be specific to some cohorts only.

Next, we estimate equations (3.1) and (3.2) with an additional dummy for the last pre-loss pregnancy to check the pre-trend of the outcome.[26] If females with a pregnancy loss had a differing spacing behaviour before their pregnancy loss episode that increases the chance of pregnancy loss, then our findings above may not capture the causal effect of pregnancy loss. Results in Appendix Figure 3.A.2 shows that the confidence interval of

[25] However, we note that these robustness findings may not generalize to data from other countries, since this may stem from the strict conditions that must be met in order to legally perform induced abortion in Uganda (Singh et al., 2018).

[26] Specifically, we estimate the equation of the form:

$$y_{ij} = \varphi_j + \gamma_1 D_{ij}^{\text{pre}} + \sum_{l \in L} \gamma_2^l D_{ij}^{\text{post},l} + x_{ij}^{\text{panel}} \gamma_3 + \nu_{ij}$$

where $D_{ij}^{\text{pre}} = 1$ if pregnancy $i$ of female $j$ is her last pregnancy preceding her loss episode, and 0 otherwise.

the pre-loss pregnancy effect includes null, which suggests that, in line with our discussion in Section 3.4.2, the pre-trend seems to be indistinguishable between females with and without a pregnancy loss experience.

## 3.6 Belief Updating Mechanism for the Change in Birth Spacing Behaviours

We now consider a plausible explanation for our main findings, pertaining to an updating of statistical belief following an experience of a rare event (Hertwig et al., 2004). That is, those who experienced a rare event may subjectively overweight the probability of the occurrence of such an event than actually is. If females with a pregnancy loss experience worry about another occurrence of a loss in the future, or even their infecundity thereafter, they may overly react to the loss, and such reactions may include shortening their subsequent spacing intervals. Updating a probabilistic expectation and changing a behaviour after an occurrence of a rare event has been examined in several contexts: Lybbert et al. (2007) report that herders in eastern Africa are found to update the expectations when they obtain a low-rainfall forecast; Oster (2018) shows evidence of, and examines the mechanism for, the increase in pertussis vaccination following local outbreaks; and some scholars (e.g., Ando et al. (2017) and Fink and Stratmann (2015)) investigate whether or not housing prices near nuclear power plants, in countries such as Sweden and U.S., changed after the nuclear plant blast in Fukushima in 2011.

To take this perspective into the context of fertility behaviours, consider a female who attempts to achieve a certain number of children during her reproductive years.[27] When she starts her reproduction at a young age, she estimates the probability that she will lose a pregnancy due to miscarriage or stillbirth at a small percentage. At some point in time, however, she experiences a pregnancy loss, and updates the loss probability to a higher percentage. Given the years of reproduction left for her and the remaining number of children to make, she re-chooses her fertility schedule with her subjective belief for the probability of pregnancy loss. This re-optimisation leads to shorter birth spacing for all the pregnancies after her pregnancy loss experience.

In order to investigate this hypothesis, we conduct an additional analysis by using the 'realised probability' of a pregnancy loss. That is, we construct a new variable, $z_{ij}$, denoting the realised probability of a pregnancy loss at the time that pregnancy $i$ of woman $j$ is conceived. It is defined as

$$z_{ij} = \frac{D_{ij}^{\text{post}}}{(i-1) + D_{ij}^{\text{post}}} \tag{3.6}$$

where $(i-1)$ is the number of pregnancies that resulted in a live birth prior to pregnancy $i$. The denominator effectively captures the total number of *attempted* pregnancies of female $j$, so the entire fraction measures the proportion of loss out of the pregnancies that she has ever conceived prior to her $i$-th pregnancy.

If all the pregnancies before the $i$-th one ended in a live birth, then $z_{ij} = 0$ as $D_{ij}^{\text{post}} = 0$. If $D_{ij}^{\text{post}} = 1$, *i.e.*, woman $j$ experienced a loss prior to her $i$-th pregnancy, then the

[27] See Appendix 3.E for a more formal discussion.

denominator of the fraction is the total number of attempted pregnancies, including live births and a loss, prior to pregnancy $i$, and $z_{ij}$ becomes the probability of a pregnancy loss that woman $j$ observes from her own fertility history. As an illustration, for female $j$ with two live births prior to a loss, the realised probability of pregnancy loss are given as $z_{1,j} = z_{2,j} = 0$, $z_{3,j} = 1/3$, $z_{4,j} = 1/4$, and the like. Thus defined $z_{ij}$ corresponds to the prior belief with an assumption that the prior is formed as the mean of a Beta distribution $\mathcal{B}(a_{ij}, b_{ij})$ where $a_{ij} = D_{ij}^{\text{post}}$ and $b_{ij} = i - 1$.[28] In addition, we allow the effect of $z_{ij}$ to differ over the subsequent fertility history, just like allowing $D_{ij}^{\text{post},l}$ to affect the spacing interval differently for $l \in L_1$, $L_2$, $L_3$, or $L_4$.

By adding $z_{ij}^l$ to the regression equation (3.2), we develop the modified model of the form:

$$y_{ij}^{\text{panel}} = \tilde{\phi}_j + \sum_{l \in L} \tilde{\delta}_1^l D_{ij}^{\text{post},l} + \underline{x}_{ij}^{\text{panel}} \tilde{\delta}_2 + \sum_{l \in L} \tilde{\delta}_3^l z_{ij}^l + \tilde{\nu}_{ij} \tag{3.7}$$

for the panel FE estimation. Conditional on $D_{ij}^{\text{post},l}$, the variation in $z_{ij}^l$ comes from the variation in the number of live births prior to pregnancy loss of female $j$.

In the new model, the parameters of interest are $\tilde{\delta}_1^l$'s and $\tilde{\delta}_3^l$'s. If, for instance, $\tilde{\delta}_1^l = 0$ and $\tilde{\delta}_3^l \neq 0$ for all $l$ in equation (3.7), then it suggests that women do not respond to the fact that they experienced a loss but do change spacing intervals according to the probability of a loss based on her own reproductive experience. That is, this is the case when the belief updating may be in place and explain the behavioural change in birth spacing. On the contrary, if $\tilde{\delta}_1^l \neq 0$ and $\tilde{\delta}_3^l = 0$ in equation (3.7), it implies that women do not respond

[28] Trivially, in case $a_{ij} = D_{ij}^{\text{post}} = 0$, one needs to take the limit $\lim_{a_{ij} \to 0} E[x | a_{ij}, b_{ij}] = 0$ where $x \sim \mathcal{B}(a_{ij}, b_{ij})$.

to the probability but do to the loss experience. This suggests that the updating of a belief may not take place, and other factors such as psychic shock or psychological trauma may trigger shortening the subsequent spacing intervals.

The estimation results for equation (3.7) are presented in Table 3.7. The main finding is that the pregnancy loss dummy indeed has a significant effect that first lengthens the spacing interval and then shortens the interval for the second pregnancy, but it no longer affects birth spacing for the subsequent pregnancies. On the other hand, for the third and subsequent post-loss pregnancies, the realised loss probability influences the spacing intervals. The estimated coefficient of $-27.79$ for the probability interacted with the third post-loss pregnancy implies that a 1 percentage point increase in the probability of a pregnancy loss leads to a 0.26 month decrease in spacing interval. To see the magnitude of the effect, let us take an example of a female who had had two successful pregnancies before her loss (two is the median number of pre-loss pregnancies). For this female, the probability of a loss at the third pregnancy after a loss is 1/5. Then, the change in spacing interval is $-27.79 \times (1/5) = -5.56$ months. Observe that this is very close to the change at the third post-loss pregnancy reported in Table 3.5, and also to the effects of infant mortality on birth intervals found in Bhalotra and van Soest (2008).[29]

The hypothesis that actual experience of pregnancy loss changed subjective belief on its probability and eventually birth spacing behaviours provides a consistent interpretation to one dimension of the heterogeneity analysis of the main results. In Figure 3.3 we plot the heterogeneous effect estimates of pregnancy loss for females with a different number

[29] Using their estimates and the mean log birth interval, their reported impact of post-natal death of the previous child is approximately -5.7 months.

of live births given by the time they had the loss. There we find that females with two or fewer live births at pregnancy loss exhibit larger shortening effect. This is a predictable result in light of the updating of subjective belief, since females with fewer live births at pregnancy loss have a higher experienced loss probability at each post-loss parity, which is likely to shorten spacing intervals in a larger magnitude compared to females with more live births at pregnancy loss.

Our results in this section suggest that the updating of belief on the probability of pregnancy loss leads to the changes in birth spacing behaviours, to the extent that the realised probability affects birth spacing through its effect on subjective belief. The growing body of literature on probabilistic beliefs suggest that people in developing countries form subjective beliefs according to their past outcomes (Delavande, 2014). Belief formation is an important aspect of decision making under uncertainty, since most generally people do not know the true probability distribution over the possible alternatives. Our results contribute a new piece of evidence to this literature, focusing on fertility behaviours in low income countries.

## 3.7 Conclusion

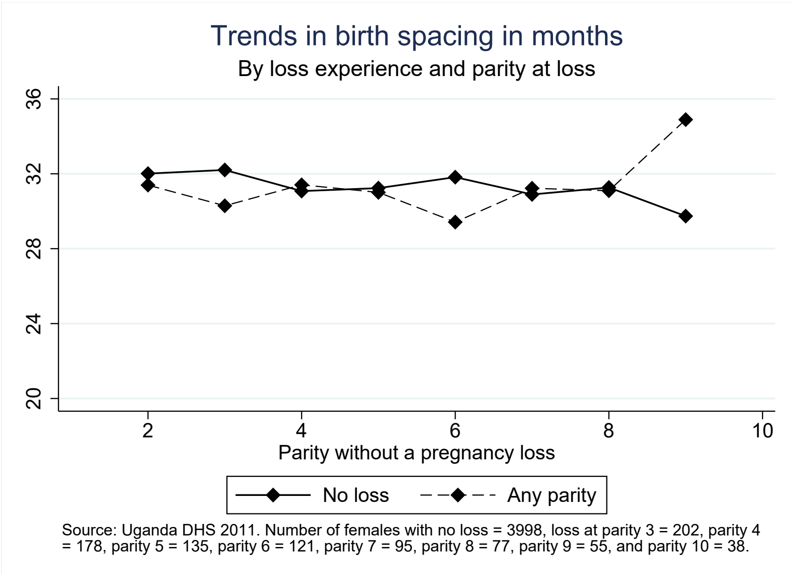
In this study, we examined the pregnancy-related behavioural response to a pregnancy loss, namely miscarriage and stillbirth. Our review of gynecology and obstetrics literature revealed that a pregnancy loss due to miscarriage and stillbirth can occur to any woman at any time due to a random genetic reason conditional on maternal age and fixed effects. Our data analysis shows that post-loss spacing intervals are first lengthened for a mechanical reason, but are later shortened persistently, a finding that most likely reflect behavioural changes. This finding is robust to the measurement of spacing intervals and estimation methods. Our additional analyses have shown that the shortening effect of pregnancy loss on birth spacing is consistent with the belief updating hypothesis, where females adjust their perceived probability of pregnancy loss and birth spacing behaviours based on their own experience.

These results provide important implications for countries with birth spacing that are shorter than ideal. For instance, we find in Uganda that birth intervals are on average shorter than recommended by World Health Organization (2007). In such countries, birth spacing may be based on the belief on the probability of pregnancy loss. As far as we know, this is the first empirical evidence which suggests that beliefs on reproductive outcomes are formed based on own past experiences and such beliefs affect future behaviours. Our estimated impact is comparable to the effect of infant mortality on birth spacing (Bhalotra and van Soest, 2008), while the estimated structural model by Mira (2007) produces a fairly small change in birth spacing due to belief updating. An exploration for why the results differ may help derive an implication for further economic analyses of reproductive behaviours and family planning policies.

Another implication for policy may pertain to a potential improvement of family planning programmes in developing countries. That is, if a change in belief on pregnancy loss probability can have such a large effect on subsequent fertility behaviours, it may seem natural to target resources to females with such experience and educate them about the true probability of pregnancy loss, which is generally smaller than the subjective overestimation. This may sound particularly plausible given that pregnancy loss due to miscarriage and stillbirth is such an obvious event for medical practitioners, and thus allows relatively easy targeting of resources. However, it has not been clear yet, both in this study and in the related literature, whether belief updating occurs in the same magnitude when the new information is provided by someone else, and not based on her own experience. Indeed, it is argued that the success of information intervention crucially hinges upon not only the contents of information but also who provides it to whom, and how (Dupas, 2011). The investigation into the effectiveness of the information intervention on pregnancy loss is thus an interesting avenue for future studies.

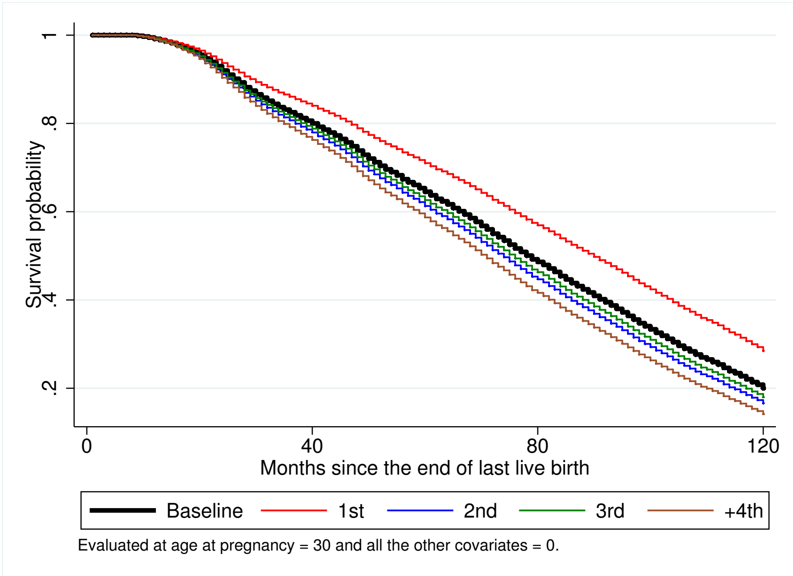
This study is not, however, free of limitations. First, improving the measurement of pregnancy loss due to miscarriage and stillbirth will help identify the behavioural changes more precisely. Our secondary data recorded multiple experiences of pregnancy loss and attempted to separate miscarraige and stillbirth from induced abortion, but the interviewees were not necessarily the females, or even if females answer the interview questions, other household members may be present; the DHS data were collected from females, but they only contain information on the timing of the latest pregnancy loss which potentially includes induced abortion. Second, in order to further understand the decision making of birth spacing relating to the beliefs on the pregnancy loss probability, it is crucial to collect probabilistic expectations, which has been increasingly demonstrated effective in the developing world (Delavande, 2014). These are the unaddressed questions left for future studies.

## 3.8 Figures.



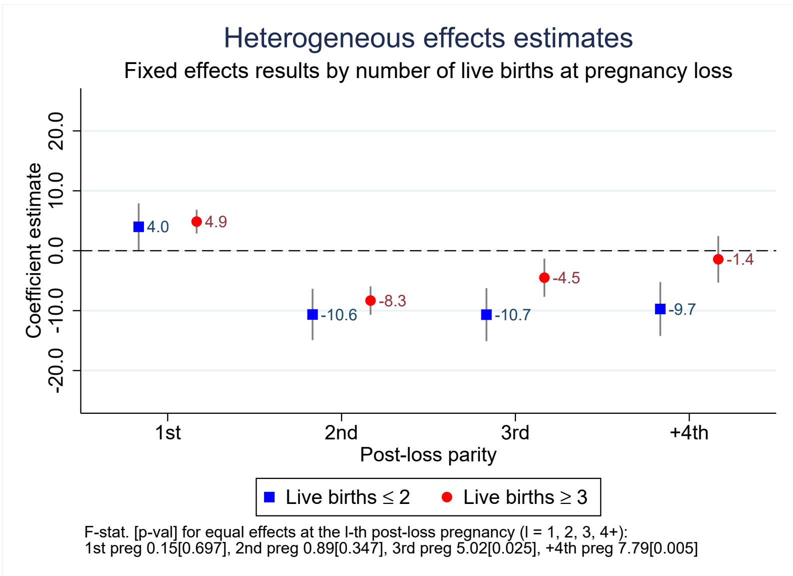

*Source*: DHS Uganda 2011. *Notes*: This figure shows spacing intervals for females with no pregnancy loss experience and pre-loss spacing intervals for females with a loss at parities two through ten.

Figure 3.1: Pre-pregnancy loss birth spacing by loss experience.




*Source*: DHS Uganda 2011. *Notes*: The baseline survivor function is evaluated with the age at pregnancy termination at 30. The others are evaluated, in addition to the above age at pregnancy termination, with the coefficient estimate for the post-loss pregnancy dummies. Post-loss pregnancy dummies indicate whether the pregnancy is the first, second, third, or fourth or any other subsequent one.

Figure 3.2: Survival time plot for the Cox proportional hazard estimation results.




*Source*: DHS Uganda 2011. *Notes*: Coefficients and 95% confidence intervals are plotted for the coefficient estimates of the post pregnancy loss dummies interacted with dummies for whether the number of live births delivered before the loss was equal to or less than 2, or above. The regression includes, as covariates, age at pregnancy termination and its square, dummies for the birth order, sex of the child, sex of the previous child, and dummies for the birth year and birth month of the child.

Figure 3.3: Regression results for spacing interval by live births delivered.

## 3.9 Tables.

Table 3.1: Summary statistics of major variables observed at pregnancy level in DHS Uganda 2011 data.

| VARIABLES | (1)<br>N | (2)<br>mean | (3)<br>sd |
|---|---|---|---|
| Panel A. Outcome variables. | | | |
| Spacing in months | 22,155 | 31.90 | 18.67 |
| 1 if spacing interval < 33 months | 22,155 | 0.656 | 0.475 |
| Panel B. Explanatory variables. | | | |
| Age at pregnancy | 28,609 | 24.22 | 6.336 |
| 1 if single birth | 28,609 | 0.971 | 0.167 |
| 1 if child male | 28,609 | 0.506 | 0.500 |
| 1 if post-loss pregnancy | 27,994 | 0.097 | 0.296 |
| 1 if 1st post-loss pregnancy | 27,994 | 0.035 | 0.183 |
| 1 if 2nd post-loss pregnancy | 27,994 | 0.021 | 0.145 |
| 1 if 3rd post-loss pregnancy | 27,994 | 0.015 | 0.120 |
| 1 if 4th post-loss pregnancy | 27,994 | 0.0098 | 0.0984 |
| 1 if 5th post-loss pregnancy | 27,994 | 0.0065 | 0.0806 |
| 1 if 6th post-loss pregnancy | 27,994 | 0.0045 | 0.0672 |
| 1 if 7th post-loss pregnancy | 27,994 | 0.0026 | 0.0510 |
| 1 if 8th post-loss pregnancy | 27,994 | 0.0018 | 0.0426 |
| 1 if 9th post-loss pregnancy | 27,994 | 0.0009 | 0.0305 |
| 1 if 10th post-loss pregnancy | 27,994 | 0.0003 | 0.0179 |
| 1 if 11th post-loss pregnancy | 27,994 | 0.0001 | 0.0120 |
| 1 if 12th post-loss pregnancy | 27,994 | 0.0001 | 0.0104 |
| 1 if +13th post-loss pregnancy | 27,994 | 0.00004 | 0.00598 |

*Source*: DHS Uganda 2011. *Notes*: This table shows the summary statistics of major variables for children born to women who were as old as 15 to 49 years as of the survey. Spacing interval is defined for one's second pregnancy and onwards, measuring the monthly interval between the ends of one pregnancy and of the previous live birth. 2 years and 9 months (33 months) corresponds to the birth spacing interval recommended by the World Health Organisation.

Table 3.2: Summary statistics of major variables observed at woman level in DHS Uganda 2011 data.

| VARIABLES | (1)<br>N | (2)<br>mean | (3)<br>sd |
|---|---|---|---|
| Panel A. Outcome variables. | | | |
| Number of pregnancies | 8,674 | 3.251 | 3.053 |
| Number of pregnancies for wome with at least one pregnancy | 6,393 | 4.411 | 2.745 |
| Number of pregnancies for women at ages of 40 and above | 1,271 | 7.041 | 2.864 |
| Number of pregnancies for women at ages of 45 and above | 562 | 7.203 | 3.140 |
| Number of children born alive | 8,674 | 3.298 | 3.105 |
| Number of children for women with at least one child | 6,393 | 4.475 | 2.795 |
| Number of children for women at ages of 40 and above | 1,271 | 7.150 | 2.925 |
| Number of children for women at ages of 45 and above | 562 | 7.310 | 3.212 |
| Panel B. Explanatory variables. | | | |
| 1 if having ever had pregnancy loss | 8,665 | 0.188 | 0.391 |
| 1 if having ever had more than one pregnancy loss | 8,665 | 0.015 | 0.122 |
| Age | 8,674 | 27.86 | 9.355 |
| Years of education | 7,342 | 6.817 | 3.630 |
| Marital status: Never married | 8,669 | 0.255 | 0.436 |
| Marital status: Married | 8,669 | 0.354 | 0.478 |
| Marital status: Cohabiting | 8,669 | 0.263 | 0.440 |
| Marital status: Widowed | 8,669 | 0.037 | 0.188 |
| Marital status: Divorced | 8,669 | 0.008 | 0.090 |
| Marital status: Separated | 8,669 | 0.083 | 0.276 |
| Decision maker for contraception use: wife | 1,545 | 0.329 | 0.470 |
| Decision maker for contraception use: husband | 1,545 | 0.101 | 0.301 |
| Decision maker for contraception use: jointly | 1,545 | 0.570 | 0.495 |
| Ideal number of children | 8,453 | 4.893 | 2.198 |

*Source*: DHS Uganda 2011. *Notes*: This table shows the summary statistics of major variables for women who were as old as 15 to 49 years as of the survey. The number of pregnancies counts multiple births as one observation, which makes a slight difference from the number of children. Years of education is the minimum years of schooling required to achieve the reported highest grade.

Table 3.3: Regression-based test for differential trend in spacing interval in months before pregnancy loss.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Spacing in months | | | | | | | |
| Loss at parity | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Ever had pregnancy loss | 0.762 | 2.567 | 3.561 | -1.183 | 1.880 | -2.443 | 6.033* | -4.373* |
| | (0.827) | (5.031) | (4.417) | (3.241) | (2.827) | (2.407) | (3.156) | (2.642) |
| ever had pregnancy loss | | -1.011 | -0.778 | 0.520 | -0.281 | 1.004* | -0.625 | 0.867 |
| × parity | | (1.989) | (1.362) | (0.968) | (0.621) | (0.567) | (0.561) | (0.536) |
| Observations | 3,386 | 5,451 | 6,103 | 6,152 | 5,392 | 4,289 | 2,915 | 1,924 |
| R-squared | 0.000 | 0.001 | 0.001 | 0.000 | 0.001 | 0.003 | 0.003 | 0.003 |
| Joint F | 0.85 | 0.13 | 0.75 | 0.27 | 0.33 | 3.30 | 3.00 | 1.40 |
| P > F | 0.357 | 0.877 | 0.473 | 0.764 | 0.718 | 0.038 | 0.052 | 0.243 |
| Number of females with no loss | 3199 | 2556 | 1915 | 1440 | 1007 | 651 | 386 | 221 |
| Number of females with loss | 129 | 123 | 80 | 63 | 47 | 48 | 21 | 13 |

*Source*: DHS Uganda 2011. *Notes*: This table shows selected estimates from the regression of birth spacing measured in months on an intercept, parity, pregnancy loss indicator, and the interaction between parity and pregnancy loss indicator. Reported in parentheses are the standard errors robust to heteroscedasticity. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1. Along with the estimates, we report the $F$-statistic for the joint test for the loss indicator and its interaction with parity. The regressions use the sample of females who have made at least as many pregnancy attempts as indicated in the column title.

Table 3.4: Falsification test: regression results of pre-determined characteristics.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Religion | | Ethnicity | | | Native Language | | |
| | Christian | Muslim | Baganda | Top 5 | Top 10 | Luganda | Top 3 | Top 5 |
| | Panel A. Single regression. | | | | | | | |
| Ever had pregnancy loss | 0.003 | -0.004 | -0.002 | 0.034** | 0.003 | -0.012 | 0.011 | 0.021 |
| | (0.010) | (0.010) | (0.011) | (0.015) | (0.014) | (0.011) | (0.015) | (0.014) |
| Observations | 6,384 | 6,384 | 6,384 | 6,384 | 6,384 | 6,384 | 6,384 | 6,384 |
| R-squared | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Panel B. Controlling for linear age. | | | | | | | |
| Ever had pregnancy loss | -0.003 | 0.002 | -0.001 | 0.034** | 0.000 | -0.010 | 0.008 | 0.021 |
| | (0.011) | (0.010) | (0.011) | (0.015) | (0.014) | (0.011) | (0.015) | (0.015) |
| Observations | 6,384 | 6,384 | 6,384 | 6,384 | 6,384 | 6,384 | 6,384 | 6,384 |
| R-squared | 0.001 | 0.001 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Panel C. Controlling for quadratic age. | | | | | | | |
| Ever had pregnancy loss | -0.003 | 0.002 | -0.001 | 0.034** | -0.001 | -0.010 | 0.007 | 0.020 |
| | (0.011) | (0.010) | (0.011) | (0.015) | (0.014) | (0.011) | (0.015) | (0.015) |
| Observations | 6,384 | 6,384 | 6,384 | 6,384 | 6,384 | 6,384 | 6,384 | 6,384 |
| R-squared | 0.001 | 0.001 | 0.000 | 0.001 | 0.001 | 0.000 | 0.000 | 0.001 |
| | Panel D. Controlling for dichotomised age. | | | | | | | |
| Ever had pregnancy loss | -0.004 | 0.003 | -0.001 | 0.035** | 0.001 | -0.010 | 0.007 | 0.023 |
| | (0.011) | (0.010) | (0.011) | (0.015) | (0.014) | (0.011) | (0.015) | (0.015) |
| Observations | 6,384 | 6,384 | 6,384 | 6,384 | 6,384 | 6,384 | 6,384 | 6,384 |
| R-squared | 0.008 | 0.007 | 0.004 | 0.007 | 0.005 | 0.006 | 0.005 | 0.006 |
| | Panel E. Mean and standard deviation of the outcome. | | | | | | | |
| Mean | 0.851 | 0.139 | 0.159 | 0.452 | 0.690 | 0.168 | 0.459 | 0.604 |
| S.D. | 0.356 | 0.346 | 0.366 | 0.498 | 0.462 | 0.374 | 0.498 | 0.489 |

*Source*: DHS Uganda 2011. *Notes*: Robust standard errors are reported in parentheses. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1. Christianity includes Catholic, Protestant, Pentecostal, and Adventist. Top 5 ethnicities include Baganda, Banyankole, Basoga, Atesa, and Langi. Top 10 ethnicities include, in addition to Top 5, Bakiga, Acholi, Ngakaramajong, Lugbara, and Mugishu. Luganda is the single largest native language in the data. Top 3 native languages include Luganda, Runyankole-Rukiga, and Luo. Top 5 native languages include, in addition to the above top 3, Lusoga and Ngakaramojong.

Table 3.5: Regression results for the effect of pregnancy loss on spacing intervals in months.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Outcome | Spacing in months | | | | | | | |
| Level of observation | Pregnancy | | | | Child | | | |
| Post-pregnancy loss dummies | $L_1$ | $L_2$ | $L_3$ | $L_4$ | $L_1$ | $L_2$ | $L_3$ | $L_4$ |
| 1 if post-loss pregnancy | -0.194 | | | | -0.059 | | | |
| | (0.773) | | | | (0.783) | | | |
| 1 if 1st post-loss pregnancy | | 4.572*** | 4.694*** | 4.725*** | | 4.761*** | 4.873*** | 4.899*** |
| | | (0.872) | (0.873) | (0.874) | | (0.882) | (0.882) | (0.883) |
| 1 if +2nd post-loss pregnancy | | -6.770*** | | | | -6.625*** | | |
| | | (1.007) | | | | (1.008) | | |
| 1 if 2nd post-loss pregnancy | | | -8.158*** | -8.107*** | | | -7.841*** | -7.798*** |
| | | | (1.029) | (1.029) | | | (1.037) | (1.037) |
| 1 if +3rd post-loss pregnancy | | | -5.131*** | | | | -5.179*** | |
| | | | (1.112) | | | | (1.109) | |
| 1 if 3rd post-loss pregnancy | | | | -5.668*** | | | | -5.624*** |
| | | | | (1.155) | | | | (1.151) |
| 1 if +4th post-loss pregnancy | | | | -4.543*** | | | | -4.701*** |
| | | | | (1.243) | | | | (1.236) |
| Observations | 20,955 | 20,955 | 20,955 | 20,955 | 21,640 | 21,640 | 21,640 | 21,640 |
| Number of pid | 5,295 | 5,295 | 5,295 | 5,295 | 5,313 | 5,313 | 5,313 | 5,313 |
| R-squared | 0.345 | 0.355 | 0.355 | 0.355 | 0.336 | 0.345 | 0.345 | 0.345 |
| Woman FE | Y | Y | Y | Y | Y | Y | Y | Y |
| Panel controls | Y | Y | Y | Y | Y | Y | Y | Y |

*Source*: DHS Uganda 2011. *Notes*: This table shows selected estimates from the regression of birth spacing measured in months on pregnancy loss indicator(s) and control covariates following equations (3.1) and (3.2). Reported in parentheses are the standard errors clustered at the woman level. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1. Sets of pregnancy loss dummies are $L_1 = \{D^{post}\}$, $L_2 = \{D^{post,1}, D^{post,2+}\}$, $L_3 = \{D^{post,1}, D^{post,2}, D^{post,3+}\}$, and $L_4 = \{D^{post,1}, D^{post,2}, D^{post,3}, D^{post,4+}\}$. Panel controls include age at pregnancy termination and its square, dummies for parity, sex of the child, sex of the previous child, and dummies for birth year and month of the child. Columns (1) through (4) use a pregnancy as the unit of observation where both single and multiple births are counted as one observation, while columns (5) through (8) use a child as the unit of observation where one pregnancy is counted more than once for multiple births.

Table 3.6: Regression results for the effect of pregnancy loss on alternative measures of spacing intervals.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Outcome | Log(Spacing in months) | | | | 1 if spacing interval < 33 months | | | |
| Post-pregnancy loss dummies | $L_1$ | $L_2$ | $L_3$ | $L_4$ | $L_1$ | $L_2$ | $L_3$ | $L_4$ |
| 1 if post-loss pregnancy | 0.003 | | | | -0.048** | | | |
| | (0.019) | | | | (0.020) | | | |
| 1 if 1st post-loss pregnancy | | 0.125*** | 0.127*** | 0.128*** | | -0.168*** | -0.170*** | -0.171*** |
| | | (0.020) | (0.020) | (0.020) | | (0.022) | (0.022) | (0.022) |
| 1 if +2nd post-loss pregnancy | | -0.165*** | | | | 0.119*** | | |
| | | (0.026) | | | | (0.025) | | |
| 1 if 2nd post-loss pregnancy | | | -0.195*** | -0.193*** | | | 0.136*** | 0.134*** |
| | | | (0.027) | (0.027) | | | (0.028) | (0.028) |
| 1 if +3rd post-loss pregnancy | | | -0.131*** | | | | 0.098*** | |
| | | | (0.029) | | | | (0.028) | |
| 1 if 3rd post-loss pregnancy | | | | -0.146*** | | | | 0.121*** |
| | | | | (0.031) | | | | (0.031) |
| 1 if +4th post-loss pregnancy | | | | -0.113*** | | | | 0.074** |
| | | | | (0.033) | | | | (0.033) |
| Observations | 20,955 | 20,955 | 20,955 | 20,955 | 20,955 | 20,955 | 20,955 | 20,955 |
| Number of pid | 5,295 | 5,295 | 5,295 | 5,295 | 5,295 | 5,295 | 5,295 | 5,295 |
| R-squared | 0.290 | 0.299 | 0.299 | 0.299 | 0.170 | 0.178 | 0.178 | 0.178 |
| Panel controls | Y | Y | Y | Y | Y | Y | Y | Y |
| Woman FE | Y | Y | Y | Y | Y | Y | Y | Y |
| Level of observation | pregnancy | pregnancy | pregnancy | pregnancy | pregnancy | pregnancy | pregnancy | pregnancy |

*Source*: DHS Uganda 2011. *Notes*: This table shows selected estimates from the regression of log of spacing intervals in months (columns (1) through (4)) or dummy for spacing interval being shorter than 33 months (columns (5) through (8)) on pregnancy loss indicator(s) and control covariates following equations (3.1) and (3.2). Reported in parentheses are the standard errors clustered at the woman level. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1. Sets of pregnancy loss dummies are $L_1 = \{D^{\text{post}}\}$, $L_2 = \{D^{\text{post},1}, D^{\text{post},2+}\}$, $L_3 = \{D^{\text{post},1}, D^{\text{post},2}, D^{\text{post},3+}\}$, and $L_4 = \{D^{\text{post},1}, D^{\text{post},2}, D^{\text{post},3}, D^{\text{post},4+}\}$. Panel controls include age at pregnancy termination and its square, dummies for parity, sex of the child, sex of the previous child, and dummies for birth year and month of the child. All the regressions use a pregnancy as the unit of observation where both single and multiple births are counted as one observation.

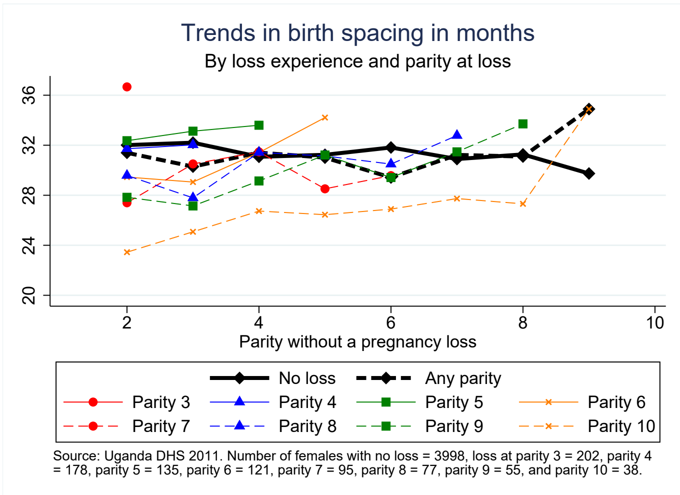
Table 3.7: Regression results for the effects of pregnancy loss status and loss probabilities on spacing intervals.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Outcome | Spacing in months | | | |
| Level of observation | Pregnancy | | | |
| Post-pregnancy loss dummies | $L_1$ | $L_2$ | $L_3$ | $L_4$ |
| post-loss pregnancy | -3.637*** | | | |
| | (1.179) | | | |
| 1st post-loss pregnancy | | 4.843** | 4.755** | 5.108** |
| | | (1.954) | (1.970) | (1.996) |
| +2nd post-loss pregnancy | | -5.096*** | | |
| | | (1.367) | | |
| 2nd post-loss pregnancy | | | -8.360*** | -7.740*** |
| | | | (1.731) | (1.849) |
| +3rd post-loss pregnancy | | | -1.023 | |
| | | | (1.929) | |
| 3rd post-loss pregnancy | | | | -1.437 |
| | | | | (2.781) |
| +4th post-loss pregnancy | | | | 3.415 |
| | | | | (2.779) |
| Pr(pregnancy loss) | 18.47*** | | | |
| | (4.197) | | | |
| Pr(pregnancy loss) | | 0.633 | 1.484 | -0.267 |
| × (1st post-loss pregnancy) | | (7.279) | (7.401) | (7.581) |
| Pr(pregnancy loss) | | -11.94*** | | |
| × (+2nd post-loss pregnancy) | | (4.364) | | |
| Pr(pregnancy loss) | | | -1.147 | -4.656 |
| × (2nd post-loss pregnancy) | | | (6.044) | (7.098) |
| Pr(pregnancy loss) | | | -29.62*** | |
| × (+3rd post-loss pregnancy) | | | (8.56) | |
| Pr(pregnancy loss) | | | | -27.79** |
| × (3rd post-loss pregnancy) | | | | (12.98) |
| Pr(pregnancy loss) | | | | -62.56*** |
| × (+4th post-loss pregnancy) | | | | (15.31) |
| Observations | 20,955 | 20,955 | 20,955 | 20,955 |
| Number of pid | 5,295 | 5,295 | 5,295 | 5,295 |
| R-squared | 0.317 | 0.327 | 0.328 | 0.328 |
| Panel controls | Y | Y | Y | Y |
| Woman FE | Y | Y | Y | Y |

*Source*: DHS Uganda 2011. *Notes*: Reported in parentheses are the standard errors clustered at the woman level. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1. Sets of pregnancy loss dummies are $L_1 = \{D^{\text{post}}\}$, $L_2 = \{D^{\text{post},1}, D^{\text{post},2+}\}$, $L_3 = \{D^{\text{post},1}, D^{\text{post},2}, D^{\text{post},3+}\}$, and $L_4 = \{D^{\text{post},1}, D^{\text{post},2}, D^{\text{post},3}, D^{\text{post},4+}\}$. Panel controls include age at pregnancy termination and its square, parity, sex of the child, sex of the previous child, and dummies for birth year and month of the child. All regressions reported in this table use a pregnancy as the unit of observation where both single and multiple births are counted as one observation.

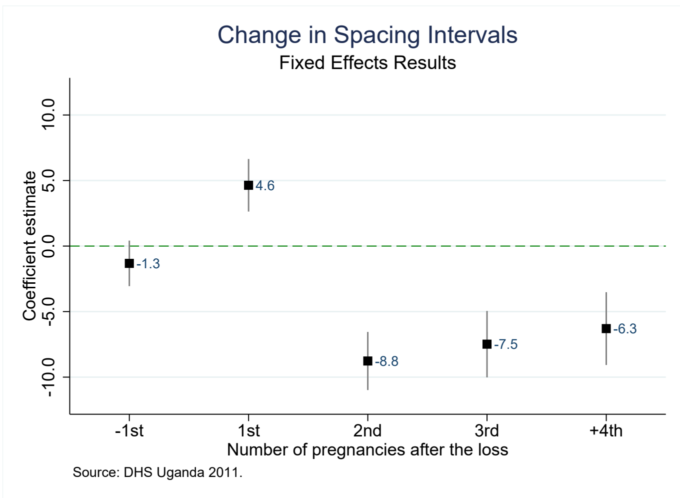
# Appendices for Chapter 2.

## Appendix 3.A Appendix figures.



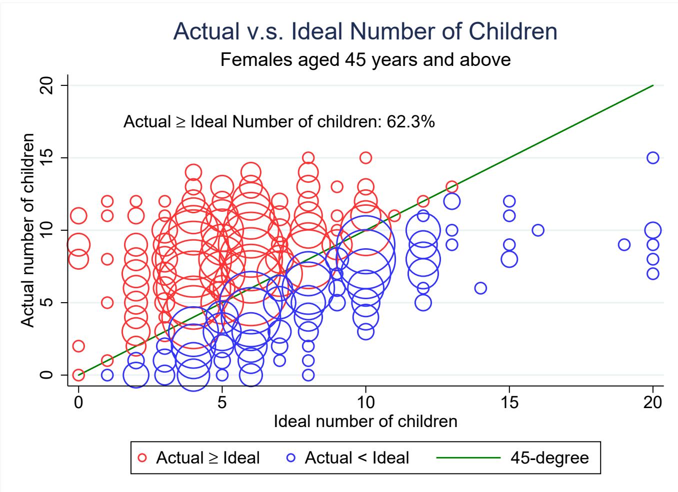

*Source*: DHS Uganda 2011. *Notes*: This figure shows spacing intervals for females with no pregnancy loss experience and pre-loss spacing intervals for females with a loss experience at parities two through ten disaggregated by the parity of the loss.

Figure 3.A.1: Pre-pregnancy loss birth spacing by loss experience and parity of the loss.




*Source*: DHS Uganda 2011. *Notes*: This figure shows the estimated coefficient estimates and corresponding confidence intervals for the pre- and post-pregnancy loss dummies. The regression includes, as covariates, age at pregnancy termination and its square, birth order, sex of the child, sex of the previous child, and the birth year and month of the child.

Figure 3.A.2: Regression results for spacing intervals with the pre-loss pregnancy dummy.




*Source*: DHS Uganda 2011. *Notes*: This figure shows the distribution of the actual and ideal number of children of females aged 45 years and above, where the size of circles correspond to the share of females at the actual and ideal numbers of children on the vertical and horizontal axes, respectively.

Figure 3.A.3: Actual v.s. ideal number of children of females aged 45 years and above.

## Appendix 3.B Appendix tables.

Table 3.B.1: Regression-based test for differential trend in spacing interval in months before pregnancy loss.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Outcome | Spacing interval in months | | | | | | | |
| Parity at pregnancy loss | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Ever had pregnancy loss | 1.524 | 0.545 | 2.628 | 0.697 | -2.526 | -1.457 | 1.618 | -2.798 |
| | (1.654) | (1.377) | (2.138) | (1.726) | (2.203) | (1.772) | (2.918) | (2.639) |
| Ever had pregnancy loss | | -1.011 | -2.649 | -1.474 | 4.971* | 1.659 | 1.806 | 1.161 |
| × (Parity = 3) | | (1.989) | (2.628) | (2.451) | (2.965) | (2.547) | (4.267) | (3.530) |
| Ever had pregnancy loss | | | -1.557 | -0.052 | 9.374** | 5.632** | 5.583 | 0.694 |
| × (Parity = 4) | | | (2.724) | (2.517) | (3.849) | (2.614) | (4.877) | (3.184) |
| Ever had pregnancy loss | | | | 1.282 | 2.204 | 4.255 | 3.517 | 2.507 |
| × (Parity = 5) | | | | (3.110) | (2.996) | (2.874) | (5.302) | (4.471) |
| Ever had pregnancy loss | | | | | -0.0312 | 3.445 | 1.190 | 4.502 |
| × (Parity = 6) | | | | | (2.709) | (2.621) | (4.934) | (4.762) |
| Ever had pregnancy loss | | | | | | 6.199* | 0.233 | 6.960 |
| × (Parity = 7) | | | | | | (3.557) | (3.934) | (6.583) |
| Ever had pregnancy loss | | | | | | | -3.318 | 5.280 |
| × (Parity = 8) | | | | | | | (4.056) | (4.876) |
| Ever had pregnancy loss | | | | | | | | 4.583 |
| × (Parity = 9) | | | | | | | | (4.692) |
| Observations | 3,386 | 5,451 | 6,103 | 6,152 | 5,392 | 4,289 | 2,915 | 1,924 |
| R-squared | 0.000 | 0.001 | 0.001 | 0.002 | 0.005 | 0.006 | 0.012 | 0.010 |
| Joint F | 0.85 | 0.13 | 0.64 | 0.27 | 2.00 | 1.70 | 1.10 | 0.53 |
| P > F | 0.357 | 0.877 | 0.590 | 0.899 | 0.070 | 0.110 | 0.368 | 0.834 |
| Number of females with no loss | 3199 | 2556 | 1915 | 1440 | 1007 | 651 | 386 | 221 |
| Number of females with loss | 129 | 123 | 80 | 63 | 47 | 48 | 21 | 13 |

*Source*: DHS Uganda 2011. *Notes*: This table shows selected estimates from the regression of birth spacing measured in months on parity dummies, pregnancy loss indicator, and the interaction between parity dummies and pregnancy loss indicator. Reported in parentheses are the standard errors robust to heteroscedasticity. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1. Along with the estimates, we report the *F*-statistic for the joint test for the loss indicator and its interaction with parity dummies. Panel A uses the full sample of females in the data, while Panel B uses the sample restricted by the number of pregnancies such that both the females with and without a pregnancy loss have attempted at least as many pregnancies as indicated in the column title.

Table 3.B.2: Regression-based test for differential trend in log spacing interval in months before pregnancy loss.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Outcome variable | Log(Spacing in months) | | | | | | | |
| Loss at parity | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Ever had pregnancy loss | 0.030 | 0.102 | 0.109 | 0.011 | 0.012 | -0.044 | 0.187* | -0.119 |
| | (0.020) | (0.148) | (0.119) | (0.099) | (0.093) | (0.076) | (0.099) | (0.104) |
| Ever had pregnancy loss | | -0.034 | -0.022 | 0.002 | 0.002 | 0.025 | -0.020 | 0.023 |
| × parity | | (0.058) | (0.038) | (0.027) | (0.021) | (0.017) | (0.018) | (0.019) |
| Observations | 3,386 | 5,451 | 6,103 | 6,152 | 5,392 | 4,289 | 2,915 | 1,924 |
| R-squared | 0.001 | 0.001 | 0.000 | 0.000 | 0.000 | 0.002 | 0.002 | 0.002 |
| Joint F | 2.3 | 0.35 | 1.20 | 0.15 | 0.17 | 3.80 | 3.00 | 0.74 |
| P > F | 0.129 | 0.702 | 0.313 | 0.862 | 0.840 | 0.022 | 0.052 | 0.476 |
| Number of females with no loss | 3199 | 2556 | 1915 | 1440 | 1007 | 651 | 386 | 221 |
| Number of females with loss | 129 | 123 | 80 | 63 | 47 | 48 | 21 | 13 |

*Source*: DHS Uganda 2011. *Notes*: This table shows selected estimates from the regression of birth spacing measured in months on an intercept, parity, pregnancy loss indicator, and the interaction between parity and pregnancy loss indicator. Reported in parentheses are the standard errors robust to heteroscedasticity. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1. Along with the estimates, we report the *F*-statistic for the joint test for the loss indicator and its interaction with parity. The regressions use the sample of females who have made at least as many pregnancy attempts as indicated in the column title.

Table 3.B.3: Regression-based test for differential trend in the indicator for short spacing intervals before pregnancy loss.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Outcome variable | 1 if spacing interval < 33 months | | | | | | | |
| Loss at parity | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Ever had pregnancy loss | -0.012 | 0.008 | -0.050 | -0.036 | -0.056 | -0.029 | -0.127 | 0.186** |
| | (0.021) | (0.153) | (0.117) | (0.098) | (0.086) | (0.076) | (0.105) | (0.095) |
| Ever had pregnancy loss | | -0.006 | 0.005 | 0.005 | 0.007 | -0.012 | 0.006 | -0.032* |
| × parity | | (0.060) | (0.038) | (0.027) | (0.020) | (0.016) | (0.020) | (0.018) |
| Observations | 3,386 | 5,451 | 6,103 | 6,152 | 5,392 | 4,289 | 2,915 | 1,924 |
| R-squared | 0.000 | 0.000 | 0.000 | 0.001 | 0.001 | 0.003 | 0.004 | 0.005 |
| Joint F | 0.34 | 0.002 | 0.68 | 0.26 | 0.50 | 4.40 | 3.10 | 2.00 |
| P > F | 0.558 | 0.976 | 0.508 | 0.773 | 0.608 | 0.013 | 0.045 | 0.138 |
| Number of females with no loss | 3199 | 2556 | 1915 | 1440 | 1007 | 651 | 386 | 221 |
| Number of females with loss | 129 | 123 | 80 | 63 | 47 | 48 | 21 | 13 |

*Source*: DHS Uganda 2011. *Notes*: This table shows selected estimates from the regression of an indicator for birth spacing being shorter than 33 months that the WHO recommends on parity dummies, pregnancy loss indicator, and the interaction between parity dummies and pregnancy loss indicator. Reported in parentheses are the standard errors robust to heteroscedasticity. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1. Along with the estimates, we report the $F$-statistic for the joint test for the loss indicator and its interaction with parity dummies. The regressions use the sample of females who have made at least as many pregnancy attempts as indicated in the column title.

Table 3.B.4: Pooled OLS regression results for the effect of pregnancy loss on spacing intervals in months.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Outcome | Spacing in months | | | | | | | |
| Controls | Basic controls | | | | Additional controls | | | |
| Post-pregnancy loss dummies | $L_1$ | $L_2$ | $L_3$ | $L_4$ | $L_1$ | $L_2$ | $L_3$ | $L_4$ |
| 1 if post-loss pregnancy | 2.152*** | | | | 0.986* | | | |
| | (0.439) | | | | (0.509) | | | |
| 1 if 1st post-loss pregnancy | | 12.01*** | 12.01*** | 12.02*** | | 11.09*** | 11.09*** | 11.10*** |
| | | (0.919) | (0.919) | (0.919) | | (1.001) | (1.001) | (1.001) |
| 1 if +2nd post-loss pregnancy | | -1.977*** | | | | -3.326*** | | |
| | | (0.441) | | | | (0.549) | | |
| 1 if 2nd post-loss pregnancy | | | -3.365*** | -3.365*** | | | -5.319*** | -5.320*** |
| | | | (0.610) | (0.610) | | | (0.690) | (0.690) |
| 1 if +3rd post-loss pregnancy | | | -1.227** | | | | -2.203*** | |
| | | | (0.535) | | | | (0.662) | |
| 1 if 3rd post-loss pregnancy | | | | -1.715** | | | | -3.083*** |
| | | | | (0.836) | | | | (0.973) |
| 1 if +4th post-loss pregnancy | | | | -0.950 | | | | -1.679** |
| | | | | (0.607) | | | | (0.765) |
| Observations | 20,955 | 20,955 | 20,955 | 20,955 | 15,305 | 15,305 | 15,305 | 15,305 |
| R-squared | 0.161 | 0.175 | 0.175 | 0.175 | 0.244 | 0.259 | 0.260 | 0.260 |
| Panel controls | Y | Y | Y | Y | Y | Y | Y | Y |
| Woman FE | N | N | N | N | N | N | N | N |

*Source*: DHS Uganda 2011. *Notes*: This table shows selected estimates from the regression of birth spacing measured in months on pregnancy loss indicator(s) and control covariates following equations (3.1) and (3.2) without woman fixed effects. Reported in parentheses are the standard errors clustered at the woman level. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1. Sets of pregnancy loss dummies are $L_1 = \{D^{\text{post}}\}$, $L_2 = \{D^{\text{post},1}, D^{\text{post},2+}\}$, $L_3 = \{D^{\text{post},1}, D^{\text{post},2}, D^{\text{post},3+}\}$, and $L_4 = \{D^{\text{post},1}, D^{\text{post},2}, D^{\text{post},3}, D^{\text{post},4+}\}$. Panel controls include age at pregnancy termination and its square, dummies for parity, sex of the child, sex of the previous child, and dummies for birth year and month of the child. Additional controls include woman's current age and its square, years of education, and age at their first sexual intercourse. All the regressions use a pregnancy as the unit of observation where both single and multiple births are counted as one observation.

Table 3.B.5: Cox proportional hazard estimation results.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Control covariates. | Age at pregnancy termination and square. | | | | Augmented panel controls. | | | |
| Pregnancy loss dummies. | $L_1$ | $L_2$ | $L_3$ | $L_4$ | $L_1$ | $L_2$ | $L_3$ | $L_4$ |
| 1 if post-loss pregnancy | -0.018 | | | | 0.005 | | | |
| | (0.023) | | | | (0.023) | | | |
| 1 if 1st post-loss pregnancy | | -0.260*** | -0.260*** | -0.260*** | | -0.247*** | -0.247*** | -0.247*** |
| | | (0.039) | (0.039) | (0.039) | | (0.039) | (0.039) | (0.039) |
| 1 if +2nd post-loss pregnancy | | 0.107*** | | | | 0.133*** | | |
| | | (0.026) | | | | (0.026) | | |
| 1 if 2nd post-loss pregnancy | | | 0.115*** | 0.115*** | | | 0.111** | 0.111** |
| | | | (0.043) | (0.043) | | | (0.043) | (0.043) |
| 1 if +3rd post-loss pregnancy | | | 0.102*** | | | | 0.146*** | |
| | | | (0.032) | | | | (0.032) | |
| 1 if 3rd post-loss pregnancy | | | | 0.064 | | | | 0.066 |
| | | | | (0.052) | | | | (0.052) |
| 1 if +4th post-loss pregnancy | | | | 0.125*** | | | | 0.196*** |
| | | | | (0.040) | | | | (0.040) |
| Panel controls. | | | | | Y | Y | Y | Y |
| Observations | 25,406 | 25,406 | 25,406 | 25,406 | 25,406 | 25,406 | 25,406 | 25,406 |

*Notes*. Estimates are for the coefficients in the log relative hazard. Standard errors are reported in parentheses. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1. Sets of pregnancy loss dummies are $L_1 = \{D^{\text{post}}\}$, $L_2 = \{D^{\text{post},1}, D^{\text{post},2+}\}$, $L_3 = \{D^{\text{post},1}, D^{\text{post},2}, D^{\text{post},3+}\}$, and $L_4 = \{D^{\text{post},1}, D^{\text{post},2}, D^{\text{post},3}, D^{\text{post},4+}\}$. All regressions include age at pregnancy termination and square. Panel controls include, additionally, birth order, dummies for birth year of the child, and dummies for birth month of the child.

Table 3.B.6: Regression results for the effects of pregnancy loss on spacing intervals for older female subsamples.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Outcome. | Spacing in months | | | | | | | |
| Sample. | Females 30 years or older. | | | | Females 40 years or older. | | | |
| Pregnancy loss dummies. | $L_1$ | $L_2$ | $L_3$ | $L_4$ | $L_1$ | $L_2$ | $L_3$ | $L_4$ |
| 1 if post-loss pregnancy | -0.410 | | | | 0.263 | | | |
| | (0.869) | | | | (1.380) | | | |
| 1 if 1st post-loss pregnancy | | 5.474*** | 5.601*** | 5.637*** | | 6.829*** | 7.098*** | 7.176*** |
| | | (1.024) | (1.028) | (1.029) | | (1.720) | (1.720) | (1.725) |
| 1 if +2nd post-loss pregnancy | | -7.425*** | | | | -6.323*** | | |
| | | (1.114) | | | | (1.630) | | |
| 1 if 2nd post-loss pregnancy | | | -8.623*** | -8.557*** | | | -8.817*** | -8.694*** |
| | | | (1.138) | (1.141) | | | (1.628) | (1.626) |
| 1 if +3rd post-loss pregnancy | | | -6.163*** | | | | -4.265** | |
| | | | (1.254) | | | | (1.832) | |
| 1 if 3rd post-loss pregnancy | | | | -6.729*** | | | | -5.358*** |
| | | | | (1.299) | | | | (2.014) |
| 1 if +4th post-loss pregnancy | | | | -5.603*** | | | | -3.362* |
| | | | | (1.405) | | | | (1.985) |
| R-squared | 0.311 | 0.322 | 0.323 | 0.323 | 0.286 | 0.296 | 0.297 | 0.297 |
| Observations | 16,076 | 16,076 | 16,076 | 16,076 | 7,286 | 7,286 | 7,286 | 7,286 |
| Number of pid | 3,092 | 3,092 | 3,092 | 3,092 | 1,173 | 1,173 | 1,173 | 1,173 |

*Notes*. Estimates are for the coefficients in the log relative hazard. Standard errors are reported in parentheses. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1. Sets of pregnancy loss dummies are $L_1 = \{D^{\text{post}}\}$, $L_2 = \{D^{\text{post},1}, D^{\text{post},2+}\}$, $L_3 = \{D^{\text{post},1}, D^{\text{post},2}, D^{\text{post},3+}\}$, and $L_4 = \{D^{\text{post},1}, D^{\text{post},2}, D^{\text{post},3}, D^{\text{post},4+}\}$. Panel controls include age at pregnancy termination and its square, parity, sex of the child, sex of the previous child, and dummies for birth year and birth month of the child. All regressions reported in this table use a pregnancy as the unit of observation where both single and multiple births are counted as one observation.

## Appendix 3.C Selection into higher-order fertility

As briefly discussed in Section 3.4.4, one of the threats to our identification based on estimation equations (3.1) and (3.2) is that females with a pregnancy loss may continue, or discontinue, their subsequent fertility with a differential probability. To see this, we estimate equation (3.3) by OLS and test the null hypotheses for the coefficient estimates in $\beta_2$ and $\beta_3$ jointly. Also, since our main analyses use at most four pregnancies after a pregnancy loss, we consider $t = \{1, 2, 3, 4\}$ simultaneously. By conducting this analysis, we attempt to examine whether there is any covariate that predict the probability of continuing fertility in a manner different between females without a pregnancy loss and those with a pregnancy loss.

In this regression estimation, there are two points to note. First, our regressions include females with at least $k$ successful births. That is, the sample of females are similar up to the $k$-th successful birth, but different only due to the loss at the next pregnancy attempt. This sample restriction effectively allows us to examine whether the fertility continuation decision for the $(k+1)$-st through $(k+4)$-th pregnancies differs only because one group of females experienced a loss at their $(k+1)$-st attempt. A differential fertility continuation decision for the two groups of females is implied if the coefficients for the interaction terms in equation (3.3) are estimated to be statistically significant.

Second, the above analysis is subject to the increased risk of making type-one errors among a large number of hypotheses jointly tested. That is, even if the probability of rejecting a null hypothesis that is actually true is controlled at some level $\alpha \in [0, 1]$, the probability of mistakenly rejecting at least one of the null hypotheses jointly tested (included in a *family* of a test) increases exponentially as $1 - (1 - \alpha)^h$ where $h$ denotes the number of hypotheses in the family (assuming independence of tests jointly considered). There are several methods that correct statistical inference for multiple hypothesis testing that control such a probability, called family-wise error rate (FWER). Easy-to-use methods include the Bonferroni's and Sidak's corrections, but these methods assume independence among the tests in the family. If the tests in the family are correlated, the conclusion from the joint tests performed with such corrections can be invalid, which is likely to be the case in our present analysis where the probability of giving the $(k + 1)$st successful birth is clearly positively correlated with that of giving the $(k + 2)$nd successful birth. Therefore, we use the bootstrap-based method proposed by Westfall and Young (1993) which allows for any arbitrary correlation structure among the tests in the family and controls the FWER below some pre-specified level of confidence $\alpha$ effectively. Jones et al. (2018) employs this method and provides a Stata code, `wyoung`, which is flexible enough to accommodate multiple equations and different covariate sets in each equation. However, their code allows testing only one parameter per equation, which is a serious limitation in our analysis that requires testing multiple coefficients of an equation. Thus, we wrote a programme that implements the MHT correction in a potentially multiple equation system involving potentially multiple parameters in a linear regression framework.[30]

The results for selected parameter estimates are presented in Tables 3.C.1 through 3.C.7. We perform the analyses for $k \in \{1, 2, ..., 7\}$ and $t \in \{1, 2, 3, 4\}$, and for two sets of covariates (basic and reduced). The basic controls include the intercept, current age, years

[30] One of the caveats of our programme is that it does not allow different equations to have different sets of covariates. Another is that it only allows least squares estimation in the current version. However, these are of little importance in the present analysis, whilst we might consider extension of the programme to overcome these issues depending upon the demand for the code.

of education, age at the first sexual intercourse, region dummies (10 categories where the base indicates Kampala, the capital), religion (6 categories where the base indicates Catholic), and ethnicity dummies (19 categories where the base indicates Baganda), thus testing 144 parameters jointly (36 parameters times 4 equations). The reduced controls excludes the dummies for region, religion, and ethnicity from the family of the test (although still include them in the regressions), thus testing 16 parameters (4 coefficients in 4 equations). The idea behind the reduced controls is that the MHT correction leads to more conservative conclusions when the family includes a larger number of parameters, and we show the robustness of our conclusion in a setting where the power of the MHT-corrected tests is arbitrarily rendered larger by reducing the number of joint tests.

As can be seen from the results tables 3.C.1 through 3.C.7, we fail to find coefficient estimates that are consistently significant across equations for the interactions between the pregnancy loss status and major covariates.[31] The non-MHT-corrected results often encounter statistically significant estimates for the coefficients of interaction terms, and such cases seem to appear at a rate larger than just by chance under the conventional level of significance such as .05. However, by applying the MHT correction to the statistical inference, the significant estimates disappear altogether. This observation remains unchanged even when we arbitrarily reduce the number of joint tests and gain the statistical power, as shown in columns (5) to (8) in the tables below. These results may be in part due to the relatively small share of females for whom the pregnancy loss indicator takes the

[31] The results in columns (1) and (5) of Table 3.C.1 are not reliable, since the data do not have any female with a pregnancy loss who has no births by the survey interview. In this case, the possible combinations of the pregnancy loss indicator and the fertility continuation indicator are (0,0), (0,1), and (1,1). Then, the coefficient for the pregnancy loss indicator becomes one minus the intercept, and all the interactions have the coefficients of the exact same magnitude with the sign opposite to the non-interacted terms.

value of one, but should not be entirely so since the standard errors are generally not that large and the coefficient estimates are smaller than those of the non-interacted variables. Therefore, we conclude from the analysis in this Appendix that the females in our data do not exhibit selective fertility continuation systematically related to pregnancy loss at least for the four post-loss parities.

Table 3.C.1: Regression results for selective fertility continuation for females with at least zero number of children v.s. females with pregnancy loss at her first pregnancy attempt.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Sample | Females with at least 0 successful births | | | | | | | |
| Covariates | Basic controls | | | | Reduced controls | | | |
| Outcome: 1 if fertility ≥ ... | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| Constant | 0.444*** | 0.206*** | 0.019 | -0.121*** | 0.521*** | 0.297*** | 0.131*** | 0.016 |
| | (0.041) | (0.043) | (0.041) | (0.039) | (0.036) | (0.037) | (0.035) | (0.033) |
| Age | 0.018*** | 0.030*** | 0.037*** | 0.037*** | 0.018*** | 0.030*** | 0.037*** | 0.037*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Years of education | -0.014*** | -0.022*** | -0.023*** | -0.022*** | -0.016*** | -0.025*** | -0.025*** | -0.025*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Age at first sex | -0.005** | -0.017*** | -0.027*** | -0.028*** | -0.005*** | -0.018*** | -0.028*** | -0.028*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Ever had pregnancy loss | 0.556***,‡ | -0.018 | -0.158 | 0.085 | 0.479***,‡ | 0.137 | 0.058 | 0.160 |
| | (0.041) | (0.213) | (0.193) | (0.166) | (0.036) | (0.202) | (0.199) | (0.147) |
| Ever had pregnancy loss | -0.018***,‡ | 0.004 | -0.001 | -0.003 | -0.018***,‡ | 0.003 | -0.001 | -0.003 |
| × Age | (0.001) | (0.003) | (0.003) | (0.004) | (0.001) | (0.003) | (0.004) | (0.004) |
| Ever had pregnancy loss | 0.014***,† | 0.018** | 0.003 | -0.007 | 0.016***,‡ | 0.012 | -0.003 | -0.009 |
| × Years of education | (0.002) | (0.008) | (0.008) | (0.008) | (0.002) | (0.008) | (0.007) | (0.007) |
| Ever had pregnancy loss | 0.005** | -0.025** | -0.003 | -0.003 | 0.005*** | -0.018 | -0.003 | -0.005 |
| × Age at first sex | (0.002) | (0.011) | (0.011) | (0.009) | (0.002) | (0.012) | (0.011) | (0.009) |
| Observations | 5,001 | 5,001 | 5,001 | 5,001 | 5,001 | 5,001 | 5,001 | 5,001 |
| R-squared | 0.219 | 0.372 | 0.482 | 0.502 | 0.210 | 0.358 | 0.468 | 0.489 |
| Share with pregnancy loss | 0.0378 | | | | | | | |

*Source*: DHS Uganda 2011. *Notes*: This table shows the estimated results of the regression of fertility continuation indicators on female characteristics and their interactions with pregnancy loss status. Reported in parentheses are the standard errors clustered at the woman level. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1, as well as ‡ if p<0.01, † if p<0.05, and ∤ if p<0.1 using the multiple-hypothesis testing adjustment proposed by Westfall and Young (1993). Basic controls include age, years of education, age at the first sexual intercourse, region dummies, religion dummies, and ethnicity dummies, while the reduced controls only include age, years of education, and age at the first sexual intercourse. The regressions use females who have never experienced pregnancy loss and those whose first attempted pregnancy ended up being lost. The number of hypotheses jointly tested as a family is 144 in columns (1) through (4) (36 variables (in addition to the four shown in the table, 9 region dummies, 5 religion dummies, and 18 ethnicity dummies) × 4 outcomes) and 16 in columns (5) through (8).

Table 3.C.2: Regression results for selective fertility continuation for females with at least one child v.s. females with pregnancy loss at her second pregnancy attempt.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Sample | Females with at least one successful birth | | | | | | | |
| Covariates | Basic controls | | | | Reduced controls | | | |
| Outcome: 1 if fertility ≥ ... | 2 | 3 | 4 | 5 | 2 | 3 | 4 | 5 |
| Constant | 0.515*** | 0.140*** | -0.119** | -0.243*** | 0.578*** | 0.259*** | 0.053 | -0.0948** |
| | (0.047) | (0.051) | (0.048) | (0.044) | (0.040) | (0.043) | (0.041) | (0.038) |
| Age | 0.0204*** | 0.0339*** | 0.0371*** | 0.0356*** | 0.0204*** | 0.0338*** | 0.0371*** | 0.0355*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Years of education | -0.0159*** | -0.0211*** | -0.0227*** | -0.0219*** | -0.0172*** | -0.0237*** | -0.0268*** | -0.0256*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Age at first sex | -0.0149*** | -0.0274*** | -0.0292*** | -0.0266*** | -0.0157*** | -0.0286*** | -0.0294*** | -0.0267*** |
| | (0.002) | (0.003) | (0.003) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| 1 if ever had pregnancy loss | -0.162 | -0.288 | -0.196 | -0.133 | -0.114 | -0.304* | -0.345** | -0.207 |
| | (0.211) | (0.192) | (0.154) | (0.157) | (0.176) | (0.164) | (0.146) | (0.137) |
| Ever had pregnancy loss | 0.00234 | 0.00233 | -0.00077 | -0.0023 | 0.00293 | 0.00125 | -0.000464 | -0.00212 |
| × Age | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| Ever had pregnancy loss | 0.0183** | -0.000561 | 0.00711 | 0.00369 | 0.0135* | 0.00139 | 0.00877 | 0.00566 |
| × Years of education | (0.009) | (0.008) | (0.007) | (0.007) | (0.008) | (0.008) | (0.008) | (0.007) |
| Ever had pregnancy loss | -0.0151 | -0.000779 | 0.00592 | 0.0103 | -0.00986 | 0.00641 | 0.00952 | 0.00771 |
| × Age at first sex | (0.010) | (0.010) | (0.008) | (0.009) | (0.009) | (0.009) | (0.008) | (0.008) |
| Observations | 4,159 | 4,159 | 4,159 | 4,159 | 4,159 | 4,159 | 4,159 | 4,159 |
| R-squared | 0.234 | 0.403 | 0.458 | 0.464 | 0.217 | 0.386 | 0.44 | 0.449 |
| Share with pregnancy loss | 0.0384 | | | | | | | |

*Source*: DHS Uganda 2011. *Notes*: This table shows the estimated results of the regression of fertility continuation indicators on female characteristics and their interactions with pregnancy loss status. Reported in parentheses are the standard errors clustered at the woman level. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1, as well as ‡ if p<0.01, † if p<0.05, and ⸸ if p<0.1 using the multiple-hypothesis testing adjustment proposed by Westfall and Young (1993). Basic controls include age, years of education, age at the first sexual intercourse, region dummies, religion dummies, and ethnicity dummies, while the reduced controls only include age, years of education, and age at the first sexual intercourse. The regressions use females who have never experienced pregnancy loss and those whose first attempted pregnancy ended up being lost. The number of hypotheses jointly tested as a family is 144 in columns (1) through (4) (36 variables (in addition to the four shown in the table, 9 region dummies, 5 religion dummies, and 18 ethnicity dummies) × 4 outcomes) and 16 in columns (5) through (8).

Table 3.C.3: Regression results for selective fertility continuation for females with at least two child v.s. females with pregnancy loss at her third pregnancy attempt.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Sample | Females with at least two successful births | | | | | | | |
| Covariates | Basic controls | | | | Reduced controls | | | |
| Outcome: 1 if fertility ≥ ... | 3 | 4 | 5 | 6 | 3 | 4 | 5 | 6 |
| Constant | 0.354*** | -0.068 | -0.287*** | -0.328*** | 0.470*** | 0.138*** | -0.102** | -0.167*** |
| | (0.058) | (0.059) | (0.054) | (0.050) | (0.048) | (0.050) | (0.047) | (0.043) |
| Age | 0.025*** | 0.035*** | 0.038*** | 0.035*** | 0.025*** | 0.035*** | 0.038*** | 0.034*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Years of education | -0.014*** | -0.021*** | -0.024*** | -0.021*** | -0.016*** | -0.026*** | -0.028*** | -0.025*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Age at first sex | -0.022*** | -0.029*** | -0.029*** | -0.028*** | -0.023*** | -0.029*** | -0.029*** | -0.027*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.002) |
| Ever had pregnancy loss | -0.438 | -0.387 | -0.223 | -0.057 | -0.468** | -0.518** | -0.288 | -0.073 |
| | (0.304) | (0.260) | (0.216) | (0.202) | (0.226) | (0.222) | (0.197) | (0.183) |
| Ever had pregnancy loss | -0.007 | -0.001 | -0.002 | -0.003 | -0.005 | 0.002 | -0.001 | -0.005 |
| × Age | (0.005) | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) |
| Ever had pregnancy loss | -0.033*** | -0.012 | -0.009 | -0.019** | -0.024** | -0.012 | -0.008 | -0.014* |
| × Years of education | (0.011) | (0.011) | (0.009) | (0.009) | (0.010) | (0.009) | (0.007) | (0.007) |
| Ever had pregnancy loss | 0.037** | 0.018 | 0.016 | 0.015 | 0.038*** | 0.022 | 0.014 | 0.011 |
| × Age at first sex | (0.015) | (0.015) | (0.013) | (0.012) | (0.015) | (0.015) | (0.013) | (0.012) |
| Observations | 3,330 | 3,330 | 3,330 | 3,330 | 3,330 | 3,330 | 3,330 | 3,330 |
| R-squared | 0.272 | 0.386 | 0.436 | 0.421 | 0.25 | 0.36 | 0.415 | 0.398 |
| Share with pregnancy loss | 0.038 | | | | | | | |

*Source*: DHS Uganda 2011. *Notes*: This table shows the estimated results of the regression of fertility continuation indicators on female characteristics and their interactions with pregnancy loss status. Reported in parentheses are the standard errors clustered at the woman level. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1, as well as ‡ if p<0.01, † if p<0.05, and ∤ if p<0.1 using the multiple-hypothesis testing adjustment proposed by Westfall and Young (1993). Basic controls include age, years of education, age at the first sexual intercourse, region dummies, religion dummies, and ethnicity dummies, while the reduced controls only include age, years of education, and age at the first sexual intercourse. The regressions use females who have never experienced pregnancy loss and those whose first attempted pregnancy ended up being lost. The number of hypotheses jointly tested as a family is 144 in columns (1) through (4) (36 variables (in addition to the four shown in the table, 9 region dummies, 5 religion dummies, and 18 ethnicity dummies) × 4 outcomes) and 16 in columns (5) through (8).

Table 3.C.4: Regression results for selective fertility continuation for females with at least three successful births v.s. females with pregnancy loss at her fourth pregnancy attempt.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Sample | Females with at least three successful births | | | | | | | |
| Covariates | Basic controls | | | | Reduced controls | | | |
| Outcome: 1 if fertility ≥ ... | 4 | 5 | 6 | 7 | 4 | 5 | 6 | 7 |
| Constant | 0.219*** | -0.229*** | -0.368*** | -0.348*** | 0.417*** | -0.0172 | -0.162*** | -0.181*** |
| | (0.071) | (0.072) | (0.067) | (0.060) | (0.060) | (0.061) | (0.058) | (0.051) |
| Age | 0.023*** | 0.036*** | 0.037*** | 0.031*** | 0.023*** | 0.036*** | 0.036*** | 0.030*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Years of education | -0.017*** | -0.026*** | -0.025*** | -0.018*** | -0.020*** | -0.030*** | -0.030*** | -0.022*** |
| | (0.003) | (0.003) | (0.003) | (0.002) | (0.003) | (0.003) | (0.002) | (0.002) |
| Age at first sex | -0.018*** | -0.027*** | -0.029*** | -0.027*** | -0.018*** | -0.026*** | -0.029*** | -0.026*** |
| | (0.003) | (0.004) | (0.004) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| Ever had pregnancy loss | -0.410 | -0.329 | -0.409* | -0.578*** | -0.424* | -0.213 | -0.284 | -0.483** |
| | (0.264) | (0.267) | (0.231) | (0.204) | (0.250) | (0.228) | (0.224) | (0.198) |
| Ever had pregnancy loss | -0.002 | 0.001 | 0.002 | -0.003 | 0.001 | 0.001 | 0.001 | -0.002 |
| × Age | (0.006) | (0.005) | (0.005) | (0.005) | (0.005) | (0.004) | (0.005) | (0.005) |
| Ever had pregnancy loss | -0.014 | -0.002 | 0.009 | 0.012 | -0.009 | -0.008 | -0.000 | 0.004 |
| × Years of education | (0.012) | (0.012) | (0.010) | (0.009) | (0.011) | (0.009) | (0.008) | (0.008) |
| Ever had pregnancy loss | 0.032** | 0.007 | 0.004 | 0.029*** | 0.022 | 0.004 | 0.008 | 0.028** |
| × Age at first sex | (0.014) | (0.015) | (0.013) | (0.011) | (0.014) | (0.014) | (0.014) | (0.012) |
| Observations | 2,581 | 2,581 | 2,581 | 2,581 | 2,581 | 2,581 | 2,581 | 2,581 |
| R-squared | 0.225 | 0.352 | 0.381 | 0.331 | 0.19 | 0.324 | 0.348 | 0.295 |
| Share with pregnancy loss | 0.045 | | | | | | | |

*Source*: DHS Uganda 2011. *Notes*: This table shows the estimated results of the regression of fertility continuation indicators on female characteristics and their interactions with pregnancy loss status. Reported in parentheses are the standard errors clustered at the woman level. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1, as well as ‡ if p<0.01, † if p<0.05, and ∤ if p<0.1 using the multiple-hypothesis testing adjustment proposed by Westfall and Young (1993). Basic controls include age, years of education, age at the first sexual intercourse, region dummies, religion dummies, and ethnicity dummies, while the reduced controls only include age, years of education, and age at the first sexual intercourse. The regressions use females who have never experienced pregnancy loss and those whose first attempted pregnancy ended up being lost. The number of hypotheses jointly tested as a family is 144 in columns (1) through (4) (36 variables (in addition to the four shown in the table, 9 region dummies, 5 religion dummies, and 18 ethnicity dummies) × 4 outcomes) and 16 in columns (5) through (8).

Table 3.C.5: Regression results for selective fertility continuation for females with at least four successful births v.s. females with pregnancy loss at her fifth pregnancy attempt.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Sample | Females with at least four successful births | | | | | | | |
| Covariates | Basic controls | | | | Reduced controls | | | |
| Outcome: 1 if fertility $\geq$ ... | 5 | 6 | 7 | 8 | 5 | 6 | 7 | 8 |
| Constant | 0.019 | -0.355*** | -0.421*** | -0.440*** | 0.167** | -0.146* | -0.235*** | -0.286*** |
| | (0.092) | (0.091) | (0.082) | (0.070) | (0.074) | (0.075) | (0.068) | (0.060) |
| Age | 0.029*** | 0.037*** | 0.034*** | 0.028*** | 0.029*** | 0.037*** | 0.034*** | 0.027*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Years of education | -0.023*** | -0.026*** | -0.021*** | -0.015*** | -0.025*** | -0.031*** | -0.025*** | -0.017*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.002) |
| Age at first sex | -0.019*** | -0.029*** | -0.029*** | -0.021*** | -0.019*** | -0.027*** | -0.028*** | -0.022*** |
| | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.003) |
| Ever had pregnancy loss | 0.163 | 0.047 | -0.162 | 0.386 | 0.320 | 0.002 | 0.003 | 0.230 |
| | (0.385) | (0.345) | (0.298) | (0.235) | (0.313) | (0.265) | (0.219) | (0.196) |
| Ever had pregnancy loss | -0.011 | -0.005 | -0.005 | -0.016** | -0.011 | -0.010 | -0.012** | -0.012** |
| × Age | (0.008) | (0.007) | (0.008) | (0.006) | (0.008) | (0.007) | (0.006) | (0.005) |
| Ever had pregnancy loss | 0.016 | 0.024 | 0.009 | 0.011 | 0.012 | 0.015 | 0.003 | 0.001 |
| × Years of education | (0.018) | (0.016) | (0.012) | (0.010) | (0.016) | (0.014) | (0.012) | (0.009) |
| Ever had pregnancy loss | -0.026* | -0.011 | 0.002 | -0.001 | -0.011 | 0.004 | 0.017* | 0.006 |
| × Age at first sex | (0.015) | (0.013) | (0.012) | (0.011) | (0.015) | (0.013) | (0.010) | (0.009) |
| Observations | 1,974 | 1,974 | 1,974 | 1,974 | 1,974 | 1,974 | 1,974 | 1,974 |
| R-squared | 0.249 | 0.331 | 0.309 | 0.268 | 0.214 | 0.29 | 0.267 | 0.229 |
| Share with pregnancy loss | 0.039 | | | | | | | |

*Source*: DHS Uganda 2011. *Notes*: This table shows the estimated results of the regression of fertility continuation indicators on female characteristics and their interactions with pregnancy loss status. Reported in parentheses are the standard errors clustered at the woman level. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1, as well as ‡ if p<0.01, † if p<0.05, and ∤ if p<0.1 using the multiple-hypothesis testing adjustment proposed by Westfall and Young (1993). Basic controls include age, years of education, age at the first sexual intercourse, region dummies, religion dummies, and ethnicity dummies, while the reduced controls only include age, years of education, and age at the first sexual intercourse. The regressions use females who have never experienced pregnancy loss and those whose first attempted pregnancy ended up being lost. The number of hypotheses jointly tested as a family is 144 in columns (1) through (4) (36 variables (in addition to the four shown in the table, 9 region dummies, 5 religion dummies, and 18 ethnicity dummies) × 4 outcomes) and 16 in columns (5) through (8).

Table 3.C.6: Regression results for selective fertility continuation for females with at least five successful births v.s. females with pregnancy loss at her sixth pregnancy attempt.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Sample | Females with at least five successful births | | | | | | | |
| Covariates | Basic controls | | | | Reduced controls | | | |
| Outcome: 1 if fertility $\geq$ ... | 6 | 7 | 8 | 9 | 6 | 7 | 8 | 9 |
| Constant | -0.027 | -0.394*** | -0.568*** | -0.496*** | 0.179* | -0.174* | -0.374*** | -0.306*** |
| | (0.119) | (0.119) | (0.103) | (0.079) | (0.095) | (0.097) | (0.087) | (0.067) |
| Age | 0.027*** | 0.034*** | 0.032*** | 0.023*** | 0.027*** | 0.033*** | 0.031*** | 0.022*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Years of education | -0.019*** | -0.019*** | -0.017*** | -0.018*** | -0.023*** | -0.024*** | -0.019*** | -0.019*** |
| | (0.004) | (0.005) | (0.004) | (0.003) | (0.004) | (0.004) | (0.004) | (0.003) |
| Age at first sex | -0.020*** | -0.028*** | -0.022*** | -0.013*** | -0.019*** | -0.027*** | -0.023*** | -0.016*** |
| | (0.005) | (0.006) | (0.005) | (0.004) | (0.005) | (0.005) | (0.005) | (0.004) |
| Ever had pregnancy loss | -0.032 | -0.263 | -0.040 | -0.095 | -0.041 | -0.433 | -0.130 | -0.228 |
| | (0.430) | (0.360) | (0.266) | (0.244) | (0.401) | (0.311) | (0.232) | (0.221) |
| Ever had pregnancy loss | -0.014 | -0.009 | -0.013* | -0.004 | -0.014 | -0.006 | -0.009 | 0.001 |
| $\times$ Age | (0.009) | (0.009) | (0.007) | (0.006) | (0.008) | (0.007) | (0.006) | (0.006) |
| Ever had pregnancy loss | -0.016 | 0.013 | 0.007 | 0.019* | -0.015 | 0.014 | 0.009 | 0.010 |
| $\times$ Years of education | (0.021) | (0.019) | (0.013) | (0.011) | (0.018) | (0.016) | (0.009) | (0.008) |
| Ever had pregnancy loss | 0.024 | 0.029 | 0.023 | 0.006 | 0.024 | 0.026 | 0.014 | 0.005 |
| $\times$ Age at first sex | (0.023) | (0.020) | (0.015) | (0.013) | (0.022) | (0.019) | (0.012) | (0.011) |
| Observations | 1,414 | 1,414 | 1,414 | 1,414 | 1,414 | 1,414 | 1,414 | 1,414 |
| R-squared | 0.222 | 0.250 | 0.254 | 0.221 | 0.164 | 0.198 | 0.206 | 0.173 |
| Share with pregnancy loss | 0.041 | | | | | | | |

*Source*: DHS Uganda 2011. *Notes*: This table shows the estimated results of the regression of fertility continuation indicators on female characteristics and their interactions with pregnancy loss status. Reported in parentheses are the standard errors clustered at the woman level. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1, as well as ‡ if p<0.01, † if p<0.05, and ⸸ if p<0.1 using the multiple-hypothesis testing adjustment proposed by Westfall and Young (1993). Basic controls include age, years of education, age at the first sexual intercourse, region dummies, religion dummies, and ethnicity dummies, while the reduced controls only include age, years of education, and age at the first sexual intercourse. The regressions use females who have never experienced pregnancy loss and those whose first attempted pregnancy ended up being lost. The number of hypotheses jointly tested as a family is 144 in columns (1) through (4) (36 variables (in addition to the four shown in the table, 9 region dummies, 5 religion dummies, and 18 ethnicity dummies) $\times$ 4 outcomes) and 16 in columns (5) through (8).

Table 3.C.7: Regression results for selective fertility continuation for females with at least six successful births v.s. females with pregnancy loss at her seventh pregnancy attempt.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Sample | Females with at least six successful births | | | | | | | |
| Covariates | Basic controls | | | | Reduced controls | | | |
| Outcome: 1 if fertility ≥ ... | 7 | 8 | 9 | 10 | 7 | 8 | 9 | 10 |
| Constant | -0.027 | -0.563*** | -0.636*** | -0.349*** | 0.116 | -0.395*** | -0.404*** | -0.241*** |
| | (0.161) | (0.148) | (0.112) | (0.090) | (0.121) | (0.117) | (0.094) | (0.076) |
| Age | 0.025*** | 0.033*** | 0.027*** | 0.018*** | 0.025*** | 0.032*** | 0.026*** | 0.017*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Years of education | -0.011* | -0.014** | -0.021*** | -0.015*** | -0.015*** | -0.015*** | -0.022*** | -0.013*** |
| | (0.006) | (0.006) | (0.005) | (0.003) | (0.006) | (0.005) | (0.004) | (0.003) |
| Age at first sex | -0.021*** | -0.021*** | -0.013** | -0.011** | -0.020*** | -0.020*** | -0.016*** | -0.014*** |
| | (0.007) | (0.007) | (0.005) | (0.004) | (0.006) | (0.006) | (0.005) | (0.004) |
| Ever had pregnancy loss | 0.532 | -0.370 | -0.415 | 0.196 | 0.019 | -0.056 | -0.213 | -0.039 |
| | (0.593) | (0.403) | (0.397) | (0.159) | (0.483) | (0.361) | (0.291) | (0.242) |
| Ever had pregnancy loss | 0.003 | 0.003 | -0.0081 | -0.012* | 0.007 | 0.005 | -0.002 | -0.008 |
| × Age | (0.011) | (0.011) | (0.010) | (0.006) | (0.011) | (0.009) | (0.010) | (0.007) |
| Ever had pregnancy loss | 0.021 | 0.012 | 0.039*** | 0.018*** | 0.005 | -0.005 | 0.021** | 0.020*** |
| × Years of education | (0.020) | (0.014) | (0.013) | (0.007) | (0.020) | (0.014) | (0.010) | (0.007) |
| Ever had pregnancy loss | -0.054*** | -0.012 | 0.023 | 0.002 | -0.027 | -0.020 | 0.004 | 0.010 |
| × Age at first sex | (0.020) | (0.019) | (0.019) | (0.014) | (0.018) | (0.018) | (0.017) | (0.016) |
| Observations | 1,018 | 1,018 | 1,018 | 1,018 | 1,018 | 1,018 | 1,018 | 1,018 |
| R-squared | 0.180 | 0.223 | 0.213 | 0.151 | 0.109 | 0.170 | 0.162 | 0.119 |
| Share with pregnancy loss | 0.038 | | | | | | | |

*Source*: DHS Uganda 2011. *Notes*: This table shows the estimated results of the regression of fertility continuation indicators on female characteristics and their interactions with pregnancy loss status. Reported in parentheses are the standard errors clustered at the woman level. Statistical significance is denoted by *** if $p<0.01$, ** if $p<0.05$, and * if $p<0.1$, as well as ‡ if $p<0.01$, † if $p<0.05$, and ∤ if $p<0.1$ using the multiple-hypothesis testing adjustment proposed by Westfall and Young (1993). Basic controls include age, years of education, age at the first sexual intercourse, region dummies, religion dummies, and ethnicity dummies, while the reduced controls only include age, years of education, and age at the first sexual intercourse. The regressions use females who have never experienced pregnancy loss and those whose first attempted pregnancy ended up being lost. The number of hypotheses jointly tested as a family is 144 in columns (1) through (4) (36 variables (in addition to the four shown in the table, 9 region dummies, 5 religion dummies, and 18 ethnicity dummies) × 4 outcomes) and 16 in columns (5) through (8).

# Appendix 3.D Research on Poverty, Environment, Agriculture, and Technology Survey in Uganda in 2015

## 3.D.1 Data description

The DHS data provide us with the detailed information to analyse the relationship between birth spacing and pregnancy loss. However, one caveat is that the survey question does not separate abortion from miscarriage and stillbirth. Since induced abortion is illegitimate in Uganda except for the case in which the mother is physically threatened to death (Singh et al., 2018), the number of abortions should not be so large. However, some females may perform unsafe abortion on their own, or fail to report their abortion experience truthfully. Another potential problem is that the DHS data do not have the precise measure of the history of pregnancy losses. Although they have information on whether the women had more than one pregnancy loss, they did not ask when the losses occurred except for the latest experience. These suggest that the measurement of miscarriage and stillbirth in the DHS data may be imperfect for the purpose of our analysis.

Therefore, we also use a different data set that allows us to analyse the effect of miscarriage and stillbirth separately from induced abortion in Uganda. The data were collected from the Research on Poverty, Environment, Agriculture, and Technology (RePEAT) survey conducted in Uganda in 2015 by the National Graduate Institute for Policy Studies and Makerere University. Its sample is not nationally representative, but it covers 1,755 households from 117 villages in 39 districts, where randomisation was at the village level, as well as the household level in each village. The RePEAT survey collected information about reproductive behaviours of females who were 15 to 59 years old, whereas we use the sub-sample of them aged 15 to 49 years to be consistent with the DHS data. From these households, we find 2,517 women, of whom 1,340 had given at least one birth in their lifetime, and 1,106 had given more than one.

The oral interview of the RePEAT survey first asked the respondents to report the years in which they had given live births. It then asked them whether they had ever had a stillbirth, and if so, the years of their stillbirths (if they had experienced more than one stillbirth, they were asked to list the years of all of their stillbirths). Once the list of years of stillbirths was completed, the survey asked about miscarriage experiences, and then induced abortions, in the same manner. Key in this interview survey is that, while being asked orally about stillbirth experiences, respondents did not know that the next questions were going to be about miscarriages, and similarly, while being asked about miscarriage experiences, they did not know that they would be asked about abortions next. This survey and questionnaire design made it fairly difficult, if not impossible, for the respondents to falsely report induced abortion experiences as miscarriages.[32] As we show in the next subsection, the ratio of unclear responses ('Do not know', 'Refuse to answer', and 'Do not recall') to valid responses is much higher with the history of abortion than that of miscarriage or stillbirth, indicating that respondents who wanted to keep secret their abortion experiences may have had to choose these unclear responses since miscarriage questions were already finished, while we still do find a few abortion cases. In addition to

[32] Singh et al. (2018) presents a summary of field studies which suggest that many women do not report abortion experiences when directly asked, and also that women tend to report miscarriages whereas they actually had abortions.

the measurement that separates miscarriage and stillbirth from abortion, we can observe more than one pregnancy loss per person if any. Using thus collected data from the RePEAT survey, we examine the extent to which the measurement of pregnancy loss alter the major findings.

### 3.D.2 Descriptive analyses

Major variables in our RePEAT data are summarised in Tables 3.D.1, 3.D.2, and 3.D.3. Below are two points worth paying attention to. First, the birth spacing variable is measured in years, not in months as in the DHS data. This is because the RePEAT survey did not record the month of pregnancy termination but only the calendar year.[33] Second, a few variables available in the DHS data are unavailable in the RePEAT data, such as an indicator for multiple births and the sex of the child born alive. Although not reported in results tables, we find these covariates have coefficient estimates that are always small and insignificant throughout our regression analyses. Moreover, our balancing test confirm the unrelatedness of the post-pregnancy loss indicator with some selected pre-determined covariates. These suggest that the omission of these variables is unlikely to significantly bias the estimated effect of pregnancy loss on birth spacing when we use the RePEAT data.

Table 3.D.1 shows that, despite the coarse measurement, the mean birth spacing interval is 2.76 years (33.1 months) in the RePEAT data, similar to the corresponding mean of 31.9 months in the DHS data. Women's age at the end of pregnancy is 25.8 years, which is also quite close to 24.2 years in the DHS data. The share of pregnancies that

[33] As the name suggests, the primary objective of the RePEAT survey is to collect information about agricultural activities of rural households, and not about reproductive behaviours in detail.

were terminated after a pregnancy loss experience is 12.5%, again similar to 9.7% in the DHS data, while the composition of post-loss parities are slightly different in the two data sets. The strengths of the RePEAT data are that they have the history of pregnancy loss if women have more than one experience and that they can differentiate miscarriage and stillbirth from induced abortion. In particular, we consider three types of post-pregnancy loss indicators and present the summary in Table 3.D.2. One is the same as the DHS data, where the indicator equals one if the pregnancy was terminated after the latest pregnancy loss experience (Panel A). Alternatively, the indicator takes unity if the pregnancy was terminated after the first pregnancy loss experience for those with more than one loss experiences (Panel B). Furthermore, we can reset the post-loss parity counter when the female experiences another pregnancy loss (Panel C). By including induced abortion (columns (1) to (3)) in, and excluding it (columns (4) to (6)) from, the measurement of pregnancy loss, we construct six different versions of post-loss indicators. Since there are extremely limited reported cases of induced abortion, the difference arising from the in- and ex-clusion of abortion is very small. On the other hand, the three post-loss indicator definitions across panels show a slight difference, although the differences are generally small. Table 3.D.3 shows the summary statistics of cross-section variables in our RePEAT data. The average number of pregnancies is 2.5, slightly smaller than the corresponding number of 3.3 in the DHS data. It increases to 4.6 if the sample is limited to females with at least one pregnancy, and further to 6.3 if the sample is limited to females aged 40 years and above, showing a similar pattern found in the DHS data.[34] The share of females

[34] The smaller average number of pregnancies for those aged 45 years and above, compared to the corresponding number for those aged 40 years and above, is likely to be by chance, given the small number of observations for these females.

who have ever had a pregnancy loss is 12.6%, smaller than 18.8% in the DHS data. This may be due to the respondent: the DHS respondents are the females themselves, while the RePEAT surveys are typically answered by the household head who usually are not necessarily the females themselves. Meanwhile, the share of females who had more than one pregnancy loss is 1.8%, which is much the same as 1.5% in the DHS data. These figures change little when excluding induced abortion from the definition of a pregnancy loss. It seems that the RePEAT sample females are slightly younger and less educated, although the differences are of almost negligible magnitude.

Panel C of Table 3.D.3 shows the share of females for whom the survey response to the question on the year of pregnancy loss includes at least one 'Do not know (DNK),' 'Refuse to answer (RTA),' and 'Do not recall (DNR),' separately for miscarriage and stillbirth, and abortion.[35] Because one female can have multiple experiences of pregnancy loss and report the exact year for the subset of her experiences but one of DNK, RTA, and DNR for the remaining subset, it is not appropriate to compare the shares of females with at least one DNK, RTA, and DNR against females with at least one pregnancy loss record. However, since there are only a few females who report such, we nonetheless take the ratio of the two and compare them for the loss due to miscarriage and stillbirth and for the loss due to abortion. The ratio is .025 / .122 = .205 for miscarriage and stillbirth, while it is .00397 / .00477 = .832 for abortion. That is, we find a larger share of females made some unclear response to the question on the exact year of abortion, given having at least one experience. This is likely to partly explain the small observations of abortion experience,

[35] There is another response category, 'Not applicable (NA).' However, NA is used mainly to mark those who are not supposed to answer the question. Therefore, we exclude the NA responses from the shares of those with at least one unclear year of pregnancy loss.

but this also increases the chance that the pregnancy loss variable does effectively exclude abortion from its measurement in our RePEAT data.

### 3.D.3 Regression results

We compare the estimated results in the change in birth spacing due to pregnancy loss using the two data sets, DHS and RePEAT, in Table 3.D.4. In all regressions in this table, we restrict the sample females to only those who live in the districts covered by the RePEAT data, and use the covariates that are available in both of the data sets. We also re-define the spacing intervals as measured in years, since the RePEAT data only have the information of the year, not month or day, of the end of each pregnancy. With these restrictions, however, the results using the DHS data (columns (1) through (4)) are virtually unchanged from the main results, a longer interval for the first post-loss pregnancy and shorter intervals for the subsequent pregnancies. Results in columns (5) through (8) show that the overall findings from the RePEAT data are similar to those from the DHS data, where the first post-loss pregnancy has a longer interval and the subsequent ones have shorter intervals. As the estimates are admittedly noisier, some coefficient estimates are weakly significant or become insignificant. However, these estimates suggest that the pattern of the changes in birth spacing is found in the RePEAT data with smaller and more restricted coverage, adding the credibility to the external validity of our main findings.

Using the RePEAT data, we demonstrate below that our main findings are robust to different measurements of the pregnancy loss. First, we show in Table 3.D.5 the effects of a pregnancy loss including and excluding induced abortion in the measurement. In columns (1) through (4), we replicate the results with the RePEAT data from Table 3.D.4.

In columns (5) through (8), we show the results when induced abortion is excluded from the measurement of pregnancy loss. It appears that the exclusion of abortion makes the estimates slightly less precise. Still, the overall findings remain qualitatively unchanged: the longer interval for the first post-loss pregnancy and the shorter intervals for the subsequent pregnancies.

Second, we show in Table 3.D.6 the results with two alternative ways to count up the post-pregnancy loss loss parities. As a reminder, the main analysis with the DHS data count the post-loss parity starting from the last experience of each woman's pregnancy loss, since the year and months of the pregnancy loss is only recorded for the last experience for each woman. In columns (1) through (4), we use the first, rather than the last, pregnancy loss experience as the start of the post-loss reproductive behaviours. Although the estimates are not always statistically significant, the qualitative findings remain the same. In columns (5) through (8), we re-define this first-experience-based post-loss indicator such that, every time a woman experiences a pregnancy loss, the post-loss parity counter is reset, and females with more than one pregnancy loss may have the first (and the subsequent orders') post-loss pregnancies more than once.[36] The results show the pattern of the changes in birth spacing similar to the main results. The estimates are largely unchanged partly because women with more than one pregnancy loss do not account for a large share of sample females (only 1.8% from Table 3.D.3).

[36] For example, suppose a woman has a successful birth, a loss, two successful births, another loss, and another successful birth in this order. Then, her second and fourth successful births are of a 'first post-loss' pregnancy.

Table 3.D.1: Summary statistics of major variables observed at pregnancy level in RePEAT Uganda 2015 data.

| VARIABLES | (1)<br>N | (2)<br>mean | (3)<br>sd |
|---|---|---|---|
| Panel A. Outcome variables. | | | |
| Spacing interval in years | 6,678 | 2.759 | 1.846 |
| Log spacing interval in years | 6,678 | 0.911 | 0.505 |
| 1 if spacing interval < 3 years | 6,678 | 0.585 | 0.493 |
| Panel B. Explanatory variables. | | | |
| Age at pregnancy | 8,317 | 25.83 | 7.537 |
| 1 if post-loss pregnancy | 8,317 | 0.125 | 0.331 |
| 1 if 1st post-loss pregnancy | 8,317 | 0.028 | 0.166 |
| 1 if 2nd post-loss pregnancy | 8,317 | 0.019 | 0.137 |
| 1 if 3rd post-loss pregnancy | 8,317 | 0.014 | 0.117 |
| 1 if +4th post-loss pregnancy | 8,317 | 0.064 | 0.245 |

*Source*: RePEAT Uganda 2015. *Notes*: This table shows the summary statistics of major variables for children born to women who were as old as 15 to 59 years as of the survey. Spacing interval is defined for one's second pregnancy and onwards, measuring the yearly interval between the ends of one pregnancy and of the previous live birth.

Table 3.D.2: Various measurement of post-pregnancy loss indicators in RePEAT Uganda 2015 data.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Any pregnancy loss | | | Excluding abortion | | |
| VARIABLES | N | mean | sd | N | mean | sd |
| Panel A. Post last pregnancy loss. | | | | | | |
| Post-loss pregnancy | 8,317 | 0.125 | 0.331 | 8,317 | 0.125 | 0.330 |
| 1st post-loss pregnancy | 8,317 | 0.0284 | 0.166 | 8,317 | 0.0281 | 0.165 |
| 2nd post-loss pregnancy | 8,317 | 0.0190 | 0.137 | 8,317 | 0.0188 | 0.136 |
| 3rd post-loss pregnancy | 8,317 | 0.0138 | 0.117 | 8,317 | 0.0137 | 0.116 |
| +4th post-loss pregnancy | 8,317 | 0.0641 | 0.245 | 8,317 | 0.0641 | 0.245 |
| Panel B. Post first pregnancy loss. | | | | | | |
| Post-loss pregnancy | 8,317 | 0.152 | 0.359 | 8,317 | 0.149 | 0.357 |
| 1st post-loss pregnancy | 8,317 | 0.0292 | 0.168 | 8,317 | 0.0290 | 0.168 |
| 2nd post-loss pregnancy | 8,317 | 0.0203 | 0.141 | 8,317 | 0.0201 | 0.140 |
| 3rd post-loss pregnancy | 8,317 | 0.0149 | 0.121 | 8,317 | 0.0148 | 0.121 |
| +4th post-loss pregnancy | 8,317 | 0.0875 | 0.283 | 8,317 | 0.0856 | 0.280 |
| Panel C. Post counter-reset pregnancy loss. | | | | | | |
| Post-loss pregnancy | 8,317 | 0.152 | 0.359 | 8,317 | 0.149 | 0.357 |
| 1st post-loss pregnancy | 8,317 | 0.0457 | 0.209 | 8,317 | 0.0459 | 0.209 |
| 2nd post-loss pregnancy | 8,317 | 0.0313 | 0.174 | 8,317 | 0.0315 | 0.175 |
| 3rd post-loss pregnancy | 8,317 | 0.0234 | 0.151 | 8,317 | 0.0238 | 0.152 |
| +4th post-loss pregnancy | 8,317 | 0.0516 | 0.221 | 8,317 | 0.0482 | 0.226 |

*Source*: RePEAT Uganda 2015. *Notes*: This table shows the share of pregnancies terminated after a pregnancy loss measured in different ways for females aged 15 to 59 years as of the survey. Panel A shows the share of pregnancies at each parity after the latest pregnancy loss. Panel B shows the share of pregnancies at each parity after the first pregnancy loss. Panel C shows the share of pregnancies at each parity after any pregnancy loss, where the parity counter is reset every time a woman loses a pregnancy. Columns 1 to 3 include miscarriage, stillbirth, and abortion in the measurement of pregnancy loss, while columns 4 to 6 exclude induced abortion from the measurement.

Table 3.D.3: Summary statistics of major variables observed at woman level in RePEAT Uganda 2015 data.

| VARIABLES | (1)<br>N | (2)<br>mean | (3)<br>sd |
|---|---|---|---|
| Panel A. Outcome variables. | | | |
| Number of pregnancies | 2,517 | 2.471 | 3.118 |
| Number of pregnancies for women with at least one pregnancy | 1,340 | 4.642 | 2.860 |
| Number of pregnancies for women at ages of 40 and above | 398 | 6.256 | 3.279 |
| Number of pregnancies for women at ages of 45 and above | 186 | 5.941 | 3.598 |
| Panel B. Explanatory variables. | | | |
| 1 if having ever had a pregnancy loss | 2,517 | 0.126 | 0.332 |
| 1 if having ever had more than one pregnancy loss | 2,517 | 0.018 | 0.133 |
| 1 if having ever had a pregnancy loss excl. abortion | 2,517 | 0.122 | 0.328 |
| 1 if having ever had more than one pregnancy loss excl. abortion | 2,517 | 0.018 | 0.133 |
| Age | 2,517 | 26.63 | 10.11 |
| Years of education | 2,485 | 6.388 | 3.525 |
| Marital status: Single | 2,508 | 0.486 | 0.500 |
| Marital status: Married | 2,508 | 0.443 | 0.497 |
| Marital status: Widowed | 2,508 | 0.026 | 0.159 |
| Marital status: Separated | 2,508 | 0.036 | 0.185 |
| Marital status: Divorced | 2,508 | 0.009 | 0.095 |
| Marital status: Other | 2,508 | 0.001 | 0.028 |
| Panel C. Clarity of survey responses. | | | |
| 1 if DNK, RTA, and DNR about year of miscarriage and stillbirth | 2,517 | 0.025 | 0.155 |
| 1 if having ever had miscarriage or stillbirth | 2,517 | 0.122 | 0.328 |
| 1 if DNK, RTA, and DNR about year of abortion | 2,517 | 0.00397 | 0.0629 |
| 1 if having ever had abortion | 2,517 | 0.00477 | 0.0689 |

*Source*: RePEAT Uganda 2015. *Notes*: This table shows the summary statistics of major variables for women who were as old as 15 to 49 years as of the survey. The number of pregnancies counts multiple births as one observation, which makes a slight difference from the number of children. Years of education is the minimum years of schooling required to achieve the reported highest grade. DNK stands for 'do not know,' RTA for 'refuse to answer,' and DNR for 'do not recall.'

Table 3.D.4: Regression results for the effect of pregnancy loss in DHS and RePEAT data.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Outcome | Spacing interval in years | | | | | | | |
| Data source | DHS | | | | RePEAT | | | |
| Post-pregnancy loss dummies | $L_1$ | $L_2$ | $L_3$ | $L_4$ | $L_1$ | $L_2$ | $L_3$ | $L_4$ |
| Post-loss pregnancy | -0.009 | | | | 0.104 | | | |
| | (0.082) | | | | (0.129) | | | |
| 1st post-loss pregnancy | | 0.387*** | 0.398*** | 0.400*** | | 0.495*** | 0.494*** | 0.497*** |
| | | (0.092) | (0.092) | (0.092) | | (0.135) | (0.135) | (0.134) |
| +2nd post-loss pregnancy | | -0.494*** | | | | -0.330* | | |
| | | (0.104) | | | | (0.168) | | |
| 2nd post-loss pregnancy | | | -0.599*** | -0.596*** | | | -0.321** | -0.318* |
| | | | (0.105) | (0.105) | | | (0.164) | (0.163) |
| +3rd post-loss pregnancy | | | -0.371*** | | | | -0.337* | |
| | | | (0.115) | | | | (0.198) | |
| 3rd post-loss pregnancy | | | | -0.400*** | | | | -0.385* |
| | | | | (0.116) | | | | (0.217) |
| +4th post-loss pregnancy | | | | -0.339*** | | | | -0.290 |
| | | | | (0.130) | | | | (0.218) |
| Observations | 12,748 | 12,748 | 12,748 | 12,748 | 4,880 | 4,880 | 4,880 | 4,880 |
| Number of pid | 3,005 | 3,005 | 3,005 | 3,005 | 1,106 | 1,106 | 1,106 | 1,106 |
| R-squared | 0.344 | 0.352 | 0.353 | 0.353 | 0.325 | 0.331 | 0.331 | 0.331 |
| Woman fixed effects | Y | Y | Y | Y | Y | Y | Y | Y |

*Source*: DHS Uganda 2011 and RePEAT Uganda 2015. *Notes*: This table shows selected estimates from the regression of birth spacing measured in years on pregnancy loss indicator(s) and control covariates following equations (3.1) and (3.2). Reported in parentheses are the standard errors clustered at the woman level. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1. The DHS specification includes as covariates the age at pregnancy termination and its square, dummies for parity, sex of the child, sex of the previous child, and dummies for birth year and month of the child, and uses the sample from all the DHS districts. The RePEAT specification includes the same covariates but the sex of the child and that of the previous child, and uses the sample from the districts covered by the RePEAT survey. In both specifications, all the regressions use a pregnancy as the unit of observation where both single and multiple births are counted as one observation. Post-pregnancy loss dummies are $L_2 = \{D^{post}\}$, $L_2 = \{D^{post,1}, D^{post,2+}\}$, $L_3 = \{D^{post,1}, D^{post,2}, D^{post,3+}\}$, and $L_4 = \{D^{post,1}, D^{post,2}, D^{post,3}, D^{post,4+}\}$. All regressions use a pregnancy as the unit of observation where both single and multiple births are counted as one observation.

Table 3.D.5: Regression results for the effect of pregnancy loss including and excluding induced abortion.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Outcome | Spacing interval in years | | | | | | | |
| Abortion | Included | | | | Excluded | | | |
| Post-pregnancy loss dummies | $L_1$ | $L_2$ | $L_3$ | $L_4$ | $L_1$ | $L_2$ | $L_3$ | $L_4$ |
| Post-loss pregnancy | 0.104 | | | | 0.143 | | | |
| | (0.129) | | | | (0.126) | | | |
| 1st post-loss pregnancy | | 0.495*** | 0.494*** | 0.497*** | | 0.519*** | 0.518*** | 0.520*** |
| | | (0.135) | (0.135) | (0.134) | | (0.134) | (0.134) | (0.133) |
| +2nd post-loss pregnancy | | -0.330* | | | | -0.276* | | |
| | | (0.168) | | | | (0.165) | | |
| 2nd post-loss pregnancy | | | -0.321** | -0.318* | | | -0.268* | -0.266* |
| | | | (0.164) | (0.163) | | | (0.161) | (0.161) |
| +3rd post-loss pregnancy | | | -0.337* | | | | -0.284 | |
| | | | (0.198) | | | | (0.195) | |
| 3rd post-loss pregnancy | | | | -0.385* | | | | -0.317 |
| | | | | (0.217) | | | | (0.212) |
| +4th post-loss pregnancy | | | | -0.290 | | | | -0.251 |
| | | | | (0.218) | | | | (0.216) |
| Observations | 4,880 | 4,880 | 4,880 | 4,880 | 4,880 | 4,880 | 4,880 | 4,880 |
| Number of ppid | 1,106 | 1,106 | 1,106 | 1,106 | 1,106 | 1,106 | 1,106 | 1,106 |
| R-squared | 0.325 | 0.331 | 0.331 | 0.331 | 0.325 | 0.331 | 0.331 | 0.331 |
| Woman fixed effects | Y | Y | Y | Y | Y | Y | Y | Y |

*Source*: RePEAT Uganda 2015. *Notes*: This table shows selected estimates from the regression of birth spacing measured in years on pregnancy loss indicator(s) and control covariates following equations (3.1) and (3.2). Reported in parentheses are the standard errors clustered at the woman level. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1. All the regressions use a pregnancy as the unit of observation where both single and multiple births are counted as one observation. Post-pregnancy loss dummies are $L_1 = \{D^{\text{post}}\}$, $L_2 = \{D^{\text{post},1}, D^{\text{post},2+}\}$, $L_3 = \{D^{\text{post},1}, D^{\text{post},2}, D^{\text{post},3+}\}$, and $L_4 = \{D^{\text{post},1}, D^{\text{post},2}, D^{\text{post},3}, D^{\text{post},4+}\}$. All regressions use a pregnancy as the unit of observation where both single and multiple births are counted as one observation.

Table 3.D.6: Regression results for the effect of pregnancy loss using different timing definitions.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Outcome | Spacing interval in years | | | | | | | |
| Timing | First | | | | Counter reset | | | |
| Post-pregnancy loss dummies | $L_1$ | $L_2$ | $L_3$ | $L_4$ | $L_1$ | $L_2$ | $L_3$ | $L_4$ |
| Post-loss pregnancy | 0.147 | | | | 0.147 | | | |
| | (0.125) | | | | (0.125) | | | |
| 1st post-loss pregnancy | | 0.570*** | 0.569*** | 0.568*** | | 0.532*** | 0.536*** | 0.535*** |
| | | (0.135) | (0.135) | (0.134) | | (0.130) | (0.130) | (0.130) |
| +2nd post-loss pregnancy | | -0.294* | | | | -0.298* | | |
| | | (0.163) | | | | (0.159) | | |
| 2nd post-loss pregnancy | | | -0.272 | -0.273 | | | -0.336** | -0.336** |
| | | | (0.169) | (0.169) | | | (0.159) | (0.159) |
| +3rd post-loss pregnancy | | | -0.313* | | | | -0.257 | |
| | | | (0.189) | | | | (0.176) | |
| 3rd post-loss pregnancy | | | | -0.292 | | | | -0.255 |
| | | | | (0.206) | | | | (0.186) |
| +4th post-loss pregnancy | | | | -0.331 | | | | -0.259 |
| | | | | (0.211) | | | | (0.192) |
| Observations | 4,880 | 4,880 | 4,880 | 4,880 | 4,880 | 4,880 | 4,880 | 4,880 |
| Number of ppid | 1,106 | 1,106 | 1,106 | 1,106 | 1,106 | 1,106 | 1,106 | 1,106 |
| R-squared | 0.325 | 0.332 | 0.332 | 0.332 | 0.325 | 0.332 | 0.333 | 0.333 |
| Woman fixed effects | Y | Y | Y | Y | Y | Y | Y | Y |
| Panel controls | Y | Y | Y | Y | Y | Y | Y | Y |

*Source*: RePEAT Uganda 2015. *Notes*: This table shows selected estimates from the regression of birth spacing measured in years on pregnancy loss indicator(s) and control covariates following equations (3.1) and (3.2). Reported in parentheses are the standard errors clustered at the woman level. Statistical significance is denoted by *** if p<0.01, ** if p<0.05, and * if p<0.1. All the regressions use a pregnancy as the unit of observation where both single and multiple births are counted as one observation. Post-pregnancy loss dummies are $L_1 = \{D^{\text{post}}\}$, $L_2 = \{D^{\text{post},1}, D^{\text{post},2+}\}$, $L_3 = \{D^{\text{post},1}, D^{\text{post},2}, D^{\text{post},3+}\}$, and $L_4 = \{D^{\text{post},1}, D^{\text{post},2}, D^{\text{post},3}, D^{\text{post},4+}\}$. In columns (1) through (4), the post-pregnancy loss indicators are defined to take the value of one if the pregnancy is terminated after the first pregnancy loss experience of the woman, while in columns (5) through (8) the parity counter used to define the order of post-loss pregnancies is reset to zero every time the woman experiences a pregnancy loss. All regressions use a pregnancy as the unit of observation where both single and multiple births are counted as one observation.

## Appendix 3.E Model of Pregnancy Timing with Subjective Belief for Pregnancy Loss Probability

In this section, we present a simple microeconomic model of analysing the timing of pregnancy that incorporates the subjective belief of the probability of pregnancy loss. Consider a female who just gave a live birth at period 0, and chooses consumption and pregnancy status for periods 1 and 2 to maximise her utility, and receives terminal utility at period 3. One condition for her lifetime fertility is to achieve at least $\bar{N}$ children. Let her utility function at period $t$ be written as

$$U_t = \lambda X_t + \kappa N_t$$

where $X_t$ represents consumption at period $t$, $N_t$ the number of children, $\lambda > 0$ and $\kappa > 0$ the marginal utility from consumption and children, respectively. Only in period 3, we assume a slightly different utility of the form:

$$U_3 = \lambda X_3 + \kappa N_3 - C\mathbf{I}\{N_3 < \bar{N}\}$$

with $C > 0$ representing the social punishment or psychic cost for failing to achieve the desired number of children while she is reproductive, where $\mathbf{I}\{\cdot\}$ denotes the indicator function that equals one if the condition in the brackets holds and zero otherwise. She faces the budget constraint:

$$Y_t = X_t + \nu N_t$$

where $\nu > 0$ denotes the cost of raising $N_t$ children, relative to the price of consumption goods that is normalized to unity.[37] The only state variable in this model, $N_t$, evolves such that:

$$N_{t+1} = N_t + (1 - \mu_t)P_t$$

where $P_t = 1$ if the female becomes pregnant at period $t$ and 0 otherwise, $\mu_t = 1$ if the pregnancy at period $t$ ends in a loss and 0 otherwise.

We consider $\bar{N} = N_0 + 1$, which implies that the female has to give at least one live birth during periods 1 and 2. In other words, she chooses whether to get pregnant and attempt to give a live birth at period 1, or period 2, or both, to achieve her total fertility equal to or larger than $\bar{N}$. To make the model more realistic and consistent with the African context, we impose the following assumption on the utility from the number of children.

***Assumption 1***. *The marginal utility of children is positive but its marginal benefit is smaller than that from consumption. That is,*

$$\lambda > \frac{\kappa}{\nu}.$$

This condition is consistent with the fertility literature using data from sub-Saharan Africa which finds that the number of children that females want to have in their lifetime is generally smaller than that desired by their partners and that females actually end up making in their lifetime (Ashraf et al., 2014). In our data, we find that approximately

[37] We assume that getting pregnant in period $t$ is costless, but this is not too unrealistic, particularly when the female is just one child to completion of her desired lifetime fertility with larger $\bar{N}$, since other older children can help her with chore works and other household activities.

two thirds of females end up having at least their desired number of children (for details, see Appendix Figure 3.A.3). This feature can be incorporated into our model herein by Assumption 1.

At period $t = 3$, the female chooses only consumption, where her optimisation problem is written as:

$$\max_{X_3} U_3 = \lambda X_3 + \kappa N_3 - C\mathbf{I}\{N_3 < \bar{N}\} \quad \text{subject to} \quad Y_3 \geq X_3 + \nu N_3.$$

Since she has only one choice variable, $X_3$, the solution is $X_3^* = Y_3 - \nu N_3$, and her value function is $V_3(N_3) = \lambda X_3^* + \kappa N_3 = \lambda Y_3 + (-\lambda\nu + \kappa)N_3$. In the case $N_3 < \bar{N}$, the value is $V_3(N_3) = \lambda Y_3 + (-\lambda\nu + \kappa)N_3 - C$.

At period $t = 2$, she faces the optimisation problem:

$$\max_{X_2, P_2} U_2 + \beta V_3(N_3)$$

subject to

$$Y_2 = \lambda X_2 + \nu N_2 \quad \text{and} \quad N_3 = N_2 + (1 - \mu_2)P_2$$

where $\beta$ represents the time discount rate between 0 and 1. To solve for the solutions, we substitute $X_2 = Y_2 - \nu N_2$ and consider the two cases as below.

*Case 1*. If $N_2 = \bar{N}$, take the utility difference when $P_2 = 1$ and 0:

$$\begin{aligned} Q_2(\bar{N}) &\equiv U_2(P_2 = 1, N_2 = \bar{N}) + \beta[(1-\mu_2)V_3(\bar{N}+1) + \mu_2 V_3(\bar{N})] \\ &\quad - [U_2(P_2 = 0, N_2 = \bar{N}) + \beta V_3(\bar{N})] \\ &= \beta(1-\mu_2)(-\lambda\nu + \kappa) \end{aligned}$$

With Assumption 1, we have that $(-\lambda\nu + \kappa) < 0$, so the solution is:

$$P_2^*(\bar{N}) \equiv \mathbf{I}\{Q_2(\bar{N}) > 0\} = 0.$$

*Case 2*. If $N_2 = \bar{N} - 1$, again take the utility difference:

$$\begin{aligned} Q_2(\bar{N}-1) &\equiv U_2(P_2 = 1, N_2 = \bar{N}-1) + \beta[(1-\mu_2)V_3(\bar{N}) + \mu_2 V_3(\bar{N}-1)] \\ &\quad - [U_2(P_2 = 0, N_2 = \bar{N}-1) + \beta V_3(\bar{N}-1)] \\ &= \beta(1-\mu_2)(-\lambda\nu + \kappa + C). \end{aligned}$$

We assume that the female attempts to get pregnant in order to avoid the terminal utility loss of $C$ at any cost. In other words, we assume that $C$ is so large that any female with $N_2 \leq \bar{N}$ attempts to avoid it by getting pregnant at period 2. The following formally states this condition:

***Assumption 2***. *The utility loss when the number of pregnancies does not reach the desired number, $\bar{N}$, is large enough to induce any female to get pregnant at period 2 and attempt* *to avoid incurring it at period 3. That is,*

$$C > \lambda\nu - \kappa.$$

By Assumption 2, we have that $Q_2(\bar{N}-1) > 0$, so the solution is:

$$P_2^*(\bar{N}-1) \equiv \mathbf{I}\{Q_2(\bar{N}-1) > 0\} = 1.$$

The value functions at period 2 depend on the number of children at the beginning of period 2 such that:

$$\begin{aligned}
V_2(\bar{N}) &= U_2(P_2 = 0, N_2 = \bar{N}) + \beta V_3(\bar{N}) \\
&= \lambda Y_2 + (-\lambda\nu + \kappa)\bar{N} + \beta V_3(\bar{N}) \\
V_2(\bar{N}-1) &= U_2(P_2 = 1, N_2 = \bar{N}-1) + \beta\left[(1-\mu_2)V_3(\bar{N}) + \mu_2 V_3(\bar{N}-1)\right] \\
&= \lambda Y_2 + (-\lambda\nu + \kappa)(\bar{N}-1) + \beta\left[(1-\mu_2)V_3(\bar{N}) + \mu_2 V_3(\bar{N}-1)\right].
\end{aligned}$$

Define the differences in the value functions as:

$$\begin{aligned}
\Delta V_3 \equiv& V_3(\bar{N}) - V_3(\bar{N}-1) = -\lambda\nu + \kappa + C \\
\Delta V_2 \equiv& V_2(\bar{N}) - V_2(\bar{N}-1) = -\lambda\nu + \kappa + \beta\mu_2\Delta V_3.
\end{aligned}$$

At period 1, she faces the optimisation problem of the form:

$$\max_{X_1, P_1} U_1 + \beta V_2(N_2) = U_1 + \beta\left[P_1(1-\mu_1)V_2(\bar{N}) + (1 - P_1(1-\mu_1))V_2(\bar{N}-1)\right]$$

subject to

$$Y_1 = X_1 + \nu N_1$$

$$N_2 = N_1 + (1 - \mu_1)P_1$$

Substitute $X_1 = Y_1 - \nu N_1$ to reduce the choice variables, and consider the utility difference if the female gets pregnant and if not:

$$\begin{aligned}
Q_1 \equiv & U_1(P_1 = 1) + \beta[(1 - \mu_1)V_2(\bar{N}) + \mu_1 V_2(\bar{N} - 1)] - U_1(P_1 = 0) - \beta V_2(\bar{N} - 1) \\
= & U_1(P_1 = 1) - U_1(P_1 = 0) + \beta(1 - \mu_1)\Delta V_2.
\end{aligned}$$

The female gets pregnant at period 1 if $Q_1 > 0$, i.e.,

$$P_1^* = \mathbf{I}\{Q_1 > 0\}.$$

For the female making the pregnancy decision at period 1, $\tilde{\mu} = \mu_1 = \mu_2$ denotes her perceived probability of pregnancy loss, which takes the same value for all the periods ahead. Consider the derivative of $Q_1$ with respect to $\tilde{\mu}$:

$$\begin{aligned}
\frac{\partial Q_1}{\partial \tilde{\mu}} = & - \beta\Delta V_2 + \beta(1 - \tilde{\mu})\frac{\partial \Delta V_2}{\partial \tilde{\mu}} \\
= & \beta(\lambda\nu - \kappa) + \beta^2(1 - 2\tilde{\mu})(-\lambda\nu + \kappa + C)
\end{aligned}$$

Here, the first term is positive by Assumption 1, and Assumption 2 implies that the second term can be positive and negative depending upon the value of $\tilde{\mu}$ as in the two cases below.

*Case 1.* If $0 \leq \tilde{\mu} \leq 1/2$, $\partial Q_1/\partial\tilde{\mu} > 0$. This suggests that females who perceives the probability of pregnancy loss being not too high always attempt to get pregnant and give a birth in period 1, rather than period 2.

That is, as long as the perceived probability of pregnancy loss is at most one half, an increase in the perceived probability always makes the female more likely to become pregnant in period 1, which leads to a shorter birth interval from period 0.

*Case 2.* If $1/2 < \tilde{\mu} \leq 1$, the sign of $\partial Q_1/\partial\tilde{\mu}$ is ambiguous and depends upon which of the positive and negative part is larger, the marginal utility from an additional child at period 2 ($\lambda\nu - \kappa$), or the difference in the value functions at period 3 ($\Delta V_3$).

Even if the perceived probability is larger than one half ($1/2 < \tilde{\mu} \leq 1$), we can further show that $\partial Q_1/\partial\tilde{\mu} > 0$ is equivalent to:

$$C < \frac{1 - \beta(1 - 2\tilde{\mu})}{-\beta(1 - 2\tilde{\mu})}(\lambda\nu - \kappa) = \frac{1 + \beta\tilde{\gamma}}{\beta\tilde{\gamma}}(\lambda\nu - \kappa)$$

where $\tilde{\gamma} \equiv -(1 - 2\tilde{\mu}) > 0$ for $1/2 < \tilde{\mu} \leq 1$. It implies that the smaller the perceived probability of pregnancy loss, the larger the value of $\tilde{\gamma}$, and the more likely this inequality is to hold. In other words, as long as the social punishment $C$ is not too large, or as long as the perceived probability of pregnancy loss is not too large, an increase in perceived probability of pregnancy loss can still lead to a larger likelihood of pregnancy in period 1, i.e., shorter birth spacing.

# Chapter 4

# CONCLUSION

In this dissertation, I showcase two empirical studies that examine a cause and consequence of human capital investment. To conclude, I summarise implications of these studies for the literature on human capital investment behaviours.

Chapter 2 focuses on a consequence, where I find that female education has reduced brideprice practice in Uganda. Based on the findings, I discuss a conjecture that altruistic parents of the bride face a trade-off between wealth transfer at their daughter's marriage in the form of brideprice and their daughter's future sound marital life and choose the latter which increases more in their daughter's education.

One implication of this study is that accumulated human capital can have a much broader impact than is directly expected. Primary motives for policies to increase educational attainments include the improvement of labour productivity, which can result in an increase in wages and welfare of households. However, the study in Chapter 2 reveals that it can change a cultural practice. Although this study does not permit drawing normative implications, it highlights the importance of a more comprehensive perspective on what would result from a policy than on its direct impact.

Nevertheless, my analysis in Chapter 2 leaves several things unaddressed. One of them is to more rigorously test the proposed mechanism for the decline in brideprice payment practices. The results imply that the cultural practice declines through a trade-off faced by the bride's parents between the instantaneous utility from brideprice and the future altruistic utility from the bride's sound marital life, an interpretation consistent with Gaspart and Platteau (2010). However, further data collection and analysis are called for to formally test this implication.

Another is to conduct a unified analysis on the effect of female education on both the extensive and intensive margins of brideprice practice. The findings from this study and in the past literature suggest that female education may opposite effects on the two margins. Due to data quality issues, I did not conduct this analysis, but this is an interesting topic that future studies can explore.

The study in Chapter 3 focuses on a cause of human capital investment in the context of reproductive behaviours in the low-income and high-fertility world. Using data from Uganda in particular, I find that an experience of miscarriage and stillbirth reduces the birth spacing intervals for all the subsequent births. I also show suggestive evidence that this behavioural change is consistent with updating of subjective belief on the probability of pregnancy loss.

One implication to the literature on human capital investment behaviours is that, unlike the studies in the literature (e.g., Mira (2007)), the effects of belief upating seem to be large. One possibility may be that, in my study, the belief updating is assumed to occur due to one's own experience. Other studies that also find large effects seem to report belief updating triggered by own experience,[1] while studies that report small to no effects on behaviours consider cases where agents are provided with third-party information.[2] This discrepancy may be related to an argument by Dupas (2011) that the effect of information intervention depends on not only the contents of information but also on who provide the information to whom and how. Thus, the investigation into belief updating and resulting behavioural change when different types of information are provided seems to be a natural extension.

Another implication is that the acquisition of new information and revision of beliefs may be an important aspect of decision making under uncertainty. There are various uncertainties when agents make significant decisions: returns to children's education, quality of drugs sold at pharmacies, earnings at destination of migration, etc. are all uncertain *ex ante*. In order to calculate expected pay-offs and choose actions, people need a pre-formed belief for the probabilities of possible outcomes. This implies that, if these beliefs change, their observable behaviours can also change. Thus, natural questions that arise are how these beliefs are formed and when they change. Moreover, lacking formal institutions for risk management, investigation into the formation and revision of beliefs is of paramount importance in developing countries, as a behavioural change may potentially have a life-changing impact. My analysis in Chapter 3 presents just one example, where people are likely to attempt to know the probability of pregnancy loss from their actual experience and adapt their fertility behaviours according to it. Obviously, there can be many more situations where people form and revise their beliefs and change

[1] Examples include college students who learn their own academic grades and switch majors accordingly (Zafar, 2011).

[2] Examples include East African pastoralists who change rainfall expectations when given computer-based rain forecasts but do not change their cattle grazing patterns (Lybbert et al., 2007).

behaviours accordingly, and more research is called for to uncover the process of decision making under uncertainty for the advancement of economic research and improvement of economic policies.

# Bibliography

Alesina, A., Algan, Y., Cahuc, P., and Giuliano, P. (2015). Family Values and the Regulation of Labor. *Journal of the European Economic Association*, 13(4):599–630.

Alesina, A., Brioschi, B., and Ferrara, E. L. (2016). Violence against women: A cross-cultural analysis for Africa. *National Bureau of Economic Research*.

Alesina, A. and Giuliano, P. (2015). Culture and Institutions. *Journal of Economic Literature*, 53(4):898–944.

Ambrus, A., Field, E., and Torero, M. (2010). Muslim family law, prenuptial agreements, and the emergence of dowry in Bangladesh. *The Quarterly Journal of Economics*, 125(3):1349–1397.

Anderson, S. (2007). The economics of dowry and brideprice. *The Journal of Economic Perspectives*, 21(4):151–174.

Ando, M., Dahlberg, M., and Engström, G. (2017). The risks of nuclear disaster and its impact on housing prices. *Economics Letters*, 154:13–16.

Arunachalam, R. and Logan, T. D. (2016). On the heterogeneity of dowry motives. *Journal of Population Economics*, 29(1):135–166.

Ashraf, N., Bau, N., Nunn, N., and Voena, A. (2019). Bride Price and Female Education. *Journal of Political Economy*, page 704572.

Ashraf, N., Field, E., and Lee, J. (2014). Household Bargaining and Excess Fertility: An Experimental Study in Zambia. *American Economic Review*, 104(7):2210–2237.

Becker, G. S. (1991). *A treatise on the family*. Harvard University Press, Cambridge, Mass, enlarged edition.

Behrman, J. A. (2015). The effect of increased primary schooling on adult women's HIV status in Malawi and Uganda: Universal Primary Education as a natural experiment. *Social Science & Medicine*, 127:108–115.

Behrman, J. R. and Rosenzweig, M. R. (2004). Returns to Birthweight. *Review of Economics and Statistics*, 86(2):586–601.

Bell, D. (1998). Wealth transfers occasioned by marriage: A comparative reconsideration. In *Schweizer, T. & D. R. White (ed.) Kinship, networks, and exchange*, number 12 in Structural analysis in the social sciences. Cambridge University Press, New York.

Bhalotra, S. (2010). Fatal fluctuations? Cyclicality in infant mortality in India. *Journal of Development Economics*, 93(1):7–19.

Bhalotra, S. and Clarke, D. (2016). The Twin Instrument. *IZA Discussion Paper*, 10405:107.

Bhalotra, S. and van Soest, A. (2008). Birth-spacing, fertility and neonatal mortality in India: Dynamics, frailty, and fecundity. *Journal of Econometrics*, 143(2):274–290.

Bharadwaj, P., Lundborg, P., and Rooth, D.-O. (2018). Birth Weight in the Long Run. *Journal of Human Resources*, 53(1):189–231.

Bishai, D. and Grossbard, S. (2010). Far above rubies: Bride price and extramarital sexual relations in Uganda. *Journal of Population Economics*, 23(4):1177–1187.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2007). From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes. *Quarterly Journal of Economics*, 122(1):409–439.

Bratti, M., Fiore, S., and Mendola, M. (2016). Family Size, Sibling Rivalry and Migration: Evidence from Mexico.

Brown, S. (2008). Miscarriage and Its Associations. *Seminars in Reproductive Medicine*, 26(05):391–400.

Buckles, K. S. and Munnich, E. L. (2012). Birth spacing and sibling outcomes. *Journal of Human Resources*, 47(3):613–642.

Card, D., Lee, S., D., Pei, Z., and Weber, A. (2015). Inference on causal effects in a generalized regression kink design. *Econometrica*, 83(6):2453–2483.

Card, D., Lee, D. S., Pei, Z., and Weber, A. (2016). Regression Kink Design: Theory and Practice. *National Bureau of Economic Research*.

Conde-Agudelo, A., Rosas-Bermúdez, A., and Kafury-Goeta, A. C. (2006). Birth spacing and risk of adverse perinatal outcomes: a meta-analysis. *Journal of American Medical Association*, 295(15):1809–1823.

Daily Monitor (2016a). Man murders father-in-law for failing to return bride price.

Daily Monitor (2016b). Man storms hospital, kills wife.

Daily Monitor (2017). A ring does not give you the right to jail your spouse.

Daily Monitor (2018a). Man fakes own kidnap to pocket bride price cash.

Daily Monitor (2018b). Parents trading daughters for cowns in Lira markets.

Daily Monitor (2018c). Selling girls in Lira market: Atrocious barter trade.

Daily Monitor (2019). Mbale cyclist killed by in-laws, several houses torched in revenge attack.

Deininger, K. (2003). Does cost of schooling affect enrollment by the poor? Universal primary education in Uganda. *Economics of Education Review*, 22(3):291–305.

Delavande, A. (2014). Probabilistic Expectations in Developing Countries. *Annual Review of Economics*, 6(1):1–20.

Demographic and Health Survey (2001). Uganda DHS EdData Survey 2001: Education Data for Decisionmaking. Technical report, Ministry of Education and Sports, Kampala.

Demographic and Health Survey (2004). DHS EdData Education Profile: Uganda. Technical report, ORC Macro, Maryland, USA.

Demographic and Health Survey (2012). Uganda Demographic and Health Survey 2011. Technical report, Kampala, Uganda: UBOS and Calverton, Maryland: ICF International Inc.

Demographic and Health Surveys (2013). Standard Recode Manual for DHS 6.

Di Tella, R., Galiant, S., and Schargrodsky, E. (2007). The formation of beliefs: evidence from the allocation of land titles to squatters. *The Quarterly Journal of Economics*, 122(1):209–241.

Doya, N. R. (2017). The Drafting History of the Uganda Penal Code (Amendment) Act and Challenges to Its Implementation. *Statute Law Review*, 38(2):226–239.

Dupas, P. (2011). Health Behavior in Developing Countries. *Annual Review of Economics*, 3(1):425–449.

Fafchamps, M. and Quisumbing, A. (2005a). Assets at marriage in rural Ethiopia. *Journal of Development Economics*, 77(1):1–25.

Fafchamps, M. and Quisumbing, A. R. (2005b). Marriage, Bequest, and Assortative Matching in Rural Ethiopia. *Economic Development and Cultural Change*, 53(2):347–380.

Fafchamps, M. and Quisumbing, A. R. (2007). Chapter 51 Household Formation and Marriage Markets in Rural Areas. In *Handbook of Development Economics*, volume 4, pages 3187–3247. Elsevier.

Fenske, J. (2015). African polygamy: Past and present. *Journal of Development Economics*, 117:58–73.

Fernández, R. (2013). Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century. *American Economic Review*, 103(1):472–500.

Fink, A. and Stratmann, T. (2015). U.S. housing prices and the Fukushima nuclear accident. *Journal of Economic Behavior & Organization*, 117:309–326.

Fletcher, J. M. (2011). The medium term schooling and health effects of low birth weight: Evidence from siblings. *Economics of Education Review*, 30(3):517–527.

Ford, H. B. and Schust, D. J. (2009). Recurrent pregnancy loss: etiology, diagnosis, and therapy. *Reviews in obstetrics and gynecology*, 2(2):76.

Francis, A. M. (2011). Sex ratios and the red dragon: using the Chinese Communist Revolution to explore the effect of the sex ratio on women and children in Taiwan. *Journal of Population Economics*, 24(3):813–837.

Frandsen, B. R. (2017). Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design When the Running Variable is discrete. In *Advances in Econometrics: Regression Discontinuity Designs*, volume 38, pages 281–315. Emerald Publishing Limited.

Gaspart, F. and Platteau, J. (2010). Strategic Behavior and Marriage Payments: Theory and Evidence from Senegal. *Economic Development and Cultural Change*, 59(1):149–185.

Gelman, A. and Imbens, G. (2019). Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, 37(3):447–456.

Gershman, B. (2016). Witchcraft beliefs and the erosion of social capital: Evidence from Sub-Saharan Africa and beyond. *Journal of Development Economics*, 120:182–208.

Goody, J. (2011). Bridewealth and Dowry in Africa and Eurasia. In *Bridewealth and Dowry*, number 7 in Cambridge papers in social anthropology, pages 1–58. Cambridge University Press, New York, paperback re-issue edition.

Gray, J. P. (1999). A corrected ethnographic atlas. *World Cultures*, 10(1):24–85.

Grogan, L. (2008). Universal Primary Education and School Entry in Uganda. *Journal of African Economies*, 18(2):183–211.

Gruber, J. and Hungerman, D. M. (2008). The church versus the mall: What happens when religion faces increased secular competition? *The Quarterly Journal of Economics*, 123(2):831–862.

Heckman, J. J., Holtz, V. J., and Walker, J. R. (1985). New Evidence on the Timing and Spacing of Births. *American Economic Review*, 75(2):179–184.

Heckman, J. J. and Walker, J. R. (1990). The Relationship Between Wages and Income and the Timing and Spacing of Births: Evidence from Swedish Longitudinal Data. *Econometrica*, 58(6):1411.

Hertwig, R., Barron, G., Weber, E. U., and Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological science*, 15(8):534–539.

Hotz, V. J., McElroy, S. W., and Sanders, S. G. (2005). Teenage childbearing and its life cycle consequences exploiting a natural experiment. *Journal of Human Resources*, 40(3):683–715.

Jacoby, H. G. (1995). The economics of polygyny in Sub-Saharan Africa: Female productivity and the demand for wives in Côte d'Ivoire. *Journal of Political Economy*, pages 938–971.

Jones, D., Molitor, D., and Reif, J. (2018). What Do Workplace Wellness Programs Do? Evidence from the Illinois Workplace Wellness Study. Technical Report w24229, National Bureau of Economic Research, Cambridge, MA.

Karimi, A. (2014). The spacing of births and women's subsequent earnings: Evidence from a natural experiment. *Institute for Evaluation of Labour Market and Education Policy*.

Kaye, D. K., Ekstrom, A. M., Kyomuhendo, G. B., and Johansson, A. (2005). Implications of bride price on domestic violence and reproductive health in Wakiso District, Uganda. *African Health Sciences*, 5(4):300–303.

Keats, A. (2018). Women's schooling, fertility, and child health outcomes: Evidence from Uganda's free primary education program. *Journal of Development Economics*, 135:142–159.

Kim, J. (2010). Women's education and fertility: an analysis of the relationship between education and birth spacing in Indonesia. *Economic Development and Cultural Change*, 58(4):739–774.

Larsen, E. C., Christiansen, O. B., Kolte, A. M., and Macklon, N. (2013). New insights into mechanisms behind miscarriage. *BMC Medicine*, 11(1):154.

Lee, D. S. and Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 48(2):281–355.

Lundqvist, H., Dahlberg, M., and Mörk, E. (2014). Stimulating Local Public Employment: Do General Grants Work? *American Economic Journal: Economic Policy*, 6(1):167–192.

Lybbert, T. J., Barrett, C. B., McPeak, J. G., and Luseno, W. K. (2007). Bayesian Herders: Updating of Rainfall Beliefs in Response to External Forecasts. *World Development*, 35(3):480–497.

Maitra, P. and Pal, S. (2008). Birth spacing, fertility selection and child survival: Analysis using a correlated hazard model. *Journal of Health Economics*, 27(3):690–705.

Manoli, D. and Turner, N. (2014). Cash on Hand and College Enrollment: Evidence from Population Tax Data and Policy Nonlinearities. Technical Report w19836, National Bureau of Economic Research, Cambridge, MA.

Masuda, K. and Yamauchi, C. (2018). How Does Female Education Reduce Adolescent Pregnancy and Improve Child Health?: Evidence from Uganda's Universal Primary Education for Fully Treated Cohorts. *The Journal of Development Studies*, forthcoming.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714.

MIFUMI Project (2009). Bride-Price, Poverty and Domestic Violence in Uganda.

Miller, A. R. (2011). The effects of motherhood timing on career path. *Journal of Population Economics*, 24(3):1071–1100.

Ministry of Education and Sports (1999). The Ugandan Experience of Universal Primary Education. Technical report, Ministry of Education and Sports, Uganda.

Mira, P. (2007). Uncertain Infant Mortality, Learning, and Life-Cycle Fertility. *International Economic Review*, 48(3):809–846.

Montenegro, C. E. and Patrinos, H. A. (2014). Comparable Estimates of Returns to Schooling around the World. Technical Report 7020, World Bank.

Moretti, E. (2004). Human Capital Externalities in Cities. In *Handbook of Regional and Urban Economics*, volume 4, pages 2243–2291. https://doi.org/10.1016/S1574-0080(04)80008-7.

Murdock, G. P. (1967). *Ethnographic Atlas*. University of Pittsburgh Press.

Museveni, Y. K. (1996a). *Tackling the tasks ahead: election manifesto*. Uganda.

Museveni, Y. K. (1996b). *Y. Kaguta Museveni's speech at the launch of his election manifesto at Uganda International Conference Centre*. Uganda.

New Vision (2016a). Man Hacks Son In Law Over Sh500,000 In Unpaid Dowry.

New Vision (2016b). Man Kills Granddad Over Bride Price.

Nishimura, M., Yamano, T., and Sasaoka, Y. (2008). Impacts of the universal primary education policy on educational attainment and private costs in rural Uganda. *International Journal of Educational Development*, 28(2):161–175.

Norton, M. (2005). New evidence on birth spacing: promising findings for improving newborn, infant, child, and maternal health. *International Journal of Gynecology & Obstetrics*, 89:S1–S6.

Okello, G. M. & Ors., v. Attorney General (2004). Constitutional Petition No. 2 of 2003.

Olson, J., Meur, C., Press, G., Felix, M., and Center, C. B. A. H. R. (1996). *The Peoples of Africa: An Ethnohistorical Dictionary*. ABC-Clio ebook. Greenwood Press.

Olsson, O. and Paik, C. (2016). Long-run cultural divergence: Evidence from the Neolithic Revolution. *Journal of Development Economics*, 122:197–213.

Oster, E. (2018). Does disease cause vaccination? Disease outbreaks and vaccination response. *Journal of Health Economics*, 57:90–101.

Papps, I., Davis, J., Grossbard-Shechtman, A., Mair, L., Pryor, F. L., Reyna, S. P., Gratton, N. E., Schneider, H. K., Tapper, N., and Wilder, W. D. (1983). The Role and Determinants of Bride-Price: The Case of a Palestinian Village [and Comments and Reply]. *Current Anthropology*, 24(2):203–215.

Platteau, J.-P. and Gaspart, F. (2007). The Perverse Effects of High Brideprices. *World Development*, 35(7):1221–1236.

Psacharopoulos, G. and Patrinos, H. A. (2018). Returns to investment in education: a decennial review of the global literature. *Education Economics*, 26(5):445–458.

Rao, V. (1993). The rising price of husbands: A hedonic analysis of dowry increases in rural India. *Journal of Political Economy*, 101(4):666–677.

Riddell, A. (2003). The introduction of free primary education in sub-Saharan Africa. *Paper commissioned for the EFA Global Monitoring Report 2003/4, The Leap to Equality*.

Rosenzweig, M. R. and Wolpin, K. I. (1988). Heterogeneity, Intrafamily Distribution, and Child Health. *The Journal of Human Resources*, 23(4):437.

Rossi, P. and Rouanet, L. (2015). Gender Preferences in Africa: A Comparative Analysis of Fertility Choices. *World Development*, 72:326–345.

Sawada, Y. and Lokshin, M. (2009). Obstacles to school progression in rural Pakistan: An analysis of gender and sibling rivalry using field survey data. *Journal of Development Economics*, 88(2):335–347.

Silver, R. M., Varner, M. W., Reddy, U., Goldenberg, R., Pinar, H., Conway, D., Bukowski, R., Carpenter, M., Hogue, C., Willinger, M., Dudley, D., Saade, G., and Stoll, B. (2007). Work-up of stillbirth: a review of the evidence. *American Journal of Obstetrics and Gynecology*, 196(5):433–444.

Simpson, J. L. (2007). Causes of fetal wastage. *Clinical obstetrics and gynecology*, 50(1):10–30.

Singh, S., Remez, L., Sedgh, G., Kwok, L., and Onda, T. (2018). Abortion Worldwide 2017: Uneven Progress and Unequal Access. Technical report, Guttmacher Institute, New York.

Stasavage, D. (2005). The role of democracy in Uganda's move to universal primary education. *The Journal of Modern African Studies*, 43(1):53–73.

Stokes, J. (2009). *Encyclopedia of the Peoples of Africa and the Middle East*. Facts On File.

Tertilt, M. (2005). Polygyny, Fertility, and Savings. *Journal of Political Economy*, 113(6):1341–1371.

The Independent (1996). Museveni lays ghosts of past dictators - Ugandan elections: Voters expected to endorse president who led them out of the darkness. *The Independent*.

Tirgil, A., Gurol-Urganci, I., and Atun, R. (2018). Early experience of universal health coverage in Turkey on access to health services for the poor: regression kink design analysis. *Journal of Global Health*, 8(2):020412.

Tomer, Y. (2014). Mechanisms of Autoimmune Thyroid Diseases: From Genetics to Epigenetics. *Annual Review of Pathology: Mechanisms of Disease*, 9(1):147–156.

Tumushabe, J. (2006). The Politics of HIV/AIDS in Uganda. *United Nations Research Institute for Social Development Social Policy and Development Programme Paper Number 28*.

Uganda Bureau of Statistics (2003). 2003 Statistical Abstract. Technical report.

Uganda Electoral Commission (1996). Ugandan Presidential Elections: Constituency Results.

Uganda Law Reform Commission (2010). *Study Report on Marriage and Divorce in Uganda*. Uganda Law Reform Commission.

van den Berg, M. M., van Maarle, M. C., van Wely, M., and Goddijn, M. (2012). Genetics of early miscarriage. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1822(12):1951–1959.

van Soest, A. and Saha, U. R. (2018). Relationships between infant mortality, birth spacing and fertility in Matlab, Bangladesh. *PLOS ONE*, 13(4):e0195940.

Vogl, T. S. (2013). Marriage Institutions and Sibling Competition: Evidence from South Asia. *The Quarterly Journal of Economics*, 128(3):1017–1072.

Wantchekon, L., Klašnja, M., and Novta, N. (2015). Education and Human Capital Externalities: Evidence from Colonial Benin *. *The Quarterly Journal of Economics*, 130(2):703–757.

Wendo, C. (2004). African women denounce bride price. *The Lancet*, 363:716.

Westfall, P. H. and Young, S. S. (1993). *Resampling-Based Multiple Testing: Examples and Methods for P-Value Adjustment*. A Wiley-Interscience publication. Wiley.

Whitworth, A. and Stephenson, R. (2002). Birth spacing, sibling rivalry and child mortality in India. *Social Science & Medicine*, 55(12):2107–2119.

Winikoff, B. (1983). The Effects of Birth Spacing on Child and Maternal Health. *Studies in Family Planning*, 14(10):231.

World Health Organization (2007). Report of a WHO Technical Consultation on Birth Spacing. Technical Report WHO/RHR/07.1, World Health Organization, Geneva, Switzerland.

Yamano, T., Sserunkuuma, D., Otsuka, K., Omiat, G., Ainembabazi, J. H., and Shimamura, Y. (2004). The 2003 REPEAT Survey in Uganda: Results. Technical report, Foundation for Advanced Studies on International Development, Tokyo.

Zafar, B. (2011). How Do College Students Form Expectations? *Journal of Labor Economics*, 29(2):301–348.