# THREE ESSAYS ON DEVELOPMENT AND HEALTH ECONOMICS

Jaehyun Jung

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2018

# ABSTRACT

Three Essays on Development and Health Economics

Jaehyun Jung

This dissertation consists of three essays on development and health economics.

In the first chapter, I study how abortion responds to drought-induced transitory income shocks and generates unintended demographic consequences under son preference. I focus on rural Vietnam where low rainfall induces a short-run downturn through a reduction in rice yields. With widely available sex-selection technologies at a low cost under son preference, Vietnamese parents can decide the quantity and the sex of child simultaneously, and it can be directly observed from rich household-level data on abortions. Linking rich microdata on fertility with droughts defined at a fine geographic unit, I first find no effects of droughts on the number and the composition of mothers who conceive. I then find compelling evidence that affected mothers were 30% more likely to get abortions, and the effect was mainly driven by the income effect because most abortions occurred in the pre-harvest season of the next rice crop when consumption smoothing is difficult. Surprisingly, droughts are associated with disproportionately more abortions of female fetuses, which exacerbated the problem of the skewed sex ratio: the affected birth cohorts become more male-biased due to the six abortions of female fetuses to one aborted male fetus, explaining up to approximately 3% of the sex ratio imbalance in rural Vietnam from 2004-2013. While a full rebound in births in approximately two years appears more consistent with the effect on the timing of fertility, the effect on the sex ratio at birth emphasizes that even transitory income shocks can have long-run demographic consequences. Thus, this study can shed light on how the gender gap can persist during a process of economic development. This study also enhances our understanding of the mechanism through which credit-constrained mothers adjust their

fertility to smooth consumption. Finally, this study can provide timely evidence to developing countries which witness demographic transitions to low infant mortality but are vulnerable to extreme weather events.

In the second chapter (joint work with Anna Choi and Semee Yoon), we study how usual economic activities can harm the health of people who are living in other countries. This study investigates the adverse effect of transboundary particulate matter on fetal health. The adverse health effects at exposures to particulate matter are evident by a handful of experimental and epidemiological studies. The health effects of $PM_{2.5}$, which has a diameter of less than 2.5 micrometers, are particularly alarming because those hazardous particles are so diminutive that they can easily enter the bloodstream to cause cardiovascular and respiratory diseases. Unlike the consensus in the United States on the negative impact of pollution on human health, the evidence for the relationship between pollution and health in developing countries is not straightforward to quantify due to the lack of accurate pollution and welfare measures as well as the difficulty of finding exogenous variables to purge other endogenous factors. However, this study can circumvent these endogeneity concerns by exploiting the unique meteorological settings which can trigger transboundary transport of particulate matter. The westerlies from heavily polluted eastern China carry pollutants to South Korea, thereby intermittently exposing the population to pollution above threshold levels. We find that conditional on local weather and pollution trends, one standard deviation increase in Beijing's $PM_{2.5}$ explains 1.1% of standard deviation of daily fetal mortality rates in South Korea. We hope that the results of this research can suggest the first accurate cost estimates of transboundary fine-particle to highlight the urgent need for regional cooperation.

In the third chapter, the unintended consequences of economic activities on human capital in developing countries can be further emphasized by a randomized control trial study (joint with Hyuncheol Bryant Kim, Booyuel Kim and Cristian Pop-Eleches). We use a four-year long follow-up of an intervention based on a two-step randomized design within classrooms in secondary schools in Malawi to understand the impact of male circumcision on risky sexual behaviors and the role that peers play in the decision and consequences of being circumcised. Although medical male circumcision can reduce HIV infections, its preventive effects may diminish if circumcised men are more likely to engage in risky sexual behaviors. Despite a number of short-term studies of risk compensation following male circumcision, there is scant rigorous evidence on how these behavioral responses change in the longer term. This study is the first evaluation of risk compensation over such a long follow-up period. Our analysis yields three main results. First, we show that the intervention substantially increased the demand for male circumcision for the students assigned to the treatment group. Second, we find evidence of positive peer effects in the decision to get circumcised among untreated students. Third, we find evidence of risk compensation using biomarkers of sexually transmitted infection for those who got circumcised due to the intervention, but not for those induced by peer effects.

# Contents

# List of Figures

# List of Tables

# Acknowledgements

First and foremost, I am deeply grateful to my committee members, Douglas Almond, Cristian Pop-Eleches, and Rodrigo Soares. I feel immensely fortunate to have their insightful advice and untiring support and guidance. I also would like to express my gratitude to Hyuncheol Bryant Kim and Booyuel Kim for providing an invaluable opportunity to participate in the Malawi projects and guiding me to become an academic researcher. Semee Yoon deserves special thanks as my research has also benefited from our long collaborations.

This dissertation has benefited enormously from comments from Eric Verhoogen, Michael Best, Jonas Hjort, Wolfram Schlenker, and participants at the sustainable development seminar and the development workshop and colloquia at Columbia University. I also extend my appreciation to those who have led and supported this exceptional Sustainable Development Ph.D. program: John Mutter for kind advice and invaluable wisdom, and Tomara Aldrich for continuous support.

Over the past six years, I have benefited enormously from the friendship and support of amazing colleagues in Sustainable Development and Economics; I want to thank my doctoral cohort, Francis Annan, Eugenie Dugoua, Eyal Frank, Ruinan Liu, Steffen Merte and my fellow doctoral students, So Yoon Ahn, Yoon Joo Jo, Suanna Oh, Anouch Missirian, and Ruiwen Lee not only for being a constant source of inspiration and learning but also for showing various important aspects of life.

Lastly, my family has provided unwavering support and optimism. Their love and encouragement made everything possible. Most of all, this achievement would not have been possible without the sacrifice and unconditional love of my parents.

*Dedicated to my parents*

# Chapter 1

# Can Abortion Mitigate Transitory Shocks?: Demographic Consequences under Son Preference

## 1.1 Introduction

Credit-constrained households in poor countries employ various coping strategies to mitigate negative aggregate shocks (Dercon, 2002). Childbearing can be subject to such arrangements in light of Becker's characterization of children as normal goods (Becker, 1960) and the observed decline in birth rates after economic recessions in both developed and developing countries (Sobotka et al., 2011; Chatterjee and Vogl, 2017). However, empirical research on how women adjust fertility to smooth consumption, particularly through abortion, has remained scarce. In fact, abortion merits careful consideration in comparison with other birth control methods; abortion not only is increasing in developing countries (Sedgh et al., 2016), but also can ensure the sex of a child once paired with ultrasound, having far-reaching demographic implications in the long run.

This paper aims to examine how negative aggregate shocks affect childbearing through abortion and the sex ratio at birth under son preference. I attempt to link adverse rainfall shocks, which credibly translate into transitory economic downturns, with a unique mother-level dataset on abortion in Vietnam. Rural Vietnam may be an ideal setting to shed light on this connection. Droughts inflict sizable damage on rural households' economic conditions through reduced rice yields, and abortion is widely available at a low cost and serves as one of the common birth control methods. Importantly, son preference in Vietnam can play a crucial role in generating differential fertility responses if the sex of a fetus is ascertained via ultrasound. By exploiting approximately 1 million rural mothers' decisions on abortion and prenatal sex determination in almost every rural district in Vietnam from 2004-2013, I provide reduced-form estimates for the effect of droughts on how rural mothers adjust fertility using abortion and whether the abortions become more sex-selective.

How do rainfall-induced aggregate shocks affect childbearing, particularly through abortion? If this procedure occurs under son preference, would it exacerbate the imbalance in the sex ratio at birth? Although the relationship between income and fertility can be diversely characterized depending on time horizons on both sides, e.g., permanent or transitory shocks on income, completed fertility or timing of childbearing, negative rainfall shocks can shed light on the effect of *transitory* income shocks on the *timing* of fertility, while holding other shadow prices constant. This is because 1) rainfall-induced economic shocks are transitory;[1] 2) transitory shocks have no effects on childbearing when credit markets are perfect (Dehejia and Lleras-Muney, 2004), but credit constraints are commonly faced by the poor, especially after aggregate shocks (Dercon, 2002); and 3) the low skill and low participation of female labor in agrarian economies plausibly limit the scope of influence exerted by the opportunity cost of childbearing (Bhalotra and Rocha, 2012).

The decision of credit-constrained rural parents facing transitory rainfall shocks is thus essentially reduced to a comparison of benefits from current consumption with those from childbearing. Choosing to give birth by forgoing current consumption would hardly be the optimal choice, especially if delaying is not costly. In particular, abortion can be preferred to other pre-emptive measures to avoid pregnancy because abortion not only allows for extra months to update the expected payoff of giving birth (Ananat et al., 2009), but also is the only method to ensure the sex of a child when used with ultrasound. Thus, if the cost of 'sex-selective' abortion is sufficiently low,[2] and a mother expects extra marginal utility by having a long-awaited son, she would rather choose to give birth to a child only if the revealed sex of a fetus is male.

To test the intertemporal substitution effects on childbearing, I draw on exogenous variations in local economic conditions by exploiting year-to-year variations in rainfall realizations. Specifically, I define droughts as seasonal rainfall occurring below the 20th percentile of the district-specific long-run rainfall distribution in 1984-2013 using Climate Hazards

[1] Rainfall shocks are assumed to be short-run and transitory by previous studies in the economic literature (Paxson, 1992; Townsend, 1994; Jayachandran, 2006; Kaur, 2014).

[2] Here, I assume the health cost of abortion is minimal given safe abortion practices in Vietnam, and that the psychological cost is also not significant, which I discuss in Section 1.2.

Group InfraRed Precipitation with Station data (CHIRPS).[3]

To confirm that the constructed drought shocks translate into negative economic shocks, I first combine rainfall with yearly rice yields. Then, I demonstrate a more direct effect on household consumption using detailed expenditure surveys. My drought measure leads to an approximately 2 percent decrease in the yield of the main rice crop and lower expenditure of non-food items by approximately 8 percent. To identify a specific time window when parents are faced with a stark trade-off between childbearing and current consumption, I further examine monthly household expenditure and find that affected households are unable to smooth consumption, especially in the pre-harvest season of the next rice crop.

To derive reduced-form estimates on fertility outcomes sequentially from conception to abortion, birth, and finally the sex ratio at birth, I link droughts with the 9 rounds of the Population Change and Family Planning Surveys (PCSs) from 2004-2013. The PCS contains rich information on fertility; in particular, it reports approximately 1 million rural mothers' decisions on abortion from almost every rural district in Vietnam.

The first main result indicates that a married woman is more likely to get abortion by approximately 30 percent in the following year after the drought event. The effect is significant both statistically and economically and remains robust to a variety of controls, including a mother's fertility history and district-specific linear time trends.

I show that the income effect is the primary pathway for explaining the abortion responses. With no changes in the number and composition of mothers who conceive, I find the birth rate drops significantly in a specific quarter, indicating that most abortions occurred in the pre-harvest season of the next rice crop when affected rural households exhibit the inability to smooth consumption. Furthermore, I find no effects on abortion among urban mothers, and substantially muted impacts on abortion for mothers living in provinces where one more

[3] The Vietnamese administrative levels have the following structure: Region $\supset$ Province $\supset$ District $\supset$ Commune. There are 674 districts in Vietnam, and 629 districts are used for the analyses in this paper. The remaining districts are mostly urban districts within major cities. The average geographic area of rural districts in my sample is 457.7 km$^2$.

rice crop can be harvested. I also find that neither the opportunity cost of a woman's labor nor the direct health impact of droughts appears dominant factors to drive the effects on abortion because of weak evidence on differential responses in a woman's labor market participation and infant mortality rates.

Importantly, droughts lead abortion to be more sex-selective, e.g., 6 female fetuses were aborted to every 1 male fetus, contributing to approximately 3 percent of the sex imbalance witnessed in rural Vietnam in 2004-2013. The evidence on sex selection can be corroborated by two further results. First, the effect of droughts on the sex ratio at birth is primarily driven by the sub-sample surveyed when the ultrasound scans for fetal sex determination were more prevalent. Second, the ultrasound scan particularly during the 12-16 weeks of pregnancy—an essential prerequisite for sex-selective abortions—was more likely to be conducted for mothers who gave birth to a son after droughts.

I argue that abortion can be considered as a coping strategy to address aggregate shocks, but at the same time, it worsened the sex imbalance to be more male-skewed. I find a rebound of fertility approximately 2 years after the shortfall in birth rates, which implies that the effects of droughts on fertility are more pertinent to the timing of fertility rather than lifetime fertility. However, if son preference interacts with the adjustment and skews the sex ratio by abortion, the resulting demographic consequences can persist in the long run.

This paper builds on and contributes to three strands of the literature. First, this paper speaks to the effects of income on fertility, particularly the fertility response to transitory economic downturns. In the previous studies focusing on developing countries, infant mortality should be taken into consideration to understand how negative income shocks affect childbearing and determine lifetime fertility (Dyson, 1993; Pitt and Sigle, 1997; Artadi, 2005). Furthermore, due to a dearth of credible mother-level data, there is scant literature providing micro-level evidence, whereas a number of studies in the US leverage exogenous variations in household-level permanent income to show its positive association with fertility (Lindo, 2010; Black et al., 2013; Lovenheim and Mumford, 2013 Dettling and Kearney, 2014). This study attempts to fill this substantial gap in the literature in recent developing country contexts where pre-emptive measures to adjust fertility become more available. Specifically, I highlight abortion as the main mechanism for adjustment in childbearing to mitigate transitory income shocks.

Second, my findings are complementary to the literature regarding how economic determinants interact with traditional cultural norms, e.g., son preference through prenatal sex selection. The influential factors discussed thus far include a decline in fertility given the importance of having at least one son (Ebenstein, 2010; Jayachandran, 2017), the availability of cheap sex-selective technologies (Bhalotra and Cochrane, 2010; Chen et al., 2013), and an increase in income when the cost of sex selection is high (Almond et al., 2017). This paper presents a new case study exploring how prenatal sex selection can be driven by short-run economic downturns. I find low-cost sex selection technologies make female 'fetuses' bear a disproportionate burden of aggregate shocks during the lean season, which used to befall female infants (Behrman, 1988).

Third, this study contributes to an emerging strand of the literature on the various coping strategies of credit-constrained households against weather-induced aggregate shocks (Dell et al., 2014). If adverse weather shocks occur, and credit and insurance markets are incomplete, poor households smooth consumption not only by changing their usual diet (Bhattacharya et al., 2003; Hou, 2010; Lohmann and Lechtenfeld, 2015), but also by migrating to urban areas (Groeger and Zylberberg, 2016), by investing less in a female infant's health (Rose, 1999; Maccini and Yang, 2009; Anttila-Hughes and Hsiang, 2013), or by gaining cash transfers through engaging in risky sexual behaviors (Burke et al., 2015) or adjusting the timing of child marriage (Corno et al., 2017). In this paper, I suggest that delaying the timing of fertility through abortion can also be a margin of adjustment to smooth consumption for poor households to cope with adverse rainfall shocks.

The remainder of the chapter proceeds as follows. I describe the relevant empirical context of Vietnam and the data in Section 1.2. Section 1.3 introduces the empirical strategy I employ. The results are presented in Section 1.4, and additional outcomes are further discussed in Section 1.5. After showing the robustness of the main results in Section 1.5, Section 1.6 concludes the paper.

## 1.2 Background and Data

### 1.2.1 Abortion, Son Preference and Sex Selection

Vietnam has one of the highest abortion rates in the world: approximately 26 abortions per 1,000 women of reproductive age or 0.6 induced abortions per woman during her lifetime in the early 2000s (Committee for Population, Family and Children [Vietnam] and ORC Macro, 2003; Sedgh et al., 2007). In addition to the most liberal laws, which allow abortion up to 22 weeks of pregnancy, this high rate is primarily attributable to the procedure's very low cost and widespread availability across all levels of public and private health facilities (Whittaker, ed, 2010). The user fees for the two main procedures performed in rural districts are considerably low: approximately US$2 and US$6.1 for early-stage (up to 8 weeks) and for later-stage abortions (up to 16 weeks), respectively (PATH and Reproductive Health Department, 2006).[4] The rapid expansion of the public health system has also lowered the barrier to abortion; not only does every rural commune have its own health clinic where early-stage abortions are provided, but 73 percent of rural residents can also pay a visit to district health centers less than a half-hour away for later-stage procedures.[5]

Abortion for married women is socially well-accepted throughout Vietnam (Whittaker,

[4] There are a number of ways for rural married women to get abortions at subsidized prices or free of charge. For example, she will not be charged if she subscribes to the family planning program, reports a malfunction in contraceptive methods or is subject to the poverty alleviation program by localities (Teerawichitchainan and Amin, 2010).

[5] Based on the author's calculation using the commune survey section of the VHLSS 2004.

ed, 2010). This anti-natalist social atmosphere has largely been established by the family planning program—the one-or-two child policy—that intensely promoted abortions along with intrauterine devices (IUDs) as the primary birth control options. Other modern contraceptives, particularly temporary methods such as condoms and pills, were frequently unavailable, and their use was discouraged by local family planning officials (Teerawichitchainan and Amin, 2010; Figure 1.A.2). These very limited options in birth control have been deemed to be the main culprits for the high abortion rates in Vietnam because abortion has replaced other modern contraceptives (Whittaker, ed, 2010).

The recent proliferation of ultrasound in the 2000s has led to the use of abortion to selectively terminate female fetuses. Vietnam has a strong son preference, influenced by Confucianism (Belanger, 2002). Having at least one son can be of primary importance because he will be responsible for the rituals to worship his ancestors, support his elderly parents and inherit family properties (Pham et al., 2008). Meanwhile, ultrasound scans, which were first introduced as part of reproductive health services, rapidly increased approximately 10-fold from 1998 to 2007 (Guilmoto et al., 2009). This rapid adoption stemmed not only from supply-side factors, such as a very low cost (20,000 VND or US$1.30 per scan) and numerous providers, but also from the easy and reliable identification of fetal sex as early as 12 weeks of pregnancy (Gammeltoft and Nguyen, 2007). Whereas prenatal sex determination using ultrasound was already prevalent in the 2000s for urban mothers, ultrasound was still expanding across rural regions in the mid-2000s; thus, the rate surpassed 80 percent for rural mothers in the late 2000s (Figure 1.1). This widespread availability of ultrasound scans undoubtedly plays a crucial role in the recent male-biased sex ratio at birth in the 2000s (Figure 1.2 and Figure 1.A.1). This demographic transformation is common among countries of India and East Asia with son preference, especially when ultrasound scans become widespread and the cost of abortion is low (Das Gupta et al., 2003).[6]

[6] Alarmed by abnormally male-skewed sex ratio in the early 2000s, the Vietnamese government has issued several official decrees to ban 'sex-selective' abortion since 2003 (Pham et al., 2011). However, similar to the

Last but not least, other relevant environments affecting aggregate fertility have been steady in the 2000s. Infant mortality rates have been relatively low in Vietnam in the 2000s in comparison with other neighboring countries or those with a comparable income level (Figure 1.A.3a). While total fertility and relevant population policies exhibit little change, the excess male infant mortality has been stable and close to the ratio found in countries without son preference in the 2000s (Figure 1.A.3b), suggesting that there was neither an immediate fertility squeeze to have a son nor active substitution between pre- and postnatal discrimination against female births in the sample period (Goodkind, 1996).

### 1.2.2 Data

I use the PCS to construct fertility measures. The Vietnam General Statistics Office (GSO) and the United Nations Population Fund designed the survey and have conducted it every year since 2000, providing repeated cross-sectional and nationally representative 3% samples.[7] In addition to socio-demographic information for all usual residents of a household, the survey asked all married women of childbearing age from 15-49 whether they had an 'induced' abortion in the last year along with requesting extensive information about contraceptive use, antenatal care, and fetal sex determination. This line of questioning makes the PCS a unique dataset containing comprehensive fertility characteristics of mothers that are representative at the national level, which is rarely available in developing countries.[8] I pool nine waves of the PCS from 2004-2008 and 2010-13, which report abortion, giving me a sample size of

previous cases in India and China, those bans had little impact because sex selection was hardly discernible from abortions for family planning (Pham et al., 2008).

[7] The reference date of the PCS is April 1st of the current year. The survey asks about events that occurred in the 12 months preceding the reference date. As a result, the survey conducted in the year n describes demographic trends for January-March in year $n$ and April-December in year $n-1$. (Figure 1.3)

[8] The Demographic and Health Surveys (DHS) Vietnam 2002 has more extensive information about fertility, including complete pregnancy histories, miscarriages, and abortion, but the sample size is only 5,665 women of reproductive age (Committee for Population, Family and Children [Vietnam] and ORC Macro, 2003). I use these data to understand the timing of abortion in Section 1.4.

approximately 1 million mothers from almost every rural district across Vietnam.[9]

To explore the effects of adverse rainfall shocks on the economic conditions of rural households, I supplement expenditure information, which is not available in the PCS, using the Vietnamese Household Living Standard Survey (VHLSS). This biennial survey has been conducted since 2000, and I use six rounds from 2004 to 2014 to match the sample period of the PCS data. Due to inconsistency in the sampling and the change of the recall period for consumption from yearly to monthly, I use the first three waves (i.e., 2004, 2006, 2008) to examine the effects on yearly expenditure and the latter three waves (i.e., 2010, 2012, 2014) to analyze the effects on monthly expenditure and labor participation. I also collect several province-level statistics from the GSO to examine the effects of droughts on yearly crop yields, monthly CPIs, and infant mortality rates.

Since my empirical strategy hinges on the relationship between rainfall and the economic condition of a household, I rely on the sample that consists of married women in rural areas, where the association is presumed to be more pronounced than urban regions. In addition, to make parental preference and cultural settings for fertility and family planning comparable across districts, I focus on those 54 provinces out of 64 where the Kinh, the major ethnicity in Vietnam, account for more than half of population in a province.[10] I further limit the sample to include mothers who had up to three children at the baseline, which is a year before the reference date of the survey, and a single childbirth at a time to have comparable results using the birth order controls.

Table 1.1 provides the summary statistics of married women aged 15-49 in the main sample. I present urban women's statistics to show how rural mothers differ from their urban

[9] There is no PCS 2009 due to the census in the same year.

[10] Using the 2009 Census of Vietnam, I chose 10 provinces where more than 50 percent of the heads of household were not ethnically Kinh. The 10 provinces are Cao Bang, Bac Kan, Ha Giang, Lang Son, Lai Chau, Son La, Dien Bien, Hoa Binh, Lao Cai, Tuyen Quang, and the locations are mapped in Figure 1.A.6. Not only are these provinces different in terms of ethnic composition, but they are also the 10 poorest provinces, thereby exhibiting lower socio-economic development and higher fertility and infant mortality rates (Glewwe et al., eds, 2004). I run the falsification test for this sample in Table 1.B.2 Panel B.

counterparts. More than half of rural mothers have completed only primary or lower secondary school, and approximately 73 percent of mothers had at least one son at the time of the survey. Rural mothers are less likely to get abortions than urban mothers (0.7 percent vs. 0.8 percent), while they are more likely to use contraceptives. Conditional on childbirth in the last year, approximately 79 percent of mothers knew the fetal sex by ultrasound at approximately 20 weeks gestation. Turning to household characteristics collected from the VHLSS 2004-2008, more than half of rural households grew rice, suggesting that their economic conditions depend substantively on rice yields.

### 1.2.3 Agriculture in Vietnam, Rainfall Data and Construction of Shocks

Agriculture in Vietnam is an important sector, and rice is of particular importance as the main source of income for a majority of rural households. Rice is cultivated in more than 50 percent of the annual crop land and accounts for more than 50 percent of the annual crop production, positioning Vietnam as the second-largest rice-exporting country. Furthermore, rice is also the main staple for the population by supplying more than 50 percent of the people's daily calorie intake (Jaffee et al., 2016). Thanks to the favorable climate, double cropping of rice is common across the country except in the northern mountainous provinces, and triple cropping is even possible, especially in the Mekong Delta region. Accordingly, the cropping season is defined by three rice crops: winter-spring rice, summer-autumn rice, and autumn-winter rice. The winter-spring crop (spring rice) is most productive and is planted across all regions, accounting for approximately 30 percent of the total annual crop production. The cropping cycles are largely determined by the rainfall patterns rather than temperature (Figure 1.3), and thus, rainfall is the most important weather variation affecting rural incomes if irrigation is ineffective (Lobell and Field, 2007).[11]

I create season-specific rainfall shocks for each district using CHIRPS version 2.0. CHIRPS

[11] In fact, according to the VHLSS 2004, only 35 percent of communes responded that their annual crop lands are fully irrigated, and only 69 percent of the land is irrigated across the country.

incorporates 0.05-degree resolution satellite imagery with station readings to provide monthly estimates of precipitation beginning in 1981.[12] In particular, CHIRPS is well-suited for the purpose of this study in that it is explicitly tailored for monitoring agricultural drought across the globe (Funk et al., 2015). To match those estimates of precipitation with the outcome variables of interest, I spatially aggregate the estimates up to the district level by taking the land area-weighted average of monthly rainfall for each district. The resolution of CHIRPS is fine enough (0.05×0.05 degree) for each overlaid administrative boundary to include several grids within a narrow geographic area to produce a weighted average, addressing potential measurement errors arising from projecting coarser grids with precipitation estimates onto smaller geographic units (Dell et al., 2014). I collapse monthly rainfall to create a season-year measure of precipitation, e.g., dry (December-March) and wet (April-November) season rainfall for each district, to precisely match rainfall realizations to agricultural cycles.

To determine rainfall shocks that capture unusually low rainfall realizations relative to a district's typical experience, I define a drought as seasonal rainfall occurring below the 20th percentile of the district-specific long-run rainfall distribution of 1984-2013. Since this study design exploits a panel structure instead of relying on cross-sectional comparisons, my estimates are not biased due to the correlations between the unobservable characteristics of different locations and their mean levels of rainfall (Dell et al., 2014).[13] As widely adopted by a host of recent studies using rainfall shocks in India and Africa (Jayachandran, 2006; Kaur, 2014; Burke et al., 2015; Shah and Steinberg, 2017; Corno et al., 2017), this measure enables me to capture significant economic impacts on a rural household while holding constant the other determinants of fertility and reproductive behaviors of a married woman. This simple specification employing one indicator variable of drought assumes that fertility decisions are

[12] 0.05 degrees is equivalent to approximately 5 kilometers at the equator.

[13] Drought and excessive rainfall can both adversely affect rice yields, but drought, in particular, inflicts heavier damage on productivity than flood does (Auffhammer et al., 2012). In addition, excessive rainfall is not adequate to be a single meaningful weather variable for proxying economic shocks, due not only to the complicated hydrologic and climatic conditions that cause flood but also to its nonlinear relationship with crop yield (Guiteras et al., 2015).

intrinsically conservative and discreet, and thus, they would respond to economic shocks in a nonlinear fashion. However, in Section 1.5, I will carefully verify the validity of the empirical specification using binned indicators for given percentiles or restricted cubic splines to explain the marginal effects between rainfall levels and abortion decisions. Figure 1.A.6 plots the districts that experienced droughts in the sample period, 2004-2013. Districts that were afflicted with droughts are relatively evenly dispersed across the country.

## 1.3 Empirical Strategy

This study aims to estimate the causal impact of economic shocks on the fertility decisions of rural married women. By exploiting arguably random year-to-year variations in seasonal rainfall within a district, I compare the childbearing decisions of mothers who are affected by droughts with those of mothers who are not affected but share similar underlying preferences for completed fertility and the gender composition of their children. I estimate the effects of droughts on various outcomes of interest using the following regression:

$$Y_{idt} = \alpha + \beta Drought_{dt} + \sum_{s}^{S}\sum_{l}^{L}\gamma_{sl}R_{s,t-l} + X'_{it}\delta + \tau_t + \mu_d + \theta_d * t + \varepsilon_{idt} \tag{1.1}$$

where $i$, $d$, $t$, and $s$ denote mother, district, year and season, respectively.

The outcome $Y_{idt}$ includes a wide array of variables related to a married woman's fertility, such as pregnancy, abortion, contraceptive use, birth and prenatal sex determination. Most variables are binary; for example, 'abortion' is coded one if a mother terminated her pregnancy within the survey year. The main coefficient of interest is $\beta$ on $Drought_{d,t}$, which becomes one if the dry or wet season rainfall occurs below the 20th percentile ('drought') of the district-specific seasonal rainfall distribution from 1984-2013. This coefficient captures the effect of negative rainfall shocks (1st quintile) compared to what would be typically observed after the other realizations of rainfall, i.e., the average effects of seasonal rainfall in the 2nd, 3rd, 4th and 5th quintiles of the historical rainfall distribution.

I primarily focus on the effect of droughts in the last dry season, $Drought_{d,t-1}$. The reference period of the PCS starts right after the realization of $Drought_{d,t-1}$; additionally, abnormally low dry season rainfall also most clearly translates into the economic conditions of rural households at that time because the crop cycle of the most productive spring rice approximately overlaps the dry season (Figure 1.3).[14]

To focus on the contemporaneous effects of $Drought_{d,t-1}$, a vector of $R_{s,t+l}$, which represent the deviations of other season-year rainfalls from the local median up to two year lags ($L = \{0,1,2\}$), enters the equation to control for the potential effects of wet and previous dry season rainfalls.[15] $X_{it}$ is a vector of covariates, which consist of a mother-specific observable characteristics that are age in year $t$ and its quadratic, and two indicator variables for woman $i$'s educational attainment and her relationship to the household head. These covariates would potentially control not only for the diverse opportunity costs of childbearing but also for the marginal utility of having an extra child.

The main specification also includes indicators for the number of children one year before the reference date of the survey (hereafter, the baseline) to account for the birth order of current conception in the survey year. I further control for the demand for a son by replacing the indicators of parity with the dummies for the gender composition of living children.[16] In particular, the inclusion of the gender composition, i.e., whether a mother has at least one son, is useful because it serves as one of the key determinants of demand for a son through sex-selective abortion. Since the PCS does not report the month of abortion but only asked whether a mother had an abortion in the last 12 months from the reference date, I include survey year fixed effects in lieu of the calendar year fixed effects to capture any sampling difference across waves over ten years. In addition to the full set of district fixed effects, I

[14] I will show that the effects of other season-year rainfalls are not significant in Section 1.5.

[15] The results are robust to the inclusion of raw precipitation values, the logs of seasonal rainfalls, and the quintile indicators.

[16] This would give me a total of 15 cases by considering all of the possible combinations of the birth order (e.g., 0 to 3) and the sex of the children at the baseline.

include district-specific linear time trends to purge the effects from differential expansions of public health services, increases in agricultural productivity, etc. by district level over time.

As robustness checks, I also include two sets of control variables pertaining to a woman's fertility: the birth spacing and the age at the first birth[17] and spousal characteristics, such as his educational attainment and age for the subsample.[18] Lastly, the $\varepsilon_{idt}$ represents mother-specific idiosyncratic factors for each outcome variable. Robust standard errors are clustered at the district level because I define rainfall shocks using district-specific rainfall distributions over 30 years.

## 1.4 Results

This section is organized as follows. Section 1.4 investigates how the drought shocks affect rice outputs and the expenditure of rural households. Then, I present in sequence the fertility results: the effect of droughts on a rural woman's conception, abortion, birth, and her newborn's sex.

### 1.4.1 Effects on Crop Yield and Expenditure

#### 1.4.1.1 Effects on Yearly Yield and Expenditure

I first assess the effects of droughts on yearly yield and expenditure using province-level yield data from the Vietnam GSO in 1995-2014 and yearly expenditure data from the three waves of the VHLSS in 2004, 2006 and 2008. The log yield of spring rice and all rice crops, which

[17] To be precise, I construct the variable of birth spacing by calculating the interval between the baseline of the survey (one year before the reference date) and the year and month of the previous birth that occurred before the baseline instead of the period between the most recent live birth that occurred in the present survey year and the previous birth. This avoids the bad control problem because the concurrent live birth that occurred in the present survey is also affected by weather shocks (Angrist and Pischke, 2009). For a married woman who had no first child as of the reference date, I use the months elapsed since the year and month of her marriage.

[18] Spousal information can be merged only if a married woman is the head of household or the spouse of the household head; additionally, the husband's educational attainment is reported only after 2006 in the PCS.

also include winter and autumn rice yields, are the outcomes for the analyses of the effects on crop outputs. For the analyses of expenditure, I use the log of total yearly expenditure and the logs of yearly expenditure on subcategories such as food items and non-food items. I aggregate the dry season rainfalls and define droughts analogously at the province level in lieu of the district level because the province is the unit of analysis for agricultural statistics from the GSO, and the VHLSS has the smaller sample size of households per district in comparison with the PCS.

Figure 1.4 presents local linear regression results that describe how crop yields and various expenditure measures are associated with rainfall levels in the dry season. Spring rice and all rice yields are plotted on rainfall percentiles after controlling for province-specific linear time trends and year and province fixed effects in Figure 1.4 (a) and Figure 1.4 (b). I find positive associations between the dry season rainfall and rice yields, but low rainfall levels in the dry season have unambiguously detrimental impacts on the yield of spring rice, particularly if rainfall is below approximately the 20th percentile. However, this correlation is weakened once rainfall is above the median, providing confidence in employing the 20th percentile cut-off as the drought indicator in the following analyses.

Figure 1.4 (c)-(f) provide local linear regression estimates of the four measures of expenditure on the dry season rainfall percentile after controlling for province, year and survey quarter fixed effects and household-level covariates.[19] Non-food item expenditure has a strong positive association with rainfall levels (Figure 1.4 (e)), but the impacts of poor rainfall to the extent of total expenditure (Figure 1.4 (c)) and food expenditure (Figure 1.4 (d)) are muted.

In Table 1.2, I provide the regression results using the drought indicator to examine how the drought measure I create determines rice yield and expenditure. Columns 1 and 2 show that adverse rainfall shocks in the dry season have negative effects on yields; there is a 2.4

[19] The household-level control covariates include the size of household, the household head's age, sex, ethnicity and educational attainment and the log of total expenditure.

percent reduction in the spring rice yield (column 1, $p < 0.01$), and a 1.3 percent reduction in all rice yields (column 2, $p < 0.05$).[20]

Given that annual rice production in Vietnam has shown robust growth at a rate of approximately 4 percent in the 2000s (Jaffee et al., 2016), a rainfall-induced decline in rice yield is difficult to anticipate and can have substantive impacts on rural households. Droughts are also associated with lower non-food expenditure by 8.5 percent ($p < 0.01$, column 4), but the effects are not precisely estimated for the other expenditure measures (columns 3 and 4), as shown in Figure 1.4, primarily because droughts result not only in slumps in agricultural income but also potentially in price surges of food, which constitute approximately half of households' real expenditure (Vu and Glewwe, 2011). These findings are also supported by Figure 1.4 (f), which shows a slight increase in the ratio of expenditure on food to total expenditure in the context of low levels of rain, suggesting that rural households allocate relatively more resources to purchasing food after droughts.

#### 1.4.1.2 Effect on Monthly Expenditure

To examine how negative aggregate shocks have intertemporal substitution effects on childbearing, e.g., postponing births by abortion, for credit-constrained parents, it is crucial to understand how lower income would affect consumption smoothing. Furthermore, describing how consumption is determined by variations in prices after droughts is of primary importance because the consumption of a rural household can be more responsive to seasonal prices than to seasonal income (Paxson, 1993; Khandker, 2012). In particular, high-frequency data on consumption is needed because low-cost abortion permits an immediate substitution between current consumption and childbearing.

I estimate the effects of droughts on monthly expenditure using the VHLSS 2010, 2012,

[20] The reduction in rice production is estimated by similar magnitudes. If the outcome is the log of province-year level production in thousand tons, the point estimates for the spring rice and all rice production are -0.055 ($p < 0.05$), and -0.031 ($p < 0.1$), respectively.

and 2014, which collect consumption information over the month preceding the date of the interviews, which occurred over the year. In addition, I examine the effects on province-level prices using the monthly CPIs of 12 provinces provided by the Vietnam GSO in 2005-2014. I estimate the impact of droughts using the following specification:

$$\begin{aligned} Y_{ipt} \quad = \quad & \alpha + \sum_{q}^{Q} \sum_{k}^{K} \beta_{qk} Drought_{p,t-k} \times Quarter_q + \sum_{q}^{Q} \theta_q Quarter_q \\ & + \sum_{s}^{S} \sum_{l}^{L} \gamma_{sl} R_{s,t+l} + X'_{it} \delta + \tau_{rt} + \mu_p + \varepsilon_{ipt} \end{aligned} \tag{1.2}$$

where $Y_{ipt}$ is the log of the monthly expenditure of household $i$ in province $p$ in year $t$.

Since enumerators' visits to rural households were mostly arranged at the end of each quarter, I include indicators for quarters, $Quarter_q$, that equal one if a household was visited in quarter $q$. To observe the lagged impacts of droughts on consumption, I interact the quarter indicators with the lagged droughts ($Drought_{p,t-1}$) and with the current drought ($Drought_{p,t}$). While maintaining the same household-level control covariates ($X_{it}$) and fixed effects used in the estimation of the effects on yearly expenditure, I replace the year fixed effects with region-year fixed effects ($\tau_{rt}$) to control for different market structures and prices across regions. The main coefficient of interest $\beta_{qk}$ represents the average effect of droughts on monthly expenditure surveyed 0-7 quarters away from the drought, relative to the usual level of monthly expenditure in the unaffected provinces.

I first estimate the effects on price levels using Equation (1.2) after replacing monthly expenditure with province-month-level CPI as the outcome variable. Figure 1.5 (a) plots the coefficients on the interaction terms, $\beta_{qk}$'s, with their 95 percent confidence intervals. Statistically significant positive point estimates on the overall CPI ($p < 0.1$) and food CPI ($p < 0.05$) are initially observed in the 3rd quarter (October-December) after droughts in the dry season, indicating that droughts have lagged effects on local prices.

Figure 1.5 (b) displays the effects of droughts on monthly expenditure in subcategories using Equation (1.2), suggesting that price surges during the 3rd quarter are associated with lower expenditure on some non-food daily items in the same quarter. To be specific, while holding the total monthly expenditure constant (columns 1-3 of Table 1.B.3), affected households maintain their consumption of rice and pork (columns 8 and 9 of Table 1.B.3), the two main food categories, by substituting away from consuming some daily non-food items (the 3rd quarter (Oct-Dec) of Figure 1.5 (b)) and toward the consumption of cheaper items, such as maize, cassava and potato (column 11 of Table 1.B.4).

Figure 1.3 shows that the 3rd quarter after droughts (October-December), when price hikes and the corresponding adjustment of consumption are conspicuous, is the pre-harvest season of winter rice, the next rice crop after spring rice for two-cropping regions.[21] Furthermore, Figure 1.A.11 shows that the adverse effects of droughts on consumption smoothing become attenuated in the triple-cropping provinces where autumn rice is also harvested between the harvests of spring and winter rice crops. Therefore, to examine whether parents want to delay childbearing to smooth consumption after droughts, this pre-harvest season would be the most likely timing for observing the intertemporal substitution effects on childbearing, which is illustrated as an increase in abortion.

### 1.4.2 Effect on Conceptions

I begin the analysis of the effect on fertility by examining how conception responds to adverse rainfall shocks. The measurement of conceptions is improved by here including mother-level abortion and current pregnancy to live birth, which has been a proxy for conception in the prior literature.[22] The indicator for conception in each survey year becomes one if a woman

[21] I find the results discussed thus far are consistent with the findings of Wainwright and Newman (2011) that existing risk-coping strategies of rural Vietnamese households are ineffective in protecting consumption against aggregate shocks. For example, rural Vietnamese households do not have large-scale facilities to store rice, which limit their capability to smooth consumption using rice as precautionary savings (Vu and Glewwe, 2011; Wainwright and Newman, 2011). The failure of perfect consumption smoothing in the pre-harvest season is also found in Bangladesh (Khandker, 2012).

[22] Since the PCS does not provide information on miscarriages, stillbirths or spontaneous abortions, my measure of conception cannot be complete. However, I assume that those cases do not bias the results if droughts do not entail severe malnutrition or alter the disease environment, for which I provide evidence in

reported one of three cases: 1) she gave birth between September and March,[23] 2) she had an induced abortion, and/or 3) she was pregnant at the time of the survey. Figure 1.3 explains the time frame of each variable. In addition, since the PCS asks married women about their current contraceptive use and their reasons for not using any contraceptives, I can ascertain whether rural mothers intend to time their pregnancies in relation to adverse rainfall shocks.

In Table 1.3, I report the estimation results for the conception and contraception responses using Equation (1.1). Panel A presents the estimates for the effects of droughts that occurred during the last 4 months of the survey year, and Panels B and C provide the estimates for 1-year and 2-year lagged effects. I expect the drought that occurred right before the start of the reference period ($Drought_{t-1}$ in Panel B) would have the largest impacts on outcomes. The estimates for the effects on conception are not statistically significant at any conventional level (columns 1-2), suggesting that concurrent or previous droughts do not lead to more or fewer pregnancies.

This finding can be further supported by exploring the effect on contraceptive use. Following the same methodology, I estimate Equation (1.1) for the outcomes of contraceptive use and report the results in Table 1.3, columns 3-8. Columns 3 and 4 do not show any significant effects of droughts on the use of any contraceptive methods, including both traditional methods such as withdrawal or abstinence, and modern methods, which include IUDs, condoms, the pill, and sterilization. Although the droughts that occurred in the dry season right before the survey year (Panel B, column 5) or a year earlier (Panel C, column 5) have statistically significant effects on the use of modern birth control methods, the estimates are not economically sizable compared with the mean (approximately 1 percent) and are not robust to controlling for spousal characteristics. Columns 7 and 8 present the estimates on an

Section 1.5.

[23] The exclusion of births from conception in the first 5 months of the survey (April-August) results from the fact that if aborted pregnancies in April were carried to full term, the earliest possible month of childbirth starts in September because abortion after 16 weeks was rarely reported by DHS 2002. I also used the 22-week threshold, which is the number of weeks of pregnancy after which abortion is prohibited by law. The results are consistent with the 16-week threshold.

indicator for demand for children, which I construct using a woman's response that no birth control was in use because she wants to have a child. These estimates also do not show any significant association with droughts, implying that contraception or abstinence, which are *ex ante* measures of birth timing, play limited roles for rural mothers.

Next, I test for balance to address the selection bias arising from the compositional changes in those mothers who conceive after droughts. In Table 1.4, I first illustrate the socio-economic and fertility characteristics of mothers who conceived after droughts, and the spousal characteristics for the subsample (columns 1,3,5,7 and 9). The balance is presented in columns 2,4,6 and 8 by calculating the differences in the means of those observable characteristics with those of unaffected mothers who conceived. The statistical significance in the difference results from the regressions of the drought indicator and the other control covariates in Equation (1.1) on those characteristics. I also cannot find significant differences in those observable characteristics between the two groups of women, suggesting that the compositional changes of mothers who conceived after being exposed to droughts would not imperil the causal link between drought-induced income shocks and abortion.

## 1.4.3 Effect on Abortion

Table 1.5 presents the main finding of this study from Equation (1.1) estimating the effect of droughts on the probability of abortion among rural married women of reproductive age (15-49) using the nine rounds of the annual PCS in 2004-2008 and 2010-2013. Married women who experienced droughts in the last dry season are approximately 0.21 percentage points more likely to get an abortion (Panel B, columns 1-8, $p < 0.01$). The magnitude of that impact is economically large; compared with a mean of the likelihood of abortion of unaffected mothers of 0.0066, abortions reported by affected mothers increase by approximately 30 percent. The effect of droughts that are concurrent with the survey period (Panel A) or one year before the reference date (Panel C) are much smaller and statistically insignificant, whereas the effects of droughts in $t-1$ are robust to the distributed lagged specification in Table 1.B.1.

The point estimates on the effects of droughts $(t-1)$ are also invariably robust to employing a wide range of control covariates. First, this specific measure of droughts is not collinear with the district-specific time trends, as shown by consistent standard errors (column 4). Furthermore, district-specific developments of relevant factors influencing the accessibility of abortion also do not bias these estimates. Second, it does not change the size of coefficient nor does it reduce statistical significance to include not only a mother's and her spouse's characteristics but also the indicators for the parity and the gender composition of previous births (columns 5-8). Those covariates would potentially control for the diverse incentives for mothers to select into pregnancy and to give birth depending on her expected payoff from having an extra child or son; thus, the effect of droughts is robust to unobservable changes in the composition of the mothers who conceived.

These results suggest that economic shocks can also have substantial effects, even to the extent of impacting a woman's decision regarding her current pregnancy, which not only reveals another onerous coping strategy when aggregate shocks befall poor rural households but also supports the theoretical prediction of the positive relationship between income and demand for children.

### 1.4.4 Effect on Births

Since I find no effects on conception, I hypothesize that the effect on abortion necessarily leads to fewer births. However, these effects may be difficult to estimate precisely if the timing of abortions is spread over a year and the number of the missing births is relatively small. To detect the effects on the likelihood of giving birth within a year, I use the district-quarter-level number of births as the outcome variable by counting infants who were born to the mothers in the sample. I restrict the sample to infants aged less than one year because the estimation using the older cohort might suffer from omitted variables bias due to postnatal discrimination against a female child.[24] I estimate the effect of droughts using the following equation:

$$Y_{dqy} = \alpha + \sum_{j}^{J} \beta_j Q_d^j + \lambda_{pq} + \lambda_{qy} + \lambda_d + \theta_d * t + \varepsilon_{dqy} \tag{1.3}$$

where $Y_{dqy}$ is the number of infants in each district-quarter cell in a survey year $y$ and $Q$ is a dummy variable that becomes one if a district is $j$ quarters ($-3 \leq J \leq 8$) before or after a drought.[25] Since I pool the nine waves of cross-sectional surveys, this is equivalent to estimating the effects of the current and up to 2-year lagged events of droughts on the number of births. In addition to district-level linear time trends, a host of fixed effects enter to control for the seasonality of births, which are shaped by the province-specific crop calendar ($\lambda_{pq}$) and national-level trends in fertility ($\lambda_{qy}$). Therefore, the coefficients, $\beta_j$'s, ascertain how the effects of an increase in abortion on births are distributed across quarters conditional on seasonality and trends in childbearing.

Figure 1.6 plots the point estimates along with their corresponding 95 percent confidence intervals. First, the coefficients before a drought (the red vertical line) tests for parallel pre-trends by comparing the number of births born in districts exposed to unusually low levels of dry season rainfall with those born in unaffected districts. While the effects on births in pre-drought quarters are not significant, the 5th quarter after the droughts (April to June) shows the first statistically significant effect on the number of births. Given the estimate of -0.364 ($p < 0.01$), more abortions in year $t+1$ result in an approximately 7 percent reduction in the number of births with a mean of 5.087 district-quarter-level number of births (Figure 1.8). Table 1.6 reports the consistent results using Equation (1.1) to estimate the effect if the outcome is an indicator for a woman's childbirth in the survey year.

---

[24] It is also important to minimize the duration of potential postnatal discrimination because this infant sample will also be analyzed to calculate the sex ratio at birth in Section 1.4.

[25] The births in January to March in a survey year $y$ would have $j = 0$. If there was a drought in a survey year $y$, then $Q_d^0 = 1$ would be matched to the number of births in January to March in the survey year.

Importantly, Figure 1.6 shows that the effects of droughts on births are not evenly distributed across quarters; instead, they are concentrated in a particular quarter, suggesting that an affected mother is very likely to terminate her pregnancy in that specific quarter of the year. Given that conception occurs 9 months before the date of birth and most abortions are performed between 8-12 weeks of pregnancy, affected mothers get abortions particularly in the 3rd quarter after droughts (October-December). Note that this quarter corresponds to the pre-harvest season of the next rice crop when a rural household's consumption smoothing is particularly difficult, as shown in Figure 1.5. Taken together, the income channel would indeed be the pivotal pathway for an increase in abortion after droughts. However, I will further investigate additional channels to explain more abortions in that quarter by examining labor participation and biological mechanisms in Section 1.5.

### 1.4.5 Effect on the Sex Ratio at Birth

Having the results on birth, I now turn to investigate the sex ratio at birth to understand whether abortion becomes more or less sex-selective. I predict that affected mothers are less likely to terminate their pregnancy if the fetus is male, given the availability of low-cost sex-selective abortion. I repeat the estimation using Equation (1.3), but to have reliable estimates on the sex ratio at birth, I aggregate the numbers of male and female infants in each district up to the province-quarter level and combine them with the province-level droughts:

$$Y_{pqy} = \alpha + \sum_{n=-3}^{8} \beta_n Q_p^n + \lambda_{rqy} + \lambda_p + \varepsilon_{pqy} \tag{1.4}$$

where $Y_{pqy}$ is the log of sex ratio at birth.[26] I include region-quarter-year fixed effects ($\lambda_{rqy}$) to flexibly control for trends in the sex ratio at birth.

Figure 1.7 (a) presents the results. I find the effect is statistically significant only in the 5th quarter after droughts when significantly fewer births are found (Figure 1.8 (c)). The

[26] The sex ratio at birth here is the number of male infants born to 100 female infants.

sex ratio at birth increases by 12.9 percent ($p < 0.05$), implying that abortions conducted in the 3rd quarter after droughts not only increase, and thus reduce the size of the birth cohort, but also become more sex-selective against female fetuses. Table 1.7 reports the consistent results using Equation (1.1) to estimate the effect when the outcome is a dummy indicator for a newborn being a boy conditional on being born in each quarter of the survey year.

To explain the more male-skewed sex ratio within the reduced birth cohort in the 5th quarter, a back-of-the-envelope calculation indicates that 6 more female fetuses were aborted for every 1 male fetus that was aborted.[27] Furthermore, I find that an approximately 3 percent increase in the sex imbalance observed in PCS 2004-2013 is attributable to the sex-selective abortions in the period before the next harvest after droughts.[28] While the effect of drought-induced transitory income shocks might be more relevant to the timing of fertility than lifetime fertility, son preference can regulate the way mothers postpone births, and thus has far-reaching and long-term implications on demographics by worsening the sex imbalances at birth.

Here, I provide two more pieces of corroborating evidence that drought-induced abortions become more sex-selective. First, since ultrasound scans for fetal sex determination became more prevalent in rural areas in the late 2000s, I split the PCS sample roughly in half, e.g., before and after the PCS 2007, and I run the same regression of Equation (1.4) for each subsample. Since sex-selective abortion can be performed at a low cost after the availability

[27] Given that the mean number of quarterly births in the unaffected districts is 5.014, the numbers of boys and girls should be 2.687 and 2.327, respectively, to explain the mean sex ratio of 115.5 at birth. In the 5th quarter after droughts, the number of births in the affected districts decreases to $4.65 (= 5.014 - 0.364)$, and the corresponding sex ratio at birth increases to $130.4 (= 115.5 \times 1.129)$, implying the numbers of boys and girls in that quarter should be 2.632 and 2.018, respectively. Thus, $0.309 (= 2.327 - 2.018)$ more girls were aborted, while $0.055 (= 2.687 - 2.632)$ boys were aborted in the affected districts in the 5th quarter after droughts.

[28] The mean sex ratio at birth found in all 1,856 quarter-province cells in the sample is 115.775. I find the mean sex ratio at birth in the 1,806 unaffected quarter-province cells to be 115.483. Thus, the 50 affected provinces in the 5th quarter after droughts increased the mean sex ratio at birth by 0.293 $(= 115.775 - 115.483)$. Given the normal rate at birth is 105 boys born to 100 girls, this suggests the drought-induced sex selection explains approximately 3 percent of the sex imbalances found in rural Vietnam in 2004-2013. $(0.293/(115.775 - 105) = 0.027)$. All of the mean sex ratios at birth are weighted by the number of mothers in each province.

of ultrasound scans has spread across rural areas, the effects on sex selection would be more conspicuous in the later periods. In Figure 1.7, Panels (b) and (c) show that the results are consistent with the prediction: the effects on the sex ratio of the full sample in Figure 1.7 (a) are mostly driven by the effects found in the later periods.

Second, given that the most plausible timing of the sex-selective abortion is 12-16 pregnancy weeks,[29] if mothers want to sex-select to have a boy, the weeks of prenatal sex determination by ultrasound should be more likely to fall within 12-16 weeks of pregnancy, which should be particularly true if a mother exposed to droughts gives birth to a boy in the 5th quarter because having a boy can compensate the forgone consumption in the pre-harvest season. In Figure 1.9a (a), I find that affected mothers who gave birth to sons in the 5th quarter were more likely to know the fetal sex between 12-16 weeks of her pregnancy, which is earlier than the period when mothers who gave birth to daughters knew the fetal sex. However, in Figure 1.9a (b), I do not find any difference in the weeks of sex determination for unaffected mothers. Furthermore, Figure 1.A.7 shows that the effect on the weeks of prenatal sex determination does not depend on the seasonality of births because the distribution is not significantly different across the quarters of births; it is only the 5th quarter after droughts when the distributions are significantly different, i.e., a mother who gave birth to a son is more likely to know the fetal sex in the 12-16 pregnancy weeks (Panel (c) and (k)).

## 1.5 Additional Results and Robustness Checks

### 1.5.1 Opportunity Cost of a Mother's Labor

The demand for a mother's labor also exerts considerable influence on the decision regarding when to time fertility. On the one hand, a woman wants to postpone childbearing when the

[29] Ultrasound can credibly detect the fetal sex as early as 12 weeks, and abortion in Vietnam is rarely performed after 16 weeks of pregnancy (Committee for Population, Family and Children [Vietnam] and ORC Macro, 2003) Table 1.1 shows that a rural mother knew the sex of child at an average of 20 weeks of pregnancy conditional on the childbirth in the survey year.

opportunity cost of her parental time is high (Schultz, 1985; Heckman and Walker, 1990). On the other hand, low demand for female labor can have ambiguous effects because it reduces the substitution effect, thereby increasing fertility; however, it can also further depress income (Dehejia and Lleras-Muney, 2004).

Figure 1.10 presents the estimated coefficients on the effects on labor participation of married women and men using Equation (1.2) on the extensive margin (e.g., whether a respondent worked last month, Panel A) and on the intensive margin (e.g., working days conditional on having worked last month, Panel B). Overall, the results suggest that droughts are not associated with higher labor participation for women and men (Panel A) although the point estimate for female working days is statistically significant in the pre-harvest season ($p < 0.05$) but is not economically large. Furthermore, Figure 1.A.13 shows little evidence of recent migration of married women or men (within 6 months), suggesting labor migration does not bias the estimates of the effect on abortion.

### 1.5.2 Biological Channel

Droughts can have direct adverse effects on the health of women or fetuses. Although the PCS specifically asks whether mothers had an 'induced' abortion, mothers might terminate their pregnancy to prevent stillbirths or miscarriages because drought-induced malnutrition or incidence of certain infectious diseases makes fetuses too feeble to be carried to term.

In columns 8 and 9 of Table 1.B.3, I first find little evidence of detrimental effects on the consumed quantity of rice and pork, the two main food categories of rural households. This inelastic calorie consumption with respect to food price can be further supported by Gibson and Kim (2013) who find that a 10 percent increase in the price of rice is associated with a less than 2 percent decrease in calorie intake.

The nutritional channel can be further dismissed by examining its effects on the sex ratio at birth; poor nutritional intake at the preconception stage increases the probability of giving birth to a daughter (Cameron, 2004; Mathews et al., 2008). Table 1.7 reports no significant effects on the sex ratio at birth between droughts (column 4 of Panel A) and the 4th quarter after droughts (columns 1-4 of Panel B). In addition, in Figure 1.A.12, I find no evidence of the effects of droughts on recent illness by estimating Equation (1.2) on an indicator for any illness or injuries among household members in the past 4 weeks using the VHLSS 2004-2008.[30]

Finally, I examine the effect of droughts on infant mortality. This investigation is doubly important because it can test not only for fetal exposure to malnutrition or drought-related diseases, which increase the risk of infant mortality, but also for postnatal discrimination against female infants after aggregate shocks. I assume that the effect of negative income shocks would be fully reflected in the margin of prenatal discrimination, e.g., sex-selective abortion, and not postnatal discrimination, such as infanticide or neglect.[31] Table 1.8 reveals that droughts are not associated with an increase in infant mortality from the estimation of Equation (1.1) for infant deaths reported in the PCS and the province-level infant mortality rate from the GSO.

## 1.5.3 Additional Evidence on the Income Channel

To corroborate the causal link between income shocks and pregnancy terminations, I investigate heterogeneous effects on abortion using local conditions that variably determine income vulnerability to adverse rainfall shocks. Panel A of Table 1.B.2 shows the imprecisely estimated effects on abortion of urban women whose income is presumably invariant to rainfall fluctuations.

Furthermore, triple cropping, the availability of irrigation and wealth level are other prox-

[30] There are no monthly reports of illness in the VHLSS from 2010.

[31] Bharadwaj and Lakdawala (2013) argue that there can be other forms of prenatal discrimination, such as lower antenatal care, once the sex of a fetus is revealed. However, I assume here that the cost of abortion is not prohibitively higher than the cost of discrimination in prenatal investment, and if it is, part of that lower investment in female fetuses would result in higher female infant mortality, which is tested in this section.

ies that are correlated with district-level income variability to the fluctuations of the rainfall in the dry season. Table 1.B.6 reports the results from estimating Equation (1.1) after augmenting it with indicators for the three proxies. Column 1 shows that the effect on abortion is mostly driven by the rural mothers living in the double-cropping provinces; they cannot grow autumn rice, which allows one extra income flow to mothers in the triple-cropping provinces. Column 2 reveals that most effects are found in the districts with the lowest coverage of irrigation, suggesting that higher volatility in the yield of rice can translate into larger effects on abortion. In column 3, I find that the coefficients for the two indicators for wealthier districts are not statistically significant, but the negative signs imply that mothers in wealthier districts are less likely to get abortions after droughts in the dry season.

### 1.5.4 Effect on Lifetime Fertility

I hypothesize that the effect on births is more relevant to short-run effects, i.e., delaying births, than to the long-term effects on a woman's completed fertility.[32] Figure 1.11 provides the results from estimating Equation (1.3) after adding the interaction terms between indicators for quarters and droughts up to six lags. After the first slump in the 5th quarter after droughts, the rebound in the 13th quarter is statistically significant ($p < 0.01$), and the size of the coefficient is comparable to the coefficient found in the 5th quarter. This full rebound in Figure 1.11 confirms my hypothesis: rainfall-induced transitory income shocks primarily affect the timing of childbearing while holding lifetime fertility unchanged. In addition, the young age of rural mothers (on average 23 years old at her first childbirth, from Table 1.1) would allow for more time for the temporary decline in fertility to recuperate.

[32]In contrast, Currie and Schwandt (2014) find that the effect of unemployment on a young woman in the US can have long-run impacts on her lifetime fertility. However, its long-run effect is driven primarily by women who remain childless, which indicates unemployment has negative impacts on her marriage markets.

### 1.5.5 Robustness Checks

#### Alternative Empirical Specifications

The main result on abortion remains intact to more flexible specifications. To test for non-linearities, I replace the indicator for drought in Equation (1.1) with multiple binned indicators which become one if rainfall falls within every 10th percentile of historical distribution and zero otherwise while making the 5th indicator the omitted category.[33] Panel (f) of Figure 1.A.10 plots the coefficients on each dummy that is constructed using the dry season rainfall in $t-1$. I find that the lowest level of the dry season rainfall in $t-1$, i.e., the 1st decile, leads to the highest rate of abortion compared to the effect of rainfall in the 5th decile. The point estimates for the other decile indicators are close to zero.

In addition, I further show that the main results are robust to replacing the rainfall estimates from CHIRPS with the modified distributions that result from fitting the historical rainfall realizations to a district-specific gamma distribution, as suggested by Burke et al. (2015) and Corno et al. (2017). Figure 1.A.8 shows consistent results with Panel (f) of Figure 1.A.10 using the historical rainfall realizations in the main analysis.

Finally, to further examine the relationship between rainfall and abortion, it is important to estimate the marginal effects of rainfall levels on the likelihood of abortion. Using restricted cubic splines with three knots, I present the estimated marginal effects of an additional unit increase in the dry season rainfall percentile on the probability of abortion. Figure 1.A.9 shows that rainfall levels below the 20th percentile have consistent and significant effects on abortion, giving confidence in using my definition of drought occurring below the 20th percentile in the main analysis.

[33] $Y_{idt} = \alpha + \sum_{k \in K} \beta_k \widetilde{R}^k_{d,t} + \sum_s^S \sum_l^L \gamma_{sl} R_{s,t-l} + X'_{it}\delta + \tau_t + \mu_d + \theta_d * t + \varepsilon_{idt}$ where $\widetilde{R}^k_{d,t}$ are the dummies for every 10th percentile.

### Other Seasonal Rainfall

I examine whether the drought shock I construct using the dry season rainfall in $t-1$ is the most crucial shock to determine economic conditions and fertility outcomes in the main analyses. Table 1.B.7 presents that the drought defined using either the wet season rainfall in $t$ (Panel A) or the calendar year rainfall in $t$ (Panel B) do not have statistically significant effects on the rice yields and yearly household expenditure measures. Next, Figure 1.A.10 shows that the coefficients resulting from repeating the estimation in Figure 1.A.10 Panel (f) using all of the combinations of other season-year-rainfall distributions. I find the drought shocks defined by rainfall levels in the preceding dry season have the largest impacts on abortion (Figure 1.A.10 Panel (f)), whereas the effects of the other drought shocks are imprecisely estimated in general. Finally, I examine whether consecutive rainfall shocks would amplify or mitigate the effect of droughts in the dry season in $t-1$ by augmenting Equation (1.1) with the indicators for positive (8th and 9th deciles) or negative shocks (1st and 2nd deciles) of the wet season rainfall in $t$ and $t-1$. Table 1.B.8 shows that the effect of droughts in the dry season in $t-1$ remains robust and the interaction terms are not statistically significant.

### Measurement Error

I calculate the average effect on yearly births by summing up the 4 point estimates from the 3rd to the 6th quarters after droughts (or equivalently from October to September) and compare it to the point estimate for yearly abortion. This exercise checks the measurement error in the reporting of abortion because mothers might be reluctant to provide their true experience of abortion if those abortions were particularly sex-selective. The average effect of droughts on the yearly birthrate is -0.0023, which is almost identical to the point estimate on yearly abortion in Table 5, implying that the under-reporting of abortion should be less of a concern.

### Spatial and Serial Correlation

I investigate whether the estimates of standard errors are biased due to spatial correlation of drought indicators. I attempt to correct for the spatial correlation by clustering the standard errors at the province level in lieu of the district level. I also show the $p$-values from clustering on both district and year to further account for serial correlation in rainfall realizations over time within districts. In Table 1.B.9, I still find that the main estimates are statistically significant at the 95 percent significance level although alternative clustering slightly reduces the significance.

## 1.6 Conclusion

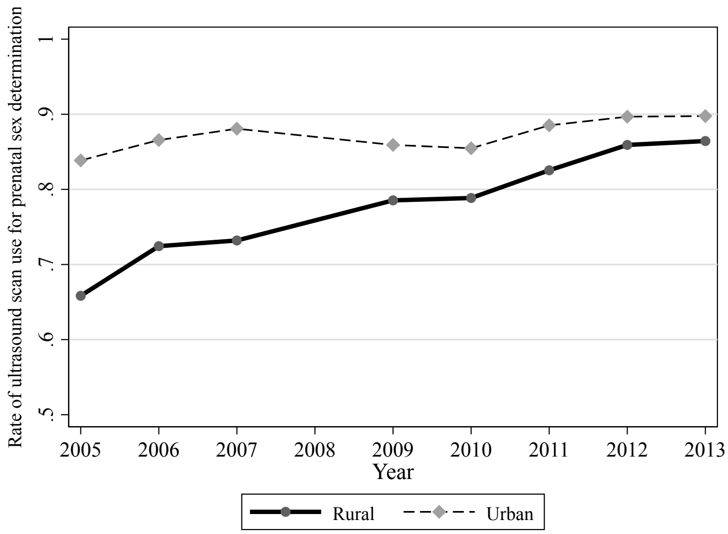
In this paper, I examine how abortion responds to drought-induced transitory income shocks and how son preference regulates such adjustments. I provide reduced-form estimates for the effects of droughts, which serve as a valid proxy for short-run slumps in agricultural income, on a series of fertility outcomes of credit-constrained rural mothers. I find that affected mothers postpone fertility using abortion when they are unable to smooth consumption in the pre-harvest season. Importantly, droughts are associated with disproportionately more abortions of female fetuses, which exacerbates the problem of the skewed sex ratio. While a full rebound in births in approximately 2 years appears more consistent with the effect on the timing of fertility, the effect on the sex ratio at birth emphasizes that even transitory income shocks can have long-run demographic consequences, which can shed light on how the gender gap can persist during the economic development process (Jayachandran, 2015).

Abortion allows parents to maximize lifetime utility by arranging the timing of childbearing at a low cost; thus, parents avoid giving birth to an unwanted child, who otherwise faces lower pre- and postnatal investment (Pop-Eleches, 2006). However, at the same time, if paired with son preference, abortion may have demographic consequences by exacerbating the "missing girls" phenomenon, which can have a detrimental influence on female human capital accumulation (Jayachandran, 2015), and can relate to social problems, such as an increase in crime by unmarried young men in the long run (Edlund et al., 2013).

My findings provide timely evidence for policy designs in developing countries. Although I focus on rural Vietnam, sex-selective abortions are increasing in the developing world (Bongaarts and Guilmoto, 2015). Furthermore, climate change would make extreme weather events more frequent, which essentially translate into transitory economic shocks to which poor economies relying on agricultural yields are especially vulnerable (IPCC, 2014). The effect on prenatal sex discrimination should be incorporated in assessing the damage functions that estimate the potential economic implications of climate change. Expanding the social safety nets or effective credit markets that help the poor mitigate aggregate shocks might minimize the resulting fertility response and therefore, the bias in sex ratios at birth.

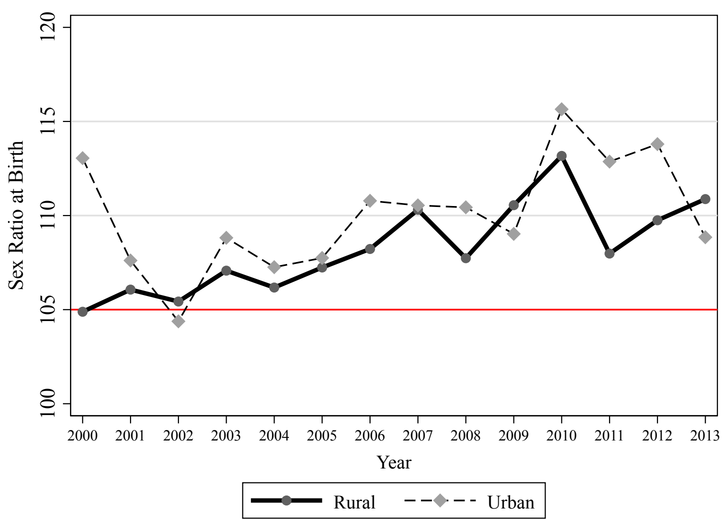
## 1.7 Figures

Figure 1.1: Rate of Ultrasound Scan Use for Prenatal Sex Determination



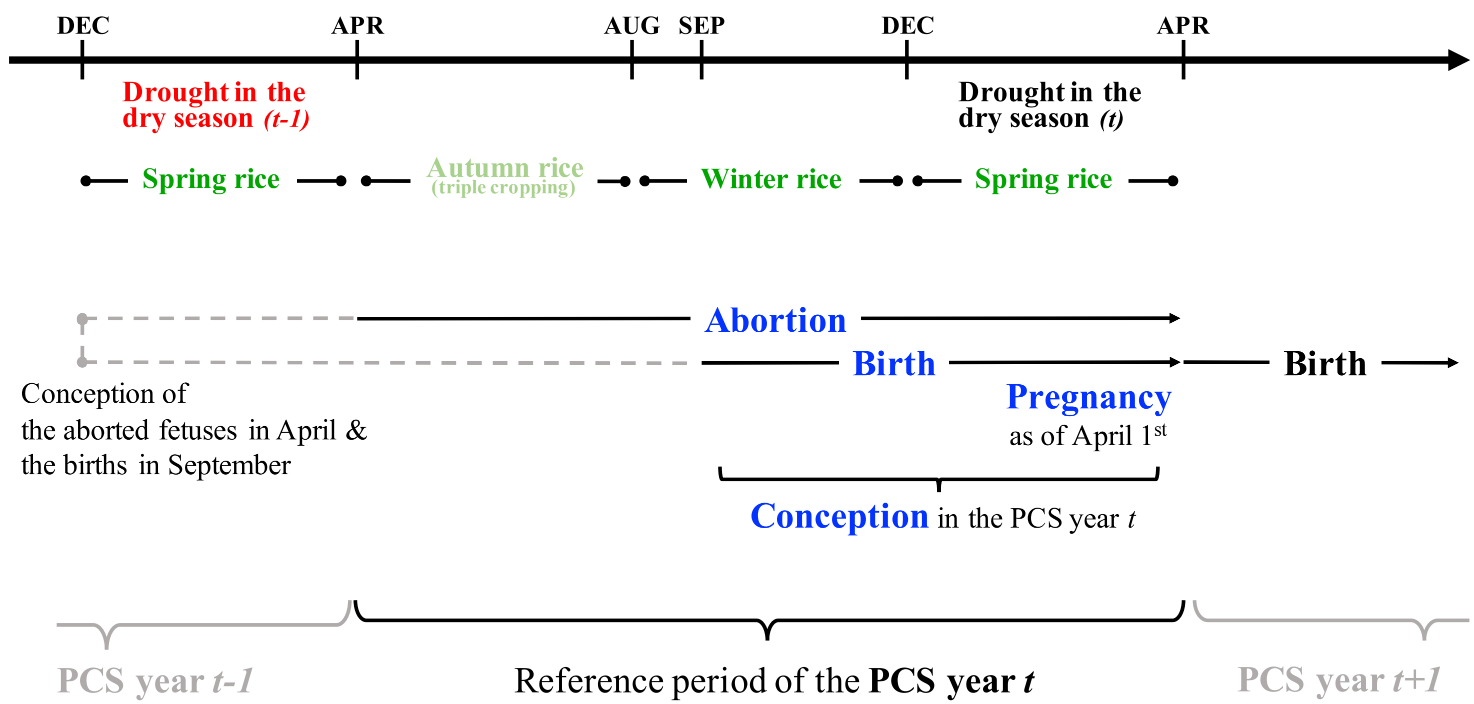

*Data*: PCS 2006, 2007, 2010-2013
*Notes*: This figure plots the rate of mothers who used ultrasound scans for prenatal sex determination given their childbirth in the survey year. The information about the ultrasound scan use is only available in the PCS 2006, 2007, and 2010-2013.

Figure 1.2: Sex Ratio at Birth in Vietnam



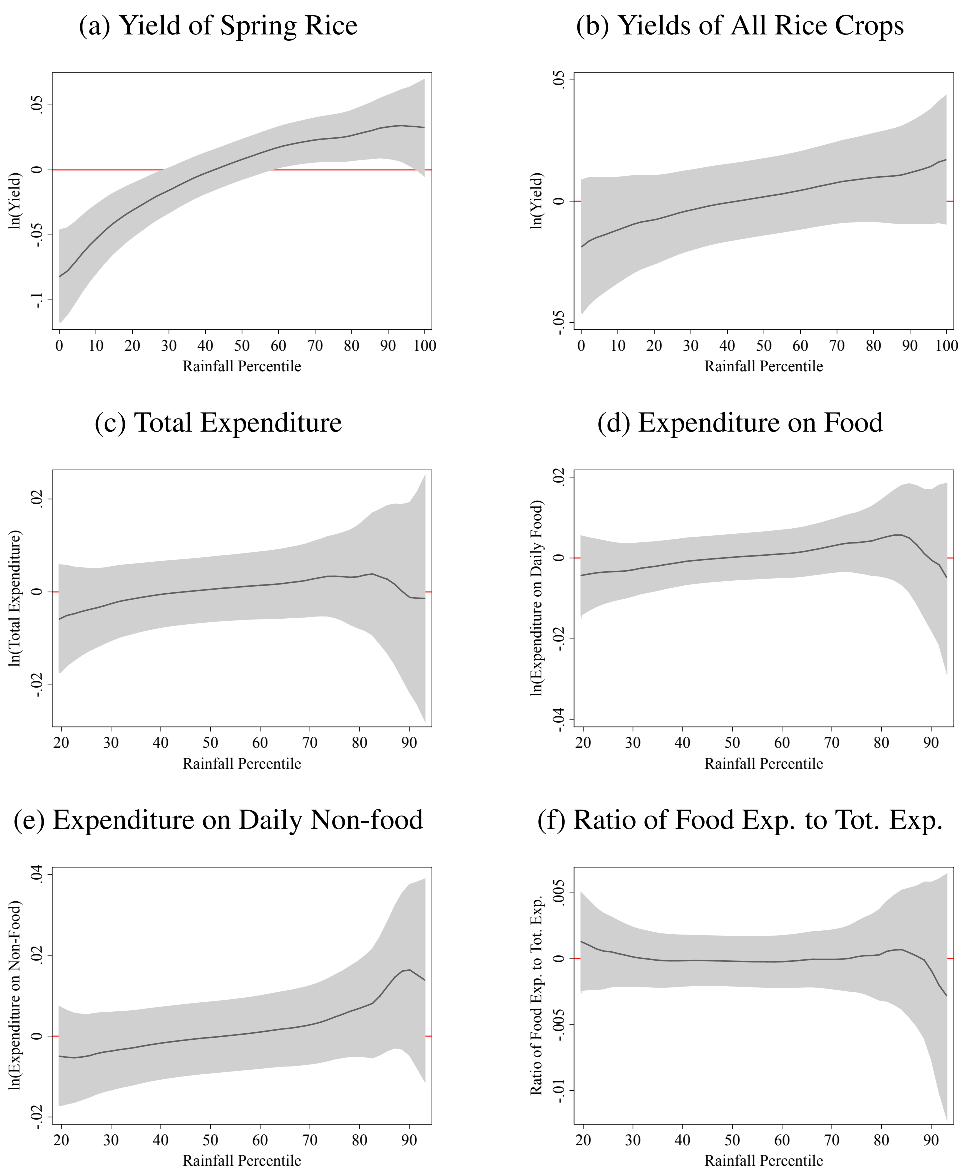

*Data*: PCS 2000-2008, 2010-2013, Census 2009
*Notes*: The figure plots the sex ratio at birth, i.e., the number of infant boys ($\leq$ 1 year old) born to one hundred infant girls in each survey year in rural and urban areas, respectively. The red horizontal line denotes the biological normal sex ratio at birth, approximately 105 male newborns to 100 female newborns.

Figure 1.3: Rice Crop Calendar and Time Frame of Drought Shocks and Fertility Variables



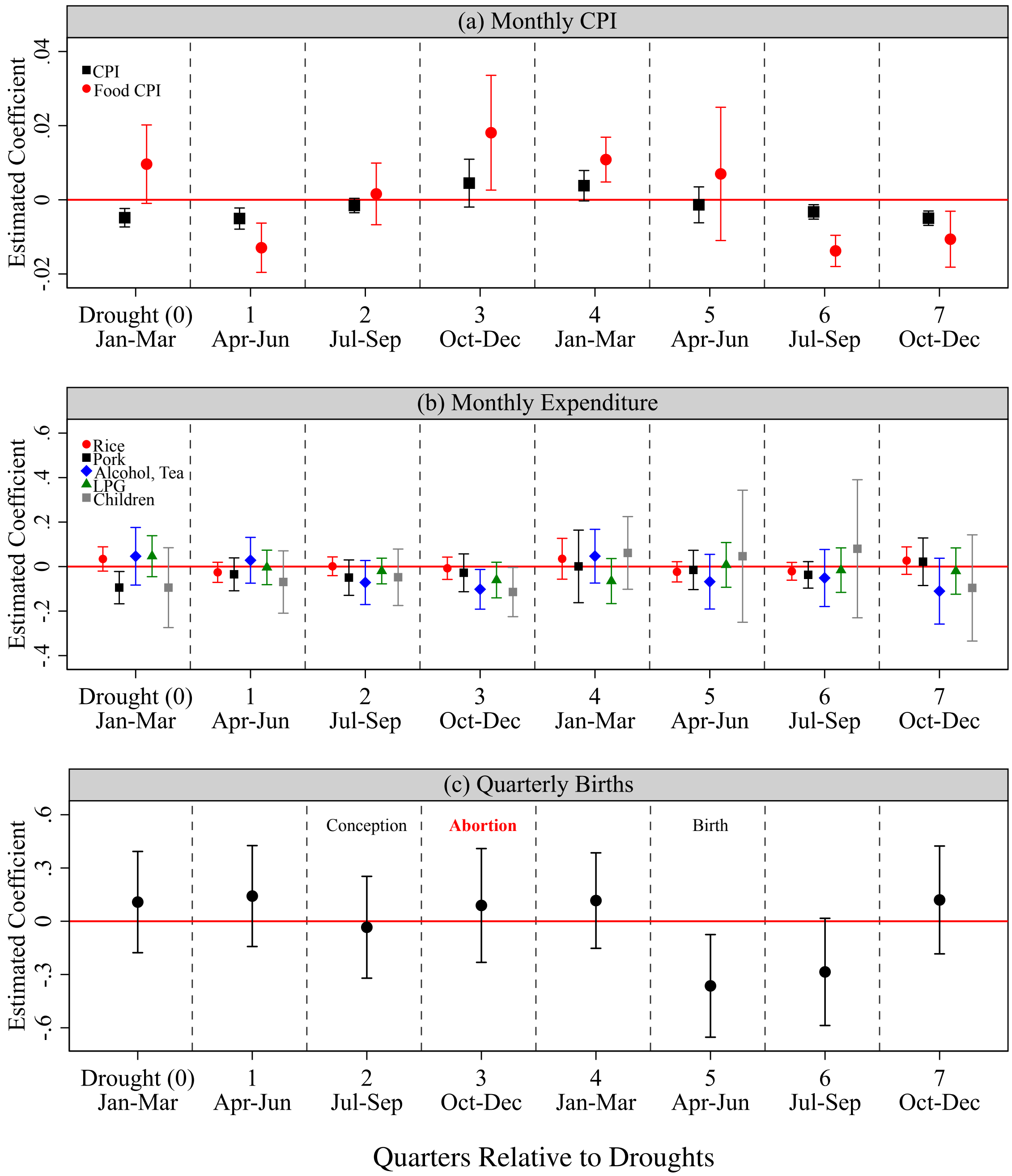

*Notes*: This plot describes the rice crop calendar across Vietnam from the FAO, and the timeframe of the drought shocks in the dry season (Dec-Apr) and fertility outcomes from the PCS. The reference period of the PCS is from April to March. Accordingly, the drought occurring right before the reference period of the PCS, $Drought_{t-1}$ written in red, would be the most relevant shock to the fertility outcomes created using the PCS in $t$. In the PCS, abortion is reported by year, and the current pregnancy is asked as of April 1st, which is the enumeration date of the PCS. The variable 'conception' denotes all the conceptions for fetuses to be potentially subject to abortion surveyed in the PCS $t$. Thus, the indicator for conception in the PCS year $t$ is coded one if a mother reports 1) abortion; 2) childbirth; or 3) pregnancy as of April 1st. In particular, a childbirth is considered as conception if it occurs from September because if aborted fetuses would otherwise have been carried to term, the earliest possible childbirth would happen in September given that abortion is rarely performed after 16 weeks of pregnancy.

Figure 1.4: Rainfall Percentiles and Yearly Rice Yields, Expenditure




*Data*: Province-level rice yields in 1995-2014 from the GSO; Yearly household expenditure from the VHLSS 2004, 2006, 2008
*Notes*: Figures provide the point estimates (line) and the corresponding 95 percent confidence intervals (shaded area) from local linear regressions of the log of yield (Quintal/Ha) and the log of rural households' yearly expenditure ('000 VND) on the percentiles of the dry season rainfall in a given year, relative to the long-run rainfall distribution in a province. The regressions for crop yields include the logs of other season-year rainfalls, province-level linear time trends, year and province fixed effects. The regressions for expenditure include household-level characteristics, survey quarter fixed effects, year and province fixed effects. All rice crops refer to Spring, Autumn and Winter rice crops. Daily non-food items include petroleum, cooking fuels, detergent, etc.

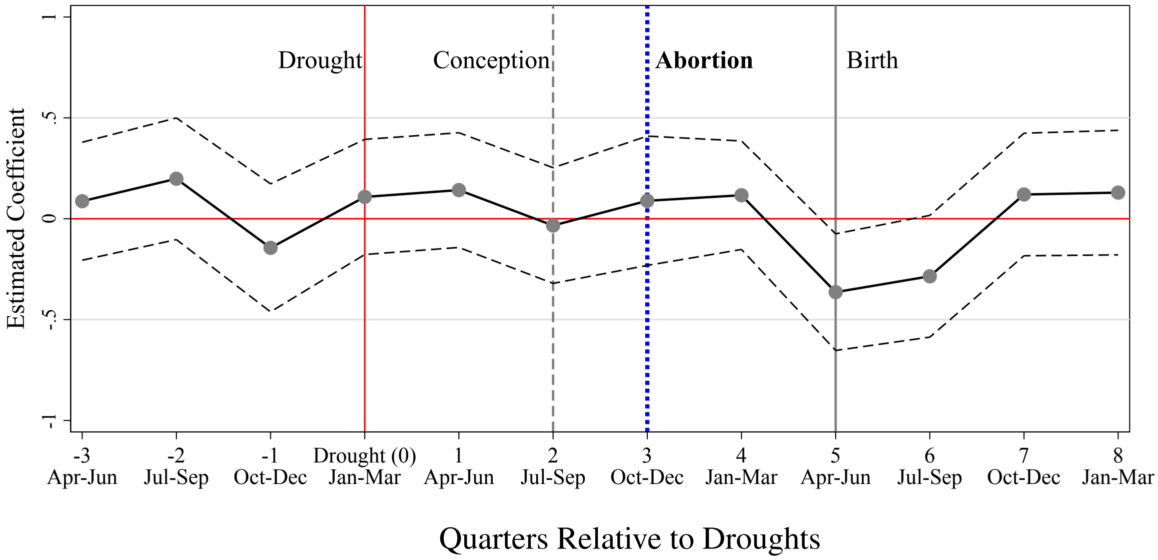
Figure 1.5: Effects of Droughts on Monthly CPI, Expenditure and Birth




*Data*: Province-level monthly CPI from the GSO; Monthly expenditure from the VHLSS 2010, 2012, 2014; Births from the PCS 2004-2008, 2010-2013

*Notes*: Figures plot the coefficients on the interaction terms between quarters and droughts in $t$ and $t-1$ from the regressions estimating the effect on the province-level monthly CPIs (Panel (a)), the log of monthly expenditures (Panel (b)) and the district-quarter level number of births (Panel (c)). Colored bars represent the 95% confidence intervals of the estimated coefficients. The 'Conception' and 'Abortion' in Panel (c) denote the timing of conception and abortion of the birth cohort in the 5th quarter after droughts.

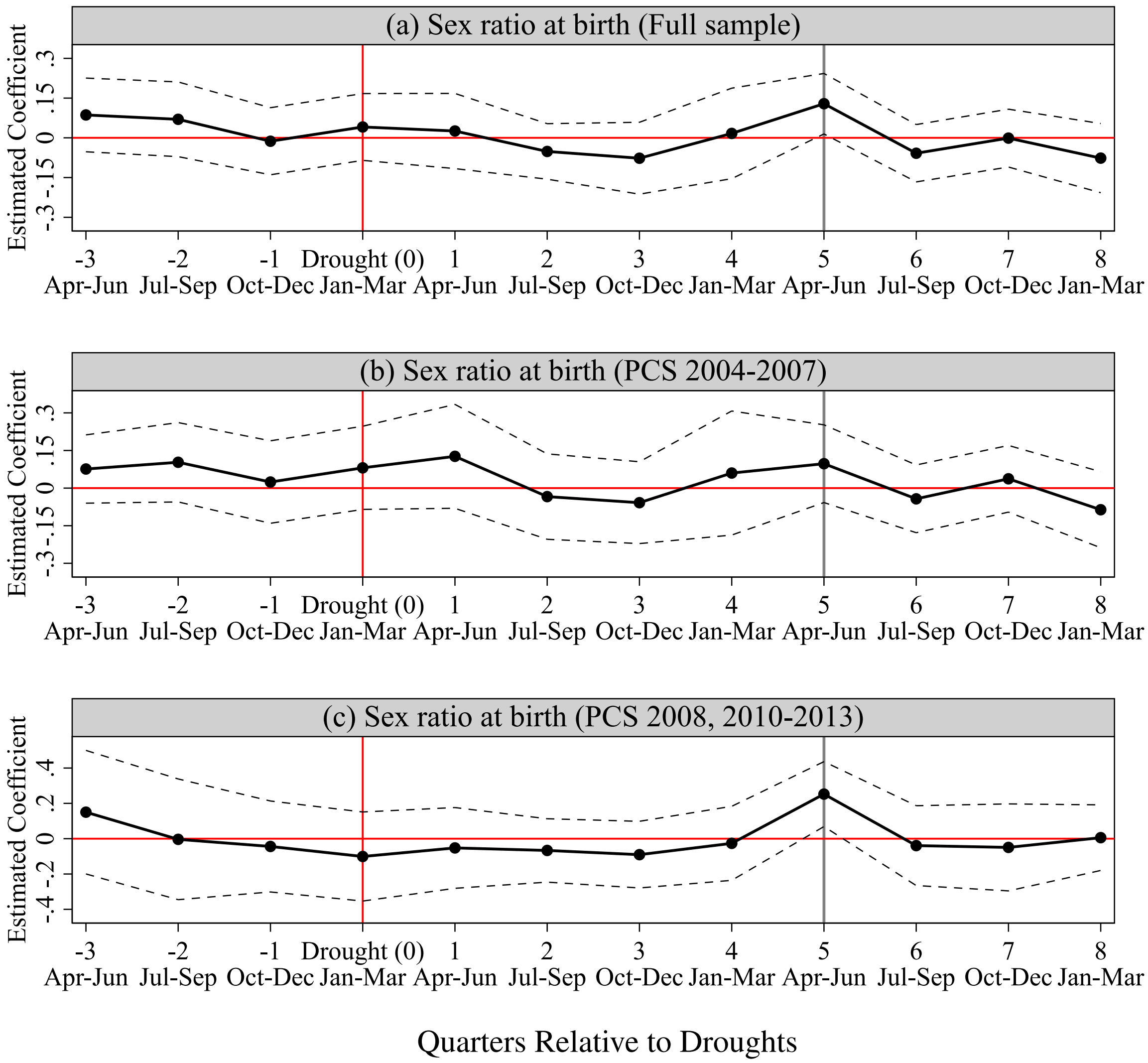
Figure 1.6: Effect of Droughts on Births




*Data*: PCS 2004-2008, 2010-2013

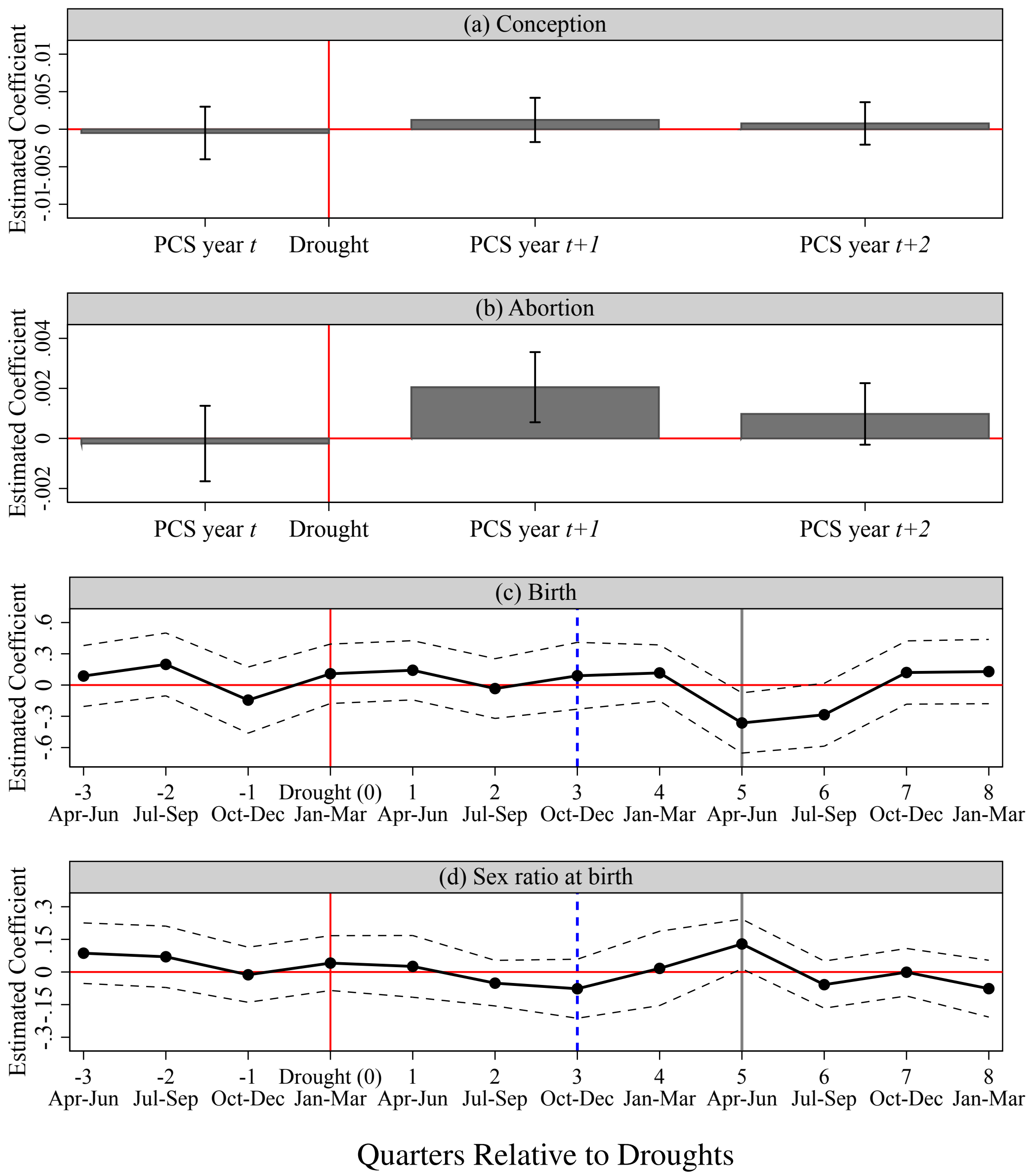
*Notes*: This figure plots the coefficients on the indicators for $n$ quarters away from the drought occurring at $n = 0$ in the regression estimating the effect on the district-quarter level number of births. The dashed black lines refer to the 95 percent confidence intervals. The gray solid line denotes the quarter when the effect on birth is statistically significant at the 5% level. Two vertical lines are additionally plotted for the birth cohort in the 5th quarter: the gray dashed line denotes the timing of conception, i.e., nine months before the birth in the 5th quarter, and the blue dotted line denotes the timing of abortion, i.e., about one quarter after the conception.

Figure 1.7: Effects of Droughts on the Sex Ratio at Birth




*Data*: PCS 2004-2008, 2010-2013
*Notes*: Panel (a) plots the coefficients on the indicators for $n$ quarters away from the drought occurring at $n = 0$ in the regression estimating the effect on the log of province-quarter level sex ratio at birth using the full PCS sample. Panel (b) and (c) plot the coefficients from repeating the same regression of Panel (a) using the first four and the latter five rounds of the PCS, respectively. The sex ratio at birth is defined by the number of infant boys ($\leq$ 1 year old) born to one hundred infant girls in each survey year. The dashed black lines refer to the 95 percent confidence intervals. The gray vertical line denotes the quarter when droughts have significant effects on births as shown in Figure 1.6.

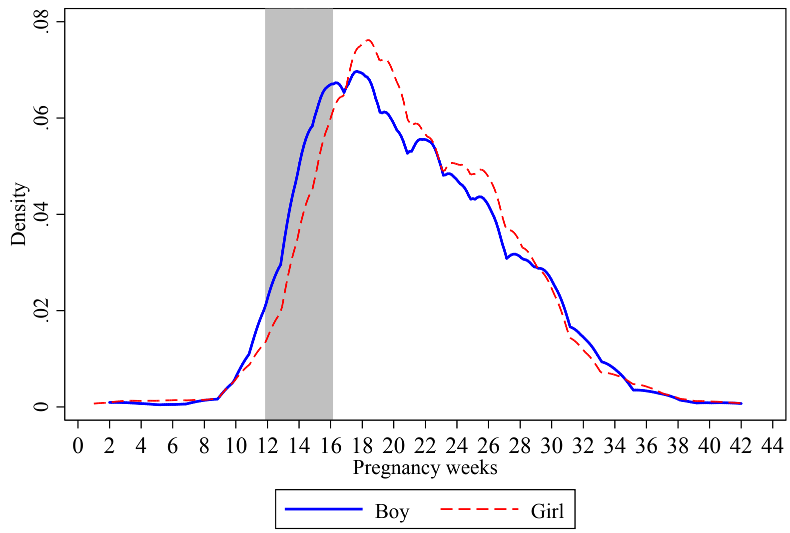
Figure 1.8: Effects of Droughts on Conception, Abortion, Birth and the Sex Ratio at Birth




*Data*: PCS 2004-2008, 2010-2013
*Notes*: Panel (a) and (b) display the effects of droughts in the dry season in $t$, $t-1$ and $t-2$ on the mother-level likelihood of yearly conception and abortion. Panel (c) and (d) plot the coefficients on the indicators for $n$ quarters away from the drought occurring at $n = 0$ in the regression estimating the effects on the district-quarter level number of births and the log of province-quarter level sex ratio at birth. The sex ratio at birth is defined by the number of infant boys ($\leq 1$ year old) born to one hundred infant girls in each survey year. The black bars (Panel (a) and (b)) and dashed lines (Panel (c) and (d)) refer to the 95 percent confidence intervals. The gray vertical lines in Panel (c) and (d) denote the quarter when the effect on births is significant at the 95% level. The blue dashed lines refer to the timing of abortion for the conception of the birth cohort in the 5th quarter (Apr-Jun) after droughts.

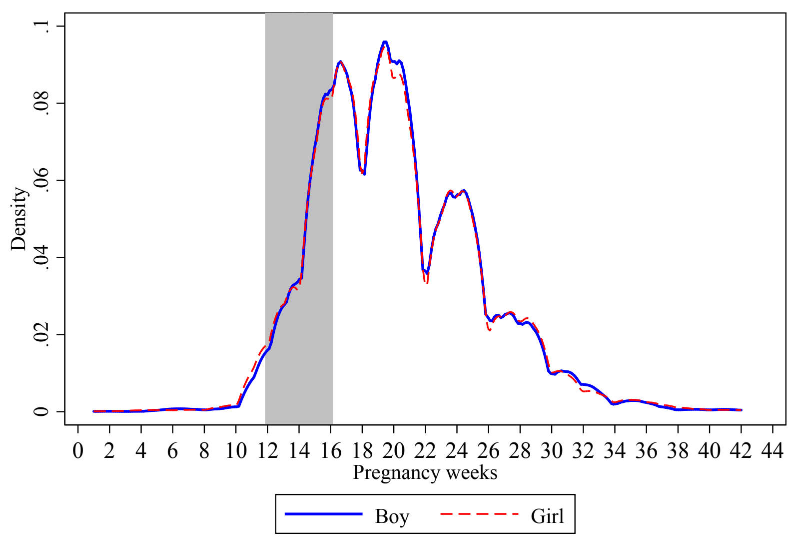
Figure 1.9: Distribution of the Pregnancy Weeks of Fetal Sex Determination

(a) By the sex of newborns born to affected mothers




*p*-value of two-sample Kolmogorov–Smirnov test: .112

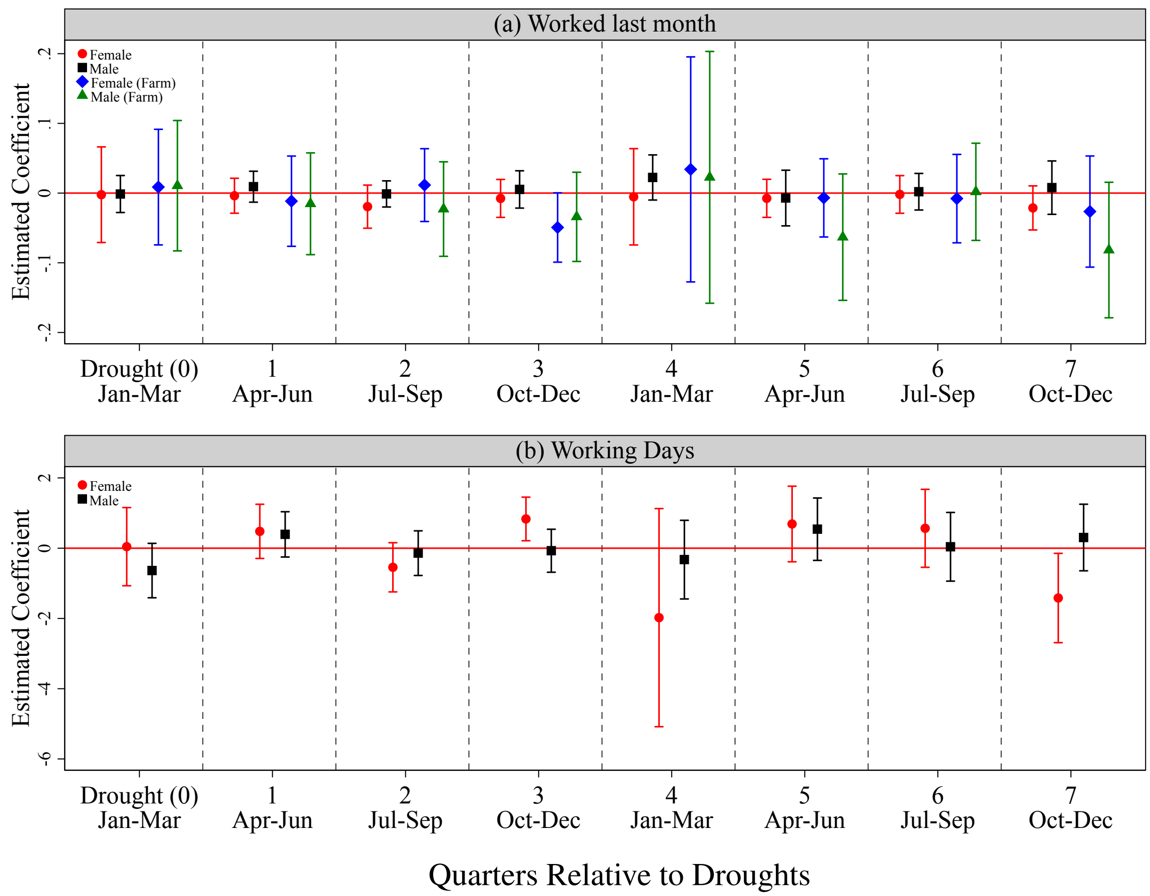
(b) By the sex of newborns born to unaffected mothers




*p*-value of two-sample Kolmogorov–Smirnov test: .999

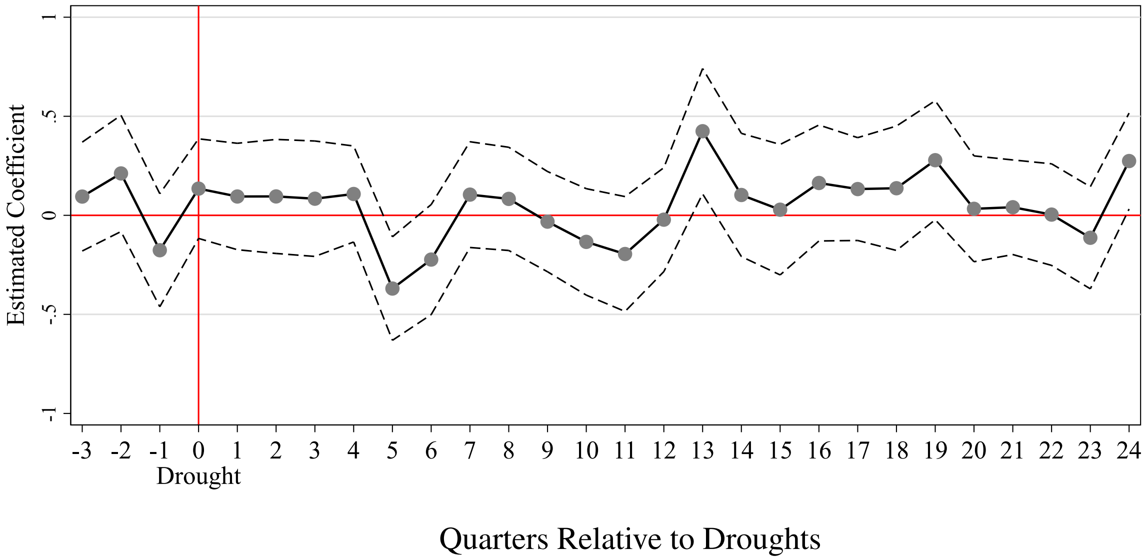
*Data*: PCS 2006, 2007, 2010-2013
*Notes*: Figures plot the kernel density estimation on the distributions of the weeks of fetal sex determination using ultrasound, conditional on the childbirth in April-June. The gray vertical bands denote the pregnancy weeks when sex-selective abortion can be performed; the 12th week is the earliest possible week when the fetal sex can be determined by ultrasound, and the 16th weeks is the latest possible week when abortion can be performed from the DHS 2002 (Committee for Population, Family and Children [Vietnam] and ORC Macro, 2003). Panel (a) plots the pregnancy weeks of fetal sex determination of affected mothers by the sex of a newborn, while Panel (b) plots those of unaffected mothers, conditional on giving births to a child in the 5th quarter after droughts (April-June) when the effect of droughts on birth is significant as shown in Figure 1.6.

Figure 1.10: Effects of Droughts on Labor Market Participation




*Data*: VHLSS 2010, 2012, 2014
*Notes*: Figures plot the coefficients on the interaction terms between quarters and droughts in $t$ and $t-1$ from the regressions estimating the effect on whether a married woman/man worked last month (Panel (a)) and the working days conditional on her/his labor market participation last month (Panel (b)). Colored bars represent the 95% confidence intervals of the estimated coefficients. Each regression includes quarter FEs, province FEs, region×year FEs, the logs of other season-year rainfalls and household-level controls such as the sex, age, ethnicity (Kinh or not) and years of schooling of the household head, the household size and the dummy for multigenerational household. Robust standard errors are clustered for the province level.

Figure 1.11: Effect of Droughts on Births in the Medium Term




*Data*: PCS 2004-2008, 2010-2013

*Notes*: This figure plots the coefficients on the interaction terms between quarters and droughts from $t$ to $t-6$ from the regression estimating the effect on the number of quarter-district-level births. Each point estimate refers to the effect on births in $n$ quarters away from the drought occurring at $n = 0$. The dashed lines refer to the 95 percent confidence intervals. The regression includes the same controls in Equation (1.3), district FEs, province×quarter FEs, quarter×year FEs, and district-level linear time trends. Robust standard errors are clustered for the district level.

## 1.8 Tables

Table 1.1: Descriptive Statistics

| | Rural | | Urban | | Diff. |
|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | |
| | (1) | (2) | (3) | (4) | (5) |
| ***Panel A. Full Sample*** | | | | | |
| Age (year) | 31.9 | (7.4) | 33.8 | (7.4) | -1.9 |
| Educ.: Primary or below | 0.506 | (0.500) | 0.285 | (0.452) | 0.221 |
| Educ.: Lower secondary | 0.355 | (0.479) | 0.294 | (0.456) | 0.061 |
| Educ.: Higher secondary or above | 0.139 | (0.346) | 0.420 | (0.494) | -0.281 |
| Age at the first birth | 22.590 | (3.659) | 24.051 | (3.938) | -1.461 |
| Number of children ever born | 1.849 | (0.837) | 1.703 | (0.744) | 0.146 |
| Number of children ever died | 0.029 | (0.180) | 0.016 | (0.134) | 0.013 |
| Have at least one son | 0.727 | (0.446) | 0.677 | (0.467) | 0.05 |
| Gave birth last year | 0.106 | (0.308) | 0.093 | (0.291) | 0.013 |
| Currently pregnant | 0.040 | (0.195) | 0.038 | (0.192) | 0.002 |
| Had abortion last year | 0.007 | (0.085) | 0.008 | (0.087) | -0.001 |
| Currently using any contraceptives | 0.769 | (0.421) | 0.758 | (0.429) | 0.011 |
| Currently using modern contraceptives | 0.670 | (0.470) | 0.628 | (0.483) | 0.042 |
| ***Panel B. Conditional on childbirth last year*** | | | | | |
| Checked pregnancy at clinics | 0.946 | (0.226) | 0.981 | (0.135) | -0.035 |
| Number of antenatal check-ups | 3.642 | (1.789) | 4.591 | (2.077) | -0.949 |
| Knew the child's sex before birth | 0.790 | (0.407) | 0.888 | (0.315) | -0.098 |
| Gestational weeks when a mother knew the fetal sex | 20.170 | (5.095) | 18.962 | (4.611) | 1.208 |
| Knew the fetal sex by ultrasound | 0.989 | (0.105) | 0.991 | (0.093) | -0.002 |
| Facility delivery | 0.934 | (0.249) | 0.985 | (0.121) | -0.051 |
| ***Panel C. Spouse Characteristics*** | | | | | |
| Age (year) | 36.6 | (7.1) | 39.0 | (7.2) | -2.5 |
| Educ.: Primary or below | 0.460 | (0.498) | 0.250 | (0.433) | 0.21 |
| Educ.: Lower secondary | 0.379 | (0.485) | 0.306 | (0.461) | 0.073 |
| Educ.: Higher secondary or above | 0.161 | (0.367) | 0.444 | (0.497) | -0.283 |
| ***Panel D. Household Characteristics (VHLSS)*** | | | | | |
| Unskilled labor worker in agri., aqua., forestry | 0.471 | (0.499) | 0.116 | (0.321) | 0.355 |
| Grow any paddy | 0.586 | (0.493) | 0.114 | (0.318) | 0.472 |
| Average monthly expenditure per capita (in '000 VND) | 9731.4 | (9608.71) | 17859.3 | (20803.2) | -8127.9 |
| Observations (PCS) | 811,117 | | 840,869 | | |
| Observations (VHLSS) | 22,253 | | 8,207 | | |

*Data*: PCS 2004-2008, 2010-2013, VHLSS 2004,2006,2008

*Notes*: This table provides the summary statistics of mothers included in the main analyses from the PCS (Panels A and B), and household characteristics from the VHLSS (Panel D). Panel C displays the spousal characteristics reported from the PCS 2006. Diff. refers to the difference between the means between rural and urban mothers. 'Full sample' includes married women of reproductive age 15-49 residing in rural or urban districts as defined in the PCS. The sample from the 10 provinces (Figure 1.A.6) are not considered for the main analyses.

Table 1.2: Effects of Droughts on Yearly Rice Yields and Expenditure

| | **Dependent variables** | | | | | |
|---|---|---|---|---|---|---|
| | Spring Rice (1) | All Rice Crops (2) | Total Expenditure (3) | Expenditure on Food (4) | Expenditure on Non-food (5) | Ratio (Food/Total) (6) |
| Drought | -0.024*** | -0.013** | 0.005 | 0.007 | -0.085*** | 0.002 |
| | (0.008) | (0.006) | (0.018) | (0.021) | (0.028) | (0.006) |
| Observations | 1,045 | 1,055 | 18,128 | 18,128 | 18,128 | 18,128 |
| R-squared | 0.804 | 0.901 | 0.530 | 0.609 | 0.530 | 0.119 |
| Mean of Dep. Var. | 3.923 | 3.783 | 9.878 | 9.078 | 7.564 | 0.471 |
| **Controls** | | | | | | |
| Province and year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Rainfall in other season-year | Yes | Yes | Yes | Yes | Yes | Yes |
| Province-specific linear time trend | Yes | Yes | | | | |
| Household Characteristics | | | Yes | Yes | Yes | Yes |
| Survey Quarter FE | | | Yes | Yes | Yes | Yes |

*Data*: Agricultural statistics from the GSO; VHLSS 2004, 2006, 2008

*Notes*: Column (1) and (2) present results from a regression of the log of annual crop yields (Quintal/Ha) on rainfall shocks. The unit of observation for crop yield is a province-year in 1995-2014. Column (3)-(6) present results from a regression of the log of expenditure (in '000 VND) on rainfall shocks. The unit of observation is a household. The sample excludes 10 poorest provinces to be consistent with the analyses using the PCS. Household characteristics controls include the sex, age, ethnicity (Kinh or not) and years of schooling of the household head, the household size and the dummy for multigenerational household. Robust standard errors, which are reported in parentheses, are clustered for the province level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 1.3: Effects of Droughts on Conception and Contraceptive Use

| | Dependent variables | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Conception | | Use any contraceptives | | Use modern contraceptives | | No contraception to have a child | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| ***Panel A. Drought in the dry season (t)*** | | | | | | | | |
| Drought | -0.0005 | -0.0003 | -0.0046 | -0.0046 | -0.0030 | -0.0094 | 0.0021 | -0.0013 |
| | (0.0018) | (0.0024) | (0.0029) | (0.0045) | (0.0040) | (0.0063) | (0.0023) | (0.0028) |
| Observations | 810,144 | 441,789 | 808,809 | 441,222 | 808,809 | 441,222 | 808,809 | 441,222 |
| R-squared | 0.148 | 0.105 | 0.326 | 0.224 | 0.233 | 0.158 | 0.237 | 0.188 |
| Mean of Dep. Var. | 0.107 | 0.107 | 0.769 | 0.769 | 0.666 | 0.666 | 0.117 | 0.117 |
| ***Panel B. Drought in the dry season*** ($t-1$) | | | | | | | | |
| Drought | 0.0012 | 0.0012 | 0.0011 | 0.0030 | 0.0070** | 0.0037 | 0.0005 | 0.0001 |
| | (0.0015) | (0.0019) | (0.0024) | (0.0031) | (0.0033) | (0.0046) | (0.0019) | (0.0022) |
| Observations | 810,144 | 441,789 | 808,809 | 441,222 | 808,809 | 441,222 | 808,809 | 441,222 |
| R-squared | 0.148 | 0.105 | 0.326 | 0.224 | 0.233 | 0.158 | 0.237 | 0.188 |
| Mean of Dep. Var. | 0.107 | 0.107 | 0.768 | 0.768 | 0.665 | 0.665 | 0.117 | 0.117 |
| ***Panel C. Drought shocks in the dry season*** ($t-2$) | | | | | | | | |
| Drought | 0.0008 | 0.0022 | -0.0003 | -0.0031 | 0.0068** | 0.0073 | 0.0033 | 0.0017 |
| | (0.0015) | (0.0018) | (0.0026) | (0.0033) | (0.0035) | (0.0046) | (0.0020) | (0.0023) |
| Observations | 810,144 | 441,789 | 808,809 | 441,222 | 808,809 | 441,222 | 808,809 | 441,222 |
| R-squared | 0.148 | 0.105 | 0.326 | 0.224 | 0.233 | 0.158 | 0.237 | 0.188 |
| Mean of Dep. Var. | 0.107 | 0.107 | 0.770 | 0.770 | 0.668 | 0.668 | 0.115 | 0.115 |
| **Controls** | | | | | | | | |
| District and year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Rainfall in other season-year | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Mother characteristics | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District-specific linear time trend | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Birth parity FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Gender composition FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Fertility characteristics | | Yes | | Yes | | Yes | | Yes |
| Spouse characteristics | | Yes | | Yes | | Yes | | Yes |

*Data*: PCS 2004-2008, 2010-2013

*Notes*: This table provides the current and the lagged effects of droughts on the conception and various measures of contraceptive use. 'Conception' is defined by the abortion occurred in the survey year and the corresponding conception cohort to this abortion, i.e., the births between September and March in the PCS and the current pregnancy. Any contraceptives include traditional methods, such as periodic abstinence or withdrawal, and modern contraceptives describes IUDs, pills, injections, condoms, diaphragm, foam or sterilization. Robust standard errors are shown in parentheses clustered at the district level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 1.4: Test for Balance in Characteristics of Affected and Unaffected Mothers

| | Full Sample | | By Parity | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 1st | | 2nd | | 3rd | | 4th | |
| Dependent Vars. | Mean/(SD) (1) | Diff. (2) | Mean/(SD) (3) | Diff. (4) | Mean/(SD) (5) | Diff. (6) | Mean/(SD) (7) | Diff. (8) | Mean/(SD) (9) | Diff. (10) |
| ***Panel A. Mother's characteristics*** | | | | | | | | | | |
| Age (Year) | 26.181 | -0.104 | 22.689 | -0.335 | 26.722 | -0.611 | 31.716 | 0.268 | 34.347 | 0.428 |
| | (5.632) | | (3.803) | | (4.249) | | (4.914) | | (5.110) | |
| Being the household head | 0.039 | -0.008 | 0.018 | -0.009 | 0.056 | -0.009** | 0.052 | -0.008 | 0.055 | 0.003 |
| | (0.193) | | (0.133) | | (0.229) | | (0.222) | | (0.228) | |
| Educ. Attain.: Primary or none | 0.284 | -0.199 | 0.210 | -0.197 | 0.303 | -0.201 | 0.376 | -0.217 | 0.487 | -0.211 |
| | (0.451) | | (0.408) | | (0.460) | | (0.484) | | (0.500) | |
| Educ. Attain.: Lower secondary | 0.493 | 0.166 | 0.480 | 0.143 | 0.506 | 0.182 | 0.513 | 0.184 | 0.455 | 0.195 |
| | (0.500) | | (0.500) | | (0.500) | | (0.500) | | (0.498) | |
| Educ. Attain.: Higher secondary or above | 0.223 | 0.033 | 0.310 | 0.053 | 0.191 | 0.019 | 0.111 | 0.032 | 0.058 | 0.016 |
| | (0.416) | | (0.463) | | (0.393) | | (0.314) | | (0.233) | |
| Age at the first birth (Year) | 22.035 | -0.289 | - | - | 22.243 | -0.365 | 21.719 | -0.147 | 21.626 | 0.116 |
| | (3.190) | | - | | (3.320) | | (2.932) | | (2.955) | |
| Number of children ever born | 0.887 | 0.016 | - | - | - | - | - | - | - | - |
| | (0.902) | | - | | - | | - | | - | |
| Having the first child | 0.434 | -0.012 | - | - | - | - | - | - | - | - |
| | (0.496) | | - | | - | | - | | - | |
| Have at least one son | 0.576 | 0.006 | - | - | 0.530 | 0.016 | 0.649 | -0.013 | 0.657 | -0.081 |
| | (0.494) | | - | | (0.499) | | (0.477) | | (0.475) | |
| ***Panel B. Spouse's characteristics*** | | | | | | | | | | |
| Age (Year) | 31.843 | 0.196 | 27.589 | 0.020 | 30.537 | -0.456 | 35.013 | 0.420 | 37.186 | 0.656 |
| | (6.018) | | (5.761) | | (4.859) | | (5.342) | | (5.188) | |
| Educ. Attain.: Primary or none | 0.311 | -0.154 | 0.237 | -0.151 | 0.296 | -0.155 | 0.348 | -0.166 | 0.432 | -0.193 |
| | (0.463) | | (0.426) | | (0.457) | | (0.477) | | (0.496) | |
| Educ. Attain.: Lower secondary | 0.493 | 0.150 | 0.470 | 0.126 | 0.491 | 0.152 | 0.515 | 0.151 | 0.484 | 0.178 |
| | (0.500) | | (0.499) | | (0.500) | | (0.500) | | (0.500) | |
| Educ. Attain.: Higher secondary or above | 0.196 | 0.004 | 0.292 | 0.025 | 0.213 | 0.002** | 0.137 | 0.016 | 0.084 | 0.015 |
| | (0.397) | | (0.455) | | (0.410) | | (0.344) | | (0.278) | |
| Observations | 14,373 | 94,995 | 6,241 | 42,239 | 5,024 | 34,227 | 2,396 | 14,298 | 712 | 4,231 |

*Data*: PCS 2004-2008, 2010-2013

*Notes*: This table is to test for balance between affected and unaffected mothers who had abortion, gave birth or were pregnant at the time of the survey. The statistics of affected mothers are presented in columns (1), (3), (5), (7) and (9). The number of observation in every first column for the full sample and for each parity refers to the number of affected mothers, while the number of unaffected mothers is presented in every second column. After regressing each dependent variable on the indicator of drought, the statistical significance of the coefficient is marked on the difference in the means between the two samples of mothers. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 1.5: Effect of Droughts on Abortion

| | Dependent variable: Abortion=1 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| ***Panel A. Drought in the dry season*** ($t$) | | | | | | | | |
| Drought | 0.0002 | 0.0001 | 0.0001 | -0.0002 | -0.0002 | -0.0002 | -0.0003 | -0.0012 |
| | (0.0007) | (0.0007) | (0.0007) | (0.0008) | (0.0008) | (0.0008) | (0.0008) | (0.0011) |
| Observations | 811,092 | 811,092 | 810,144 | 810,144 | 810,144 | 810,144 | 802,589 | 441,789 |
| R-squared | 0.008 | 0.008 | 0.009 | 0.012 | 0.012 | 0.012 | 0.012 | 0.013 |
| Mean of Dep. Var. | 0.0066 | 0.0066 | 0.0066 | 0.0066 | 0.0066 | 0.0066 | 0.0066 | 0.0066 |
| ***Panel B. Drought in the dry season*** ($t-1$) | | | | | | | | |
| Drought | 0.0020*** | 0.0021*** | 0.0022*** | 0.0020*** | 0.0021*** | 0.0020*** | 0.0021*** | 0.0023** |
| | (0.0006) | (0.0007) | (0.0007) | (0.0007) | (0.0007) | (0.0007) | (0.0007) | (0.0010) |
| Observations | 811,092 | 811,092 | 810,144 | 810,144 | 810,144 | 810,144 | 802,589 | 441,789 |
| R-squared | 0.008 | 0.008 | 0.009 | 0.012 | 0.012 | 0.012 | 0.012 | 0.013 |
| Mean of Dep. Var. | 0.0066 | 0.0066 | 0.0066 | 0.0066 | 0.0066 | 0.0066 | 0.0066 | 0.0066 |
| ***Panel C. Drought in the dry season*** ($t-2$) | | | | | | | | |
| Drought | 0.0009 | 0.0009 | 0.0009 | 0.0010 | 0.0010 | 0.0010 | 0.0010 | 0.0008 |
| | (0.0006) | (0.0006) | (0.0006) | (0.0006) | (0.0006) | (0.0006) | (0.0006) | (0.0009) |
| Observations | 811,092 | 811,092 | 810,144 | 810,144 | 810,144 | 810,144 | 802,589 | 441,789 |
| R-squared | 0.008 | 0.008 | 0.009 | 0.012 | 0.012 | 0.012 | 0.012 | 0.013 |
| Mean of Dep. Var. | 0.0069 | 0.0069 | 0.0069 | 0.0069 | 0.0069 | 0.0069 | 0.0069 | 0.0069 |
| **Controls** | | | | | | | | |
| District and year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Rainfall in other season-year | | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Mother characteristics | | | Yes | Yes | Yes | Yes | Yes | Yes |
| District-specific linear time trend | | | | Yes | Yes | Yes | Yes | Yes |
| Birth parity FE | | | | | Yes | Yes | Yes | Yes |
| Gender composition FE | | | | | | Yes | Yes | Yes |
| Fertility characteristics | | | | | | | Yes | Yes |
| Spouse characteristics | | | | | | | | Yes |

*Data*: PCS 2004-2008, 2010-2013

*Notes*: The dependent variable is the indicator for the experience of abortion during the survey year. Fertility characteristics control consists of her age at the first birth and the birth spacing referred to as the months between the most recent childbirth and the starting month of the survey period. Spouse characteristics include her spouse age, age squared and his educational attainment. The mean of dependent variable is the mean abortion rate of mothers living in the districts that were not inflicted with droughts. Robust standard errors are shown in parentheses clustered at the district level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 1.6: Effects of Droughts on Childbirth

| | Dependent variable: Giving Birth=1 | | | |
|---|---|---|---|---|
| | Birth in Apr-Jun | Birth in Jul-Sep | Birth in Oct-Dec | Birth in Jan-Mar |
| | (1) | (2) | (3) | (4) |
| ***Panel A. Drought in the dry season (t)*** | | | | |
| Drought | 0.0003 | 0.0008 | -0.0014* | 0.0004 |
| | (0.0008) | (0.0008) | (0.0008) | (0.0007) |
| Observations | 810,144 | 810,144 | 810,144 | 810,144 |
| R-squared | 0.035 | 0.040 | 0.042 | 0.035 |
| Mean of Dep. Var. | 0.0258 | 0.0281 | 0.0202 | 0.0323 |
| ***Panel B. Drought in the dry season ($t-1$)*** | | | | |
| Drought | 0.0005 | 0.0001 | 0.0002 | 0.0002 |
| | (0.0007) | (0.0007) | (0.0007) | (0.0006) |
| Observations | 810,144 | 810,144 | 810,144 | 810,144 |
| R-squared | 0.035 | 0.040 | 0.042 | 0.035 |
| Mean of Dep. Var. | 0.0258 | 0.0281 | 0.0201 | 0.0321 |
| ***Panel C. Drought in the dry season ($t-2$)*** | | | | |
| Drought | -0.0023*** | -0.0012* | 0.0010 | 0.0009 |
| | (0.0006) | (0.0007) | (0.0006) | (0.0006) |
| Observations | 810,144 | 810,144 | 810,144 | 810,144 |
| R-squared | 0.035 | 0.040 | 0.042 | 0.035 |
| Mean of Dep. Var. | 0.0258 | 0.0283 | 0.0200 | 0.0318 |
| **Controls** | | | | |
| District and year FE | Yes | Yes | Yes | Yes |
| Rainfall in other season-year | Yes | Yes | Yes | Yes |
| Mother characteristics | Yes | Yes | Yes | Yes |
| District-specific linear time trend | Yes | Yes | Yes | Yes |
| Birth parity FE | Yes | Yes | Yes | Yes |
| Gender composition FE | Yes | Yes | Yes | Yes |

*Data*: PCS 2004-2008, 2010-2013

*Notes*: This table show OLS regressions for the effects of droughts on the likelihood of giving birth. The outcome is the indicator for giving birth in a given quarter in the survey year. Robust standard errors are shown in parentheses clustered at the district level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 1.7: Effects of Droughts on a Child's Sex

| | Dependent variable: Newborn is a boy=1 | | | |
|---|---|---|---|---|
| | Born in Apr-Jun | Born in Jul-Sep | Born in Oct-Dec | Born in Jan-Mar |
| | (1) | (2) | (3) | (4) |
| ***Panel A. Drought in the dry season ($t$)*** | | | | |
| Drought | 0.0068 | 0.0143 | -0.0071 | 0.0261 |
| | (0.0182) | (0.0199) | (0.0148) | (0.0181) |
| Observations | 20,683 | 22,791 | 23,212 | 19,005 |
| R-squared | 0.009 | 0.007 | 0.009 | 0.010 |
| Mean of Dep. Var. | 0.5244 | 0.5143 | 0.5267 | 0.5262 |
| ***Panel B. Drought in the dry season ($t-1$)*** | | | | |
| Drought | 0.0006 | -0.0271 | -0.0156 | 0.0137 |
| | (0.0153) | (0.0170) | (0.0135) | (0.0197) |
| Observations | 20,683 | 22,791 | 23,212 | 19,005 |
| R-squared | 0.009 | 0.007 | 0.009 | 0.010 |
| Mean of Dep. Var. | 0.5232 | 0.5190 | 0.5267 | 0.5250 |
| ***Panel C. Drought in the dry season ($t-2$)*** | | | | |
| Drought | 0.0234* | -0.0046 | 0.0058 | -0.0134 |
| | (0.0134) | (0.0152) | (0.0128) | (0.0157) |
| Observations | 20,683 | 22,791 | 23,212 | 19,005 |
| R-squared | 0.009 | 0.007 | 0.008 | 0.010 |
| Mean of Dep. Var. | 0.5219 | 0.5179 | 0.5272 | 0.5264 |
| **Controls** | | | | |
| District and year FE | Yes | Yes | Yes | Yes |
| Rainfall in other season-year | Yes | Yes | Yes | Yes |
| Mother characteristics | Yes | Yes | Yes | Yes |
| District-specific linear time trend | Yes | Yes | Yes | Yes |
| Birth parity FE | Yes | Yes | Yes | Yes |
| Gender composition FE | Yes | Yes | Yes | Yes |

*Data*: PCS 2004-2008, 2010-2013
*Notes*: This table provides the effects of droughts on the likelihood that a newborn is a boy. The outcome is the indicator for a newborn being a boy conditional on being born in each quarter of the survey year. Robust standard errors are shown in parentheses clustered at the district level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 1.8: Effects of Droughts on Infant Mortality

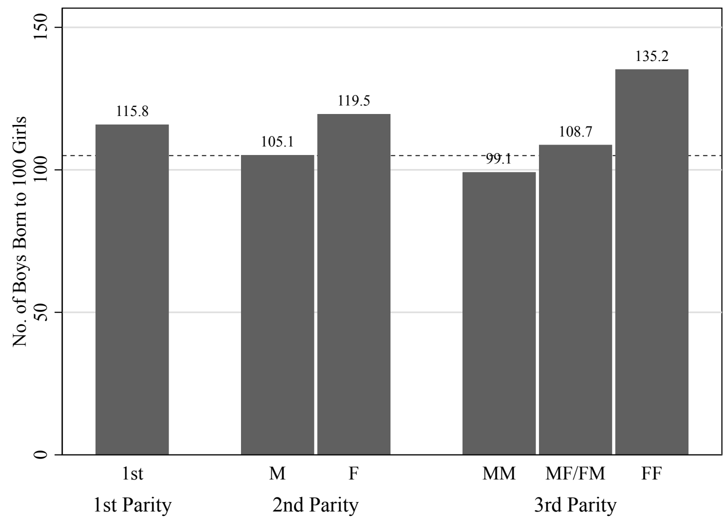
| | Dependent variable | | | | | |
|---|---|---|---|---|---|---|
| | **Newborn is dead=1** | | | | | **ln(IMR)** |
| | Born in Apr-Mar | Born in Apr-Jun | Born in Jul-Sep | Born in Oct-Nov | Born in Dec-Mar | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| ***Panel A. Drought in the dry season (t)*** | | | | | | |
| Drought | -0.0010 | 0.0032 | 0.0005 | -0.0047 | -0.0026 | -0.0057 |
| | (0.0014) | (0.0030) | (0.0029) | (0.0032) | (0.0029) | (0.0144) |
| Observations | 85,691 | 20,683 | 22,791 | 23,212 | 19,005 | 477 |
| R-squared | 0.008 | 0.014 | 0.013 | 0.014 | 0.017 | 0.789 |
| Mean of Dep. Var. | 0.0081 | 0.0083 | 0.0081 | 0.0084 | 0.0077 | 2.7273 |
| ***Panel B. Drought in the dry season ($t-1$)*** | | | | | | |
| Drought | -0.0001 | -0.0043 | 0.0018 | 0.0044 | -0.0047* | -0.0045 |
| | (0.0014) | (0.0026) | (0.0029) | (0.0028) | (0.0024) | (0.0201) |
| Observations | 85,691 | 20,683 | 22,791 | 23,212 | 19,005 | 477 |
| R-squared | 0.008 | 0.014 | 0.013 | 0.014 | 0.017 | 0.789 |
| Mean of Dep. Var. | 0.0082 | 0.0082 | 0.0090 | 0.0079 | 0.0081 | 2.7245 |
| ***Panel C. Drought in the dry season ($t-2$)*** | | | | | | |
| Drought | -0.0006 | -0.0024 | -0.0025 | 0.0002 | 0.0026 | -0.0066 |
| | (0.0013) | (0.0023) | (0.0026) | (0.0032) | (0.0023) | (0.0161) |
| Observations | 85,691 | 20,683 | 22,791 | 23,212 | 19,005 | 477 |
| R-squared | 0.008 | 0.014 | 0.013 | 0.014 | 0.017 | 0.788 |
| Mean of Dep. Var. | 0.0082 | 0.0082 | 0.0084 | 0.0082 | 0.0086 | 2.7309 |
| **Controls** | | | | | | |
| Province FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Region×Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Rainfall in other season-year | Yes | Yes | Yes | Yes | Yes | Yes |
| Mother characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| Province-specific linear time trend | Yes | Yes | Yes | Yes | Yes | Yes |
| Birth parity FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Gender composition FE | Yes | Yes | Yes | Yes | Yes | Yes |

*Data*: PCS 2004-2008, 2010-2013; Province-level IMR from the GSO
*Notes*: This table presents the results from regressions estimating the effect of droughts on infant mortality. Columns (1)-(5) report the effect on the likelihood that a newborn is dead in the survey year (column (1)) and in a given quarter in the survey year (columns (2)-(5)) from the PCS. Column (6) presents the effect on the log of province-level infant mortality rates from the GSO. Robust standard errors are shown in parentheses clustered at the province level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

# Appendix to Chapter 1

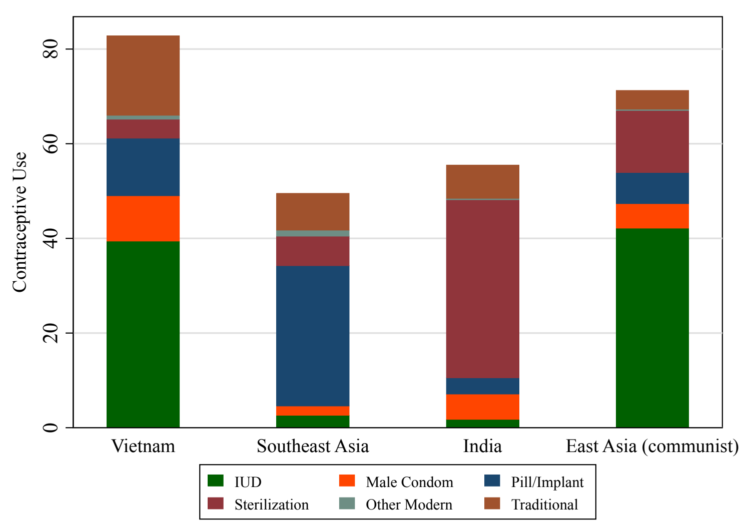
## 1.A Appendix Figures

Figure 1.A.1: Sex Ratio at Birth by Birth Parity




*Data*: Census 2009
*Notes*: This figure describes the sex ratio at birth by birth parity depending on the sex of previous births. 'M' of the 2nd parity means the firstborn is son, and 'F' means a daughter. Likewise, for the third parity, 'MM' means the first two births are boys, whereas 'FF' means there is no son in the previous two births.

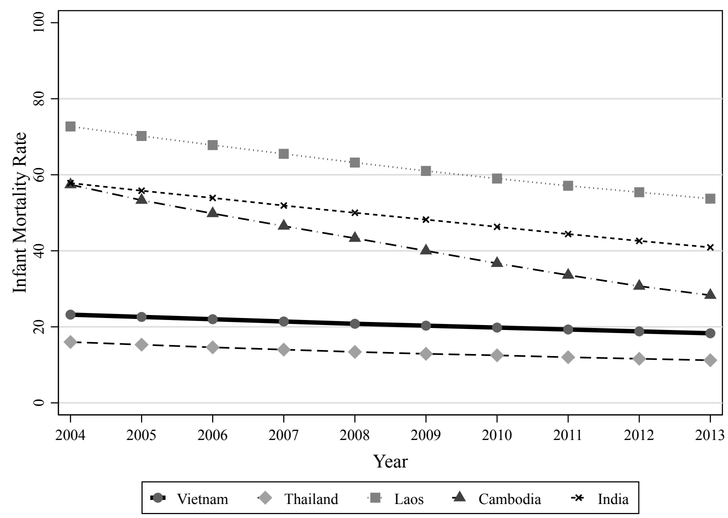
Figure 1.A.2: Contraceptive Use




*Data*: The World Contraceptive Use 2016, the United Nations
*Notes*: This figure describes average contraceptive use by methods across region and country in 2004-2013. Communist East Asian countries include China, North Korea and Mongolia. Southeast Asia statistics exclude Vietnam.

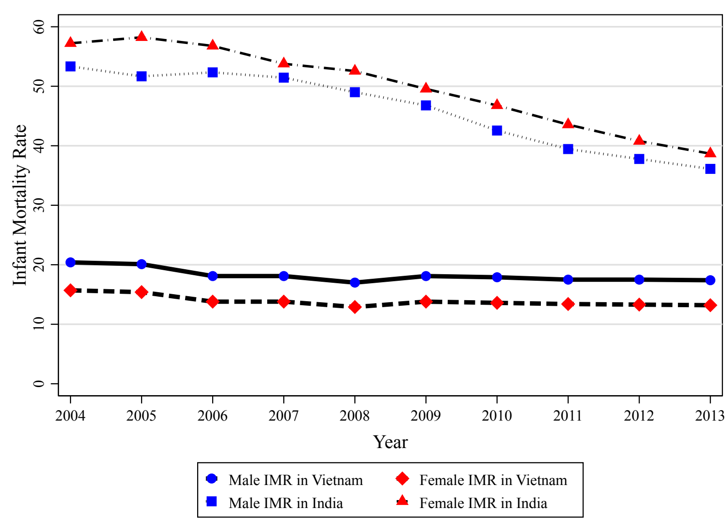
Figure 1.A.3: Infant Mortality Rate

(a) Infant Mortality Rate by Country



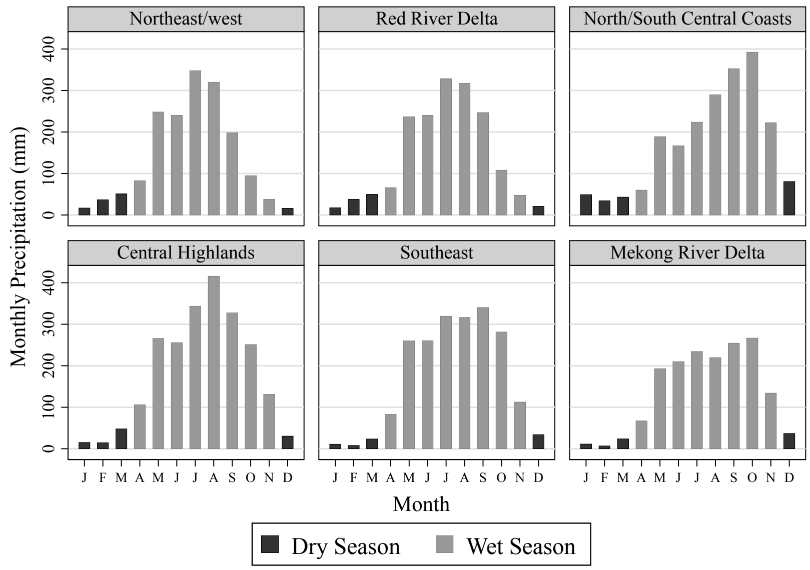

*Notes*: This figure describes the infant mortality rates of Vietnam and neighboring countries in Southeast Asia using the official statistics from the World Bank.

(b) Infant Mortality Rate by Sex



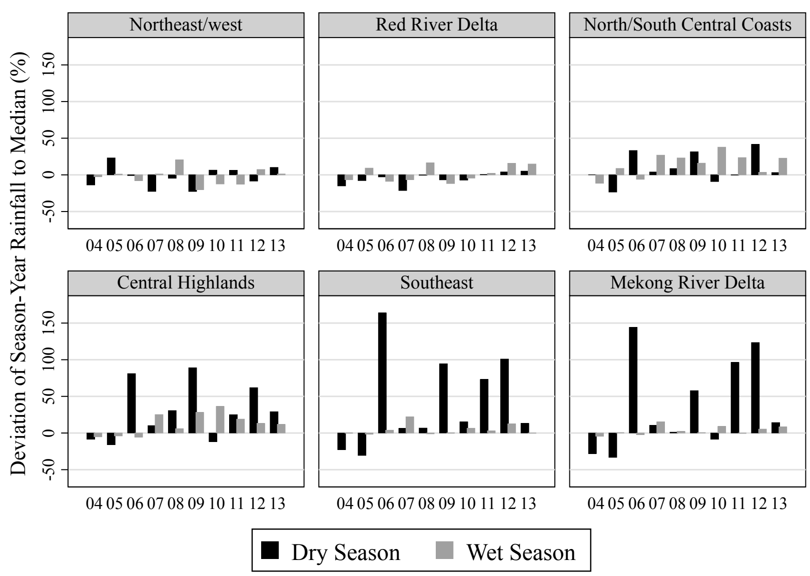

*Notes*: This figure describes infant mortality rates by sex using the official statistics from the Vietnam GSO. For the IMR in India, I choose nine states in northwestern India showing strong son preference (Anukriti (2017)). They consist of Rajasthan, Himachal Pradesh, Delhi, Gujarat, Uttar Pradesh, Madhya Pradesh, Maharashtra and Haryana.

Figure 1.A.4: Monthly Precipitation by Region



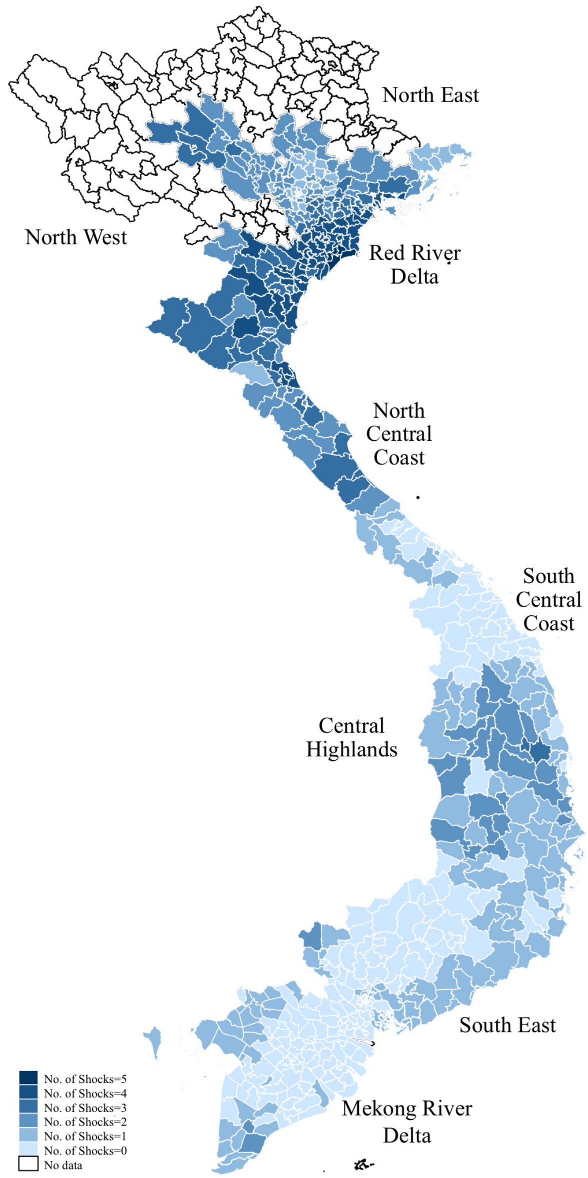

*Notes*: The figure describes the monthly precipitation averaged over the years in 1984-2013 by region. The dry season is from December to March.

Figure 1.A.5: Rainfall Deviation to 30-Year Median by Region



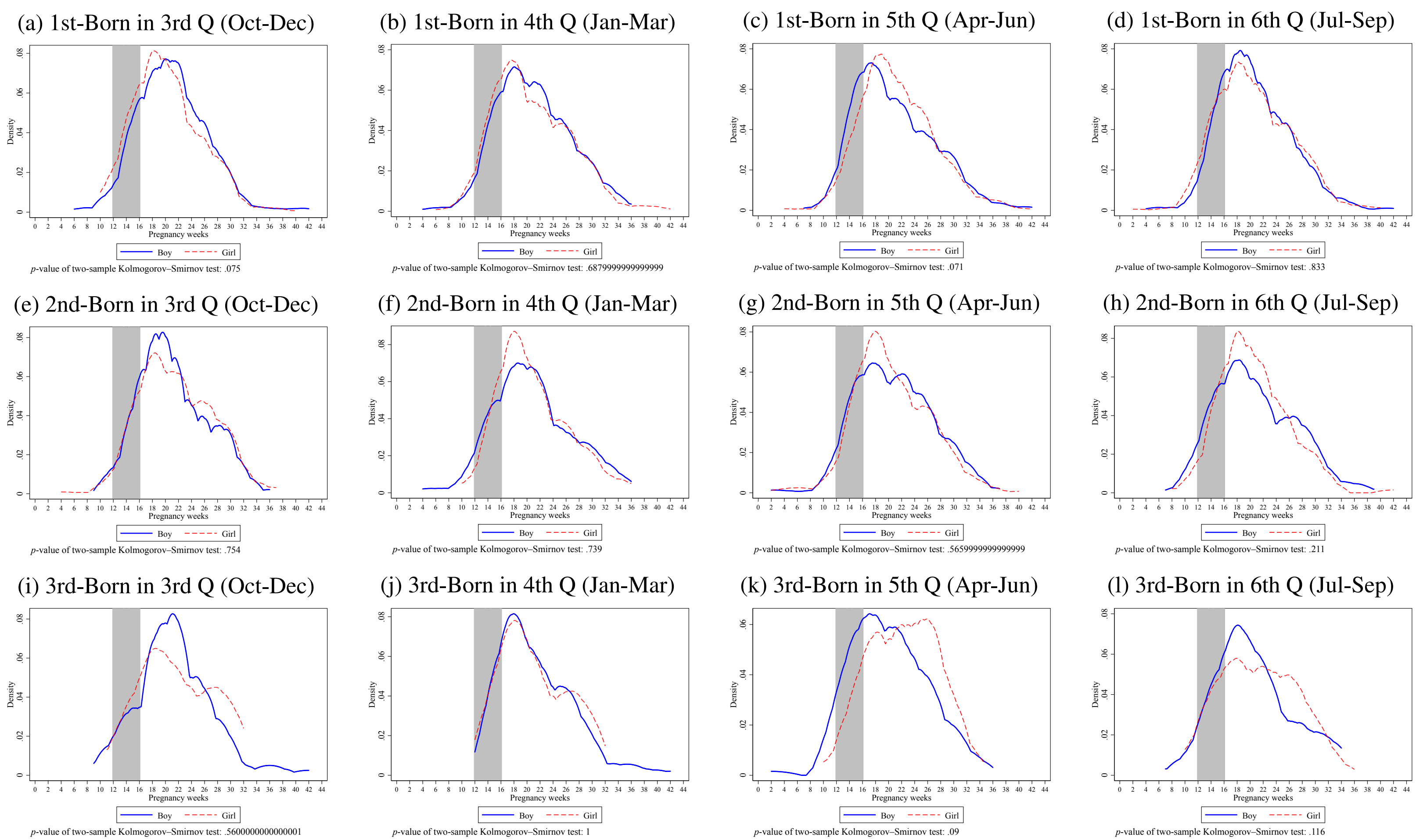

*Notes*: This figure provides the rainfall deviation (in percentage) relative to the long-run average season-year rainfall for each region.

Figure 1.A.6: Distribution of Droughts across Districts in Vietnam




*Notes*: This map describes the number of droughts experienced by each district in the sample period from 2004-2013. Droughts are defined as seasonal rainfall occurring below the 20th percentile of the district-specific dry season rainfall distribution in 1984-2013 from Climate Hazards Group InfraRed Precipitation with Station version 2.0. The districts with 'No data' are excluded in the analysis because more than 50 percent of the heads of household are not ethnically Kinh in the 10 provinces. The 10 provinces are Cao Bang, Bac Kan, Ha Giang, Lang Son, Lai Chau, Son La, Dien Bien, Hoa Binh, Lao Cai, and Tuyen Quang.

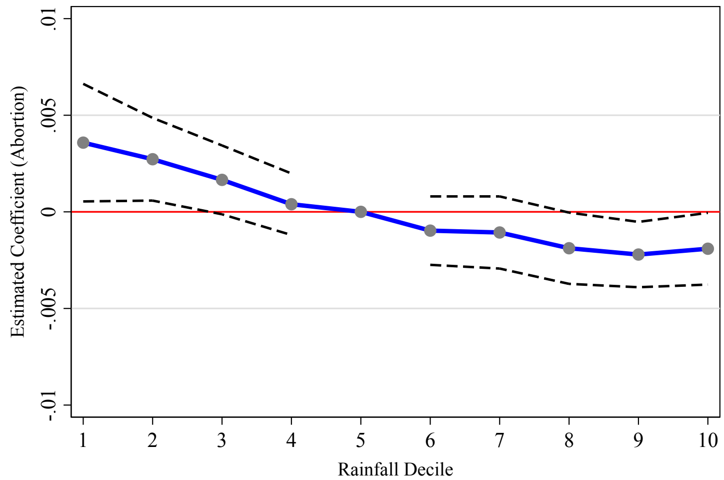
Figure 1.A.7: Distribution of the Weeks of Sex Determination by the Sex of Newborn and by the Birth Order




*Data*: PCS 2004-2008, 2010-2013

*Notes*: Figures plot the kernel density estimation on the distributions of the weeks of fetal sex determination of affected mothers using ultrasound, conditional on the childbirth of each parity from the 3rd (Panel (a), (e), (i)) to the 6th quarter (Panel (d), (h), (l)) after droughts. It is from the 3rd to the 6th quarter after droughts when the aborted fetuses in the PCS ($t$) would have been born otherwise. The gray vertical bands denote the pregnancy weeks when sex-selective abortion can be performed; the 12th week is the earliest possible week when the fetal sex can be determined by ultrasound, and the 16th weeks is the latest possible week when abortion can be performed from the DHS 2002 (Committee for Population, Family and Children [Vietnam] and ORC Macro, 2003). It is the 5th quarter after droughts (April-June) when the effect of droughts on birth is significant as shown in Figure 1.6.

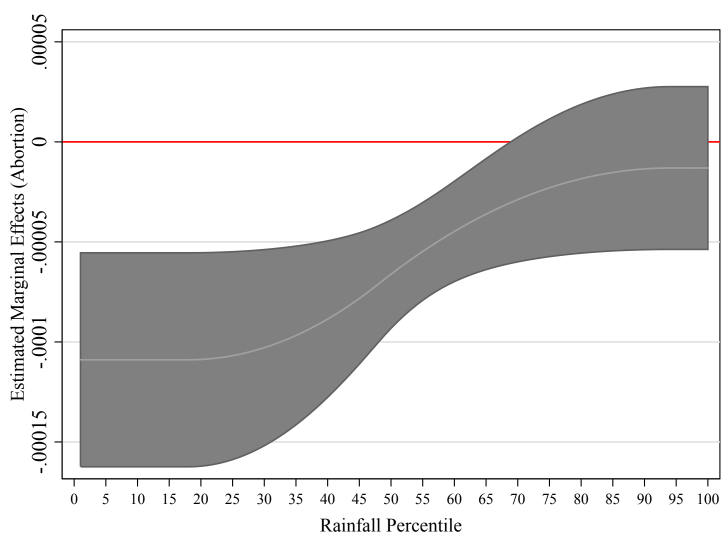
Figure 1.A.8: Effects of Rainfall Decile on Abortion




Mean at the 5th Decile: .0088

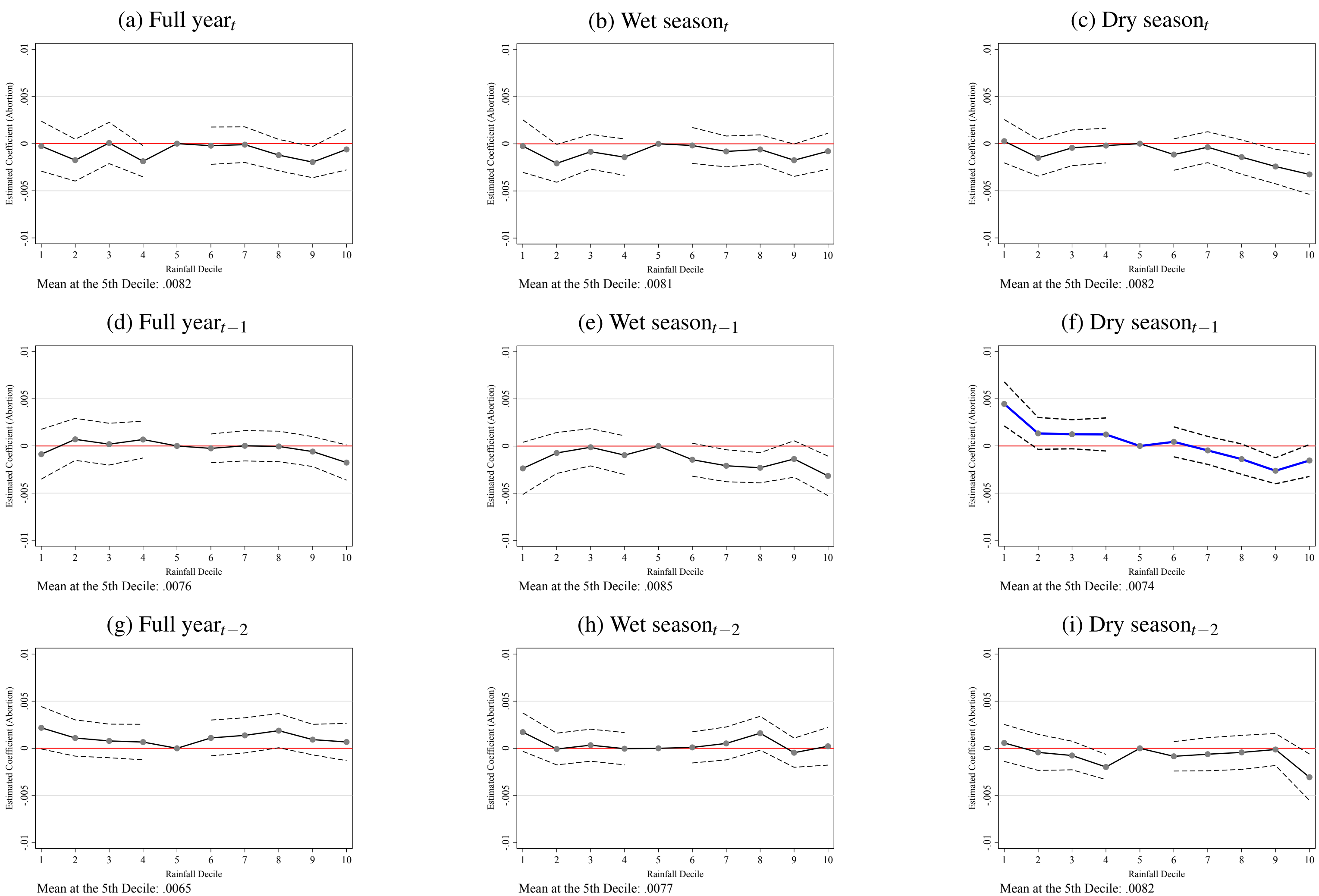
*Data*: PCS 2004-2008, 2010-2013
*Notes*: The figure plots coefficients and 95% confidence intervals from a regression of abortion on the dummies for each 10th percentile (decile) of the gamma distribution fitted by the district-specific dry season rainfalls in 1984-2013. The omitted category is the 5th decile.

Figure 1.A.9: Marginal Effects of Rainfall Percentiles on Abortion




*Data*: PCS 2004-2008, 2010-2013
*Notes*: The figure plots marginal effects of rainfall percentile on the indicator for abortion along with its 95% confidence intervals. The marginal effects are estimated using a restricted cubic spline with the knots at 18, 48 and 98 which are chosen by Harrell's procedure.

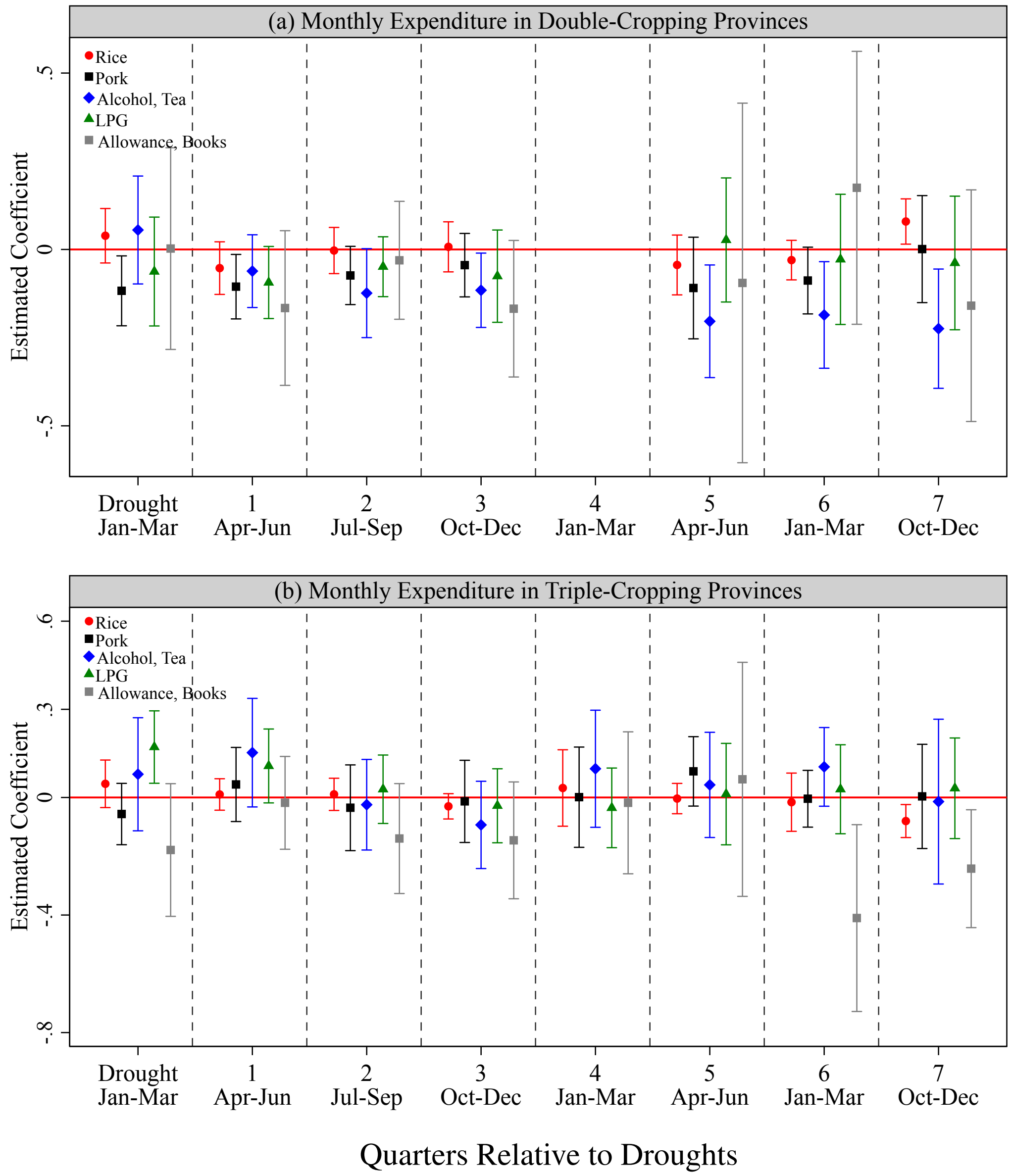
Figure 1.A.10: Effects of Other Season-Year Rainfalls on Abortion




*Data*: PCS 2004-2008, 2010-2013

*Notes*: Figures plot coefficients and 95% confidence intervals from regressions of abortion on the dummies for each 10th percentile (decile) of the district-specific rainfalls of the full year (Dec-Nov), dry season (Dec-Mar) and wet season (Apr-Nov) in 1984-2013. The omitted category is the 5th decile. Panel (f) plots the coefficients on each dummy (connected with blue solid lines) that is constructed using the dry season rainfall in $t-1$, the drought shocks used in the main analyses.

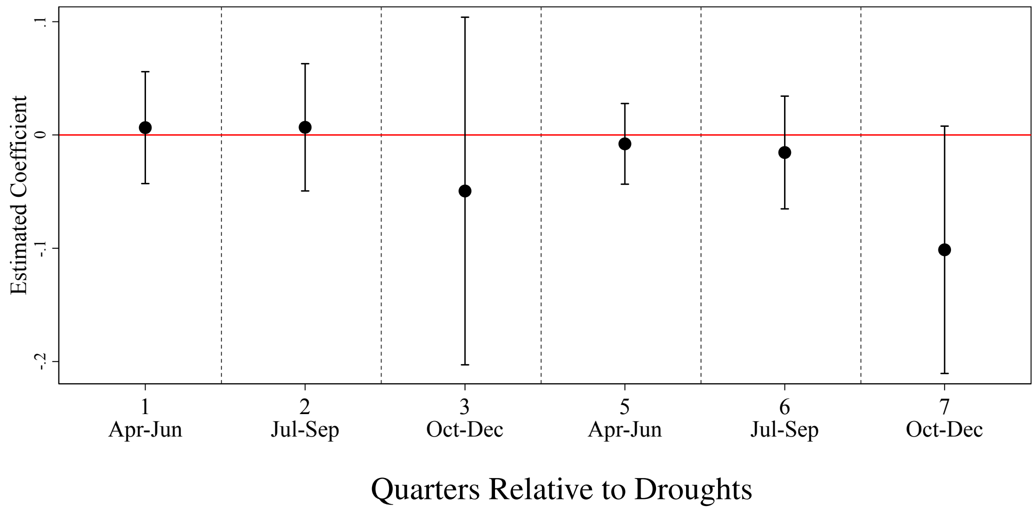
Figure 1.A.11: Effects of droughts on monthly expenditure by rice-cropping patterns




*Data*: VHLSS 2010,2012,2014

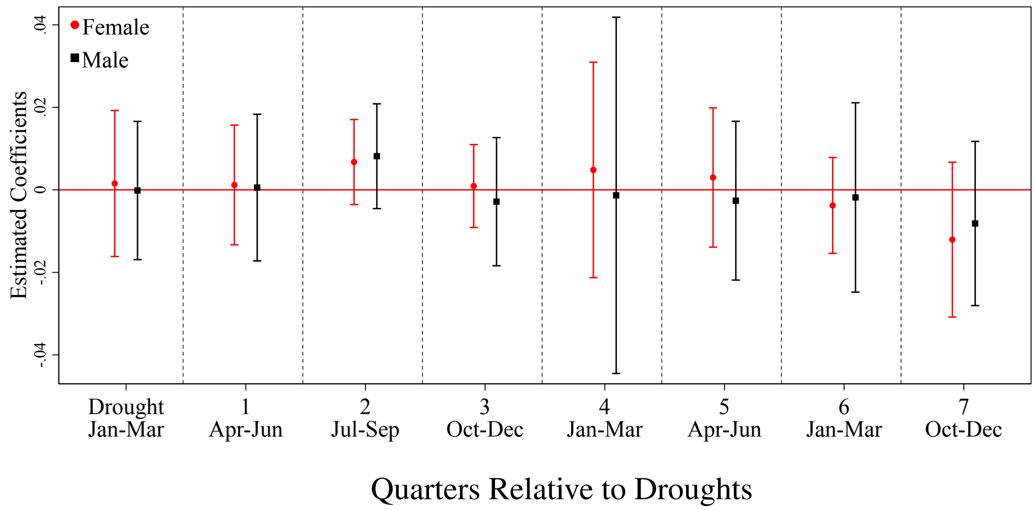
*Notes*: Figures plot the coefficients on the interaction terms between quarters and droughts in $t$ and $t-1$ from the regressions estimating the effect on the log of monthly household expenditures in the double-cropping provinces (Panel (a)) and in the triple-cropping provinces (Panel (b)). Colored bars represent the 95% confidence intervals of the estimated coefficients. There are no point estimates in the 4th quarter after droughts in Panel (a) due to no observations of households surveyed in that quarter in the double-cropping provinces.

Figure 1.A.12: Effect of Droughts on Recent Illness




*Data*: VHLSS 2004, 2006, 2008
*Notes*: This figure plots the coefficients on the interaction terms between quarters and droughts in $t$ and $t-1$ from the regression estimating the effect on recent illness of any rural household member. The indicator for the recent illness becomes one if any household members suffered from any illness or injuries for the past 4 weeks from the date of the survey. The regression includes quarter FEs, province FEs, region×year FEs, the logs of other season-year rainfalls and household-level controls such as the sex, age, ethnicity (Kinh or not) and years of schooling of the household head, the household size and the dummy for multigenerational household. Robust standard errors are clustered for the province level. There are no point estimates in the quarter with droughts, and the 4th quarter after droughts due to no observations of households surveyed in that quarters.

Figure 1.A.13: Effect of Droughts on Recent Migration




*Data*: VHLSS 2010, 2012, 2014
*Notes*: This figure plots the coefficients on the interaction terms between quarters and droughts in $t$ and $t-1$ from the regression estimating the effect on recent migration of a rural household. The indicator for the recent migration becomes one if a married woman or man had been away from home for less than 6 months at the time of the survey. Black bars represent the 95% confidence intervals of the estimated coefficients. The regression includes quarter FEs, province FEs, region×year FEs, the logs of other season-year rainfalls and household-level controls such as the sex, age, ethnicity (Kinh or not) and years of schooling of the household head, the household size and the dummy for multigenerational household. Robust standard errors are clustered for the province level.

## 1.B Appendix Tables

Table 1.B.1: Effect of Droughts on Abortion using Distributed Lagged Model

| | Dependent variable: Abortion=1 | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Low rainfall in the dry season ($t$) | 0.0008 | 0.0008 | 0.0006 | 0.0006 | 0.0006 | 0.0006 | 0.0006 |
| | (0.0008) | (0.0008) | (0.0008) | (0.0008) | (0.0008) | (0.0008) | (0.0012) |
| Low rainfall in the dry season ($t-1$) | 0.0022*** | 0.0023*** | 0.0021*** | 0.0021*** | 0.0021*** | 0.0021*** | 0.0030** |
| | (0.0007) | (0.0007) | (0.0008) | (0.0008) | (0.0008) | (0.0008) | (0.0012) |
| Low rainfall in the dry season ($t-2$) | 0.0013** | 0.0013** | 0.0013* | 0.0013* | 0.0013* | 0.0013* | 0.0016 |
| | (0.0006) | (0.0006) | (0.0007) | (0.0007) | (0.0007) | (0.0007) | (0.0011) |
| Low rainfall in the wet season ($t$) | -0.0006 | -0.0006 | -0.0008 | -0.0008 | -0.0008 | -0.0007 | 0.0007 |
| | (0.0007) | (0.0007) | (0.0008) | (0.0008) | (0.0008) | (0.0008) | (0.0013) |
| Low rainfall in the wet season ($t-1$) | -0.0007 | -0.0006 | -0.0009 | -0.0009 | -0.0009 | -0.0009 | -0.0007 |
| | (0.0008) | (0.0008) | (0.0008) | (0.0008) | (0.0008) | (0.0008) | (0.0014) |
| Low rainfall in the wet season ($t-2$) | -0.0004 | -0.0004 | -0.0002 | -0.0002 | -0.0002 | -0.0002 | -0.0000 |
| | (0.0006) | (0.0006) | (0.0006) | (0.0006) | (0.0006) | (0.0006) | (0.0011) |
| Observations | 811,092 | 810,144 | 810,144 | 810,144 | 810,144 | 802,589 | 441,789 |
| R-squared | 0.008 | 0.009 | 0.012 | 0.012 | 0.012 | 0.012 | 0.013 |
| Mean of Dep. Var. | | | | 0.0066 | | | |
| **Controls** | | | | | | | |
| District and year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Mother characteristics | | Yes | Yes | Yes | Yes | Yes | Yes |
| District-specific linear time trend | | | Yes | Yes | Yes | Yes | Yes |
| Birth parity FE | | | | Yes | Yes | Yes | Yes |
| Gender composition FE | | | | | Yes | Yes | Yes |
| Fertility characteristics | | | | | | Yes | Yes |
| Spouse characteristics | | | | | | | Yes |

*Data*: PCS 2004-2008, 2010-2013

*Notes*: Low rainfall shocks refer to the realization of rainfall in the wet season (April-November) or in the dry season (December-March) below the 20th percentile of historical distribution of district-specific seasonal rainfall in 1984-2013. The dependent variable is the indicator for the experience of abortion during the survey year. Fertility characteristics control consists of her age at the first birth and the birth spacing referred to as the months between the most recent childbirth and the starting month of the survey period. Spouse characteristics include her spouse age, age squared and his educational attainment. The mean of dependent variable is the mean abortion rate of mothers living in the districts that were not inflicted with droughts. Robust standard errors are shown in parentheses clustered at the district level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 1.B.2: Effect of Droughts on Abortion of Women in Urban and 10 Northern Provinces

| | Dependent variable: Abortion=1 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| ***Panel A. Urban Sample*** | | | | | | | | |
| Drought | 0.0016* | 0.0014 | 0.0014 | 0.0013 | 0.0013 | 0.0013 | 0.0012 | 0.0008 |
| | (0.0009) | (0.0009) | (0.0009) | (0.0009) | (0.0009) | (0.0009) | (0.0009) | (0.0010) |
| Observations | 840,836 | 840,836 | 839,551 | 839,551 | 839,551 | 839,551 | 834,191 | 490,732 |
| R-squared | 0.008 | 0.008 | 0.009 | 0.012 | 0.012 | 0.012 | 0.013 | 0.014 |
| Mean of Dep. Var. | 0.0069 | 0.0069 | 0.0069 | 0.0069 | 0.0069 | 0.0069 | 0.0069 | 0.0069 |
| ***Panel B. 10 Northern Provinces*** | | | | | | | | |
| Drought | -0.0008 | -0.0012 | -0.0012 | -0.0009 | -0.0009 | -0.0009 | -0.0009 | 0.0007 |
| | (0.0017) | (0.0018) | (0.0018) | (0.0019) | (0.0019) | (0.0019) | (0.0019) | (0.0021) |
| Observations | 169,410 | 169,410 | 169,196 | 169,196 | 169,196 | 169,196 | 167,805 | 93,159 |
| R-squared | 0.014 | 0.015 | 0.016 | 0.020 | 0.021 | 0.021 | 0.021 | 0.019 |
| Mean of Dep. Var. | 0.0129 | 0.0129 | 0.0129 | 0.0129 | 0.0129 | 0.0129 | 0.0129 | 0.0129 |
| **Controls** | | | | | | | | |
| District and year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Rainfall in other season-year | | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Mother characteristics | | | Yes | Yes | Yes | Yes | Yes | Yes |
| District-specific linear time trend | | | | Yes | Yes | Yes | Yes | Yes |
| Birth parity FE | | | | | Yes | Yes | Yes | Yes |
| Gender composition FE | | | | | | Yes | Yes | Yes |
| Fertility characteristics | | | | | | | Yes | Yes |
| Spouse characteristics | | | | | | | | Yes |

*Data*: PCS 2004-2008, 2010-2013

*Notes*: This table reports the results of regressions for the urban and the 10 northern-province samples. In the 10 provinces, more than 50 percent of the heads of household are not ethnically Kinh. The 10 provinces are Cao Bang, Bac Kan, Ha Giang, Lang Son, Lai Chau, Son La, Dien Bien, Hoa Binh, Lao Cai, Tuyen Quang, and the locations are mapped in Figure 1.A.6. The dependent variable is the indicator for the experience of abortion in the survey year. Fertility characteristics control consists of her age at the first birth and the birth spacing referred to as the months between the most recent childbirth and the starting month of the survey period. Spouse characteristics include her spouse age, age squared and his educational attainment. The mean of dependent variable is the mean abortion rate of mothers living in the districts that were not inflicted with droughts. Robust standard errors are shown in parentheses clustered at the district level. *** Significant at the 1 percent level.

Table 1.B.3: Effect of Droughts on Monthly Household Expenditure

| | **Dependent variables** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | ln(Exp. in ’000VND) | | | **1**(Expenditure¿0) | | | | ln(Quantity in Kg) | |
| | Total | Excl. gift | Excl. gift & self | Rice | Pork | Gas | Child | Rice | Pork |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Q1 (Jan-Mar)×Drought($t$) | -0.017 | -0.017 | -0.035 | 0.000 | -0.006 | -0.047 | 0.019 | 0.074*** | -0.062* |
| | (0.029) | (0.030) | (0.036) | (0.005) | (0.011) | (0.029) | (0.018) | (0.026) | (0.037) |
| Q2 (Apr-Jun)×Drought($t$) | 0.021 | 0.023 | 0.022 | 0.005 | 0.009 | -0.006 | 0.017 | -0.001 | -0.017 |
| | (0.020) | (0.020) | (0.024) | (0.004) | (0.008) | (0.021) | (0.017) | (0.017) | (0.036) |
| Q3 (Jul-Sep)× Drought($t$) | 0.031 | 0.033 | 0.028 | 0.006 | 0.017 | -0.028 | -0.013 | 0.013 | -0.046 |
| | (0.021) | (0.020) | (0.025) | (0.004) | (0.011) | (0.021) | (0.018) | (0.022) | (0.039) |
| Q4 (Oct-Dec)× Drought($t$) | 0.014 | 0.020 | 0.015 | 0.002 | -0.014 | -0.036* | 0.021 | 0.029 | -0.054 |
| | (0.022) | (0.022) | (0.027) | (0.006) | (0.011) | (0.019) | (0.017) | (0.022) | (0.037) |
| Q1 (Jan-Mar)× Drought($t-1$) | 0.017 | 0.020 | 0.071 | 0.002 | 0.002 | -0.028 | 0.001 | 0.053 | 0.012 |
| | (0.036) | (0.035) | (0.054) | (0.017) | (0.017) | (0.054) | (0.055) | (0.046) | (0.077) |
| Q2 (Apr-Jun)× Drought($t-1$) | -0.029 | -0.028 | -0.057 | 0.011* | 0.003 | -0.054* | -0.013 | 0.017 | 0.007 |
| | (0.041) | (0.042) | (0.051) | (0.006) | (0.010) | (0.028) | (0.026) | (0.019) | (0.042) |
| Q3 (Jul-Sep)× Drought($t-1$) | -0.011 | -0.008 | -0.016 | 0.009 | 0.024** | -0.009 | -0.028 | -0.009 | -0.007 |
| | (0.033) | (0.033) | (0.044) | (0.008) | (0.011) | (0.030) | (0.025) | (0.021) | (0.027) |
| Q4 (Oct-Dec)× Drought($t-1$) | 0.023 | 0.030 | 0.017 | 0.011 | -0.000 | -0.004 | -0.064*** | -0.003 | -0.056 |
| | (0.030) | (0.029) | (0.049) | (0.007) | (0.011) | (0.031) | (0.022) | (0.021) | (0.050) |
| Observations | 17,448 | 17,433 | 17,432 | 17,448 | 17,448 | 17,448 | 17,448 | 17,297 | 16,665 |
| R-squared | 0.536 | 0.531 | 0.499 | 0.016 | 0.063 | 0.287 | 0.217 | 0.591 | 0.350 |
| Mean of Dep. Var. | 7.860 | 7.834 | 7.654 | 0.991 | 0.948 | 0.590 | 0.367 | 3.545 | 1.058 |
| **Controls** | | | | | | | | | |
| Province and year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Region × Quarter FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Rainfall in other season-year | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Household characteristics | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| ln(Total Expenditure) | | | | Yes | Yes | Yes | Yes | Yes | Yes |

*Data*: VHLSS 2010, 2012, 2014

*Notes*: This table presents the coefficients on the interaction terms between quarters and droughts in $t$ and $t-1$ from the regressions estimating the effect on the log of total expenditure (column (1)-(3)), the indicator for the consumption of each good (columns (4)-(7)), and the log of consumed quantity (columns (8)-(9)). Column (2) excludes the consumption of gifts from total expenditure, and column (3) further excludes the consumption of self-generated goods. Robust standard errors are shown in parentheses clustered at the province level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 1.B.4: Effect of Droughts on Monthly Household Expenditure on Food Items

| | Dependent vars. (ln(Expenditure in '000 VND)) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | FAFH | Rice | Pork | Veget., Fruit | Other Meat | Seafood | Dairy | Alcohol Tea | Other Starch | Tobacco | Staple | ETC |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Q1 (Jan-Mar)×Drought($t$) | 0.027 | 0.036 | -0.091** | -0.056 | 0.014 | -0.114** | -0.101* | 0.047 | 0.087 | 0.100* | 0.056 | -0.065** |
| | (0.065) | (0.027) | (0.036) | (0.037) | (0.054) | (0.048) | (0.057) | (0.063) | (0.054) | (0.053) | (0.065) | (0.031) |
| Q2 (Apr-Jun)×Drought($t$) | -0.054 | -0.023 | -0.039 | 0.056 | 0.011 | 0.002 | 0.073 | 0.048 | 0.061 | -0.016 | 0.032 | 0.078** |
| | (0.060) | (0.022) | (0.036) | (0.044) | (0.042) | (0.047) | (0.050) | (0.051) | (0.044) | (0.064) | (0.056) | (0.032) |
| Q3 (Jul-Sep)×Drought($t$) | 0.068 | 0.007 | -0.046 | -0.019 | 0.066 | -0.049 | 0.065 | -0.045 | 0.058 | -0.002 | 0.046 | 0.031 |
| | (0.068) | (0.022) | (0.038) | (0.039) | (0.045) | (0.049) | (0.056) | (0.051) | (0.046) | (0.055) | (0.058) | (0.035) |
| Q4 (Oct-Dec)×Drought($t$) | -0.008 | -0.005 | -0.035 | -0.012 | 0.012 | -0.028 | 0.057 | -0.100** | 0.078 | 0.052 | 0.099* | 0.000 |
| | (0.065) | (0.027) | (0.041) | (0.042) | (0.039) | (0.046) | (0.063) | (0.043) | (0.047) | (0.065) | (0.057) | (0.036) |
| Q1 (Jan-Mar)×Drought($t-1$) | -0.116 | 0.035 | 0.002 | 0.123** | -0.077 | -0.103* | -0.200 | 0.052 | -0.152 | 0.080 | 0.355* | 0.062 |
| | (0.103) | (0.048) | (0.083) | (0.047) | (0.079) | (0.059) | (0.155) | (0.059) | (0.096) | (0.086) | (0.211) | (0.050) |
| Q2 (Apr-Jun)×Drought($t-1$) | -0.110 | -0.024 | -0.025 | -0.013 | 0.017 | -0.066 | -0.022 | -0.047 | 0.118 | -0.149* | 0.005 | -0.068 |
| | (0.078) | (0.023) | (0.045) | (0.056) | (0.077) | (0.043) | (0.063) | (0.058) | (0.080) | (0.082) | (0.099) | (0.056) |
| Q3 (Jul-Sep)×Drought($t-1$) | -0.070 | -0.024 | -0.038 | -0.032 | -0.106** | -0.035 | 0.042 | -0.030 | 0.054 | 0.019 | -0.087 | 0.024 |
| | (0.052) | (0.021) | (0.028) | (0.043) | (0.046) | (0.060) | (0.068) | (0.063) | (0.066) | (0.074) | (0.064) | (0.045) |
| Q4 (Oct-Dec)×Drought($t-1$) | 0.106 | 0.017 | 0.005 | -0.018 | -0.007 | 0.009 | 0.080 | -0.107 | 0.026 | -0.005 | -0.051 | 0.008 |
| | (0.074) | (0.030) | (0.056) | (0.046) | (0.067) | (0.063) | (0.072) | (0.073) | (0.038) | (0.099) | (0.062) | (0.048) |
| Observations | 12,643 | 17,297 | 16,665 | 17,384 | 13,437 | 16,827 | 15,547 | 14,239 | 15,373 | 8,987 | 6,526 | 17,422 |
| R-squared | 0.317 | 0.570 | 0.392 | 0.378 | 0.311 | 0.434 | 0.274 | 0.212 | 0.214 | 0.380 | 0.176 | 0.419 |
| Mean of Dep. Var. | 5.721 | 5.867 | 5.303 | 5.205 | 5.265 | 5.209 | 4.242 | 4.182 | 4.062 | 4.407 | 3.140 | 5.194 |
| **Controls** | | | | | | | | | | | | |
| Province and year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Region × Quarter FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Rainfall in other season-year | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Household characteristics | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| ln(Total Expenditure) | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

*Data*: VHLSS 2010, 2012, 2014

*Notes*: This table presents the coefficients on the interaction terms between quarters and droughts in $t$ and $t-1$ from the regressions estimating the effect on the log of monthly household expenditure on each food category. The regression columns are sorted by the share of expenditure on the item to the total expenditure from the largest to the smallest. FAFH denotes the food away from home. Robust standard errors are shown in parentheses clustered at the province level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 1.B.5: Effect of Droughts on Monthly Household Expenditure on Non-food Items

| | **Dependent vars. (ln(Expenditure in '000 VND))** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Petro | Hygiene | Gas (LPG) | Biomass Fuel | Child | Detergent | Female Goods | Coal Kerosene | ETC |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Q1 (Jan-Mar)×Drought($t$) | 0.015 | -0.038 | 0.043 | 0.104* | -0.079 | -0.040 | -0.035 | 0.118 | 0.137* |
| | (0.043) | (0.041) | (0.044) | (0.062) | (0.091) | (0.033) | (0.094) | (0.151) | (0.077) |
| Q2 (Apr-Jun)×Drought($t$) | 0.079** | 0.045 | -0.008 | -0.006 | -0.065 | 0.040 | 0.079 | -0.103 | 0.047 |
| | (0.035) | (0.030) | (0.037) | (0.056) | (0.068) | (0.027) | (0.095) | (0.154) | (0.078) |
| Q3 (Jul-Sep)× Drought($t$) | 0.024 | -0.005 | -0.023 | 0.060 | -0.040 | -0.006 | -0.036 | 0.063 | 0.073 |
| | (0.030) | (0.031) | (0.027) | (0.042) | (0.060) | (0.033) | (0.097) | (0.134) | (0.068) |
| Q4 (Oct-Dec)× Drought($t$) | 0.002 | -0.000 | -0.066* | 0.038 | -0.103* | 0.010 | -0.014 | -0.010 | 0.040 |
| | (0.038) | (0.031) | (0.039) | (0.052) | (0.055) | (0.038) | (0.084) | (0.110) | (0.078) |
| Q1 (Jan-Mar)× Drought($t-1$) | 0.109** | 0.103 | -0.068 | 0.124 | 0.065 | -0.015 | -0.024 | 0.069 | -0.030 |
| | (0.045) | (0.068) | (0.051) | (0.107) | (0.087) | (0.068) | (0.159) | (0.231) | (0.084) |
| Q2 (Apr-Jun)× Drought($t-1$) | -0.006 | -0.047 | 0.001 | -0.044 | 0.054 | -0.040 | 0.085 | 0.045 | 0.013 |
| | (0.058) | (0.037) | (0.049) | (0.076) | (0.148) | (0.038) | (0.137) | (0.142) | (0.073) |
| Q3 (Jul-Sep)× Drought($t-1$) | 0.051 | 0.010 | -0.023 | -0.074 | 0.101 | -0.029 | 0.263*** | -0.081 | 0.092 |
| | (0.047) | (0.037) | (0.049) | (0.076) | (0.152) | (0.033) | (0.096) | (0.131) | (0.099) |
| Q4 (Oct-Dec)× Drought($t-1$) | 0.068 | 0.066 | -0.032 | -0.027 | -0.085 | 0.022 | 0.129 | 0.081 | 0.060 |
| | (0.041) | (0.042) | (0.052) | (0.046) | (0.114) | (0.037) | (0.122) | (0.104) | (0.107) |
| Observations | 12,831 | 17,322 | 10,117 | 11,051 | 5,831 | 17,399 | 4,405 | 2,752 | 17,068 |
| R-squared | 0.253 | 0.526 | 0.235 | 0.216 | 0.337 | 0.430 | 0.226 | 0.389 | 0.285 |
| Mean of Dep. Var. | 5.340 | 4.379 | 4.708 | 4.200 | 4.585 | 3.748 | 3.612 | 2.995 | 4.155 |
| **Controls** | | | | | | | | | |
| Province and year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Region × Quarter FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Rainfall in other season-year | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Household characteristics | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| ln(Total Expenditure) | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

*Data*: VHLSS 2010, 2012, 2014
*Notes*: This table presents the coefficients on the interaction terms between quarters and droughts in $t$ and $t-1$ from the regressions estimating the effect on the log of monthly household expenditure on each non-food category. The regression columns are sorted by the share of expenditure on the item to the total expenditure from the largest to the smallest. 'Child' denotes the expenditure on allowance and books for children. Robust standard errors are shown in parentheses clustered at the province level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 1.B.6: Heterogeneous Effects of Droughts on Abortion

| | Dependent variable: Abortion=1 | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Drought | 0.0030*** | 0.0030*** | 0.0030* |
| | (0.0010) | (0.0010) | (0.0015) |
| Drought × Triple cropping | -0.0026** | | |
| | (0.0013) | | |
| Drought × Irrigation (2nd) | | -0.0032** | |
| | | (0.0016) | |
| Drought × Irrigation (3rd) | | 0.0008 | |
| | | (0.0018) | |
| Drought × Wealth (2nd) | | | -0.0010 |
| | | | (0.0019) |
| Drought × Wealth (3rd) | | | -0.0016 |
| | | | (0.0020) |
| Observations | 811,089 | 811,089 | 811,089 |
| R-squared | 0.012 | 0.012 | 0.012 |
| Mean of Dep. Var. | 0.0066 | 0.0066 | 0.0066 |
| **Controls** | | | |
| District and year FE | Yes | Yes | Yes |
| Rainfall in other season-year | Yes | Yes | Yes |
| Mother characteristics | Yes | Yes | Yes |
| District-specific linear time trend | Yes | Yes | Yes |
| Birth parity FE | Yes | Yes | Yes |
| Gender composition FE | Yes | Yes | Yes |

*Data*: PCS 2004-2008, 2010-2013
*Notes*: This table presents results from regressions of the dummy for abortion on the interaction terms between the drought in the dry season and the indicator for mothers residing in triple-cropping provinces (column (1)), and the tercile indicators for district-level irrigation coverage (column (2)), and for district-level wealth index (column (3)). Triple-cropping provinces are defined as if a province produces all three rice crops: spring, autumn and winter rice. Irrigation coverage is the area-weighted irrigation coverage found in the VHLSS 2004. The district-level wealth index is created by aggregating the household-level wealth index from the principal component analysis of 16 asset and residence characteristics found in the 2009 census. Robust standard errors, which are reported in parentheses, are clustered for the district level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 1.B.7: Effects of Alternative Rainfall Shocks on Yearly Rice Yields and Expenditure

| | Dependent variables | | | | | |
|---|---|---|---|---|---|---|
| | Spring Rice (1) | All Rice (2) | Total Expenditure (3) | Expenditure on Food (4) | Expenditure on Non-food (5) | Ratio (Food/Total) (6) |
| ***Panel A. Low rainfall shocks in the wet season*** | | | | | | |
| Low rainfall | 0.005 | 0.001 | -0.018 | -0.018 | 0.024 | -0.000 |
| | (0.007) | (0.006) | (0.016) | (0.015) | (0.021) | (0.003) |
| Observations | 1,045 | 1,055 | 18,128 | 18,128 | 18,128 | 18,128 |
| R-squared | 0.8010 | .900 | 0.531 | 0.610 | 0.530 | 0.119 |
| Mean of Dep. Var. | 3.922 | 3.788 | 9.861 | 9.060 | 7.505 | 0.471 |
| ***Panel B. Low rainfall shocks in the calendar year*** | | | | | | |
| Low rainfall | 0.005 | 0.003 | -0.018 | -0.013 | 0.016 | 0.002 |
| | (0.006) | (0.006) | (0.015) | (0.014) | (0.021) | (0.003) |
| Observations | 1,045 | 1,055 | 18,128 | 18,128 | 18,128 | 18,128 |
| R-squared | 0.800 | 0.898 | 0.530 | 0.609 | 0.529 | 0.119 |
| Mean of Dep. Var. | 3.921 | 3.787 | 9.874 | 9.070 | 7.516 | 0.469 |
| **Controls** | | | | | | |
| Province and year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Rainfall in other season-year | Yes | Yes | Yes | Yes | Yes | Yes |
| Province-specific linear time trend | Yes | Yes | | | | |
| Household Characteristics | | | Yes | Yes | Yes | Yes |
| Survey Quarter FE | | | Yes | Yes | Yes | Yes |

*Data*: Agricultural statistics from the Vietnam GSO and the VHLSS 2004, 2006, 2008
*Notes*: This table presents results from regressions of the log of annual crop yields (Quintal/Ha) and the log of expenditure (in ’000 VND) on low rainfall shocks in the wet season (Panel A) and in the calendar year (Panel B), respectively. Low rainfall shocks refer to the realization of rainfall in the wet season (April-November) or in the calendar year (January-December) below the 20th percentile of historical distribution of district-specific rainfall in 1984-2013. The sample excludes the 10 poorest provinces to be consistent with the analyses using the PCS. Household characteristics controls include the sex, age, ethnicity (Kinh or not) and years of schooling of the household head, the household size and the dummy for multigenerational households. Robust standard errors, which are reported in parentheses, are clustered for the province level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 1.B.8: Effects of Multiple Rainfall Shocks on Abortion

| | Dependent variable: Abortion=1 | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Drought$_{(t-1)}$ | 0.0021** | 0.0022*** | 0.0015* | 0.0022*** |
| | (0.0009) | (0.0008) | (0.0008) | (0.0008) |
| Drought$_{(t-1)}$ $\times$Low rainfall in the wet season $(t-1)$ | -0.0001 | | | |
| | (0.0012) | | | |
| Drought$_{(t-1)}$ $\times$High rainfall in the wet season $(t-1)$ | | -0.0008 | | |
| | | (0.0015) | | |
| Drought$_{(t-1)}$ $\times$Low rainfall in the wet season $(t)$ | | | 0.0031 | |
| | | | (0.0022) | |
| Drought$_{(t-1)}$ $\times$High rainfall in the wet season $(t)$ | | | | -0.0009 |
| | | | | (0.0018) |
| Mean of Dep. Var. | 0.0066 | 0.0066 | 0.0066 | 0.0066 |
| **Controls** | | | | |
| District and year FE | Yes | Yes | Yes | Yes |
| Rainfall in other season-year | Yes | Yes | Yes | Yes |
| Mother characteristics | Yes | Yes | Yes | Yes |
| District-specific linear time trend | Yes | Yes | Yes | Yes |
| Birth parity FE | Yes | Yes | Yes | Yes |
| Gender composition FE | Yes | Yes | Yes | Yes |

*Data*: PCS 2004-2008, 2010-2013
*Notes*: This table presents results from regressions of the indicator for abortion on various interaction terms between the drought in the dry season and a high or low level of wet season rainfall. 'Low' and 'high' level rainfall in the wet season refer to the realization of rainfall in the wet season (April-November) below the 20th percentile or in the 8th or 9th decile of historical distribution of district-season-level rainfall in 1984-2013. Robust standard errors, which are reported in parentheses, are clustered for the district level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 1.B.9: Robustness for Alternative Clustering of Standard Errors

| | Dependent variable: Abortion=1 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Drought (coefficient) | 0.0020 | 0.0021 | 0.0022 | 0.0020 | 0.0021 | 0.0020 | 0.0021 | 0.0023 |
| *p*-value (clustered by district) | 0.003 | 0.001 | 0.001 | 0.004 | 0.004 | 0.004 | 0.003 | 0.024 |
| No. of Clusters (district) | 502 | | | | | | | |
| *p*-value (clustered by province) | 0.018 | 0.007 | 0.006 | 0.017 | 0.017 | 0.017 | 0.015 | 0.056 |
| No. of Clusters (province) | 51 | | | | | | | |
| *p*-value (two-way by district & year) | 0.041 | 0.015 | 0.014 | 0.042 | 0.042 | 0.042 | 0.037 | 0.101 |
| **Controls** | | | | | | | | |
| District and year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Rainfall in other season-year | | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Mother characteristics | | | Yes | Yes | Yes | Yes | Yes | Yes |
| District-specific linear time trend | | | | Yes | Yes | Yes | Yes | Yes |
| Birth parity FE | | | | | Yes | Yes | Yes | Yes |
| Gender composition FE | | | | | | Yes | Yes | Yes |
| Fertility characteristics | | | | | | | Yes | Yes |
| Spouse characteristics | | | | | | | | Yes |

*Data*: PCS 2004-2008, 2010-2013
*Notes*: This table shows *p*-values from alternative clustering for the regressions reported in Table 1.5. The first *p*-values in the 2nd row are derived from the standard errors clustered by district. The second series of *p*-values in the 3th row are derived from the standard errors clustered by province. The last series of *p*-values in the 4th row are derived from the two-way clustering of district and year.

# Chapter 2

# Transboundary Air Pollution and Health: The Impact of PM Concentration on Fetal Deaths

with Anna Choi[1] and Semee Yoon[2]

[1]OECD, Email: Anna.Choi@oecd.org

[2]Underwood International College, Yonsei University, Email: yoonsemee@yonsei.ac.kr

## 2.1 Introduction

As countries experience economic growth, one of the key environmental challenges that they face is the maintenance of ambient air quality. Due to the increase in outdoor air pollution in the less developed Asian countries, the demand for rigorous evaluation of the health effects of air pollution for the region has increased (HEI International Scientific Oversight Committee, 2010). In the East Asian context, the dynamics of strong westerly winds in the middle latitudes provide a situation worthy of microeconomic analysis of health impacts due to the air pollution traveling from China to Korea. For example, daily fine particulate matter ($PM_{2.5}$)[3] concentration in Seoul abruptly increased to $86\mu g/m^3$ on February 25, 2014, right after the daily $PM_{2.5}$ level in Beijing reached 12 times the WHO safe limit (25 $\mu g/m^3$) from February 20 to 25 (Figure 2.1 and Figure 2.2). As Korea had the highest average exposure to $PM_{2.5}$ among all the OECD countries in 2013 (Figure 2.A.1), people in South Korea became aware of the potential adverse effect of exposure to such high levels of air pollution, which may have originated from China.

This situation evokes following research questions: (1) what is the effect of unprecedentedly high levels of ambient air pollution on the health of the Chinese population? and (2) does a measurable part of pollution, which South Koreans are exposed to, come from transboundary pollutants from China? If so, what is the impact of the pollution on the South Korean population? These questions are important but difficult to answer because of two potential sources of endogeneity: residential sorting and avoidance behaviors (Currie et al., 2014). Since housing prices reflect the environmental qualities of the neighborhood (Chay and Greenstone, 2005; Currie et al., 2015), the public's response to pollution levels is hardly assigned randomly (Banzhaf and Walsh, 2008). This non-random assignment of pollution

[3]Particulate matter (PM) is one of major hazardous air pollutants, consisting of a mixture of solid and liquid particles. PM can be categorized by the particle size: $PM_{10}$, which has a diameter between 2.5 and 10 $\mu$m, and the 'fine' particulate matter $PM_{2.5}$, with a diameter of less than 2.5 $\mu$m. $PM_{2.5}$ is produced mainly by the combustion of fossil fuel. $PM_{2.5}$ can remain in the atmosphere for days or weeks and thus be subject to long-range transboundary transport in the air (National Research Council, ed, 2010).

can confound the causal estimation of the effect of pollution on the welfare of residents because unobservable characteristics that covary with the level of exposure to pollution may also determine the health of the residents (Graff Zivin and Neidell, 2013). Particularly in less developed countries, the relationship between pollution and health is not straightforward to quantify due to the lack of accurate pollution and welfare measures (Graff Zivin and Neidell, 2013).[4] Moreover, it can also be challenging to credibly disentangle the contribution of transboundary pollution from locally emitted pollution.[5]

To circumvent the aforementioned empirical challenges, we link daily variations in $PM_{2.5}$ concentration measured in Beijing, China with daily fetal mortality rates in South Korea. Although correlations between ambient air pollution and health within a given location might not imply the causal link because of the non-random assignment of pollution (Chay and Greenstone, 2003), our study design can purge such type of confounders by using arguably exogenous variations derived from the 'outside' of a country. To be specific, we take advantage of the daily variations in particulate matter concentrations in Beijing, which are not only substantial enough to affect pollution levels and the health of the exposed population but also arguably orthogonal to the daily determinants of health of the residents in neighboring countries. That is, given that a sudden increase in "daily" pollution levels in South Korea can be driven by the transboundary transport of pollutants from China, we assume that day-to-day changes in Chinese pollution can affect the health of the South Korean population, whereas socio-economic characteristics of the population at risk or local economic activities are not affected in the same manner in such a short time period. Moreover, factors detrimental to the health of fetus are unlikely to change on a day-to-day basis, except for sudden environmental

[4]In addition to the difficulties of obtaining health data for China, possible bunching at the threshold has made researchers question the credibility of the pollution data even though hundreds of monitors were installed across China and some of the readings have been released to public (Chen et al., 2012; Ghanem and Zhang, 2014).

[5]Complete models to attribute parts of pollution to remote sources have yet been specified in the field of environmental science, because numerous precursor pollutants and their interactions under certain atmospheric conditions substantially affect the level of local pollution through an overwhelmingly complex mechanism (EPA, 2009).

changes and inexplicable anomalies (Rasmussen et al., 2003).

In this paper, we analyze the impact of pollution on the fetal health using rich micro-data on fetal deaths. Fetal mortality can be a better metric compared to infant mortality because the exposure to pollution can be more precisely identified by focusing on the "indirect" exposure through the mother, whereas shedding light on the effect of pollution on infants further requires researchers to parse out the direct exposure of newborns to pollution. Thus, in addition to a pregnant woman's very high level of avoidance behaviors, such as refraining from migrations during pregnancy, the indirect effect of PM pollution on a fetus will allow us to have more conservative estimates of health cost. Furthermore, in the context of more developed countries, fetal mortality can be a more relevant measure for policymakers provided that infant mortality rates are very low and low fertility rates are one of the major pressing issues (Woods, 2008).

For the first-stage analyses, we find one standard deviation increase in Beijing's daily $PM_{2.5}$ is correlated with an increase in the daily levels of six common pollution ($PM_{2.5}$, $PM_{10}$, CO, $NO_2$, $SO_2$, $O_3$) by 7.7% of a standard deviation in South Korea. In case of $PM_{2.5}$, one standard deviation of Beijing's $PM_{2.5}$ can explain about 16% of the standard deviation of daily $PM_{2.5}$ level in Seoul. Then, we find that one standard deviation increase in Beijing's $PM_{2.5}$ during the previous day $(t-1)$ statistically significantly explains 1.1% of the standard deviation of daily fetal mortality rates at 16 weeks or more of gestational age across cities in South Korea. We find statistically significant results on the fetal deaths at 16 weeks and 20 weeks or more gestational age, but not at 28 weeks or more. This result is mainly driven by the fetal deaths that occurred in the cities located close to Beijing, the source region. And the magnitude of estimates on Beijing's $PM_{2.5}$ is largest in one day lag $(t-1)$ and attenuates when moving from two to six-day lags $(t-2)$ to $(t-6)$. Moreover, statistically insignificant estimates on forwarded coefficients on Beijing's $PM_{2.5}$ on the fetal deaths in South Korea indicate that the estimates are primarily driven by the transboundary pollution from China.

Among several new attempts made in this study, one of the notable contributions to existing literature will be the direct investigation of negative spillovers between countries manifested on the health of fetuses. Almond et al. (2009a) and Jayachandran (2009) exploit unusual events in the source regions, Chernobyl and Indonesia to investigate the adverse health effects caused by the long-range transport of hazardous matters. In this study, we focus on how usual economic activities can exert negative externalities on the health of the people in neighboring countries. In particular, our analysis of the micro-level health census data from South Korea can provide lower bound estimates for the health cost of the Chinese population. Compared to Jia and Ku (2018) who also assessed the impact of Chinese pollution on district-month mortality rates in South Korea, we attempt to examine the detailed mechanisms of adverse health effects using a narrower window for mortality to react to daily pollution, providing another informative evidence on the adverse impact of air pollution spillover.[6]

Secondly, our study contributes to the literature studying the impact of pollution on fetuses in various development stages. By taking advantage of fetal deaths at different gestational ages, we can precisely match the length of fetal exposure to the level of ambient air pollution. Moreover, we also provide rather rare evidence on the effects of pollution on fetal deaths, whereas a vast majority of the existing literature in both more and less developed countries focus on infant mortality. Thus, our findings can be complementary to the previous studies suggesting that *in utero* exposure to ambient air pollution can be of more critical importance than postnatal exposure in determining not only infant health (Chay and Greenstone, 2003; Sanders and Stoecker, 2015) but also labor market outcomes in the long run (Isen et al., 2017; Lavaine and Neidell, 2017).[7]

[6]Jia and Ku (2018) use the incidence of Asian dust recorded in 28 stations across Korea from 2000 to 2011 as the main source of variation in pollution and examine the effects on respiratory and cardiovascular mortality rates. The Asian dust (yellow dust) originates from the deserts of Northern China, Mongolia and Kazakhstan and can transport different pollutants from China to Korea by the westerly winds.

[7]Isen et al. (2017) examined the long term effects of air pollution exposure *in utero* and found that the cohorts who were exposed to higher pollution levels during the year of their birth had lower labor force par-

Last but not least, this study contributes to a burgeoning strand of the literature using a variety of noble instruments for ambient air pollution to address the endogeneity problem from residential sorting. For example, there are abrupt changes in the TSP levels in counties after the enforcement of the Clean Air Act in the U.S. (Chay and Greenstone, 2003 and Isen et al., 2017), the spatial discontinuity of $PM_{10}$ due to the Huai River heating policy in China (Almond et al., 2009b; Ito and Zhang, 2016). In addition to legislative changes, some have used phenomena in the transportation sector, such as the opening of a subway network across the world (Gendron-Carrier et al., 2018), and the flight taxi time and wind patterns (Schlenker and Walker, 2016). Others have exploited unexpected social changes, including the temporary pollution reduction due to oil-refinery strikes in France (Lavaine and Neidell, 2017). As Deryugina et al. (2016) use wind-driven daily variations in fine particulate matter to causally estimate the effect on health, if we can find significant associations between Beijing's PM levels and the daily fetal mortality in South Korea, our estimates can provide evidence for the causal link between the two.

The rest of the paper is organized as follows. Section 2.2 discusses the association between air pollution and fetal death and the background knowledge of transboundary transport of pollution in East Asia. Section 2.3 describes the data. Section 2.4 and 2.5 provide our empirical methodologies and the results. We conclude in Section 2.6.

## 2.2 Background

### 2.2.1 Air Pollution and Fetal Death

The adverse health effects of exposures to particulate matter have been reported in the extensive body of epidemiological and toxicological studies. In particular, respiratory and cardio-

---

ticipation and lower earnings at 30 than those cohorts who did not have such exposure *in utero*. Lavaine and Neidell (2017) find a decrease in sulfur dioxide levels after the temporary strike in oil refineries is associated with higher birth weight and gestational ages of newborns and the effect is stronger for those exposed to the shock during the first and third trimesters of pregnancy.

vascular systems are known to be the most sensitive to the ambient pollution (Dockery et al., 1993; Pope et al., 1995; Samet et al., 2000; Dominici et al., 2006) and a handful of studies document some of its pathophysiologic mechanisms(Brook, 2002, Kim et al., 2012; Hoek et al., 2013). Both $PM_{10}$ and $PM_{2.5}$ are hazardous if exposed for a short-term period, but the health effects of long-term exposure to $PM_{2.5}$ is particularly alarming because it can remain airborne for a longer period of time and easily enter the bloodstream to cause cardiovascular and respiratory diseases (Pope et al., 2002; Hoek et al., 2013).

Numerous studies have tried to investigate the relationship between ambient pollution and negative postnatal outcomes, such as infant mortality, low birth weight and/or preterm birth (Ritz and Yu, 1999, Chay and Greenstone, 2003, Currie and Neidell, 2005, Ritz et al., 2007, Currie et al., 2009, Currie and Walker, 2011, Knittel et al., 2011 , Ha et al., 2014). However, although there is accumulating evidence on fetal deaths using fetal death data (Faiz et al., 2012; Faiz et al., 2013; DeFranco et al., 2015) or the sex ratio of live birth (Sanders and Stoecker, 2015), studies to elucidate mechanisms for the particulate matter to affect fetuses are still scarce, except for some suggestive evidence on the biological pathways (Pereira et al., 1998; Erickson and Arbour, 2014). Recently, there is a new emerging literature that suggest pollutants, including particulate matter, could be causing “sudden intrauterine unexplained death syndrome (SIUDS),” by two mechanisms: pollutants causing hypoxic status, which leads to impairments of the central nervous system of the fetus, as found with cigarette smoke, and endocrine disruptor chemicals that cause developmental alterations, as found with pesticides. Roncati et al. (2016) suggest although these studies have not directly looked at the impact of polluted air, particulate matter could also be having a similar influence on SIUDS.

### 2.2.2 Transboundary Transport of Pollutants

The background of the transboundary transport of pollutants from China to South Korea can be explained based on two distinctive features: characteristics of pollutants and meteorological conditions in the region. First, because of its diminutive size, $PM_{2.5}$ can remain in the atmosphere for days or weeks and thus be subject to long-range transboundary transport (National Research Council, ed, 2010). In addition to the physical characteristics of $PM_{2.5}$, very high levels of $PM_{2.5}$ concentration in the source region deserve attention because China is the world's largest emitter of air pollutants. Measurable amounts of pollutants are produced and are subject to be transported via the consistent air flows heading to Korea and even to the United States (Lin et al., 2014). Not only are the transported primary particles affecting the PM concentration in South Korea, but also precursor particles, such as sulfate or nitrate, are also transported and potentially form secondary particulate matter or gaseous pollutant when chemicals from remote and local sources react in the atmosphere (EPA, 2009).

Secondly, seasonal wind patterns in the East Asian region influences the movement of pollutants from China to the eastern region. In conjunction with the Westerlies as the prevailing wind pattern in the mid-latitudes, seasonal differences in the location of high and low-pressure systems can either strengthen or weaken such wind pattern. For example, high-pressure systems are located over northwest China and lows are in the Pacific Ocean, which creates the strongest wind of all seasonal winds to blow from China to Korea and Japan during the winter season. The direction rarely changes unless local geography changes the prevailing direction (Figure 2.A.2). On the contrary, the wind speed attenuates and the direction changes as temperature rises because the highs over the southern Pacific become stronger. This seasonal wind pattern allows the pollution emitted from the heavily industrialized eastern China to be blown to Korea through the strong wind flows except for the summer. It takes one day on average for the pollutants from Beijing to arrive in cities in South Korea (Lee et al., 2011).[8]

## 2.3 Data

### 2.3.1 Pollution in South Korea

We collect the hourly readings of five major pollutants ($PM_{10}$, CO, $NO_2$, $SO_2$, $O_3$) from the Air Korea,[9] the information center for ambient air pollution in South Korea. It releases real-time readings of pollution measured by the monitors located across South Korea. Since each census unit (district) in our study has its own pollution monitors, we do not need to go through remedial procedures to calculate the level of potential exposure of the residents to pollution. This richness in spatial resolution minimizes a potential measurement error in this study because the estimation of exposure to pollution is not robust to the interpolation techniques used in analyses (Lleras-Muney, 2010).

Of 256 monitors as of 2013 across cities in South Korea, we primarily focus on 140 monitors in the cities with a high population density ($\geq$3,000 people/km$^2$), which corresponds to 140 districts in 21 cities. Since unobservable characteristics of the location in which pregnant women live can significantly affect the likelihood for the women to have negative pregnancy outcomes[10], we exclude those cities having low population density from our sample because they presumably have different characteristics than those with high population densities. Locations of the district-level pollution monitors are marked in Figure 2.A.4. In our analysis,

[8]The identification of transboundary source and the air flows can be more clearly illustrated by using the wind back trajectory analysis method shown in Figure 2.A.3 (Lee et al., 2011; Lee et al., 2013). The eastern region of China can be identified as the source region for the events in November and January, but for July, the high $PM_{2.5}$ is very likely to be attributable to local sources in 2013. Using this back trajectory analysis, among 254 high $PM_{10}$ episodes ($\geq 100\mu$g/m$^3$ for 24h mean) in South Korea in 2001-2008, 178 events could be identified as events influenced by an external source of pollution (Lee et al., 2011).

[9]http://www.airkorea.or.kr

[10]For example, if we assume that there is negative impact on a fetus after a mother's exposure to the high level of ambient air pollution, the distance between a hospital and her residence might critically affect the likelihood of having stillbirth.

each district in a city has at least one monitor, and it is mostly located in the residential area.

The daily mean of each pollutant is the duration-weighted average of hourly readings (Schlenker and Walker, 2016). Since PM measurements from the monitors in China and South Korea have a relatively high frequency of missing values compared to the other pollutants and weather covariates, we keep the observations only if the monitor has at least 5 readings per day. The summary statistics of daily pollution levels are provided in Table 2.1.

To study the heterogeneous effects of transboundary pollution, we group cities into two regions depending on the distance from China. The local pollution levels in cities in South Korea can be different depending on a variety of characteristics, such as geography, topography, and the road network. In our settings, the distance from China can be one of the key explanatory variables determining the intensity of the city-level exposure to the transboundary air pollution. Region 1 includes Seoul and its surrounding 13 metropolitan cities that closely interact with Seoul. This region also has the highest population density in the country and is the closest to Beijing (950 *km*). Region 2 includes the remaining 7 cities approximately 1030-1200 *km* away from Beijing. Although the cities in Region 1 is twice the number of cities in Region 2, the categorization of the cities into the two regions solely depends on the distance from Beijing. One benefit of this categorization is that we can exploit the differential response to the Chinese pollution depending on the distance from Beijing. If it were not for the transboundary pollution, the effects of Beijing's air pollution on health in South Korea will not differ by the distance from China.

## 2.3.2 Pollution in China

We assume China's pollution as one of the major sources of pollution in South Korea, so we are mainly interested in how much pollutants are accumulated in the source region[11]. By

[11] Although the local weather conditions are also important to lift pollution from the ground level to up in the atmosphere to be in the air flows heading toward South Korea, we are using the level of pollution in this analysis to simplify the empirical specification.

the nature of the Westerlies, we focus on the concentration of pollution in the Northeastern region in China, which is highly industrialized and heavily populated. Chinese government operates hundreds of pollution monitors and releases some of the current readings to the public. However, due to access restrictions to historical readings of $PM_{2.5}$ and the possibility of non-classical measurement error caused by under-reporting and bunching at thresholds (Chen et al. (2012), Ghanem and Zhang (2014)), we use the $PM_{2.5}$ readings collected by the monitors located in the five U.S. missions to China, as shown in Figure 2.A.5. There are two U.S. monitors, one in Beijing and one in Shenyang in Northeast China, which are presumed to be the most influential source region. However, since we are covering the period of 2009-2013, we choose $PM_{2.5}$ readings from the monitor located at the U.S. embassy in Beijing which fully covers our study period[12]. We assume that the readings from the Beijing monitor can be a proxy for the average level of pollution over the Northeastern China where most transboundary air pollution is originated. If the readings of the Beijing monitor systemically underestimates or overestimates the true level of $PM_{2.5}$ concentration over the source region, then our estimates are biased accordingly. PM readings in Beijing have a high frequency of missing values, so we only choose the dates with at least more than five hourly readings per day. To compute daily averages, we also use duration-weighted averages. The level of $PM_{2.5}$ pollution in Beijing is so high that the daily average $PM_{2.5}$ levels are rarely below the WHO guideline for daily exposure to $PM_{2.5}$, which is $25\mu g/m^3$ per day (Panel (a) in Figure 2.3). The summary statistics of the daily level of $PM_{2.5}$ pollution is presented in Table 2.1.

A particular period to pay attention in terms of Chinese pollution is between November and March, which is publicly administered as the heating season by the Chinese government in northern Chinese cities (Almond et al., 2009b). Since the public heating is primarily powered by the inefficient combustion of coal, the PM pollution may increase discontinuously

[12]We are not including Shenyang station because the readings started from 2011. We checked how much the pollution in the other three cities in China (Shanghai, Chendu and Guangzhou) is correlated with the level of particulate matter in South Korea and found that it is not significant or significant but not strong as pollution in Beijing.

during the heating season (Xiao et al., 2015). However, the pollution level is very high, irrespective of seasons along with high variability in day-to-day concentrations (Panel (a) in Figure 2.3 and Panel (a) in Figure 2.A.6).

### 2.3.3 Weather in South Korea

The weather data in South Korea is collected at the stations administered by the Korea Meteorological Administration (KMA). Each station reports hourly observations of temperature, wet-bulb temperature, wind direction and speed, precipitation, pressure, absolute and relative humidity and sun hour. Since local weather conditions critically affect the formation of local pollutants (EPA, 2009) and can directly influence health (Deschenes et al., 2009; Deschenes and Moretti, 2009), it is important to include flexible controls of all the observable weather variables. Seven metropolitan areas and four cities use the weather observations measured at their own weather stations. For cities that do not have a weather station within the city boundary, observations from a weather station to the nearest cities are assigned. In total, hourly observations from 11 weather stations are matched to hourly readings of pollution in 21 cities. The daily value of weather observations is calculated using the duration-weighted average, but in most cases, there are very few missing values. The summary statistics for the city level weather data are given in Table 2.1.

### 2.3.4 Fetal Mortality

We use fetus-level data matched by the mother information in 2009-2013. The Statistics Korea has provided the fetal death data since 2009 by merging information from hospitals and the cremation certificates. All the fetal deaths at the gestation age after 16 weeks are reported with the cause of death, sex, and mothers' characteristics, such as age, weight, pregnancy history, smoking, marital status, and education. The outcome of our interest is the city-level daily fetal mortality rate, which is the city-day level counts of the fetal deaths weighted by the monthly live births in a city and the number of fetal death counts. We also use daily prospective fetal mortality rate which is the number of fetal deaths at a given gestational age per 1,000 live births and fetal deaths at that gestational age or greater on a given day.[13] This can represent the population at risk of the event more precisely, but it requires the birth certificate data with the exact birthday and pregnant weeks. To have comparable results, our main outcome uses the fetal deaths after 20 weeks of gestational age as defined by the CDC (MacDorman and Kirmeyer, 2009). The city-day level fetal mortality rates of 16 and 28 weeks of gestational age and the perinatal mortality rates which additionally use infant deaths within 7 and 28 days after births in addition to the fetal deaths at 28 weeks or more of gestational age are also computed for each day to reveal heterogeneous effects of pollution on fetuses and newborns (MacDorman and Kirmeyer, 2009).[14]

Since none of the fetal losses is provided before 16 weeks of gestational age, the estimates of the true effect of pollution on fetuses may be biased due to truncated samples. However, even though the majority of fetal losses happens within the first trimester, the causes of death are largely left to be unknown or attributable to underlying maternal medical conditions in the early stage of pregnancy (Rasmussen et al., 2003). Thus, confining the sample to be in the later stages of pregnancy can be more relevant to the studies investigating the effect of pollution as one of the external stressors on fetal health.

The summary statistics regarding fetal deaths are presented in Table 2.1. There is no significant difference in the characteristics of the mother depending on the regions of her residence, indicating there is no potential bias driven by omitting mothers' covariates in the specification using aggregated fetal mortality rates at a city-day level. In addition, the plot of the daily fetal mortality rates by day indicates that there is little evidence on seasonal trends in fetal deaths (Figure 2.A.7).

[13]In the regressions, we multiply fetal mortality rates by 1,000 to have readable coefficients.

[14]We are less concerned about potential selection bias or measurement error in the fetal deaths after 16 or more weeks of gestation given that South Korea's public health insurance requires almost every pregnant woman to be registered in the public health systems for prenatal care.

## 2.4 Empirical Methodology

The purpose of this study is to provide the quantifiable estimation of the adverse effect of the Chinese $PM_{2.5}$ pollution on the fetal health in South Korea. First, we investigate how much of the daily particulate matter pollution in South Korea can be explained by the Chinese $PM_{2.5}$ and how much of the daily fetal mortality can be explained by the day-to-day variations of local pollution in South Korea.

### 2.4.1 Chinese PM Level and Local Pollution Levels

To provide a basis for the relationship between Chinese pollution and South Korean pollution ('first stage'), the intensity of pollution in the source region (Beijing) is analyzed in conjunction with a vector of atmospheric variables in the receptor location under the assumption that they are interacting in an additive and linear fashion (Chay and Greenstone, 2003). We estimate the association using the following equation:

$$Pollution_{ct} = \alpha_0 + \alpha_1 PM_{2.5t} + \mathbf{Weather}_{ct}\Pi + \mathbf{Z}_t\Lambda + \mathbf{Trend}_{ct}\Psi + \nu_c + \varepsilon_{dt} \tag{2.1}$$

where $Pollution_{ct}$, one of five pollutants ($PM_{10}$, CO, $SO_2$, $NO_2$ and $O_3$) in city $c$ in South Korea on day $t$, is a function of a lead or a lag up to 6 days of daily mean $PM_{2.5}$ level in Beijing conditional on weather, temporal controls and time trends. Since the estimates of the impact of pollution on health can be substantially influenced by the inclusion of higher order terms of temperature, precipitation and second order terms of the other weather variables (Knittel et al., 2011), we include the weather controls in flexible polynomials and cross-terms interacted with year, month, and day of the week dummies.[15] Temporal controls $\mathbf{Z}_t$

[15] We followed the approach of Auffhammer and Kellogg (2011) and Schlenker and Walker (2016); a vector of $\mathbf{Weather}_{ct}$ include cubic polynomials in minimum and maximum temperature, a quadratic in precipitation, cross-terms between lagged max and min temperatures with rain, wind speed and direction, humidity, sea surface pressure and sun hours. In addition, all the weather variables are interacted with day of year dummies, and max and min temp, precipitation as well as wind speed and directions are interacted with day of week dummies.

include year, month and day of week dummies to indirectly control for pollution-generating activities in a city. In addition, we include city-specific linear and quadratic time trends in the main specification. Since not only infrastructure and economic characteristics of the city but also the geography and topography of a city largely determine the daily pollution level (EPA, 2009), the city level fixed effects are included to control for the time-invariant unobserved determinants of pollution. To purge the concern of serial correlation of variables depending on the public heating season in China each year, robust-standard errors are clustered on the region-year-season level.

The coefficient of our primary interest is $\alpha_1$, which quantifies the contribution of Beijing's $PM_{2.5}$ from six-day lags $(t-6)$ to six-day leads $(t+6)$ on the level of five pollutants in a city $c$ on day $t$ in South Korea. In particular, the coefficients $\alpha$ on Beijing's $PM_{2.5}$ from $t+1$ to $t+6$ confirm whether the transboundary transport of pollution occurs in one direction from China to South Korea, not the other way around from South Korea to Beijing.

### 2.4.2 Daily PM Level in Beijing and Fetal Mortality in South Korea

To estimate the link between daily average $PM_{2.5}$ levels in Beijing and daily fetal mortality rates in South Korea ('reduced-form'), we follow the same approach used in Equation (2.1).

$$Fetrate_{ct} = \beta_0 + \beta_1 PM_{2.5t} + \mathbf{Weather}_{ct}\Gamma + \mathbf{Z}_t\Theta + \mathbf{Trend}_{ct}\Omega + v_c + e_{ct} \tag{2.2}$$

where $Fetrate_{dt}$ is the daily fetal mortality rate in a city $c$ in South Korea on day $t$. We focus on the fetal death after 20 weeks of gestational age as the baseline, but we also use diverse measures of fetal mortality to examine heterogeneous effects of pollution on different stages of fetal development. $PM_{2.5t}$ enters in a lead or lag up to 6 day. The parameter of our primary interest is $\beta_1$, which captures the effect of exposure to transboundary particulate on fetal death. The consistent estimation of $\beta$ requires $\mathbb{E}[PM_t \cdot e_{ct} | \mathbf{Weather}_{ct}, \mathbf{Z}_t, \mathbf{Trend}_{ct}] = 0$. We also use local $PM_{10}$ level in a city $c$ instead of $PM_{2.5t}$ in Beijing to understand the impact of local pollution on fetal health. We run a weighted regression using the count of newborns in each city.

## 2.5 Results

### 2.5.1 Effects of Daily Beijing's $PM_{2.5}$ on Daily Local Pollution Levels in South Korea

We start by looking at how much pollution in Beijing can explain the daily level of local pollution in cities in South Korea on average. Table 2.2 presents the first-stage estimates using Equation (2.1), explaining how much yesterday$(t-1)$'s $PM_{2.5}$ level in Beijing can be associated with today$(t)$'s levels of $PM_{10}$ in cities in South Korea. Although $PM_{2.5}$ can be a more suitable pollutant for the transboundary transport, we focus on $PM_{10}$ as the baseline estimates due to the data availability. The parameter of our primary interest is shown in Column (5), Panel A of Table 2.2, which shows that one standard deviation increase in Beijing's $PM_{2.5}$ level in $t-1$ can explain 3% of the standard deviation of daily $PM_{10}$ levels in $t$ in 21 cities in South Korea. This is driven primarily by the effect found in Region 1 (Panel B); it explains 5% of one standard deviation of daily $PM_{10}$ in $t$ in Region 1. However, the coefficient in the other region is not significant at any conventional levels of significance. If we exclude the observations in June to August during the summer, one standard deviation of Beijing's $PM_{2.5}$ in $t-1$ can explain 6% for all regions and 8% for Region 1 of a standard deviation of daily $PM_{10}$ in $t$, respectively.

Since $PM_{2.5}$ can be a more relevant pollutant to be subject to transboundary transport compared to $PM_{10}$ due to its minuscule size, we estimate the effect of Beijing's $PM_{2.5}$ in $t-1$ on Seoul's $PM_{2.5}$ in $t$ using the Seoul $PM_{2.5}$ readings from 2010-2013. We find that one standard deviation increase in Beijing's $PM_{2.5}$ in $t-1$ leads to about $2.22\mu g/m^3$ increase in the level of today's $PM_{2.5}$ in Seoul, which is about 16% increase of a standard deviation of Seoul's daily $PM_{2.5}$ level in $t$. This means that the more minuscule the particle size is, the farther it is likely to be transported to reach neighboring countries. Consequently, if we can find any adverse effect of transboundary pollution, it can be very likely to be attributable to the particulate pollution as fine as $PM_{2.5}$. Our first-stage resutls are also in line with the findings from scientific studies that report $PM_{10}$ is more relevant to local sources such as emissions from vehicles rather than external source from which long-range pollutants are originated (EPA (2009), National Research Council, ed (2010)).

Next, the coefficients on 6 day lags and leads of Beijing's $PM_{2.5}$ level under the specification in Equation (2.1) are plotted to find the exact timing of arrival of transboundary pollutants from Beijing onto cities in South Korea and to find the key time frame to capture the contemporaneous effect of pollution on fetal health. In addition, the coefficients on the leads from $t+1$ to $t+6$ provide information on whether the transboundary pollution works in one way, not in the other way around. In Figure 2.4, Panel (a), (b) and (c) show the coefficients plots on local $PM_{10}$ by regions. In general, for $PM_{10}$, Beijing's $PM_{2.5}$ in $t-1$ is the most influential on $PM_{10}$ level in $t$ in South Korea. On the regional level, however, the coefficient $t-2$ is the highest for Region 2, while the coefficient on Beijing's $PM_{2.5}$ level in $t-1$ has the largest magnitude and the narrowest confidence interval for the cities in Region 1. These results are in line with our knowledge in that the farther the cities are located away from the source region (Beijing), the longer it takes for the pollutants to affect the level of pollution in those cities. The magnitudes of the coefficients in Region 2 are also smaller compared to those of Region 1, which indicate that the mixing process and deposition happen as the pollution moves to eastward.

Finally, we check how the daily variation in $PM_{2.5}$ level in Beijing affect the other gaseous pollution levels (CO, $SO_2$, $NO_2$ and $O_3$) in South Korea. The other major pollutants can provide precursor particles such as sulfate or nitrate for the secondary pollution in the electrochemical reactions under certain local atmospheric conditions. Figure 2.4 presents that all the other pollutants are reacting to the Beijing's $PM_{2.5}$ level as expected; Beijing's $PM_{2.5}$ in $t-1$ is the most influential factor to explain the level of pollution conditional on weather and temporal covariates. One standard deviation of Beijing's $PM_{2.5}$ in $t-1$ explains 10.3%, 7.8%, 5.7% and 3.5% of a standard deviation of $NO_2$, CO, and $SO_2$ respectively[16].

### 2.5.2 Effects of Beijing's Daily $PM_{2.5}$ Level on Daily Fetal Mortality in South Korea

We begin by investigating the relationship between Beijing's $PM_{2.5}$ in $t-1$ and the daily fetal mortality rates across 21 cities in South Korea. Table 2.3 presents the regression estimates of Equation (2.2), showing the effect of yesterday's Beijing's $PM_{2.5}$ on the number of fetal deaths ($\geq$20 gestational weeks) per 1 million live births on the next day in South Korea. The results of our main specification are provided in Panel A, Column (5), displaying one standard deviation of Beijing's $PM_{2.5}$ level in $t-1$ can explain 0.8% of the day-to-day standard deviation of daily fetal mortality rate on average across all regions. To be specific, a $10\mu g/m^3$ increase in mean $PM_{2.5}$ levels in Beijing may lead to 0.001085 more daily death of fetuses at more than 20 weeks of gestational age in South Korea.[17]

Since the Korea Statistics department also collects data on fetal death that occurred more than 16 gestational weeks, we further investigate the contemporaneous response of fetuses to one-day lagged Beijing's $PM_{2.5}$ levels using diverse metrics of the daily fetal mortality rates with 16 gestational weeks and 28 gestational weeks. The mean of the fetal death of 16 gestational weeks or older is greater than that of 20 weeks or older of gestational age, but the former is more likely to be due to inexplicable causes of death that happen during

[16]Since the $O_3$ is the result of photochemical reaction which is active during summer, it shows the negative correlation with $PM_{2.5}$ which has high concentration during winter because of heating and lower mixing layer (EPA, 2009).

[17]Considering that the average daily fetal death is 0.55 per 1,000 live births, and this mortality rate is weighted by monthly live birth in each city, we will have more sensible magnitude of the estimates by using daily live births in a city.

the early stage of pregnancy (Rasmussen et al., 2003). Fetal deaths that occur at 28 weeks of gestational age or older is helpful to analyze the effect of transboundary pollution on the fetus in the third trimester. We also show the effects of Beijing's particulate matter pollution on the prospective fetal mortality rates, which takes at-risk fetuses into consideration using all of the women who are still pregnant at that gestational age as denominators. The estimates for the daily fetal deaths with 16 weeks and 28 weeks of gestational ages are provided in Table 2.B.1 and Table 2.B.2, respectively. We have statistically significant results only for 16 weeks. The magnitude of the coefficient is larger, which means that one standard deviation increase in Beijing's $PM_{2.5}$ in $t-1$ can be associated with 1.1% of a standard deviation increase in daily fetal mortality rates of more than 16 weeks of gestational age across cities in South Korea.

To investigate the source of pollution, we examine heterogeneous effects of Beijing's $PM_{2.5}$ depending on the distance from Beijing by splitting the sample by region. In Column (5) of Panel B and C in Table 2.3, Table 2.B.1, and Table 2.B.2, we find that adverse effects of pollution on fetuses are mainly driven by the fetal deaths occurred in those cities in Region 1 for all the three definitions of fetal mortality rate. Furthermore, instead of splitting the sample, Table 2.B.3 reports the results from estimating Equation (2.2) after augmenting it with the interaction term of Beijing's $PM_{2.5}$ level with the indicator for cities in Region 2 in Panel A1 and B1, and with each city's distance from Beijing (in km) in Panel A2 and B2. We find that most coefficients on the interaction terms are not precisely estimated, but the negative signs with statistical significance in Panel A2, Column (6) and Panel B2, Column (6) are noteworthy. That is, those cities in Region 1, which are located closer to Beijing, witnessed higher fetal mortality rates in response to the Beijing's $PM_{2.5}$ level in $t-1$, compared to the cities in Region 2 located farther from Eastern China. Since those cities are chosen for their comparable high population densities, and the fetal deaths are aggregated at the city level to calculate fetal mortality rates, the results are robust to the mother's individual characteristics. Furthermore, since we use day-to-day variations, it is unlikely for the cities to respond to the transboundary pollution differently depending on the location. Thus, the larger effects on fetal deaths in Region 1 are attributable to the close distance to the source location of transboundary pollution even after controlling for time-invariant characteristics of a city.

One way to strengthen our causal estimation of transboundary pollution on fetal death is to provide estimates using a series of time lags and leads Beijing's $PM_{2.5}$ value. The estimates of 6-day forwarded and lagged values of Beijing's $PM_{2.5}$ on various measures of daily fetal mortality rates by region are provided in Figure 2.5. First of all, in Panel (a) and (b), the largest magnitude can be observed for the coefficients on $PM_{2.5}$ in $t-1$ and those effects are primarily driven by the estimates of Region 1. The size of coefficient decreases in magnitude with the time lag moving backward from $t-1$ to $t-6$, suggesting the previous finding that yesterday's level of $PM_{2.5}$ in Beijing is the most important on today's pollution level in South Korea. In addition, most of the coefficients on forwarded values of Beijing's $PM_{2.5}$ are not significantly different from zero, indicating previous pollution in Beijing can adversely affect fetuses in South Korea, but not the other way around. Similar patterns can be also observed when using prospective fetal mortality rates as the outcome in Figure 2.6, meaning the results are robust to the changes in the denominator that better reflects the fetuses at risk. Overall, the results suggest the possibility of a negative spillover effect of pollution from one country to the other, and these are causal estimates which are derived by wind and meteorological patterns exogenous to local or mother-level determinants of fetal deaths.

In addition, we split the sample and run the same regressions to examine whether fetal mortality in South Korea reacts differently to the public heating season in northern China. In Figure 2.A.9 and Figure 2.A.10, we can find that most effects of pollution on fetal mortality rates can be explained by those found in the non-heating season. This result can be somewhat surprising because mean pollution levels are higher during the heating season in Korea as shown in Panel (b) and (c) in Figure 2.A.6. However, the pollution levels are not necessarily higher during the heating season in China (Panel (a) in Figure 2.A.6). Furthermore, avoidance behaviors of pregnant women can be different depending on the ambient temperature and the public awareness of the mean level of pollution by seasons.

Lastly, it is important to note that even though we use the fine particulate matter pollution in Beijing as a proxy for the transboundary pollution, we do not claim that these results show the adverse impact of the 'particulate matter' on fetuses. As shown in our first stage results as well as evidence from by a handful of scientific studies, Beijing's $PM_{2.5}$ can also contribute to a measurable amount of precursor elements for the secondary reactions, not only for PM but also for the other gaseous pollution (EPA, 2009). Thus, we here provide the combined effects of transboundary pollution on fetuses rather than the narrowly defined health effects of particulate matter.

### 2.5.3 Nonlinear Effects of Pollution on Fetal Deaths

To address biases resulting from linear specifications between pollution and health, we delve into the nonlinear relationship between Chinese particulate matter and its impact on fetuses. In particular, since we do not find statistically significant linear associations between local pollution on fetal mortality rates, the adverse effect of pollution found in Section 2.5.3 can be primarily driven by high PM episodes in South Korea, which are less likely to occur without very high PM events in China (Lee et al., 2011).

Figure 2.7 shows how the local pollution levels respond to high pollution episodes in Beijing. Panel (a) and (b) show the raw and residualized $PM_{10}$ levels following the 90th percentile high $PM_{2.5}$ event ($199\mu g/m^3$) in Beijing, and Panel (c) and (d) display the responses after the 95th percentile event ($246\mu g/m^3$). Since weather conditions also play crucial roles in local pollution, residualized $PM_{10}$ after accounting for the weather and temporal conditions specified in Equation (2.1) can be a more relevant measure to reveal the contribution of pollutants from external sources. First, we can clearly find that Region 1 responds rapidly to the high PM events and Region 2 follows. In Panel (b) and (d), Region 1 has additional pollutants right after the high $PM_{2.5}$ episodes in Beijing. Moreover, the average level of $PM_{10}$ in Region 1 one day after the event in Beijing is $66\mu g/m^3$, which is 32% higher than the WHO guideline for daily exposure to $PM_{10}$. Thus, by exploiting differential time lags for cities to expect transboundary pollutants, an event study-style empirical specification can better capture excess fetal deaths driven by transboundary air pollution.

Before moving onto the event study, we first determine whether high PM concentration in Beijing necessarily translate into more transboundary pollutants for cities in South Korea. That is, there can be certain weather conditions that are indispensable factors for very high pollution events in Beijing, but at the same time, they also can be unfavorable conditions for transbondary transport of pollutants (e.g., extremely stable atmosphere). To determine the most relevant high PM events in Beijing that can affect local pollution levels in Korea, we replace $PM_{2.5}$ in Equation (2.2) with multiple binned indicators which become one if daily $PM_{2.5}$ levels in $t-1$ Beijing falls within every 10th percentile of historical distribution and zero otherwise while making the 5th indicator the omitted category. In Figure 2.8 Panel (a), we find that Beijing's $PM_{2.5}$ level in the the 9th decile of the distribution leads to the highest rate of fetal deaths compared to the effect of $PM_{2.5}$ levels in the 5th decile. The point estimates for the other decile indicators are close to zero or imprecisely estimated in general.

We then demonstrate the lagged impacts of high PM episode in China on fetal deaths in Korea using an event-study specification. We again replace the continuous variable $PM_{2.5}$ in Equation (2.2) with indicators which become one if a day is $n$ day before or after the high $PM_{2.5}$ event in Beijing. Figure 2.9 plots the coefficients on the indicators with their 95 percent confidence intervals. Statistically significant positive point estimates on the one-day lagged indicator after the 9th decile event (i.e., 80th-90th percentile in the distribution of Beijing's daily mean $PM_{2.5}$ distribution) are initially observed for the outcome of fetal mortality rates ($\geq$ 16 weeks) in Panel (a) and (b), indicating that the high pollution event in Beijing have lagged effects on fetal deaths in South Korea. However, the 10th decile event does not show consistent results with those found after the 9th decile event in Figure 2.A.11.

### 2.5.4 Effects of Daily Local $PM_{10}$ Levels on Daily Fetal Mortality

Table 2.B.4 presents how much the local pollution in $t$ can explain the daily fetal mortality (Table 2.B.5 for the effects of local pollution in $t-1$). We find that the estimates of local $PM_{10}$ on daily fetal mortality rates are statistically insignificant. Considering the magnitude of day-to-day variation of local $PM_{10}$ is much lower than those of Beijing's $PM_{2.5}$, the effect of local $PM_{10}$ on fetuses is close to zero. For this little evidence on the effects of local $PM_{10}$ level on fetal mortality, omitted variables in this cross-sectional analysis can lead to biased estimates on fetal deaths due to endogeneity in the potential exposure to pollution. Secondly, with the modest levels of local pollution on average, this regression specification formulating covariates in additive and linear terms may be misspecified to capture any nonlinear effect of pollution on fetal health. To test for nonlinearities, we replace the local $PM_{2.5}$ level in Equation (2.2) with multiple binned indicators which become one if rainfall falls within every 10th percentile of daily mean $PM_{10}$ distribution in the sample period and zero otherwise. Figure 2.A.12 plots the coefficients on each dummy that is constructed using local $PM_{10}$ levels in $t$ while making the 5th indicator the omitted category. We still do not find significant associations between local PM pollution and fetal mortality rates using different gestational ages.

### 2.5.5 Discussion on the Effects on Fetal and Infant Mortality

It is worthwhile to compare our estimates to the results from relevant literature studying the effects of *in utero* exposure to ambient air pollution on fetal or infant deaths. Using the live birth and fetal death data in New Jersey from 1998 to 2004, Faiz et al. (2013) argued that exposure during the third trimester to $NO_2$, $SO_2$ and CO can be associated with an increase in the relative odds of stillbirth. DeFranco et al. (2015)'s cohort study using Ohio birth records

(2006-2010) reports the exposure to high levels of $PM_{2.5}$ in the third trimester of pregnancy ($\geq$16.22$\mu$g/m$^3$) was associated with a 42% increase in the risk of stillbirth, while they could not find any statistically significant association between the exposure in the first and second trimesters and the risk of stillbirth. In this study, however, we do not have significant results in the effects of pollution on fetuses in the third trimester and infant mortlaity (Panel (c), (f), and (i) in Figure 2.5 and Figure 2.A.8).

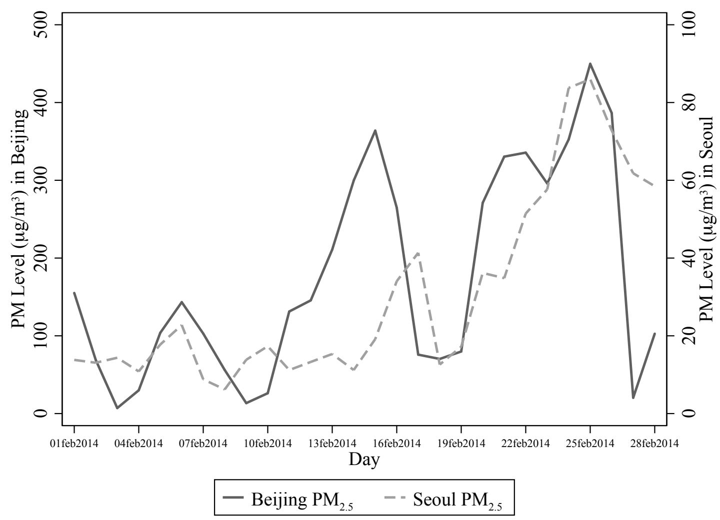
About fetal deaths, Sanders and Stoecker (2015) suggest that one-unit TSP decrease results in 100-110 fewer fetal losses per 100,000 live births using changes in the observed live sex ratio. Their estimates are much larger in magnitude compared to our estimates; their estimates reflect all fetal losses during the whole period of pregnancy while we focus on the fetal deaths after 16 gestational weeks. In addition, our estimates may be underestimated due to an increase in avoidance behavior of pregnant mothers by an extensive and/or intensive margin following an increase in public awareness of the ambient air quality in South Korea.

## 2.6 Conclusion

We exploit the unique transboundary setting to study the impact of the increase in PM concentration, resulting from the excessive fossil fuel combustion in China, on fetal health in South Korea, where large-sample microdata on health are available. Due to the westerly winds blowing from China to Korea, residents in South Korea are intermittently exposed to high pollution levels depending on wind and pressure patterns. Thus, in this study, we provide lower-bound estimates for the impact of pollution on the Chinese population and also contribute to the growing literature on the negative externalities of pollution between countries. Conditional on weather conditions and trends of local pollution, we find that daily variations in pollution in China, measured by Beijing's $PM_{2.5}$ level, negatively affect the fetal health in South Korea. Our results present accurate cost estimates of fine-particle pollution carried from China to South Korea, calling for more rigorous evidence on negative externalities of air pollution and highlighting the urgent need for regional cooperation in Northeast Asia. Considering the severity of pollution which even reaches the west coast of the United States (Lin et al., 2014), quantifying the magnitude of health impact is crucially needed, not only for reshaping the domestic health policy, but also for initiating environmental negotiations in the East Asian region. For example, the estimated cost of burden of transboundary pollution may be used as concrete evidence on whether South Korea and Japan to invest in emission control technologies for coal-fired power plants in China (beneficiary pays), or to negotiate an agreement with China to do so (polluter pays).

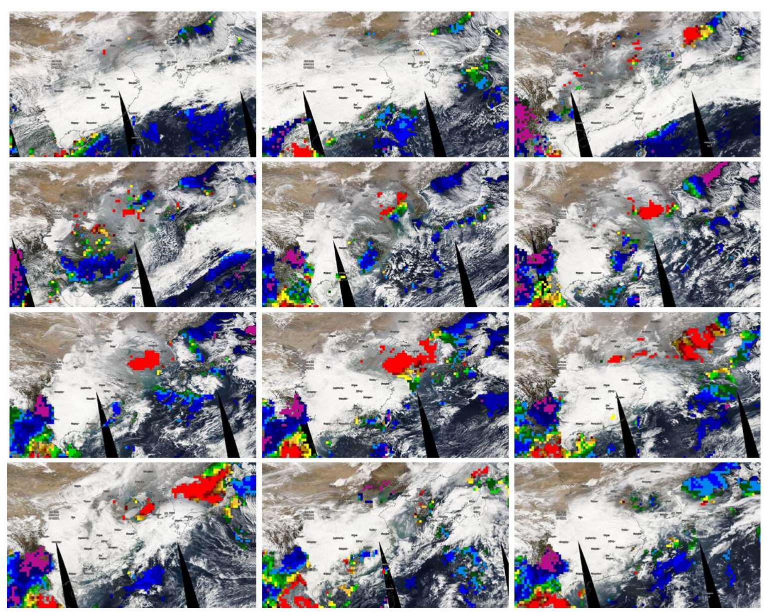
## 2.7 Figures

Figure 2.1: Daily $PM_{2.5}$ Concentration in Beijing and Seoul in Feb. 2014




*Notes*: The dashed line refers to the WHO guideline for daily mean exposure to $PM_{2.5}$, $25\mu g/m^3$

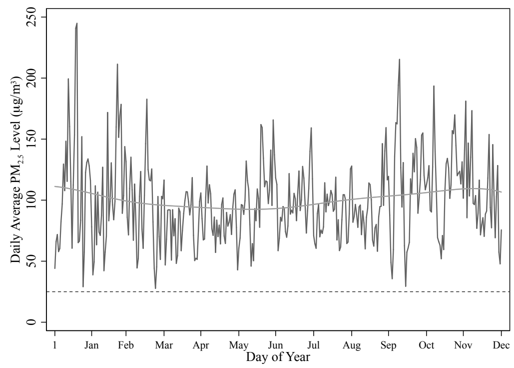
Figure 2.2: Satellite Images from Feb 17 to 28, 2014 in East Asia



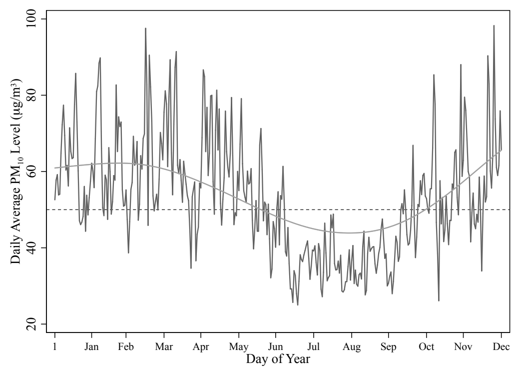
*Notes*: This figure displays MODIS Combined Value-Added Aerosol Optical Depth images. Red and blue show high and low concentration of aerosol, respectively. Aerosol levels are not estimated under cloud covers.
Source: https://earthdata.nasa.gov/labs/worldview/

Figure 2.3: Seasonality of PM levels in Beijing and in Korea

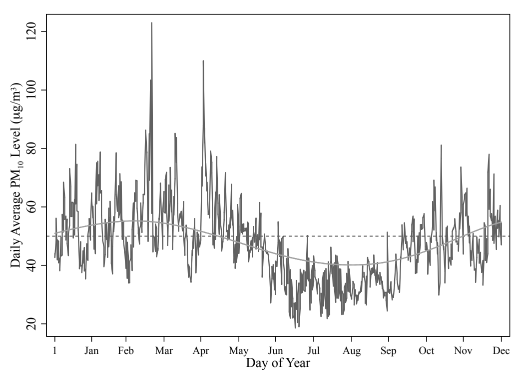
(a) $PM_{2.5}$ in Beijing




(b) $PM_{10}$ in Region 1, South Korea

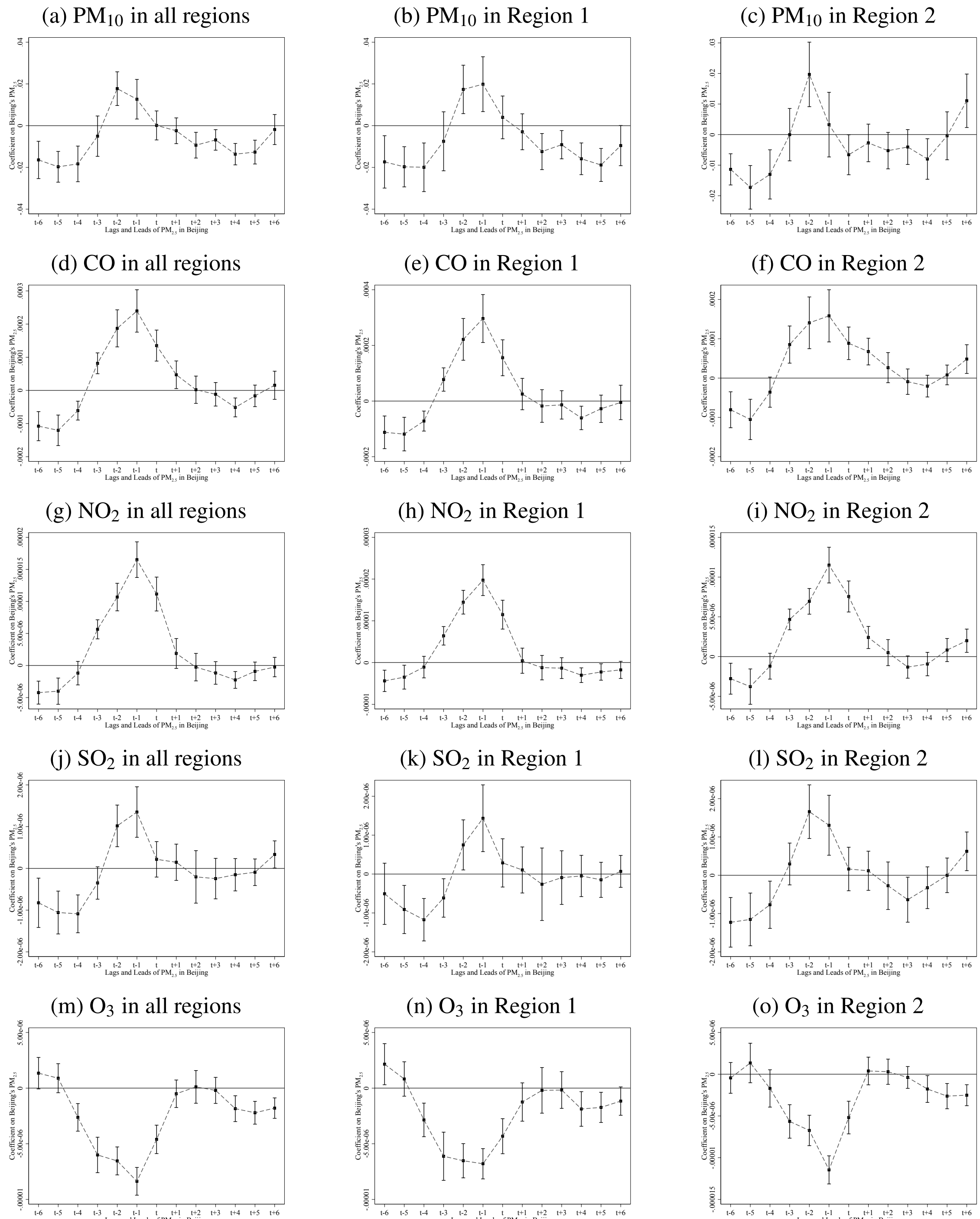



(c) $PM_{10}$ in Region 2, South Korea



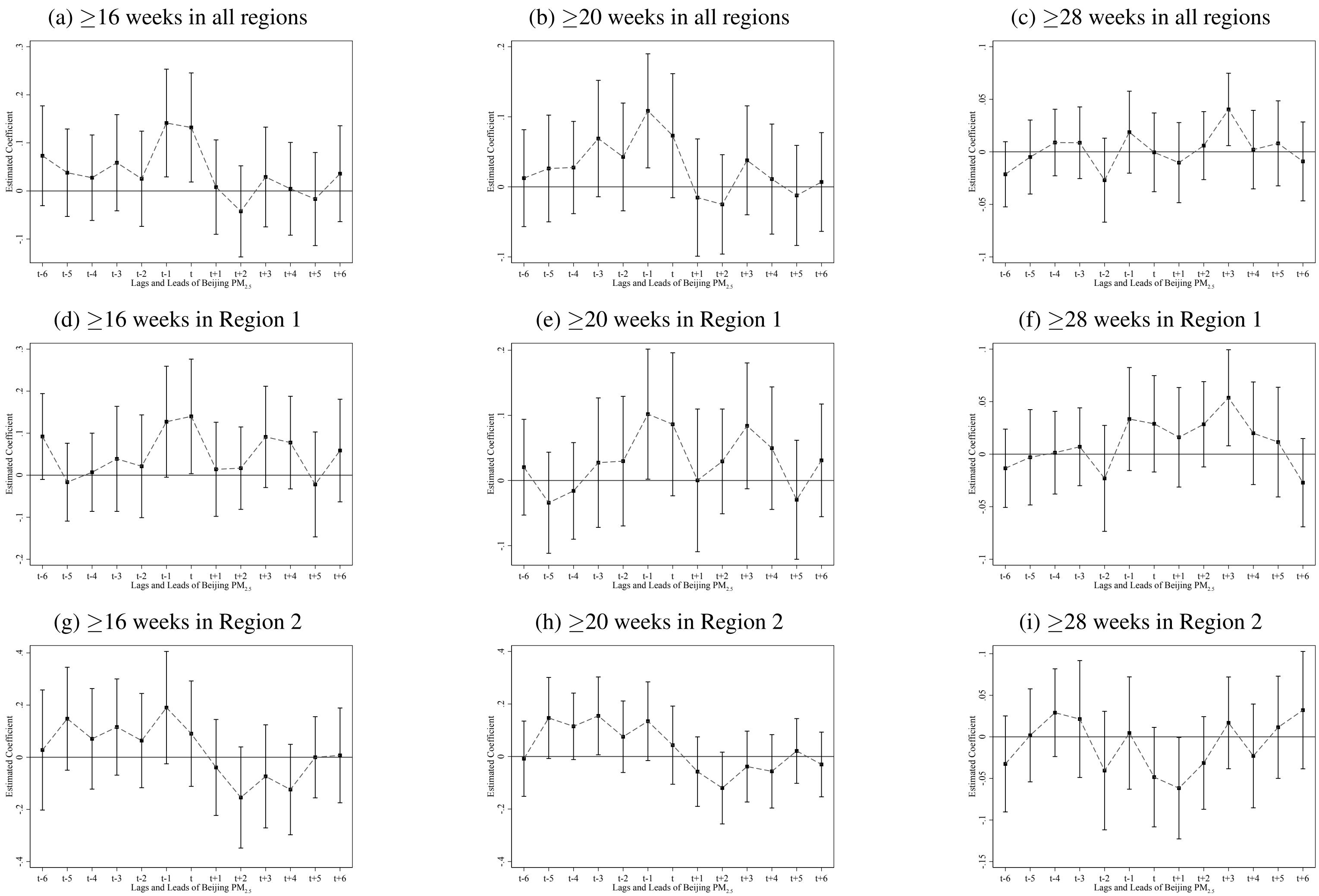

*Notes*: Dashed lines refer to the WHO guidelines for daily exposure to $PM_{2.5}$ ($25\mu g/m^3$) and $PM_{10}$ ($50\mu g/m^3$), respectively. Lowess lines are overlaid onto the line plots of daily mean PM concentration.

Figure 2.4: Effect of Beijing's $PM_{2.5}$ on Local Pollution Levels by Region in South Korea



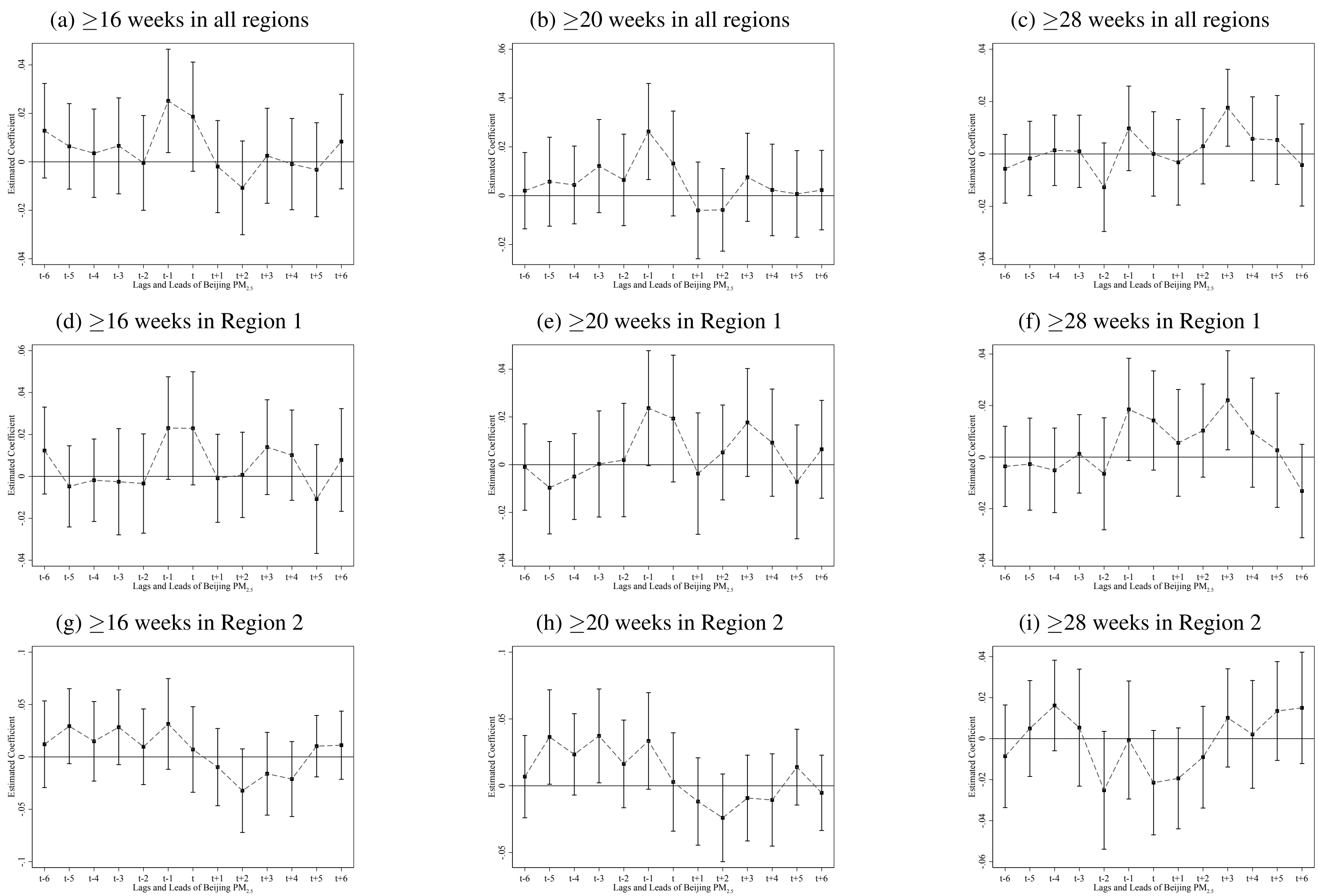

*Notes*: Figures provide the point estimates and the corresponding 95 percent confidence intervals (bars) from linear regressions of daily mean of five major pollution levels in cities in South Korea on lags and leads of Beijing $PM_{2.5}$ level from $t-6$ to $t+6$.

Figure 2.5: Effect of Beijing's $PM_{2.5}$ on Daily Fetal Mortality Rates by Region in South Korea




*Notes*: Figures provide the point estimates and the corresponding 95 percent confidence intervals (bars) from linear regressions of daily fetal mortality rates ($\geq$ 16, 20 and 28 pregnancy weeks) per 1 million live births on lags and leads of Beijing $PM_{2.5}$ level from $t-6$ to $t+6$.

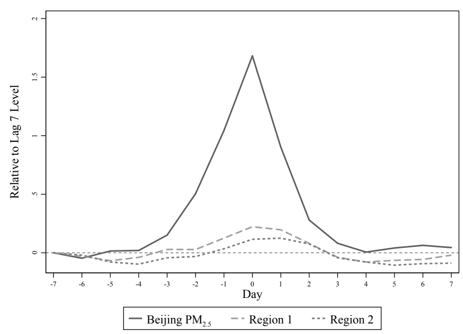
Figure 2.6: Effect of Beijing's $PM_{2.5}$ on Daily Prospective Fetal Mortality Rates by Region in South Korea



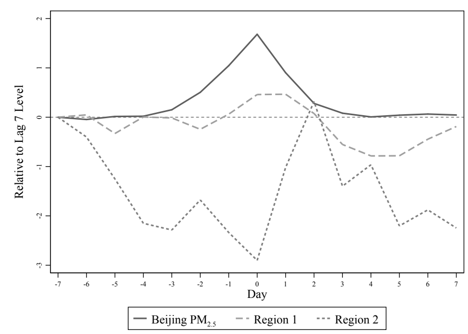

*Notes*: Figures provide the point estimates and the corresponding 95 percent confidence intervals (bars) from linear regressions of mortality rates of fetuses ($\geq$ 16, 20 and 28 pregnancy weeks) per 1 million fetuses at risk on lags and leads of Beijing $PM_{2.5}$ level from $t-6$ to $t+6$.

Figure 2.7: Relative Change in Raw and Residualized $PM_{10}$ Levels in South Korea 7 Days before and after High $PM_{2.5}$ Events in Beijing

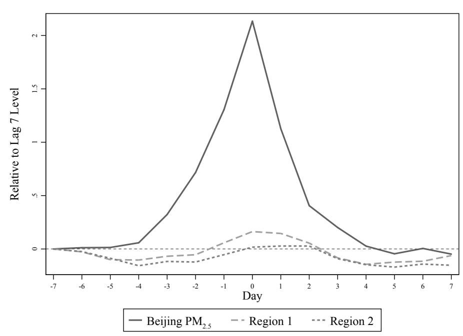
(a) Raw, 90th percentile



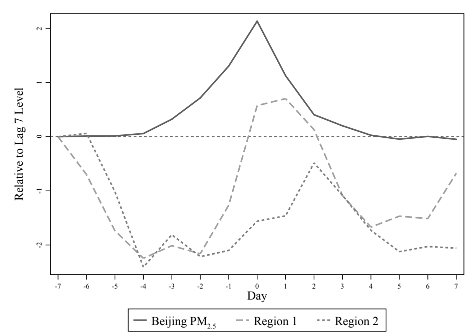

(b) Residualized, 90th percentile




(c) Raw, 95th percentile

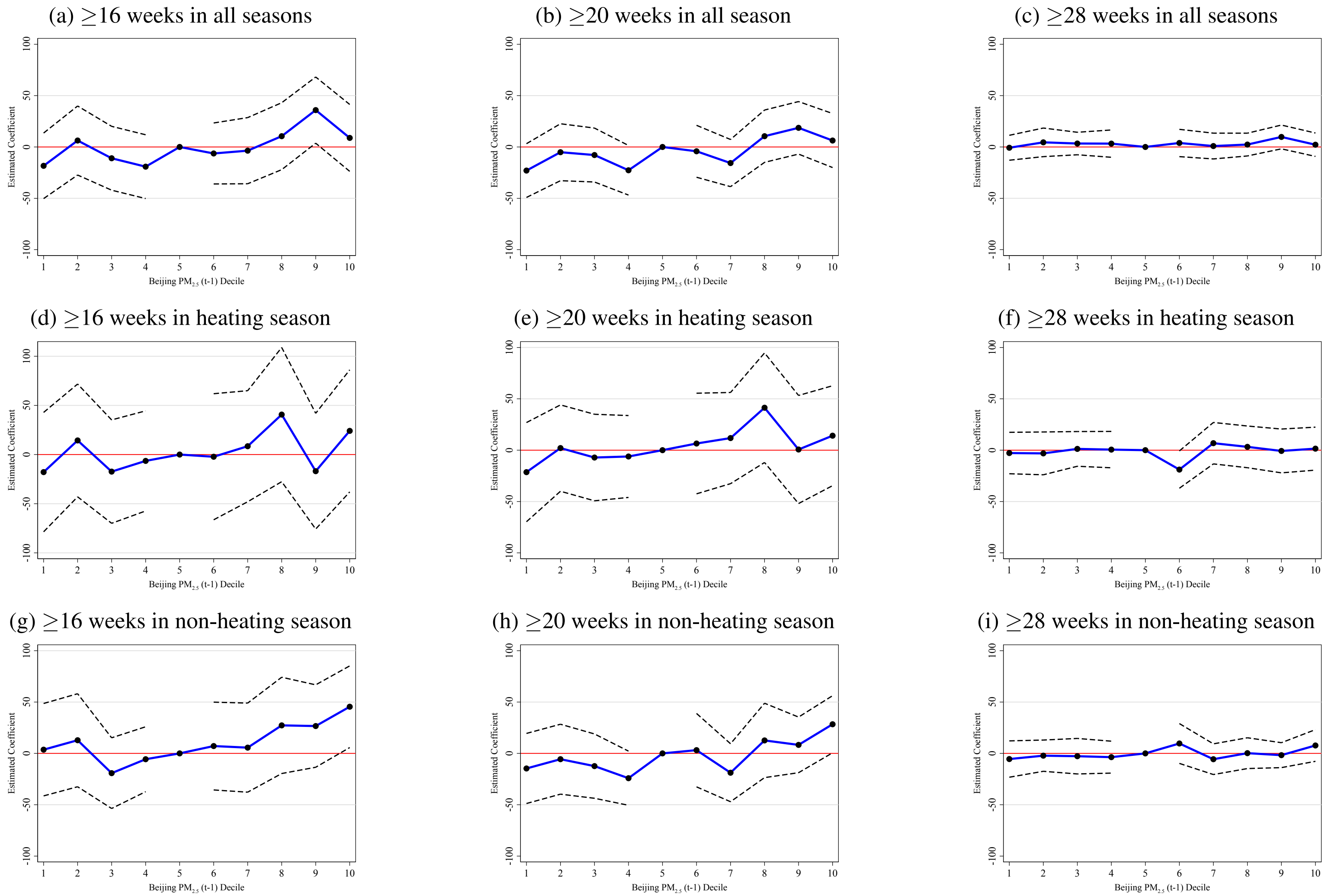



(d) Residualized, 95th percentile



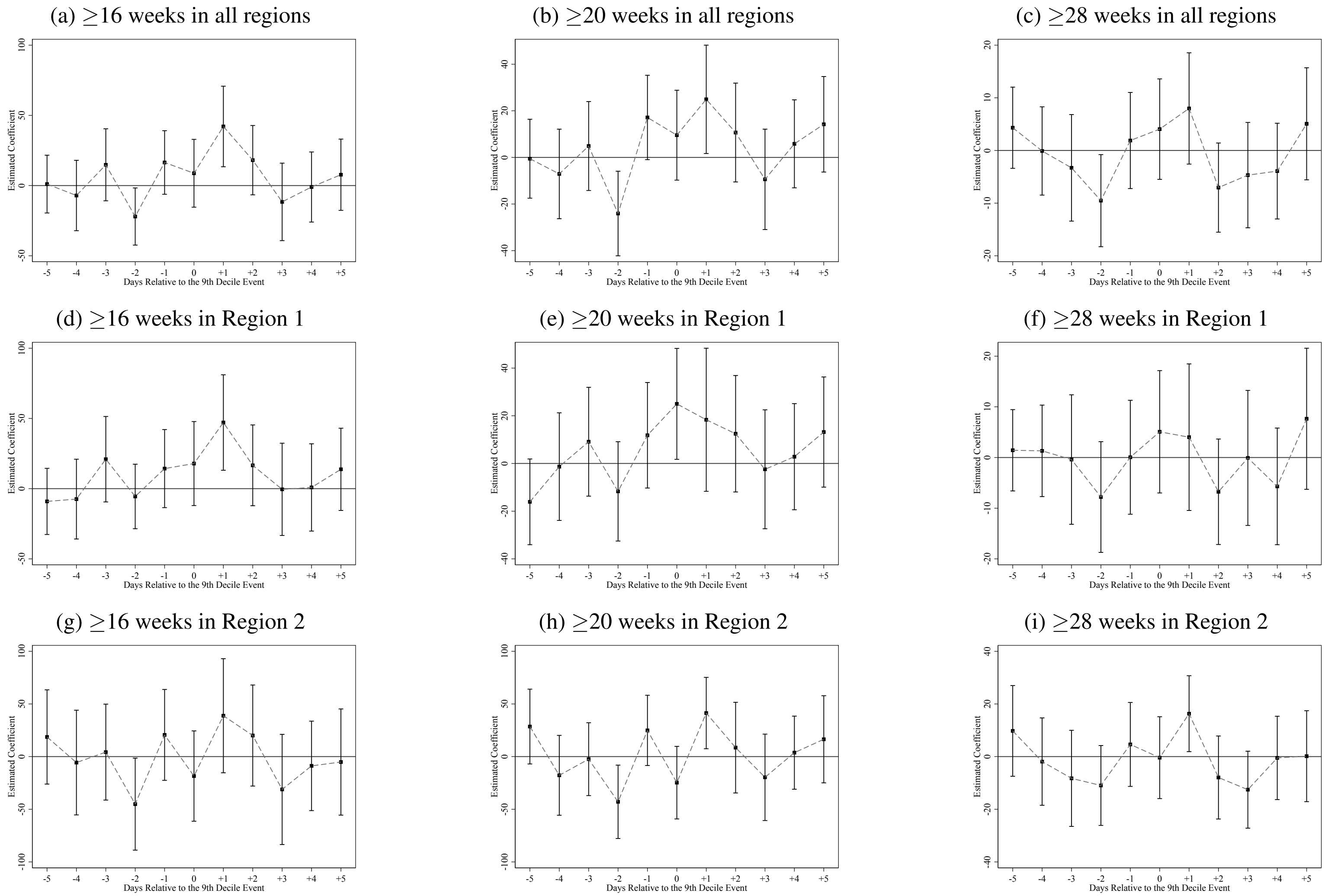

*Notes*: Panel (a) and (b) show the raw and residualized $PM_{10}$ levels in cities of South Korea after the 90th percentile high $PM_{2.5}$ event ($199\mu g/m^3$) in Beijing. Panel (c) and (d) display the same response after the 95th percentile event ($246\mu g/m^3$), respectively. The residualized $PM_{10}$ level is obtained after regressing daily $PM_{10}$ on weather and temporal covariates specified in Equation (2.1).

Figure 2.8: Test for Non-linear Effects of Beijing's $PM_{2.5}$ in $t-1$ on Daily Fetal Mortality Rates in South Korea




*Notes*: The figure plots coefficients and 95% confidence intervals from a regression of fetal mortality rates on the dummies for each 10th percentile (decile) of the Beijing's $PM_{2.5}$ $(t-1)$ distribution in 2009-2013. The omitted category is the 5th decile. Heating and non-heating seasons refer to the period from November to March, and from April to October, respectively by the public heating season in northern China from November to March.

Figure 2.9: Event Study on Daily Fetal Mortality Rates in South Korea after High $PM_{2.5}$ Episodes in Beijing (9th decile)




*Notes*: The figures plot the coefficients on the indicators for $n$ days away from high $PM_{2.5}$ episodes in Beijing (9th decile) at $n = 0$ in the regression estimating the effect on daily fetal mortality rates. The bars refer to the 95 percent confidence intervals.

## 2.8 Tables

Table 2.1: Summary Statistics

| Sample | All Regions | | | | Region 1 | | | | Region 2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Statistics | Mean | Std.Dev. | Min | Max | Mean | Std.Dev. | Min | Max | Mean | Std.Dev. | Min | Max |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| ***Panel A. Daily Mean Air Pollutant Concentration in South Korea*** | | | | | | | | | | | | |
| $PM_{10}$ ($\mu$g/m$^3$) | 51.212 | (29.479) | 812.483 | 3.992 | 52.971 | (29.821) | 356.000 | 4.917 | 47.691 | (28.459) | 812.483 | 3.992 |
| CO (ppm) | 0.544 | (0.232) | 2.250 | 0.000 | 0.574 | (0.242) | 2.250 | 0.000 | 0.485 | (0.196) | 2.250 | 0.100 |
| NO2 (ppm) | 0.027 | (0.012) | 0.094 | 0.002 | 0.031 | (0.013) | 0.094 | 0.002 | 0.021 | (0.009) | 0.070 | 0.003 |
| O3 (ppm) | 0.022 | (0.011) | 0.084 | 0.001 | 0.021 | (0.011) | 0.084 | 0.001 | 0.025 | (0.011) | 0.073 | 0.002 |
| SO2 (ppm) | 0.005 | (0.003) | 0.038 | 0.001 | 0.005 | (0.002) | 0.024 | 0.001 | 0.005 | (0.003) | 0.038 | 0.001 |
| $PM_{2.5}$ ($\mu$g/m$^3$) | 25.006 | (12.773) | 121.587 | 3.325 | 24.180 | (13.490) | 121.587 | 3.325 | 25.868 | (11.923) | 93.406 | 5.885 |
| ***Panel B. Daily Mean Air Pollutant Concentration in Beijing*** | | | | | | | | | | | | |
| $PM_{2.5}$ ($\mu$g/m$^3$) | 99.246 | (75.434) | 556.458 | 2.917 | 99.246 | (75.434) | 556.458 | 2.917 | 99.246 | (75.434) | 556.458 | 2.917 |
| ***Panel C. Daily Mean Local Weather*** | | | | | | | | | | | | |
| Temperature (Celsius Degree) | 12.781 | (10.620) | 32.975 | -14.825 | 12.242 | (10.891) | 32.021 | -14.825 | 13.859 | (9.971) | 32.975 | -12.238 |
| Precipitation (mm) | 4.265 | (16.560) | 436.000 | 0.000 | 4.549 | (17.848) | 436.000 | 0.000 | 3.698 | (13.608) | 310.000 | 0.000 |
| Wind speed (m/s) | 2.300 | (1.067) | 10.113 | 0.029 | 2.383 | (1.075) | 8.971 | 0.054 | 2.135 | (1.032) | 10.113 | 0.029 |
| ***Panel C. Birth and Fetal Death*** | | | | | | | | | | | | |
| No. of daily live births | 37 | (52.450) | 491 | 1 | 37 | (62.340) | 491 | 1 | 39 | (21.931) | 148 | 1 |
| No. of daily fetal deaths | 0.550 | (1.055) | 13 | 0 | 0.502 | (1.119) | 13 | 0 | 0.646 | (0.906) | 9 | 0 |
| Daily fetal mortality rate ($\geq$16 weeks) | 0.498 | (1.004) | 12.579 | 0.000 | 0.460 | (1.065) | 12.579 | 0.000 | 0.575 | (0.867) | 8.316 | 0.000 |
| Daily fetal mortality rate ($\geq$20 weeks) | 0.303 | (0.776) | 12.579 | 0.000 | 0.289 | (0.836) | 12.579 | 0.000 | 0.330 | (0.638) | 6.522 | 0.000 |
| Daily fetal mortality rate ($\geq$28 weeks) | 0.070 | (0.372) | 8.621 | 0.000 | 0.067 | (0.404) | 8.621 | 0.000 | 0.076 | (0.300) | 4.202 | 0.000 |
| Observations | 38,338 | | | | 25,556 | | | | 12,782 | | | |

*Notes*: This table reports summary statistics of selected day-city level pollution, weather and outcome variables. $PM_{2.5}$ observations in Region 1 and Region 2 come from Seoul and Busan in 2010-2013, respectively.

Table 2.2: Effects of Beijing's $PM_{2.5}$ Level in $_{t-1}$ on Local $PM_{10}$ Level in $t$ in South Korea

| | **$PM_{10}$ Concentration in South Korea ($\mu m/m^3$)** | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| ***Panel A: All Regions*** | | | | | |
| Beijing $PM_{2.5}$ $(t-1)$ | 0.0646*** | 0.0640*** | 0.0163*** | 0.0161*** | 0.0127*** |
| | (0.00725) | (0.00607) | (0.00469) | (0.00469) | (0.00481) |
| Observations | 36,007 | 36,007 | 36,007 | 36,007 | 36,007 |
| R-squared | 0.031 | 0.176 | 0.369 | 0.370 | 0.372 |
| Mean of Dep. Var. | 51.2 | 51.2 | 51.2 | 51.2 | 51.2 |
| ***Panel B: Region 1*** | | | | | |
| Beijing $PM_{2.5}$ $(t-1)$ | 0.0729*** | 0.0717*** | 0.0231*** | 0.0229*** | 0.0199*** |
| | (0.00980) | (0.00804) | (0.00639) | (0.00639) | (0.00665) |
| Observations | 24,006 | 24,006 | 24,006 | 24,006 | 24,006 |
| R-squared | 0.038 | 0.193 | 0.397 | 0.398 | 0.399 |
| Mean of Dep. Var. | 53.0 | 53.0 | 53.0 | 53.0 | 53.0 |
| ***Panel C: Region 2*** | | | | | |
| Beijing $PM_{2.5}$ $(t-1)$ | 0.0489*** | 0.0496*** | 0.00755 | 0.00744 | 0.00327 |
| | (0.00862) | (0.00771) | (0.00553) | (0.00550) | (0.00530) |
| Observations | 12,001 | 12,001 | 12,001 | 12,001 | 12,001 |
| R-squared | 0.019 | 0.143 | 0.357 | 0.358 | 0.363 |
| Mean of Dep. Var. | 47.7 | 47.7 | 47.7 | 47.7 | 47.7 |
| **Controls** | | | | | |
| City FE | Yes | Yes | Yes | Yes | Yes |
| Temporal Controls | | Yes | Yes | Yes | Yes |
| Weather Controls | | | Yes | Yes | Yes |
| Linear time trend | | | | Yes | Yes |
| Quadratic time trend | | | | | Yes |

*Notes*: Regressions are weighted by the number of live births in a city. Robust standard errors are shown in parentheses clustered at the region-year-season level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 2.3: Effects of Beijing $PM_{2.5\ t-1}$ on Daily Fetal Mortality Rates ($\geq$ 20 Gestational Weeks) in South Korea

| | Daily Fetal Mortality Rates (≥20 Weeks) | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| ***Panel A: All Regions*** | | | | | |
| Beijing $PM_{2.5}$ $(t-1)$ | 0.0905** | 0.0928*** | 0.106*** | 0.104*** | 0.109*** |
| | (0.0367) | (0.0319) | (0.0385) | (0.0390) | (0.0413) |
| Observations | 36,020 | 36,020 | 36,020 | 36,020 | 36,020 |
| R-squared | 0.000 | 0.037 | 0.039 | 0.040 | 0.041 |
| Mean of Dep. Var. | 0.303 | 0.303 | 0.303 | 0.303 | 0.303 |
| ***Panel B: Region 1*** | | | | | |
| Beijing $PM_{2.5}$ $(t-1)$ | 0.0847* | 0.0859** | 0.0966** | 0.0946** | 0.102** |
| | (0.0468) | (0.0397) | (0.0471) | (0.0476) | (0.0505) |
| Observations | 24,008 | 24,008 | 24,008 | 24,008 | 24,008 |
| R-squared | 0.000 | 0.031 | 0.034 | 0.034 | 0.035 |
| Mean of Dep. Var. | 0.289 | 0.289 | 0.289 | 0.289 | 0.289 |
| ***Panel C: Region 2*** | | | | | |
| Beijing $PM_{2.5}$ $(t-1)$ | 0.101* | 0.106* | 0.137* | 0.138* | 0.134* |
| | (0.0589) | (0.0535) | (0.0707) | (0.0711) | (0.0751) |
| Observations | 12,012 | 12,012 | 12,012 | 12,012 | 12,012 |
| R-squared | 0.000 | 0.048 | 0.054 | 0.055 | 0.055 |
| Mean of Dep. Var. | 0.330 | 0.330 | 0.330 | 0.330 | 0.330 |
| **Controls** | | | | | |
| City FE | Yes | Yes | Yes | Yes | Yes |
| Temporal Controls | | Yes | Yes | Yes | Yes |
| Weather Controls | | | Yes | Yes | Yes |
| Linear time trend | | | | Yes | Yes |
| Quadratic time trend | | | | | Yes |

*Notes*: Regressions are weighted by the number of live births in a city. Robust standard errors are shown in parentheses clustered at the region-year-season level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

# Appendix to Chapter 2

## 2.A Appendix Figures

Figure 2.A.1: Average $PM_{2.5}$ exposure across OECD countries (2013)

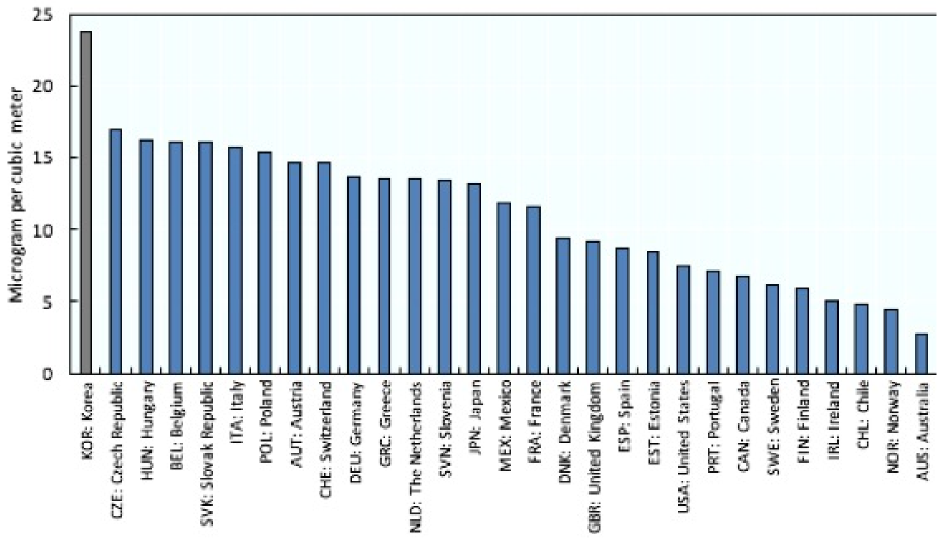
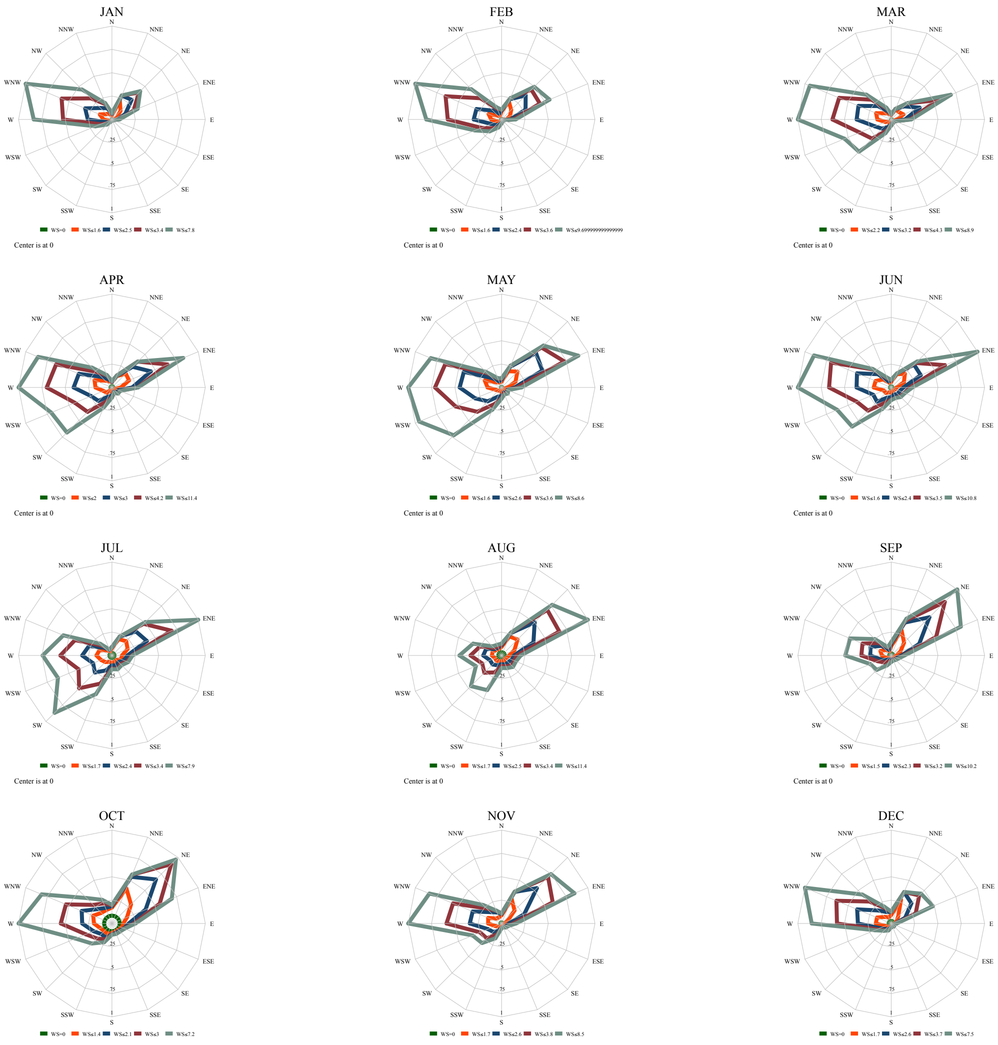


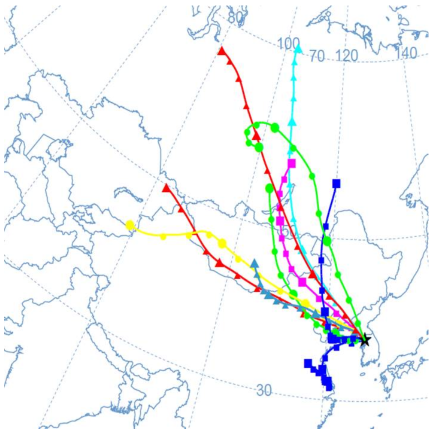
Figure 2.A.2: Daily Wind Patterns by Month in Seoul, South Korea in 2010-2012



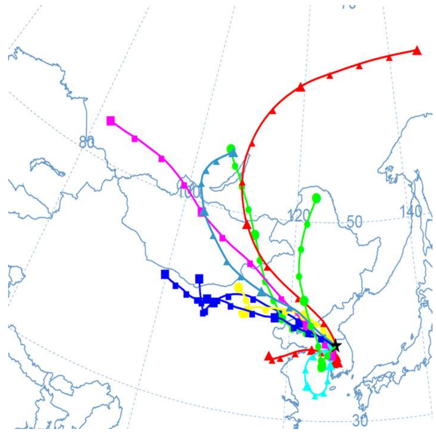

*Notes*: This figure plots the histogram of daily wind directions and speeds. Winds blow from the direction at which the shapes are pointing. For example, in January, winds mostly blow from WNW, and the wind is strongest among all seasons.

Figure 2.A.3: Backtrajectory Analysis on High $PM_{10}$ Day in Each Season in 2013

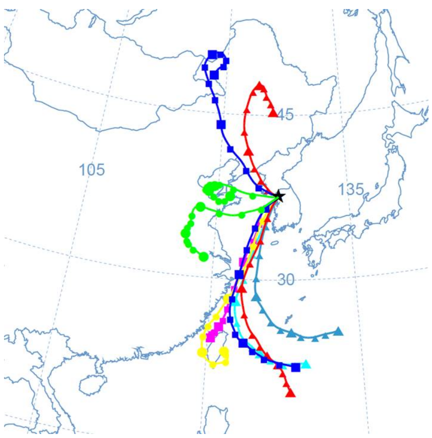
(a) Jan 13, 2013



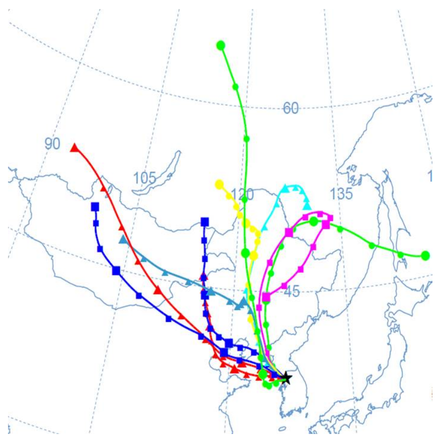

(b) Apr 5, 2013




(c) Jul 26, 2013

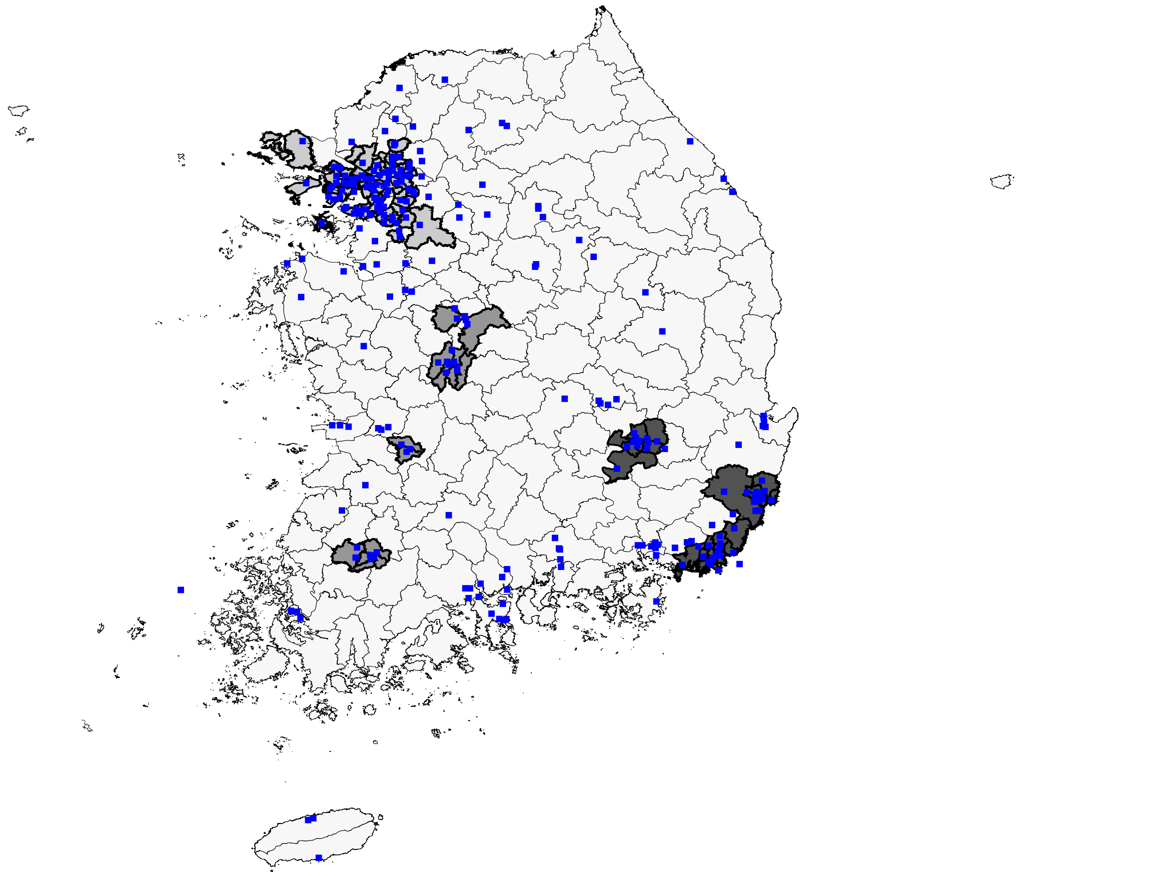



(d) Nov 23, 2013



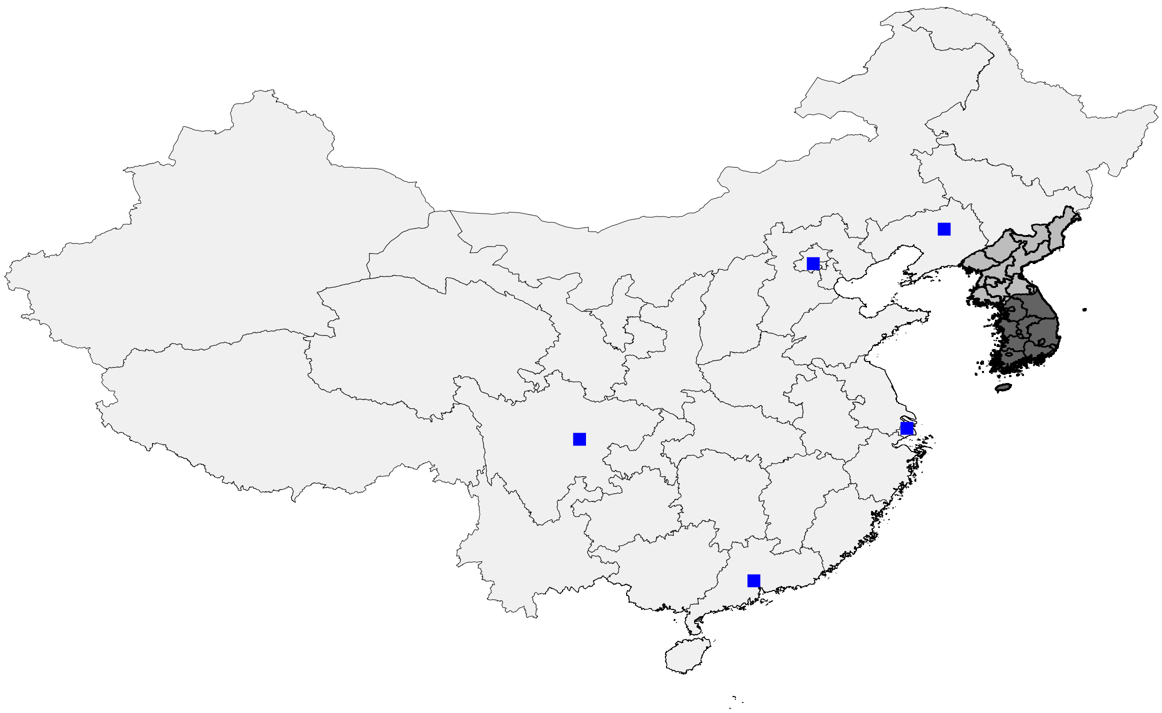

*Notes*: Using the readings from one station in Seoul, South Korea, we choose the days with the highest readings of $PM_{2.5}$ for each season in 2013. Then, we calculate 72-hour back trajectories for every hour on that day to identify air mass pathways using NOAA HYSPLIT MODEL Backward trajectories ending at 0000 UTC on each day in panels. The arrival point at 00 UTC on the day of a high-$PM_{2.5}$ episode is Seoul (37.57 N, 126.97 E). Ten 72-hour back trajectories for every 24 hour are drawn. Each dot represents six-hour trajectory. The levels of $PM_{10}$ are as follows: for Jan 13, Apr 5, Jul 26 and Nov 23 2013, the readings are $76.5\mu g/m^3$, $71.9\mu g/m^3$, $62.5\mu g/m^3$, $51.75\mu g/m^3$, respectively. For given boundary layers, 1000m above the surface is settled as the arrival height.

Figure 2.A.4: Cities and Pollution Monitors in South Korea



*Notes*: Blue squares denote the pollution monitors in residential areas. Region 1 (light gray) includes 14 cities. Region 2 (darker grays) includes the remaining 7 cities.

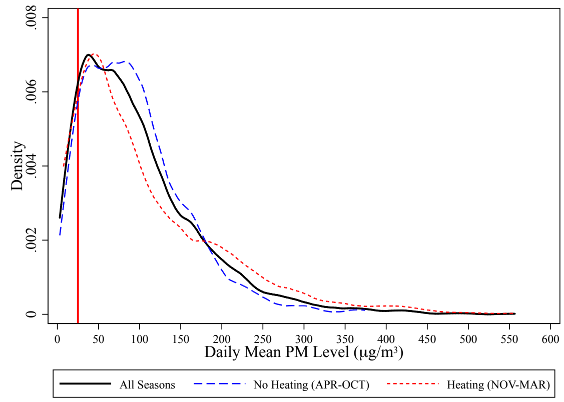
Figure 2.A.5: Locations of Pollution Monitors in China



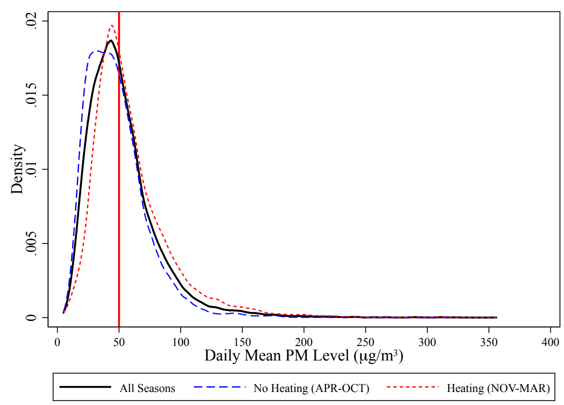
*Notes*: The blue dot at 12 o'clock represents the monitor located at the U.S. embassy in Beijing. Shenyang is at 1 o'clock, Shanghai at 5, Guangzhou is at 7, and Chengdu is at 8, respectively.

Figure 2.A.6: Distributions of Daily Mean PM Level in Beijing and in Korea

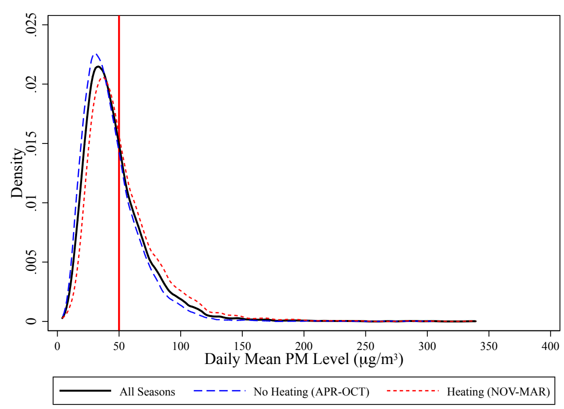
(a) $PM_{2.5}$ in Beijing




(b) $PM_{10}$ in Region 1, Korea

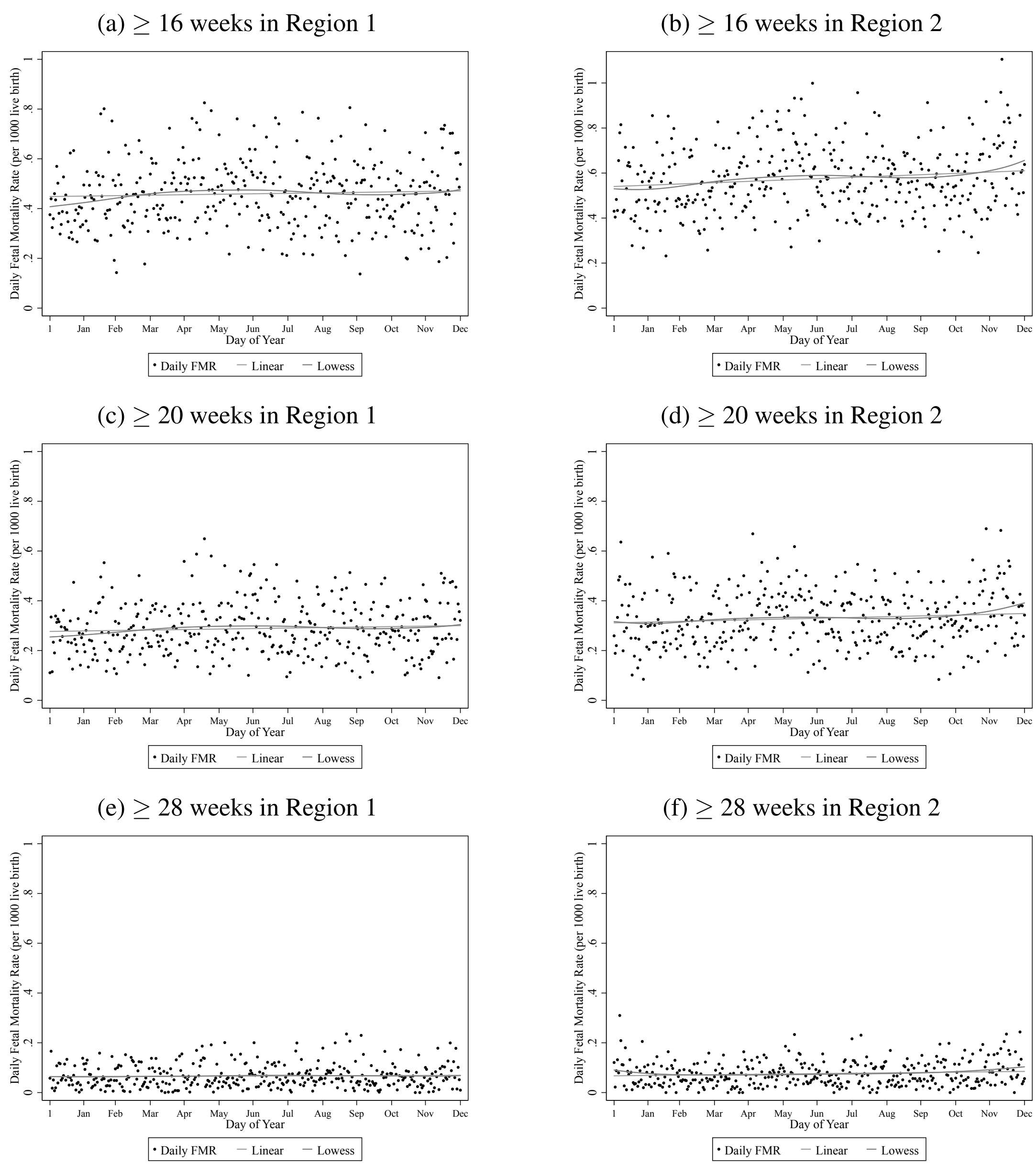



(c) $PM_{10}$ in Region 2, Korea



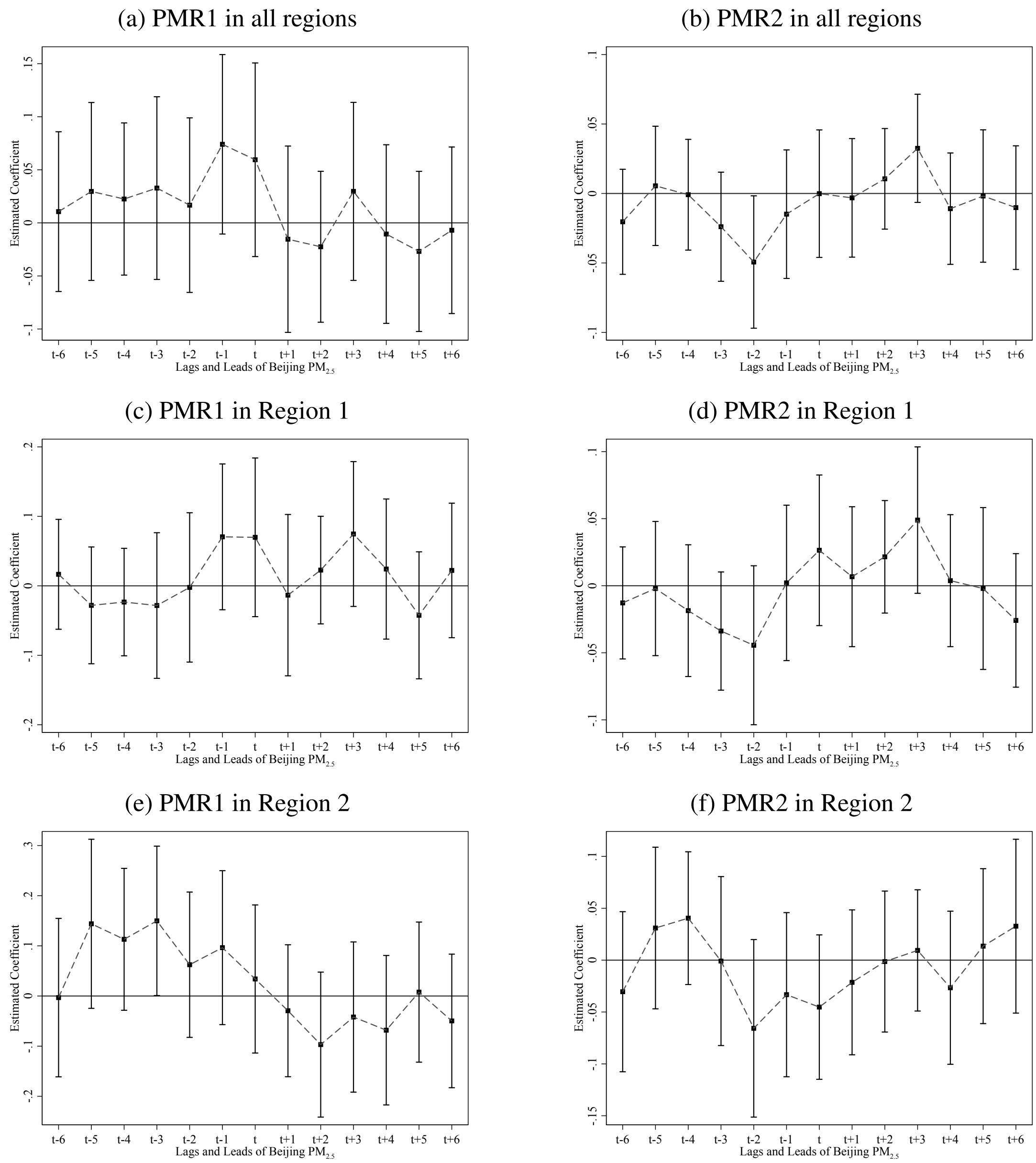

*Notes*: Figures plot the kernel density estimation on the distributions of daily mean PM levels by season in Beijing and in Korea. Red solid lines denote the WHO guidelines for daily exposure to $PM_{2.5}$ ($25\mu g/m^3$) and $PM_{10}$ ($50\mu g/m^3$), respectively. We can see that Region 1 has high density weights over the guideline.

Figure 2.A.7: Seasonality of Daily Fetal Mortality Rates by Region in South Korea



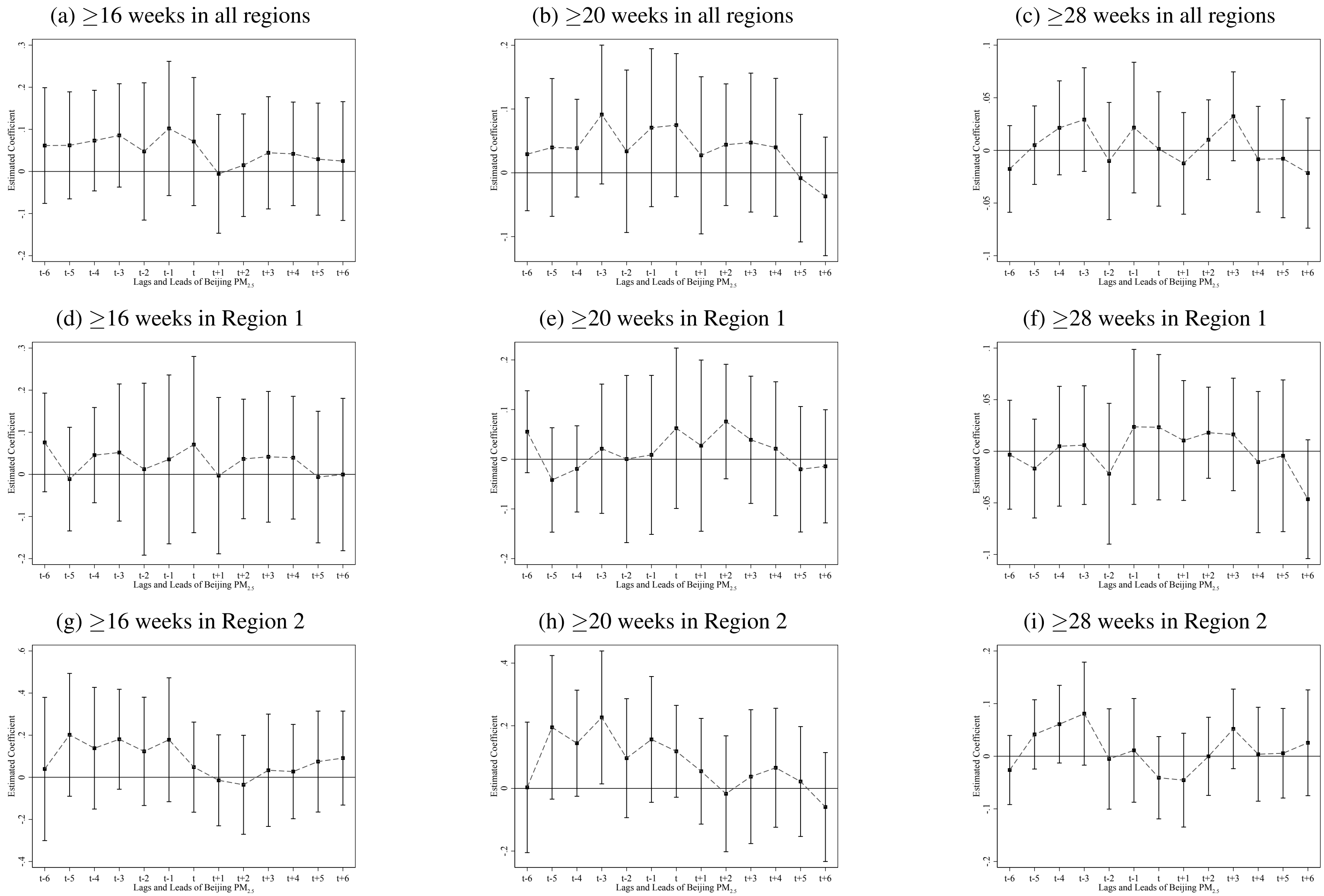

*Notes*: Liner fits and Lowess are overlaid to detect possible trends and seasonalities over the scatter plots of daily fetal mortality rates.

Figure 2.A.8: Effect of Beijing's $PM_{2.5}$ on Daily Perinatal Mortality Rates by Region in South Korea



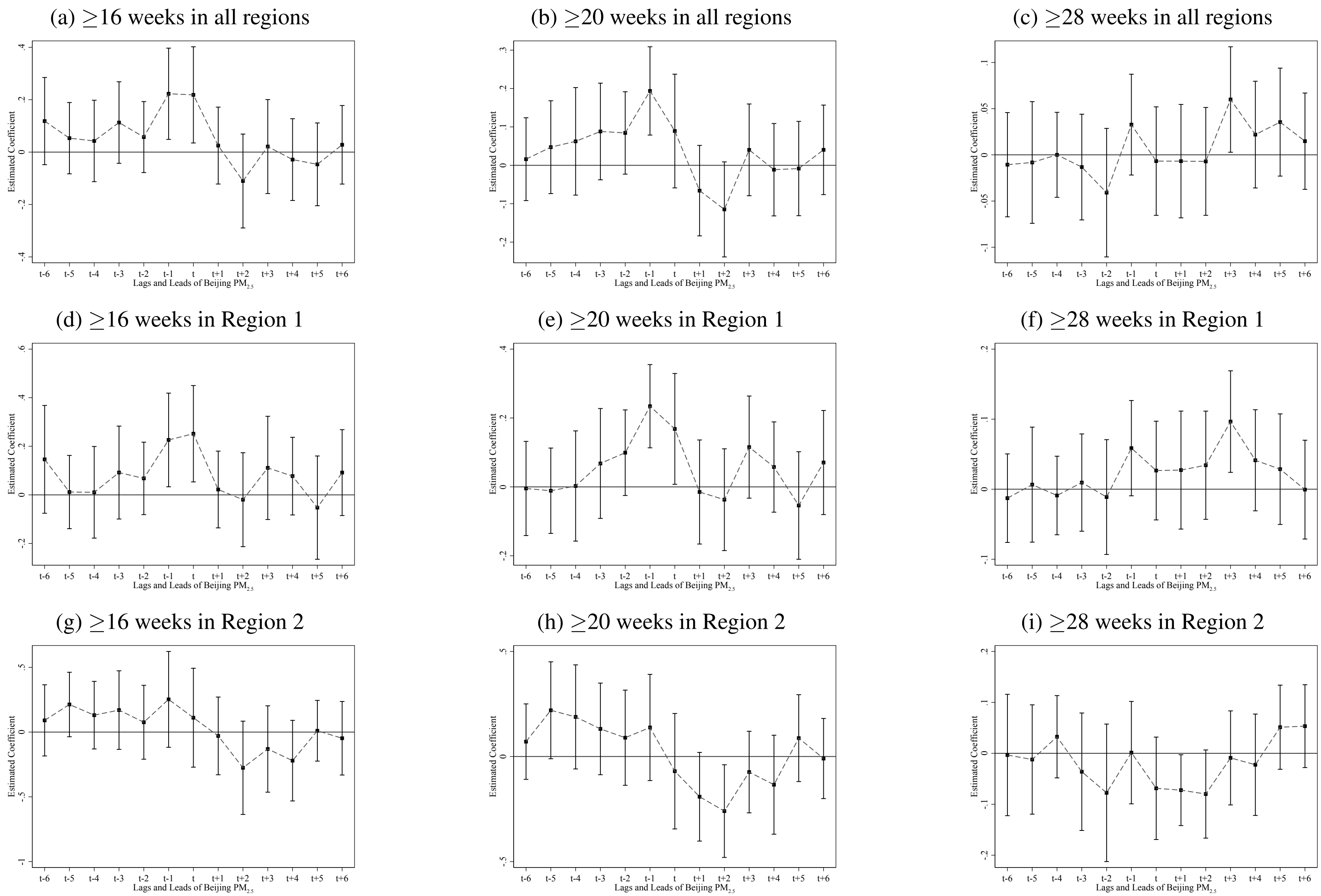

*Notes*: Figures provide the point estimates and the corresponding 95 percent confidence intervals (bars) from linear regressions of perinatal mortality rates per 1 million on lags and leads of Beijing $PM_{2.5}$ level from $t-6$ to $t+6$. PMR1 means perinatal mortality rates using the fetal deaths at gestational age at 20 weeks or later and the infant deaths within 28 days after birth. PMR2 uses the fetal deaths at gestational age at 28 weeks or later and the infant deaths within 7 days after birth.

Figure 2.A.9: Effect of Beijing's $PM_{2.5}$ on Daily Fetal Mortality Rates during Heating Season by Region in South Korea



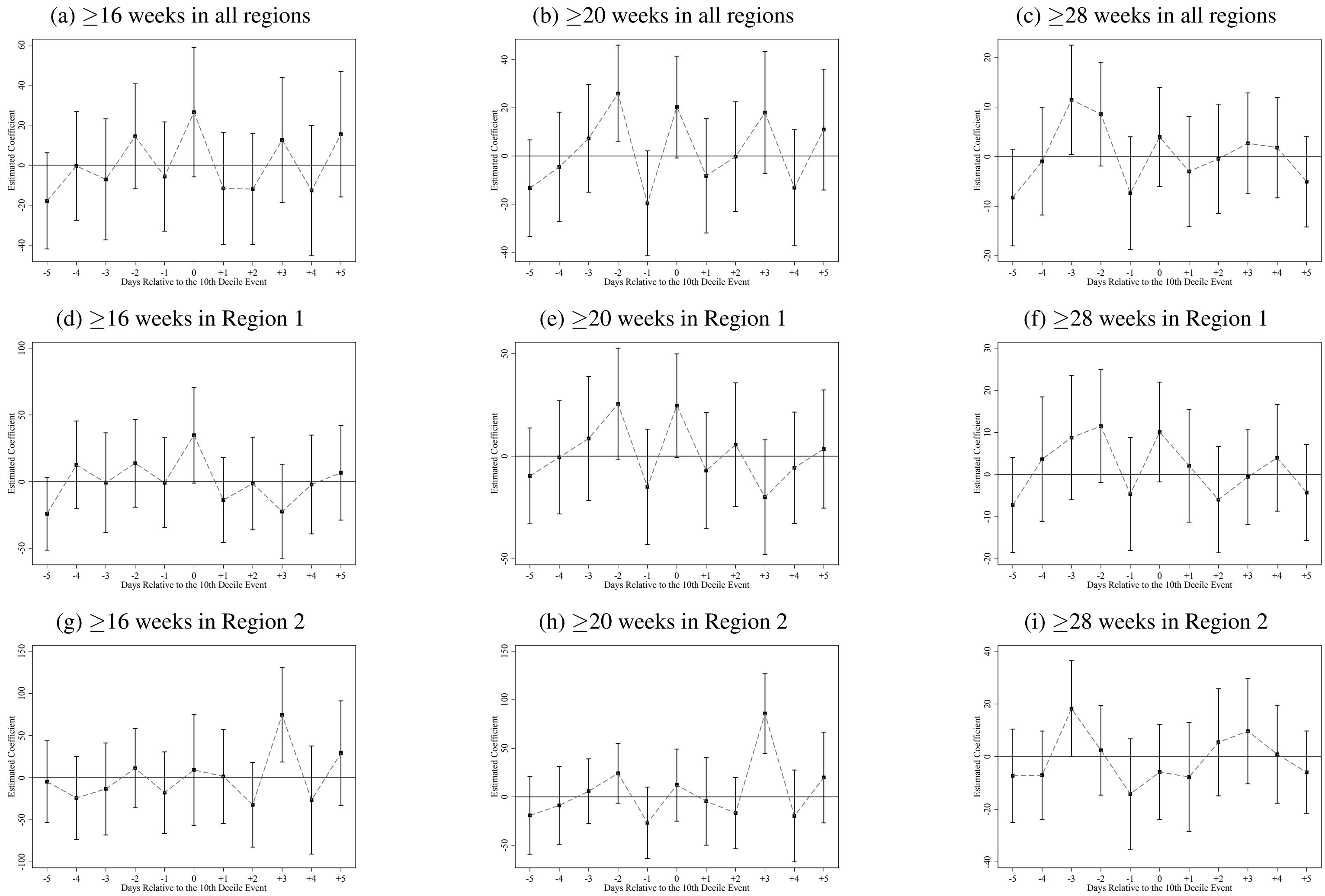

*Notes*: Figures provide the point estimates and the corresponding 95 percent confidence intervals (bars) from linear regressions of mortality rates of fetuses ($\geq$ 16, 20 and 28 pregnancy weeks) per 1 million live births on lags and leads of Beijing $PM_{2.5}$ level from $t-6$ to $t+6$. The public heating season in northern China lasts from November to March.

Figure 2.A.10: Effect of Beijing's $PM_{2.5}$ on Daily Fetal Mortality Rates during Non-Heating Season by Region in South Korea



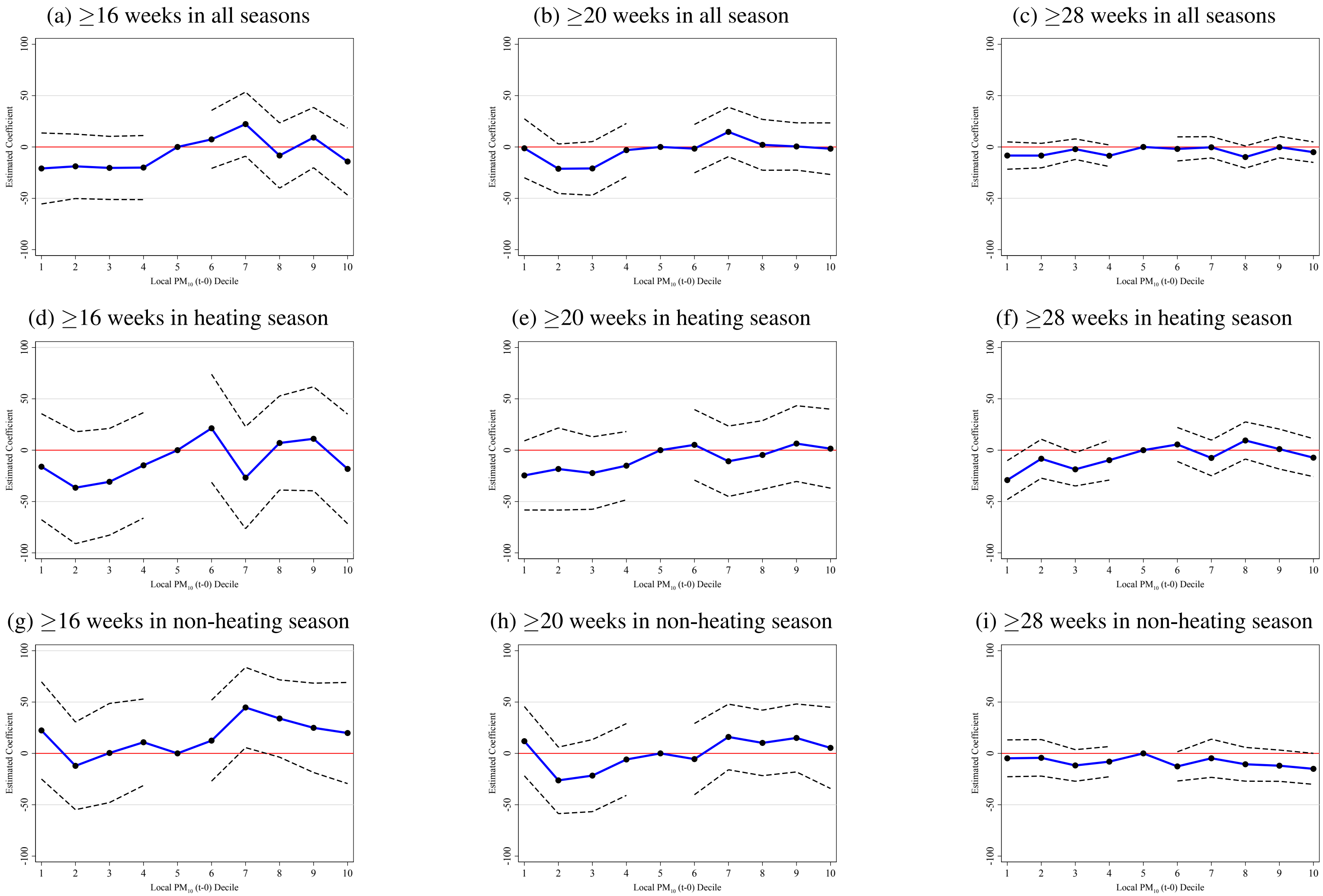

*Notes*: Figures provide the point estimates and the corresponding 95 percent confidence intervals (bars) from linear regressions of mortality rates of fetuses ($\geq$ 16, 20 and 28 pregnancy weeks) per 1 million live births on lags and leads of Beijing $PM_{2.5}$ level from $t-6$ to $t+6$. There is no public heating in northern China from April to October.

Figure 2.A.11: Event Study on Daily Fetal Mortality Rates in South Korea after High $PM_{2.5}$ Episodes in Beijing (10th decile)




*Notes*: The figures plot the coefficients on the indicators for $n$ days away from high $PM_{2.5}$ episodes in Beijing (10th decile) at $n = 0$ in the regression estimating the effect on daily fetal mortality rates. The bars refer to the 95 percent confidence intervals.

Figure 2.A.12: Test for Non-linear Effects of Local $PM_{10}$ in $t$ on Daily Fetal Mortality Rates in South Korea




*Notes*: The figure plots coefficients and 95% confidence intervals from a regression of fetal mortality rates on the dummies for each 10th percentile (decile) of the Beijing's $PM_{2.5}$ $(t)$ distribution in 2009-2013. The omitted category is the 5th decile. Heating and non-heating seasons refer to the period from November to March, and from April to October, respectively by the public heating season in northern China from November to March.

## 2.B Appendix Tables

Table 2.B.1: Effects of Beijing $PM_{2.5\ t-1}$ on Daily Fetal Mortality Rates ($\geq$ 16 Weeks) in South Korea

| | **Daily Fetal Mortality Rates (≥16 Gestational Weeks)** | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| ***Panel A: All Regions*** | | | | | |
| Beijing $PM_{2.5}$ $(t-1)$ | 0.0920 | 0.102** | 0.128** | 0.128** | 0.141** |
| | (0.0561) | (0.0479) | (0.0548) | (0.0550) | (0.0569) |
| Observations | 36,020 | 36,020 | 36,020 | 36,020 | 36,020 |
| R-squared | 0.000 | 0.053 | 0.055 | 0.055 | 0.056 |
| Mean of Dep. Var. | 0.498 | 0.498 | 0.498 | 0.498 | 0.498 |
| ***Panel B: Region 1*** | | | | | |
| Beijing $PM_{2.5}$ $(t-1)$ | 0.0909 | 0.0992* | 0.107* | 0.107* | 0.127* |
| | (0.0683) | (0.0567) | (0.0633) | (0.0639) | (0.0668) |
| Observations | 24,008 | 24,008 | 24,008 | 24,008 | 24,008 |
| R-squared | 0.000 | 0.041 | 0.044 | 0.044 | 0.045 |
| Mean of Dep. Var. | 0.460 | 0.460 | 0.460 | 0.460 | 0.460 |
| ***Panel C: Region 2*** | | | | | |
| Beijing $PM_{2.5}$ $(t-1)$ | 0.0939 | 0.106 | 0.191* | 0.193* | 0.190* |
| | (0.0985) | (0.0882) | (0.107) | (0.106) | (0.108) |
| Observations | 12,012 | 12,012 | 12,012 | 12,012 | 12,012 |
| R-squared | 0.000 | 0.071 | 0.077 | 0.078 | 0.078 |
| Mean of Dep. Var. | 0.575 | 0.575 | 0.575 | 0.575 | 0.575 |
| **Controls** | | | | | |
| City FE | Yes | Yes | Yes | Yes | Yes |
| Temporal Controls | | Yes | Yes | Yes | Yes |
| Weather Controls | | | Yes | Yes | Yes |
| Linear time trend | | | | Yes | Yes |
| Quadratic time trend | | | | | Yes |

*Notes*: Regressions are weighted by the number of live births in a city. Robust standard errors are shown in parentheses clustered at the region-year-season level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 2.B.2: Effects of Beijing $PM_{2.5\ t-1}$ on Daily Fetal Mortality Rates ($\geq$ 28 Weeks) in South Korea

| | Daily Fetal Mortality Rates (≥28 Gestational Weeks) | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| ***Panel A: All Regions*** | | | | | |
| Beijing $PM_{2.5}$ $(t-1)$ | 0.00260 | -0.00121 | 0.0146 | 0.0145 | 0.0187 |
| | (0.0155) | (0.0154) | (0.0191) | (0.0192) | (0.0198) |
| Observations | 36,020 | 36,020 | 36,020 | 36,020 | 36,020 |
| R-squared | 0.000 | 0.012 | 0.015 | 0.015 | 0.016 |
| Mean of Dep. Var. | 0.070 | 0.070 | 0.070 | 0.070 | 0.070 |
| ***Panel B: Region 1*** | | | | | |
| Beijing $PM_{2.5}$ $(t-1)$ | 0.0140 | 0.0117 | 0.0289 | 0.0289 | 0.0334 |
| | (0.0199) | (0.0198) | (0.0237) | (0.0239) | (0.0248) |
| Observations | 24,008 | 24,008 | 24,008 | 24,008 | 24,008 |
| R-squared | 0.000 | 0.012 | 0.015 | 0.016 | 0.016 |
| Mean of Dep. Var. | 0.067 | 0.067 | 0.067 | 0.067 | 0.067 |
| ***Panel C: Region 2*** | | | | | |
| Beijing $PM_{2.5}$ $(t-1)$ | -0.0188 | -0.0255 | 0.0012 | 0.0015 | 0.0046 |
| | (0.0242) | (0.0242) | (0.0332) | (0.0331) | (0.0339) |
| Observations | 12,012 | 12,012 | 12,012 | 12,012 | 12,012 |
| R-squared | 0.000 | 0.014 | 0.021 | 0.022 | 0.022 |
| Mean of Dep. Var. | 0.076 | 0.076 | 0.076 | 0.076 | 0.076 |
| **Controls** | | | | | |
| City FE | Yes | Yes | Yes | Yes | Yes |
| Temporal Controls | | Yes | Yes | Yes | Yes |
| Weather Controls | | | Yes | Yes | Yes |
| Linear time trend | | | | Yes | Yes |
| Quadratic time trend | | | | | Yes |

*Notes*: Regressions are weighted by the number of live births in a city. Robust standard errors are shown in parentheses clustered at the region-year-season level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 2.B.3: Heterogeneous Effects of Beijing $PM_{2.5}$ on Daily Fetal Mortality Rates in South Korea by Distance from Beijing

| Dependent vars. | Fetal Mortality Rate | | | Prospective Fetal Mortality Rate | | |
|---|---|---|---|---|---|---|
| | $\geq$ 16 weeks | $\geq$ 20 weeks | $\geq$ 28 weeks | $\geq$ 16 weeks | $\geq$ 20 weeks | $\geq$ 28 weeks |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| ***Panel A1*** | | | | | | |
| Beijing $PM_{2.5}$ $(t-1)$ | 0.137** | 0.100** | 0.0335 | 0.0263** | 0.0254** | 0.0174* |
| | (0.0668) | (0.0493) | (0.0240) | (0.0123) | (0.0118) | (0.00969) |
| Beijing PM × Region 2 | 0.0131 | 0.0230 | -0.0412 | -0.00322 | 0.00241 | -0.0215 |
| (Distance>1000km) | (0.110) | (0.0721) | (0.0330) | (0.0212) | (0.0172) | (0.0140) |
| Observations | 36,020 | 36,020 | 36,020 | 33,668 | 33,668 | 33,668 |
| R-squared | 0.056 | 0.041 | 0.016 | 0.054 | 0.040 | 0.017 |
| Mean of Dep. Var. | 0.498 | 0.303 | 0.070 | 96.171 | 70.393 | 28.025 |
| ***Panel A2*** | | | | | | |
| Beijing $PM_{2.5}$ $(t-1)$ | -0.443 | -0.129 | 0.235* | -0.0670 | -0.0198 | 0.115* |
| | (0.503) | (0.310) | (0.139) | (0.0949) | (0.0736) | (0.0588) |
| Beijing PM × Distance (km) | 0.000573 | 0.000232 | -0.000212 | 0.0000903 | 0.0000452 | -0.0001034* |
| | (0.000492) | (0.000297) | (0.000132) | (0.000093) | (0.00007) | (0.000056) |
| Observations | 36,020 | 36,020 | 36,020 | 33,668 | 33,668 | 33,668 |
| R-squared | 0.056 | 0.041 | 0.016 | 0.054 | 0.040 | 0.017 |
| Mean of Dep. Var. | 0.498 | 0.303 | 0.070 | 96.171 | 70.393 | 28.025 |
| ***Panel B1*** | | | | | | |
| Beijing $PM_{2.5}$ $(t-2)$ | 0.0229 | 0.0338 | -0.0133 | -0.00164 | 0.00449 | -0.00467 |
| | (0.0588) | (0.0468) | (0.0244) | (0.0114) | (0.0114) | (0.0106) |
| Beijing PM × Region 2 | 0.00660 | 0.0244 | -0.0376 | 0.00329 | 0.00541 | -0.0221 |
| (Distance>1000km) | (0.0945) | (0.0776) | (0.0365) | (0.0187) | (0.0187) | (0.0155) |
| Observations | 35,991 | 35,991 | 35,991 | 33,639 | 33,639 | 33,639 |
| R-squared | 0.056 | 0.041 | 0.016 | 0.054 | 0.039 | 0.017 |
| Mean of Dep. Var. | 0.498 | 0.303 | 0.070 | 96.171 | 70.393 | 28.025 |
| ***Panel B2*** | | | | | | |
| Beijing $PM_{2.5}$ $(t-2)$ | -0.490 | -0.112 | 0.253 | -0.102 | -0.0339 | 0.130* |
| | (0.394) | (0.362) | (0.167) | (0.0767) | (0.0883) | (0.0702) |
| Beijing PM × Distance (km) | 0.000505 | 0.000152 | -0.000274* | 0.0000998 | 0.0000396 | -0.000140** |
| | (0.000379) | (0.000353) | (0.000162) | (0.000074) | (0.000086) | (0.000067) |
| Observations | 35,991 | 35,991 | 35,991 | 33,639 | 33,639 | 33,639 |
| R-squared | 0.056 | 0.041 | 0.016 | 0.054 | 0.039 | 0.017 |
| Mean of Dep. Var. | 0.498 | 0.303 | 0.070 | 96.171 | 70.393 | 28.025 |
| **Controls** | | | | | | |
| City FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Temporal Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Weather Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Linear time trend | Yes | Yes | Yes | Yes | Yes | Yes |
| Quadratic time trend | Yes | Yes | Yes | Yes | Yes | Yes |

*Notes*: Regressions are weighted by the number of live births in a city. Robust standard errors are shown in parentheses clustered at the region-year-season level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 2.B.4: Effects of Local $PM_{10\,t}$ on Daily Fetal Mortality Rates ($\geq$ 20 Gestational Weeks) in South Korea

| | Daily Fetal Mortality Rates (≥20 Weeks) | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| ***Panel A: All Regions*** | | | | | |
| Local $PM_{10}$ ($t$) | 0.225** | 0.00720 | -0.0132 | -0.0102 | -0.00884 |
| | (0.105) | (0.0998) | (0.117) | (0.118) | (0.119) |
| Observations | 38,325 | 38,325 | 38,296 | 38,296 | 38,296 |
| R-squared | 0.000 | 0.037 | 0.039 | 0.039 | 0.040 |
| Mean of Dep. Var. | 0.303 | 0.303 | 0.303 | 0.303 | 0.303 |
| ***Panel B: Region 1*** | | | | | |
| Local $PM_{10}$ ($t$) | 0.264* | 0.0810 | 0.00309 | 0.00433 | 0.0168 |
| | (0.135) | (0.126) | (0.149) | (0.149) | (0.151) |
| Observations | 25,554 | 25,554 | 25,532 | 25,532 | 25,532 |
| R-squared | 0.000 | 0.031 | 0.034 | 0.034 | 0.035 |
| Mean of Dep. Var. | 0.289 | 0.289 | 0.289 | 0.289 | 0.289 |
| ***Panel C: Region 2*** | | | | | |
| Local $PM_{10}$ ($t$) | 0.137 | -0.153 | -0.0157 | -0.0191 | -0.0387 |
| | (0.162) | (0.167) | (0.187) | (0.185) | (0.188) |
| Observations | 12,771 | 12,771 | 12,764 | 12,764 | 12,764 |
| R-squared | 0.000 | 0.047 | 0.052 | 0.053 | 0.054 |
| Mean of Dep. Var. | 0.330 | 0.330 | 0.330 | 0.330 | 0.330 |
| **Controls** | | | | | |
| City FE | Yes | Yes | Yes | Yes | Yes |
| Temporal Controls | | Yes | Yes | Yes | Yes |
| Weather Controls | | | Yes | Yes | Yes |
| Linear time trend | | | | Yes | Yes |
| Quadratic time trend | | | | | Yes |

*Notes*: Regressions are weighted by the number of live births in a city. Robust standard errors are shown in parentheses clustered at the region-year-season level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 2.B.5: Effects of Local $PM_{10\ t-1}$ on Daily Fetal Mortality Rates ($\geq$ 20 Gestational Weeks) in South Korea

| | Daily Fetal Mortality Rates (≥20 Weeks) | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| ***Panel A: All Regions*** | | | | | |
| Local $PM_{10}$ $(t-1)$ | 0.127 | 0.00494 | -0.0254 | -0.0255 | -0.0242 |
| | (0.0973) | (0.0998) | (0.104) | (0.104) | (0.104) |
| Observations | 38,296 | 38,296 | 38,296 | 38,296 | 38,296 |
| R-squared | 0.000 | 0.037 | 0.039 | 0.039 | 0.040 |
| Mean of Dep. Var. | 0.303 | 0.303 | 0.303 | 0.303 | 0.303 |
| ***Panel B: Region 1*** | | | | | |
| Local $PM_{10}$ $(t-1)$ | 0.165 | 0.0763 | 0.00130 | -0.000252 | 0.0117 |
| | (0.112) | (0.117) | (0.121) | (0.120) | (0.122) |
| Observations | 25,532 | 25,532 | 25,532 | 25,532 | 25,532 |
| R-squared | 0.000 | 0.031 | 0.034 | 0.034 | 0.035 |
| Mean of Dep. Var. | 0.289 | 0.289 | 0.289 | 0.289 | 0.289 |
| ***Panel C: Region 2*** | | | | | |
| Local $PM_{10}$ $(t-1)$ | 0.0407 | -0.150 | -0.0705 | -0.0761 | -0.0951 |
| | (0.189) | (0.185) | (0.207) | (0.205) | (0.204) |
| Observations | 12,764 | 12,764 | 12,764 | 12,764 | 12,764 |
| R-squared | 0.000 | 0.047 | 0.052 | 0.053 | 0.054 |
| Mean of Dep. Var. | 0.330 | 0.330 | 0.330 | 0.330 | 0.330 |
| **Controls** | | | | | |
| City FE | Yes | Yes | Yes | Yes | Yes |
| Temporal Controls | | Yes | Yes | Yes | Yes |
| Weather Controls | | | Yes | Yes | Yes |
| Linear time trend | | | | Yes | Yes |
| Quadratic time trend | | | | | Yes |

*Notes*: Regressions are weighted by the number of live births in a city. Robust standard errors are shown in parentheses clustered at the region-year-season level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

# Chapter 3

# Male Circumcision, Peer Effects, and Risk Compensation

with Hyuncheol Bryant Kim[1], Booyuel Kim[2], and Cristian Pop-Eleches[3]

[1]Department of Policy Analysis and Management, Cornell University, Email: hk788@cornell.edu

[2]KDI School of Public Policy and Management, Email: bkim@kdischool.ac.kr

[3]School of International and Public Affairs, Columbia University, Email: cp2124@columbia.edu

## 3.1 Introduction

HIV/AIDS is one of the world's most serious health challenges. Although HIV/AIDS treatment reached 8 million out of 34 million people living with HIV by the end of 2011, a 20-fold increase since 2003, HIV/AIDS prevention remains an important challenge since the number of new infections in 2011 was 2.5 million, only 20% lower than in 2001 (UNAIDS, 2013).

Recently, male circumcision has received much attention as an HIV prevention strategy after three efficacy trials showed that male circumcision can reduce HIV transmission risk by 50 percent (Auvert et al., 2005; Bailey et al., 2007; Gray et al., 2007). In addition, male circumcision also reduces herpes simplex virus type 2 (HSV-2) and human papillomavirus (HPV) infection (Tobian et al., 2009). To promote demand for male circumcision, the World Health Organization (WHO) strongly recommends male circumcision as a key strategy for reducing female to male transmission of HIV (WHO/UNAIDS, 2007), and there is a global mobilization for scaling up male circumcision especially in countries with high HIV incidence of heterosexually acquired HIV infection and low male circumcision rates.

However, there are two major concerns related to scaling up male circumcision: weak demand and potential risk compensation. First, the demand for male circumcision is still very low even with a heavily subsidized price and proper information (Chinkhumba et al., 2014). Among the major barriers discussed in the literature are financial constraints, lack of information, awareness, and accessibility, fear of pain, as well as religious and cultural norms. Second, even if a scale-up project is successful in increasing the take-up of male circumcision, such programs might have limited impacts if circumcised men are more likely to engage in risky sex behaviors (Cassell et al., 2006; Kalichman et al., 2007; Sawires et al., 2007; Mattson et al., 2008; Padian et al., 2008; Bingenheimer and Geronimus, 2009; Eaton and Kalichman, 2009; Brooks et al., 2010; Weiss et al., 2010).

This paper attempts to understand the role that peers play in the decision to get circumcised as well as the long-term impact of male circumcision on risky sexual behaviors. The motivation for this paper is twofold. One motivating factor is to understand peer effects, which have been shown to affect behavior in a wide range of areas including education (Sacerdote, 2001; Zimmerman, 2003; Foster, 2006; Lyle, 2007; Kremer and Levy, 2008), health (Miguel and Kremer, 2004; Godlonton and Thornton, 2012; Oster and Thornton, 2012; Chong et al., 2013) and labor (Mas and Moretti, 2009; Bandiera et al., 2010). More specifically there is interest among researchers and policy makers to understand the role that peer effects may play in the take-up of certain interventions or the adoption of certain technologies. This is because takeup is often considered suboptimal or because potential spillovers to peers might justify subsidizing an intervention (Miguel and Kremer, 2004).

The second motivation is to understand whether in addition to the direct biological effects of male circumcision on the probability of transmission of HIV, there are also behavioral responses that could reduce or amplify the impact of getting circumcised. One concern is that circumcision leads to risk compensation if individuals overestimate the protection that male circumcision provides and therefore engage in riskier sexual practices (WHO/UNAIDS, 2007).[4] Moreover, the settings of scale-up projects are different from those in the original efficacy trials in many ways, and thus behavioral responses could also be also different. First, current beliefs and attitudes to circumcision could be substantially different from those in the efficacy trials since biological impacts of male circumcision on HIV and HSV-2 infections were not established. Second, efficacy trials were combined with intensive health education

[4]Efficacy trials do not consistently show a relationship between male circumcision and risk compensation behaviors. For example, the studies by Bailey et al. (2007) in Kenya and Auvert et al. (2005) in South Africa find some evidence of risk compensation while Gray et al. (2007) in Uganda do not find such evidence. More importantly, impacts of male circumcision on HIV and HSV-2 infection in the scale-up setting could be different from those in efficacy trials due to the following reasons. First, a long-term change in risk compensation behaviors may mitigate a protective effect against HIV and HSV-2 infection. Actually, these studies were closed early by the safety monitoring board and provided male circumcision to the control group when an interim analysis proved the protective effect of male circumcision against HIV infection, and therefore risk compensation behaviors in the long-term cannot be evaluated. Risk compensation in these studies could be underestimated. Several studies try to explore risk compensation in the long-term by comparing risk sexual behaviors between circumcised and uncircumcised men in the post-trial follow-up of the trial in Uganda (Kong et al., 2012; Gray et al., 2012) and Kenya (Mattson et al., 2008; Wilson et al., 2014). These do not find compelling evidence of increased risky sexual behaviors among circumcised men, however, these results may be driven by confounders that may affect both circumcision take-up and risky sexual behaviors.

and individual HIV counselling throughout all the follow-up visits up to 24 months. This suggests that risk compensation in original RCT may be underestimated compared to the current male circumcision scale-up projects where knowledge of the benefits of male circumcision is widely accepted and follow-up HIV counseling is often limited. Third, as take-up of male circumcision is not as high as that in the clinical trial, circumcision takers are self-selected in non-study scale up project setting. As a result, those who decided to take-up male circumcision could be different from those who do not, and therefore the impacts of male circumcision in scale up project could be also different from efficacy trials.

In our analyses, we use a four-year long follow-up of an intervention based on a two-step randomized design within classrooms in secondary schools in Malawi in order to: (1) estimate the take-up of male circumcision from the provision of free male circumcision and transportation vouchers, (2) understand the role that peers within classrooms play in the take-up of male circumcision among students not assigned initially to treatment, (3) measure risk compensation from a 4 year follow-up that collected biomarkers of sexually transmitted diseases and (4) understand if the risk compensation is different for males who take up circumcision because of the direct inducement of the intervention or the indirect peer effect.

We take advantage of a large-scale HIV prevention program implemented in 124 classrooms in 33 public secondary schools in Malawi by the Africa Future Foundation (AFF), an international non-governmental organization (NGO). AFF provided free male circumcision at the assigned clinic, and randomized the intensity of the transportation support for the surgery. Classrooms were randomly assigned into three groups: *100% Treatment*, *50% Treatment*, or *No Treatment* classrooms. All male students in *100% Treatment* classrooms received a free male circumcision offer with *stronger* transportation support in the first round, while no students in *No Treatment* classroom received the offer in the first round, but received *weaker* transportation support offer in the second round. In 50% Treatment classrooms, half of the students were randomly selected to receive the *stronger* transportation support in the first round. In the first round of the intervention, free male circumcision surgery and transportation subsidies were provided to a total of 1,972 male students in 2012. During the first round, all students including those untreated are allowed to take-up free male circumcision at the assigned hospital. In the second round, the remaining 2,002 male students who were temporarily untreated in the first round received the offer with weaker transportation support. This was due to funding constraints of the collaborating NGO and it resulted in less intense intervention.

We attempted to interview all students for a short-term follow-up at the end of the first round and prior to the start of the second round. In addition, the two youngest cohorts (9th and 10th graders in 2011) were selected for the long-term follow-up that was implemented after four years from the baseline survey. The main advantages of our setting and data are the relatively long follow-up period, the randomized two-step design that allows the estimation of the direct and peer effects of the intervention, an administrative dataset of all the circumcisions performed at the only provider in the study area, and the collection of biomarkers for sexually transmitted diseases (HIV and HSV-2).

We begin our analysis using the administrative data collected at the end of the first round to understand take-up of male circumcision. We find that male students who received a transportation incentive are around three times more likely to take-up male circumcision. Next, we also use our administrative data to understand the influence of peers on the decision to get circumcised. Untreated students in 50% Treatment classrooms were 79% more likely to get circumcised than students in No treatment classrooms, suggesting a positive externality in the demand for male circumcision. Since at baseline we also collected a roster of friends within the classroom, we also use an alternative empirical strategy that uses the experimental variation in the fraction of peers that was offered treatments to further understand peer effects. The results using this method to capture peer effects are also positive but less precise. However, we find evidence of important reinforcement effects when close friends within the same classroom receive the intervention together. The effects described above persist even in the long run after treatment is offered to everybody in the study, a finding that is consistent with the shorter and less intense intervention in the second stage.

Having established that there are positive direct and (indirect) peer effects in the decision to adopt the take-up of male circumcision, we analyze the long-term impact of male circumcision on biomarkers as well as risky sexual behaviors. In a nutshell, we find evidence that is consistent with risk compensation among those who received the more intense transportation support. Our main results show an increase in STIs when measuring specific IgG and IgM antibodies to HSV-2. These results are corroborated with other measures of risky sexual behaviors that we have implemented, including self-reports of unprotected sex, the purchase of condoms when offered during follow-up (as in Thornton, 2008) and the measurement of reported sexual behavior using the item count technique (as in Coffman et al., 2013).

Our third main finding is that risky sexual behavior among those who decide to get circumcised because of peer effects are different from those who take-up circumcision because they received were allocated to the treatment group. In our setting, when a boy gets circumcised as a result of peer pressure, we do not observe any evidence of compensating behavior. We further explore this striking difference by analyzing observable characteristics of compliers in the two groups that get circumcised and somewhat surprisingly do not find a consistent pattern of observable characteristics (such as self-declared safe sex practices) of the two groups of male students who select into treatment based on the direct intervention channel or the indirect peer effect channel.

Our findings contribute to four main strands of the literature. First, our paper contributes to the literature on peer effects and in particular to the literature on the role of peer effects in health intervention programs in developing countries. For example, Godlonton and Thornton (2012) find significant peer effect in the take up of HIV testing in Malawi. A 10 percent point increase in the probability of having a neighbor within 0.5 km learning his/her HIV result leads to a 1.1 percent point increase in learning one's own HIV result. Chong et al. (2013) also find that the treatment effects of online sexual education are largest when the peers were treated together. Oster and Thornton (2012) also find a strong positive peer effect on take-up of new health technology (menstruation cup) in the short term.

Second, our work contributes to the literature on how to increase male circumcision take-up. For example, previous trials show that financial compensation (Thirumurthy et al., 2014) or educating religious leaders can promote circumcision take-up significantly (Downs et al., 2017).

Third, our paper contributes to the risk compensation literature. Risk compensation behaviors are well known in the economic literature as the "Pelzman effect", an example of moral hazard. For example, risk compensation in seat belt (Peltzman, 1975; Evans and Graham, 1991) and bicycle helmet (Thompson et al., 2001) are well documented. As mentioned earlier, the early efficiency trials (Auvert et al., 2005; Bailey et al., 2007; Gray et al., 2007) find mixed evidence on risk compensation but these studies are based on short term and self-reported outcomes. To the best of our knowledge, our study is the first to show causal effects of male circumcision on risk compensation using biomarkers and a long term follow-up.

Finally, and maybe most importantly are the implications of different long-term behaviors for male students who take up circumcision directly through the program or indirectly through peer effects. Our reading of the literature on the role of peer effects in the take-up of technologies or interventions is that most studies aim to understand short-term take-up decisions but then do not perform long-term follow-up studies to validate that such interventions have the desired outcomes. Our results, imply that at least in our setting, this is an important issue and that takeup of male circumcision through two different mechanisms (direct inducement of transport intervention and indirect inducement through peer effect) can lead to very different behavioral responses.

The chapter is organized as follows. Section 3.2 provides background information on the

Malawian context and the male circumcision intervention. In Section 3.3 we describe the data followed in section 3.4 by the empirical strategies. Section 3.5 presents the main results. The final section presents conclusions.

## 3.2 Background and Experimental Design

### 3.2.1 HIV and Male Circumcision in Malawi

Malawi has been heavily affected by the HIV pandemic with an HIV prevalence rate of 10.6% for people aged 15 to 54 (National Statistical Office Malawi and ICF Macro, 2011) and a life expectancy at birth of 55 years (United Nations Population Fund, 2012). The prevalence of male circumcision in Malawi is relatively low with 21.6% of men being circumcised (National Statistical Office Malawi and ICF Macro, 2011).[5] Male circumcision is practiced mainly for religious and cultural reasons: 93.3% of Muslims are circumcised while only 11.6% of Christians practice circumcision (Bengo et al., 2010). Culturally, 86.8% of those belonging to the Yao tribe practice male circumcision while other tribes have low levels of male circumcision (National Statistical Office Malawi and ICF Macro, 2011). The baseline circumcision rate in our sample is 10.5% which reflects the fact that most residents in the catchment area of our study are non-muslim from non-Yao tribes.[6]

Medical male circumcision for HIV prevention has become an important component of Malawi's national HIV prevention program since 2010 (NAC, 2012). Nevertheless, despite recent growth, the number of medical male circumcisions was still small at the time when our study was implemented in 2012.[7] The estimated number of male circumcisions between

[5]The true prevalence of complete male circumcision could be lower because many of those reporting being circumcised practice incomplete circumcision which only removes part of the foreskin. Incomplete circumcision may not have the full protective benefits of male circumcision (Bengo et al., 2010).

[6]Most people belonging to the Yao tribe live in the southern region of Malawi, while the majority of ethnicity in the central region to which the project catchment is belonging is the Chewa. Chewa people consist of 34.1% of the total population and only 6.2% of them practice male circumcision according to MDHS (2011).

[7]Starting from very small numbers (589 in 2008, 1,234 in 2009, and 1,296 in 2010), there was a sizeable

2008 and 2011 was only 0.7% of the target number (2.1 million) needed to achieve 80% of male circumcision prevalence in Malawi.

### 3.2.2 Study Setting and Experimental Design

This study was conducted in the context of larger HIV prevention project implemented in public secondary schools in four catchment districts of Lilongwe by the Africa Future Foundation.[8] The target population of the study is male secondary school students identified through a school-based baseline survey. Our sample includes 3,970 male students enrolled in 9th to 11th grade at 124 classrooms in 33 secondary schools in the catchment area. Due to the budget constraints of the implementing NGO, for the long-term follow-up survey, we focus only on the 2,663 students who were 9th and 10th graders at the time of enrollment. In terms of their organizational structure, most schools (28 out of 33) have one class per grade and there are limited cross-grade school activities.

In order to increase the demand for male circumcision among teenagers, AFF began an initiative to encourage the take-up of male circumcision in the fall of 2011. First, it established a medical male circumcision clinic within the premises of the Daeyang Luke Hospital located in the catchment area of the study. The clinic was equipped with a self-contained surgical unit and surgical beds to perform modern and safe medical circumcisions performed by appropriately trained medical personnel.[9]

Table 3.1 summarizes the two-stage randomization phase-in design of the study. Figure 3.A.1 presents detailed information on each randomization process. 124 available classrooms within the 33 schools were stratified by grade and randomly assigned into three groups: *100% Treatment*, *50% Treatment*, and *No Treatment* classrooms. All male students in the

increase in 2011, when 11,881 people became circumcised (World Health Orgarnization , WHO).

[8] AFF's catchment area includes the four districts are Chimutu, Chitukula, Tsbango, and Kalumba.

[9] The medical male circumcision surgeries were performed under local anesthesia in the assigned clinic by medical personnel using the standard forceps-guided method.

100% Treatment classrooms received the male circumcision offer with transportation subsidies during the first round of treatment (Group 1). No students in the No Treatment classrooms received an offer during the first round of the treatment (Group 4). In 50% Treatment classroom, we randomly selected half of students for the treatment (Group 2), and the remaining students were not treated during the first round of the treatment (Group 3). Assignment to the treatment in 50% Treatment classrooms was done at the individual level.[10] This two stage experimental design that randomizes treatment both across and within classes allows us to measure not only the direct effect but also peer effect of the male circumcision offer. As summarized in Table 3.1, 41 classrooms (across 24 schools) were assigned to 100% Treatment, 41 classrooms (across 25 schools) were assigned to 50% Treatment, and 42 classrooms (across 28 schools) were assigned to the No Treatment group.[11]

The timeline of the project was as follows: after the baseline survey that was collected between October 2011 and May 2012, male students assigned to treatment in Groups 1 and 2 received the transportation subsidies during December 2011 to April 2013 (Round 1). The short-term followup survey was conducted between January and June of 2013. Since our research design was based on a phase-in design, at the time of the short-term follow-up survey, the untreated students during Round 1 (Groups 3 and 4) had not received the transportation support yet. Groups 3 and 4 received less intensive transportation subsidies (than Groups 1 and 2) from July to December 2013 (Round 2). A long-term follow-up survey for 9th and 10th graders was implemented approximately 4 years after the baseline survey between October 2015 and August 2016.

After the baseline survey, AFF provided to the selected students (Groups 1 and 2) a male circumcision offer that consisted of free access to the surgery, complication check-ups, and

[10]AFF wanted to keep the demand constant due to the capacity constraints. During the study period, maximum number of surgery per day was 19.

[11]This standard two-step randomized design to estimate peer effects was first proposed in Duflo and Saez (2003).

transportation support to the clinic.[12] Students receiving the offer could choose either a direct pick-up service provided by AFF from the school to the clinic or a transportation voucher that was reimbursed after the circumcision surgery at the hospital was performed. The amount of the voucher varied according to the distance between the clinic and student's school.[13] These vouchers were student specific and student participants were not allowed to trade these vouchers among themselves.

The intensity of the transportation support in Round 1 and 2 is different. The transportation vouchers for independent travel to the clinic was give out twice during Round 1. The first voucher did not specify an expiration date but specified six possible dates available to use for the pick-up service to the clinic (Panel A of Figure 3.A.3). The second voucher specified an expiration date (the end date of Round 1) and one possible pick-up date (Panel B of Figure 3.A.3). In the beginning of Round 2, Groups 3 and 4 were also given the offer. However, the Round 2 treatment was less intense due to budgetary constraints that were not anticipated at the start of the project. The treatment lasted only for 6 month and the travel vouchers were offered only once (Panel C of Figure 3.A.3).[14]

The male circumcision surgeries were performed in the assigned clinic by medical personnel. Post-circumcision care was provided after three days and one week at the student's school to check and disinfect the wound, and record complications. In pre- and post-surgical counseling sessions, participants were advised that male circumcision provides only partial protection against HIV infection and thus practicing safe sex after the surgery was encour-

[12]Those already circumcised and those with HIV, severe anemia, or penis abnormality such as hypospadias were not eligible for the medical male circumcision procedure. During the study period, 394 study participants received male circumcision at the assigned clinic and 12 people were refused to receive the procedure; Seven were already circumcised; four had penis abnormality, and one was HIV positive.

[13]Although we set the transportation voucher amounts to reflect the minimum public transportation fees, many of rural areas do not have access to public transportation and students who live in rural areas often walked to the hospital when they chose the voucher option.

[14]One may worry that the timing of the circumcision offers are also different between the treatment and control group, and it could be problematic if risk compensation pattern is non-linear over time. We address this concern using biomarkers that captures lifetime and recent infections separately, which is discussed in Section 5.

aged and recommended (Figure 3.A.2).

## 3.3 Data

We use four main sources of data for this research: baseline, short term and long term follow survey data as well as administrative data from the circumcision clinic. At baseline, we started with a list of 10,715 enrolled students at the 33 participating schools. We managed to successfully get consent and complete the baseline survey for 74.4% of the students in the school roll-call lists. Of the 7,971 secondary students who completed the baseline, 3,997 were girls and 3,974 were male students.

The baseline survey was designed to measure detailed background characteristics including student information about HIV knowledge, sexual behaviors, and their friendship network. At the end of the survey we gave students 10 kwacha (6 cents), and offered them a chance to buy condoms at the subsidized price of 5 kwacha to measure the demand for safe sex as in Thornton (2008).

In Table 3.2 we present summary statistics and balance tests of baseline observables. Columns (1) to (4) and (5) to (9) of Table 3.2 refer to the full sample (9th to 11th graders) and the long-term follow-up sample (9th and 10th graders), respectively. In Columns (1) and (5), we show the mean characteristics for those in the control group. As shown in Column (1), the average age of study participants in the control group is 16.8 years and 16% belong to ethnic groups that practice circumcision and 10% are already circumcised. In general, study participants showed a high level of HIV/AIDS knowledge: the average number of correct answers to the HIV/AIDS knowledge questionnaire was 17.4 out of 20 questions. Nevertheless, they have a relatively low knowledge on the medical benefit of male circumcision (63.9%). In addition, 36% of control group study participants believed that male circumcision is painful and 15% believe that that male circumcision is only for Muslim. 31% ever had sex and 9% are currently engaged in a sexual relationship. Columns (2) to (4) and (6) to (9) present mean differences between treatment the treatment groups (Groups 1 to 3) and the control group (Group 4). The results confirm that classroom randomization was well implemented: the proportion of statistically significant mean difference at the 10% significance level is 9 out of 60 (15%) in full sample (Columns (2) to (4)) and 9 out of 60 (15%) in long-term follow-up sample (Columns (6) to (9)).

We also asked students to list three best friends within the classroom regardless of sex, but in the analysis, we reconstructed friendship data by reordering best male friends after excluding friends without baseline survey and female friends. Table 3.B.1 shows summary statistics for the friendship networks. In Panel A, Column (1) includes the original friendship network data, and Column (3) presents reordered male friends after excluding female friends and those who did not participate in baseline survey. Almost all (3,832 out of 3,844) have at least one male friend, and around 80% (3,135 out of 3,844) have two male friends among their three best friends. Panel B presents the network treatment distribution among eligible male friends. It shows substantial variation in the fraction of best male friends who got treated and the fraction of treated friends is well balanced across the baseline characteristics (Table 3.B.2).

Our main outcome variables are male circumcision take-up rates, self-reported sexual behaviors, and HIV and HSV2 biomarker test results.[15,16] One advantage of our setting is that we do not need to rely on self-reported data, which is easily affected by social desirability bias, on circumcision take-up since we collected a hospital administrative dataset containing the complete records of all the medical circumcisions performed at the assigned clinic.[17] It

[15]HIV serostatus was measured with Determine HIV 1/2 and Unigold Recombigen HIV test. Different results between the two HIV tests were not found in this study. Participants who tested HIV positive were advised to receive proper treatment at the collaborating hospital.

[16]Biological data was not collected at baseline because the NGO was concerned that students who tested HIV positive at school may affect their ability to continue their studies. Instead, HIV testing were done when students came to the clinic for the circumcision.

[17]It is unlikely for students to get circumcised in other medical facilities because there are few facilities nearby that provide male circumcision on a regular basis. One exception is the Banja La Mtsogolo (BLM) clinic located in the Chitukula district, which is one of our four catchment districts. However, this clinic

records the name and age of the patient as well as the medical record related to the procedure, which includes the date of the surgery, the dates of any follow-up visits as well as information on possible side effects related to the procedure. In addition, and very important for our study, the administrative record also records the voucher used for either the direct pick-up service or the transport voucher used for public transportation. During the study period, 502 were circumcised in the first round and 124 in the second, which likely reflects the lower intensity of treatment in the second round.[18]

We also use the short- and long-term follow-up surveys implemented one and four years after the baseline survey to measure sexual behaviors, and sexually transmitted diseases (HIV and HSV-2) infection. The long-term follow-up survey, focusing only 9th and 10th graders, included several novel measures of sexual behaviors and STDs infection, which will be the main outcomes variables used to analyze the impact of the circumcision intervention. In terms of biomarkers, HSV-2 and HIV infections were evaluated using rapid test kits. For HSV-2, we use two measures of infection: the IgG is a permanent marker of HSV-2 infection that has occurred at any point during a person's lifetime (Obasi et al., 1999) while IgM shows recent infection (Workowski and Bolan, 2015).[19,20]

Our main outcome for measuring risky behavior is the HSV-2 test for Sexually Transmitted Diseases (STDs), a test that has been used successfully in a number of related studies (Baird et al., 2012; Duflo et al., 2015). Since HSV-2 is almost exclusively sexually transmitted through sexual, our rapid test results could be a good measure for risky sexual behaviors.

---

charged around $10 for the surgery and complication check-ups, which should seriously limit the demand for the circumcision procedure, especially for the population of secondary school students.

[18]In the second round, 2,002 male students from Group 3 and Group 4 were offered the transportation treatment, a number that is roughly similar to the 1,972 students from Group 1 and 2 who received it in the first round.

[19]Although we have implemented a test for HIV during the follow-up, given the low incidence of HIV, our power calculations indicated that our study was not powered to detect differences in HIV rates between treatment and control groups.

[20]Misclassification of HSV-2 status might have occurred since there was no supplementary testing. However, all IgM positive cases are also IgG positive in our study sample. Moreover, there is no reason that misclassification of HSV-2 would be systematically different across the treatment status.

Our preferred test is the one for IgG antibodies because the prevalence in our setting is high compared to other types of STIs. One potential complication with using the results from an HSV-2 and HIV test to capture changes in risky sexual behavior after circumcision is that male circumcision prevents HSV-2 and HIV infections. Therefore, the increase of HSV-2 infection we estimate is a lower bound of the true effects induced by risk compensation.

In addition, we implemented an alternative measure of risky sexual behavior using the item count technique (ICT) (Miller, 1984; Coffman et al., 2013). This is to further account for potential measurement error in self-reported responses, since numerous studies point out that self-reported answers are poor proxies for true attitudes toward private and sensitive subjects like sexual activities (Palen et al., 2008; Minnis et al., 2009) The ICT methodology asks respondents to report the total number of true statements in a set of questions that may include a sensitive item, instead of directly endorsing it. Participants were randomly given one of two sets of questions. One set included only four non-sensitive items (*short-form*) and the other set included four nonsensitive items and an extra sensitive statement of our interest (*long-form*). We used two sensitive questions of interest: "I think I have to use a condom in case of sex with somebody that I do not know well" (Using condoms for casual sex) and "I had sex with more than two people in last 12 months" (Multiple sexual partners). We can measure the impact of treatment on the sensitive statement by estimating a differential proportion of respondents who agree with the sensitive statement between the treatment and control group. The details of the ICT methodology are described in the Appendix Figure 3.A.4. Lastly, at the end of the survey we again gave students 10 kwacha (6 cents), and offered to sell them condoms at a subsidized price of 5 kwacha to measure the demand for condoms.

In the baseline and two follow-up surveys we also included questions that are typically asked in surveys that are based on self-reported sexual behaviors. These include questions about attitude toward condom, inconsistent use of condom, unprotected sex with recent a partner and multiple sexual partner.

We conclude this section with an analysis of attrition rates in the follow-up surveys and biomarker test participation. We were able to track about 91.9% and 86.8% of study subjects during the first and second follow-up surveys, respectively.[21] A smaller percentage (78.2%) participated in the bio-marker test. Table 3.B.3 suggests that attrition from the sample is not correlated with treatment assignment, with one exception in G2 at the first follow-up. In addition, there is no statistically significant differences between the treatment group (F-tests results).

## 3.4 Estimation Strategy

We employ a number of empirical strategies to capture the direct effect of being assigned to the more intense male circumcision intervention as well as possible peer effects in this setting. Our first empirical strategy estimates the following model:

$$Y_{ijk} = \beta_0 + \beta_1 G1_{ijk} + \beta_2 G2_{ijk} + \beta_3 G3_{ijk} + \gamma' X_{ij} + \delta_k + \varepsilon_{ijk} \quad (3.1)$$

where $Y_{ijk}$ denotes an outcome of interest such as male circumcision take-up, sexual behaviors, and bio-marker test results for individual student $i$ in classroom $j$ at school $k$. $G1$, $G2$, and $G3$ refer Group 1, Group 2, and Group 3, respectively. The control vector, $X$, includes age, circumcising ethnicity, circumcising religion (Muslim), orphan status, parent's education, parent's job, household assets and school type, and the assignments to HIV/AIDS

[21] While the long-term follow-up survey were implemented for entire baseline 9th and 10th graders, the data collection for the short-run follow-up was conducted in two stages. In the first stage, we revisited the schools in our sample and attempted to re-interview all the students in our initial sample. We successfully interviewed 67.9% of the baseline sample students (or 2,698 students) using interviews conducted at school. During the second stage, we implemented a more extensive tracking exercise as in Thomas et al. (2001) and Thomas et al. (2012). First, we selected a subsample of 15% of the students (191 students) among those who did not participate in the school follow-up survey and attempted to visit them at home to complete the survey. The home survey follow-up rate was 74.9%. Combining the school and home visits results in an effective survey follow-up rate of 91.9%. The effective survey rate (ESR) is a function of the regular school follow-up rate (RFR) and intensive home-visit follow-up rate (HFR) as follows: ESR = RFR + (1-RFR)×HFR. Overall, ESR is 91.9% (67.9% + 32.1%×74.9%). We run a weighted regression with a weight 6.67 for home-visit survey since we randomly selected 15% students from the attrition sample (Baird et al., 2012).

education and girl's CCT program. $\delta$ is school fixed effects, and $\varepsilon$ is a random error.[22] Errors are clustered at the classroom level. We also present heterogeneous treatment effects by three different priors such as knowledge on the medical benefit of male circumcision, fear of pain, and religious norms.

In these specifications, both $\beta_1$ and $\beta_2$ capture the direct effects of being assigned in Group 1 and Group 2. The difference between $\beta_1$ and $\beta_2$ also captures possible peer effects, given that Group 2 has peer who did not received the treatment in the classroom (Group 3) while everybody in Group 1 received the treatment. In addition, $\beta_3$ might capture possible peer effects, given that a key difference between Group 3 and Group 4 is existence of peer who received the treatment (Group 2) within the classroom.

In some of our specifications, we also restrict ourselves only to the 50% Treatment classrooms. An alternative way to estimate the main direct effect of the intervention is by comparing the difference in outcomes between students in group G2 who received the intervention and students in group G3 who did not receive the intervention during the first round. Both of these groups (G2 and G3) are within the same classrooms and contain students who are exposed to the same peer effects since 50% of their peers are treated.

Next, we extend the analysis of peer effects by using our experimental variation in treatment intensity with the data from the friendship rosters. We try to measure these effects in the restricted sample of 50% Treatment classrooms because there is no within class variation in 100% Treatment and No Treatment classrooms. These following linear regressions are estimated:

$$Y_{ijk} = \alpha_0 + \alpha_1 G2_{ijk} + \alpha_2 Peer_{ijk} + \gamma' X_{ij} + \Psi_k + \varepsilon_{ijk} \tag{3.2}$$

$$Y_{ijk} = \kappa_0 + \kappa_1 G2_{ijk} + \kappa_2 Peer_{ijk} + \kappa_3 (G2_{ijk} \cdot Peer_{ijk}) + \gamma' X_{ij} + \Psi_k + \varepsilon_{ijk} \tag{3.3}$$

where $Peer_{ijk}$ is an indicator variable taking value 1 if a male friend on the friendship roster was offered treatment in Round 1. The unit of observation is a student-friend pair, so a student

[22] Our results are almost identical when we include grade fixed effects instead of school fixed effects.

who indicated three male friends in his classroom as his best friends enters the analysis three times. $\Psi$ is classroom fixed effect and errors are clustered at the classroom level. Since receiving the offer within our 50% Treatment classrooms is randomly assigned, whether a best received the treatment in the classroom is also random. It is worth noting that the peer effect that $\alpha_2$ in Equation (2) captures is different from the peer effect $\beta_3$ in Equation (1) since one captures possible peer effects coming from the smaller group of best friends and the other reflects peer effects from the larger group of classroom students. Even though we did not explicitly provide information on free male circumcision opportunity to both Group 3 and Group 4 in the first round, Group 3 can find this opportunity more easily than Group 4 due to their peer (Group 2).

In Equation (3), we estimate another type of peer effect, resulting from potential complementarities between your offer and his friend's offers. This type of peer effect could also be defined as a reinforcement effect, and is captured by the coefficient of the interaction term $\kappa_3$. In our setting, it is certainly possible for such reinforcement effects to be present, if peers make a decision to get circumcised jointly.

In our setting, the role of peer effects in the demand for male circumcision is theoretically ambiguous. It may be positive if friends provide emotional support that reduces the psychological cost or share private information about the benefits of the circumcision procedure. Alternatively, a negative experience following male circumcision (i.e., complication or pain) might decrease a friend's demand for male circumcision.

## 3.5 Results

### 3.5.1 Impacts of intervention on the take-up of male circumcision

Figure 3.1 shows the cumulative prevalence of male circumcision over time based on hospital records data and self-reported data in the follow-up. The solid lines present take-up based on hospital administrative data and the dotted lines are based on self-reported data. We rely more on hospital records data since self-reported data may suffer from recall bias.[23] Hospital administration data is not available after December 2013 because the clinic was closed after the program was ended by the NGO. At the beginning of the project, male circumcision prevalence was about 11.0%. Due to the difference in intensity of transportation support, take-up among treated students in 100% Treatment (G1) and 50% Treatment classrooms (G2) significantly increased during Round 1. Untreated students in 50% Treatment classroom (G3) were also more likely to take-up male circumcision during Round 1 than the control group (G4). 249 (18.6%) out of 1,342 students in G1 and G2 and 64 (7.5%) out of 851 students in G3 and G4 were additionally circumcised at the assigned hospital during the study period.

In Panel A of Table 3.3 we present estimates of the impact on the intervention on the demand for male circumcision that are based on Equation (1), while in Panel B we combine the two treatment arms (G1 and G2) in order to increase statistical power. The dependent variables are male circumcision take-up from hospital administration data at the end of the first round (Columns (1)-(2) and (5)-(6)), and the second round (Columns (3)-(4) and (7)-(8)). In Columns (1)-(4) we use the full sample of 9-11th graders, although our preferred estimates are those in Columns (5)-(8) which are restricted to the long term follow-up cohort of 9th and 10th graders. The results confirm that the intensive support intervention for male circumcision significantly increased the demand for male circumcision. At the end of Round 1, G1 and G2 were 14.1 and 16.9 percentage points (290% and 363%) more likely to get circumcised than those in control classrooms (G4), respectively (Columns (5)-(6)). These effects are sustained until the end of Round 2 (Columns (7)-(8)), although they become smaller, which is not surprising given the phase-in design. We find similar pattern for full sample (Columns (1)-(4)). In Panel B of Table 3 we show that for our preferred specification when we use the pooled treatment groups (G1 and G2), the demand for circumcision for being assigned

[23]For example, only 17% and 43 % of those who circumcised at the study hospital recall the month-year and year of the circumcision at the time of the follow-up survey, respectively.

to treatment in Round 1 increases by 15% at the end of Round 1 and by 10% at the end of Round 2.

We also test whether the take-up of male circumcision is heterogeneous by prior beliefs such as knowledge about the benefits of male circumcision, beliefs about how painful male circumcision is, and beliefs that male circumcision is only appropriate for Muslim. Columns (1) to (6) of Table 3.B.4 compares the G1, G2, and G3 groups with the control group. In general, we find that those who think that male circumcision is painful or that it is only for appropriate for Muslim are less likely to receive MC in general. The interactions of these variables and the male circumcision variable are also negative, suggesting that individuals with these prior beliefs are less responsive to the circumcision intervention, although we note that many of the coefficients in this table are imprecisely estimated.

### 3.5.2 Peer effects on male circumcision take-up

As discussed in Section 4, we estimate different types of peer effects based on Equations (1), (2), and (3). We start with results presented in Panel A of Table 3.3 that use Equation (1). The G3 untreated students in 50% Treatment classrooms, were depending on specifications between 3.2% and 4.2% points more likely to receive male circumcision than the G4 (No treatment) classroom students. Moreover, this effect is getting larger after Round 2 an increase by 6-7%. We interpret this result as the spillover effects within the classroom from the other half of classmates who received the offer of male circumcision in Round 1. Moreover, this increase persists (and even increases) at the end of the study period. (Panel B1). A second way to look at the existence of spillovers is to test whether take-up rates of G1 and G2 are different given that everybody in G1's classroom are treated in Round 1 while G2 had peer who did not received an offer initially. Somewhat surprisingly, we do not evidence of spillovers using these two groups. The difference in take up is not statistically significant and if anything, the take-up rate is larger in G2 than G1 by 2.6% in Round 1 and .8% in Round

2.[24]

Another way to test for the existence of peer effects is based on Equation (2) and (3) using the restricted sample of 50% Treatment classroom (G2 and G3). In this analysis, we take advantage of the details of the friend networks. As Table 3.4 shows, we find that having a higher proportion of friends who are treated increases one's circumcision take up in general (Panel B and C). However, the coefficient is not statistically significant. Panel D provides evidence of a complementarity between a student's offer and his friend's offer that increase a student's take-up of circumcision. The impacts are large, statistically significant, and robust across the specifications, which suggests the existence of important reinforcement effects among peers in school.

Lastly, we study the role of popular kids in the classroom on male circumcision take-up. Most popular kids are defined as a person who had the largest number of classmates who claim him to be one of the top three closest friend. Column (1) of Table 3.B.5 shows that an experimentally induced male circumcision offer to the most popular students in the classroom decreases male circumcision take-up (Panel A), especially when the most popular one thinks male circumcision is painful.

### 3.5.3 Mechanism of peer effects

In this section, we try understand the mechanisms through which peer effects promote the demand for male circumcision by taking advantage of our detailed hospital administrative data. We explore two channels through which peer effects might work in our setting. The first is that students might organize each other to come to the hospital together. We capture this channel when students are coming to the hospital on the same day as their friend. Another possibility which we call the social learning channel, arises in a situation when a student who

[24]One potential explanation of this (non-significant) difference might be the scarcity heuristic argument, which states that when a resource is less readily available people are more likely to perceive it as more valuable (Cialdini, 2009).

has experienced the male circumcision procedure influences his friend to take-up circumcision. We capture this channel when students and their friends receive the male circumcision on different days.[25]

Table 3.5 uses the stacked 50% treatment classroom sample. In this analysis, the unit of observation is a single friendship relationship. The dependent variable in Column 1 is male circumcision take-up, and the dependent variables in Columns (2) to (5) are take up of male circumcision without a friend's take up, concurrently with a friend, and before or after my friend respectively. These last three categories are mutually exclusive, and thus the sum of these coefficient is the same to the coefficient in Column 1: Columns (3)-(4) show results when a friend did not receive an offer; Columns (5)-(6) show results when one takes-up male circumcision with a friend on the same date, and; (7)-(8) show results when one takes-up male circumcision separately from a friend. Panel A confirms that one is more likely to take-up male circumcision when your best friend also takes it up. The coefficients in Columns (5)-(8) are similar, which implies that two mechanisms discussed above are playing an important role. Panel B1 and C1 further disaggregate the above responses depending on whether a friend did or did not receive an offer. As expected, peer effects are particularly significant when both a student and his friend received an offer, which corresponds to the results shown at Panel D of Table 3.4.

### 3.5.4 Long-term impacts of male circumcision on STD infections and sexual behaviors

In this section, we study the long-term impacts of male circumcision on HSV2 and HIV infections, sexual behaviors measured by ICT, and a range of self-reported sexual behaviors. As explained previously, one of the major concern related to the scale up of male circumci-

[25] It is possible that students might discuss in a way that "I will take it if you take it". This case would capture social learning channel in our analysis but the nature of discussion is the other channel. It would lead us to overestimate social learning channel. The results should be interpreted with this caveat.

sion relates to possible risk compensation resulting from circumcised men engaging in risky sexual behaviors. In this analysis, we focus only on the baseline 9th and 10th graders who were included in the long-term follow-up survey.

Panel A of Table 3.6 shows that the cumulative probability of HSV-2 infection measured by IgG in G1 and G2 increases by about 3.8% and 2.8% points after four years, respectively, and the corresponding increases for IgM are 1.9% and 2.2% points. Interestingly, we do not find such evidence for G3 even though circumcision take up rate of G3 are comparable at the end of the experiment and the long-term follow-up survey. Lastly, we do not find a long-term impact of the intervention on testing HIV positive, although we not that our study was not powered to detect HIV impacts.

Table 3.7 presents results using the Item Count Technique (ICT) to analyze responses to two sensitive risky sexual behavior: "I think I have to use a condom in case of sex with somebody that I do not know well" (Columns (1) and (2)) and "I had sex with more than two people in last 12 months" (Columns (3) and (4)). As a reminder, half of the survey respondents were randomly given a set of straightforward true/false statements, while the other half also received the additional sensitive statement about risky sexual behavior.

Table 3.7 shows the regression results of the following estimation:

$$Y_{ijk} = \beta_0 + \sum_{a=1}^{3} \beta_a Ga_{ijk} + \kappa long_{ijk} + \sum_{a=1}^{3} \lambda_a \cdot Ga_{ijk} \cdot long_{ijk} + \gamma' X_{ij} + \delta_k + \varepsilon_{ijk} \tag{3.4}$$

where $long_{ijk}$ is a binary indicator if individual $i$ assigned to a questionnaire with sensitive questions (*long form*). $\lambda$ is s coefficients of interest that captures the difference in the response of subjects who agree to the sensitive question between the treatment and control group.

The ICT results suggest that the difference in prevalence is about 22% between G1 and G4, and 23% between G2 and G4 (Panel A). Among treated group (G1 and G2), only 38% of students think that they have to use a condom in case of a casual sexual encounter while the corresponding number for the control group is 60% (Panel B). We do not find statistically significant changes in the multiple sex partners question. Table 3.7 confirms that the changes in sexual behaviors for group G3 which increased take-up through the peer channel is not significant.

Next, we measure the impact of the intervention on condom purchases and self-reported sexual behaviors (Table 3.8). First, in Columns (1) and (2) we do not find statistically significant difference between the groups in the likelihood of purchasing a condom. Second, risky sexual behaviors, measured by using an index of eight major risky sexual behavior indicators, increases among males in the intervention group compared to the control group even though it is not statistically significant. We also find increases in each component of risky sexual behaviors but these changes are not statistically significant.

In sum, we find evidence of risk compensation that diminishes the preventive effect of male circumcision on HIV-1 and HSV-2 infection. Specifically, we find a long term increase of HSV-2 infection among the male circumcision treatment group while there is no significant change in HIV-1 infection. Risky sexual behaviors measured using the ICT technique provide complementary evidence to the results using the biomarkers. Secondly, these impacts are driven by individuals who took up circumcision through the direct intervention (G1 and G2) and not through the peer effect channel (G3).

In order to study why biomarker outcomes of G1 and G2 (G12) and G3 are different, we compare complier characteristics of G12 and G3 by restricting the sample to circumcision takers following Kim and Lee (2017) (Table 3.B.6). Since everyone has undergone circumcision in the restricted sample, any difference between G12 and G3 is due to the compositional change of circumcision takers. In this sample, circumcision takers of G12 are always takers and compliers driven by the transportation support while circumcision takers of G3 are always takers and compliers driven by peer effects. Thus, using the analysis with the restricted sample allows us to compare the characteristics of two different types of compliers. Our results in Table 3.B.6 do not show significant differences in the complier characteristics, suggesting that other unobserved characteristics between these groups are driving the differences in sexual behavior and HSV2 infections. However, it is worth noting that our analysis of complier characteristics is based on fairly small sample sizes and as a result we are not powered to detect meaningful differences between these groups.

## 3.6 Conclusion

This paper addresses questions on demand for male circumcision and its long-term consequences. Specifically, we study how to promote demand for male circumcision and what is the role of peer effects in demand for male circumcision. In addition, this study provides the first experimental evidence of the long-term impacts of community based male circumcision scale up project. To do this, we implemented a randomized controlled trial that randomly provided free male circumcision and transportation voucher to male students across 33 public secondary schools in Malawi. Classrooms are assigned into three groups: 100% Treatment, 50% Treatment, or No Treatment classrooms. Randomly selected half of male students in 50% Treatment classrooms were treated.

We find that our school based intervention substantially increases the demand for male circumcision by on average of 15.0 percentage points. Moreover, we find evidence consistent with important positive peer effects among classroom peers and the peer effects are particularly strong when both one and one's friend are treated. In addition, our results show that the preventive effects of male circumcision against HIV and HSV-2 could be mitigated in the long-run through risk compensation. We detect a significant increase of HSV-2 infection measured by IgG and IgM after four years from the treatment even though male circumcision decreases HSV- 2 infection biologically holding sexual activities constant. Interestingly, we find evidence of risk compensation only among those who were induced to take-up circumcision through transportation support program, but not among those who were induced through the peer effect.

To the best of our knowledge, our study is the first to measure the impact of male circumcision through randomized trial using a scale-up project setting where the preventive effect against HIV is widely known and people can voluntarily accept or decline free male circumcision service. It is also unique because we are able to study risk compensation behaviors in the long-term using biomarkers. Our findings are surprisingly different from results in several efficacy trials that showed 51-60% protective effect against HIV acquisition (Auvert et al., 2005; Bailey et al., 2007; Gray et al., 2007) and 25% protective effect against HSV-2 infection (Tobian et al., 2009). Our community based trial might more closely resemble what the situation is likely to be under non-study conditions. For example, our program carefully provides benefits and risks of the medical male circumcision to study participants, which could be closer to the scale-up settings.

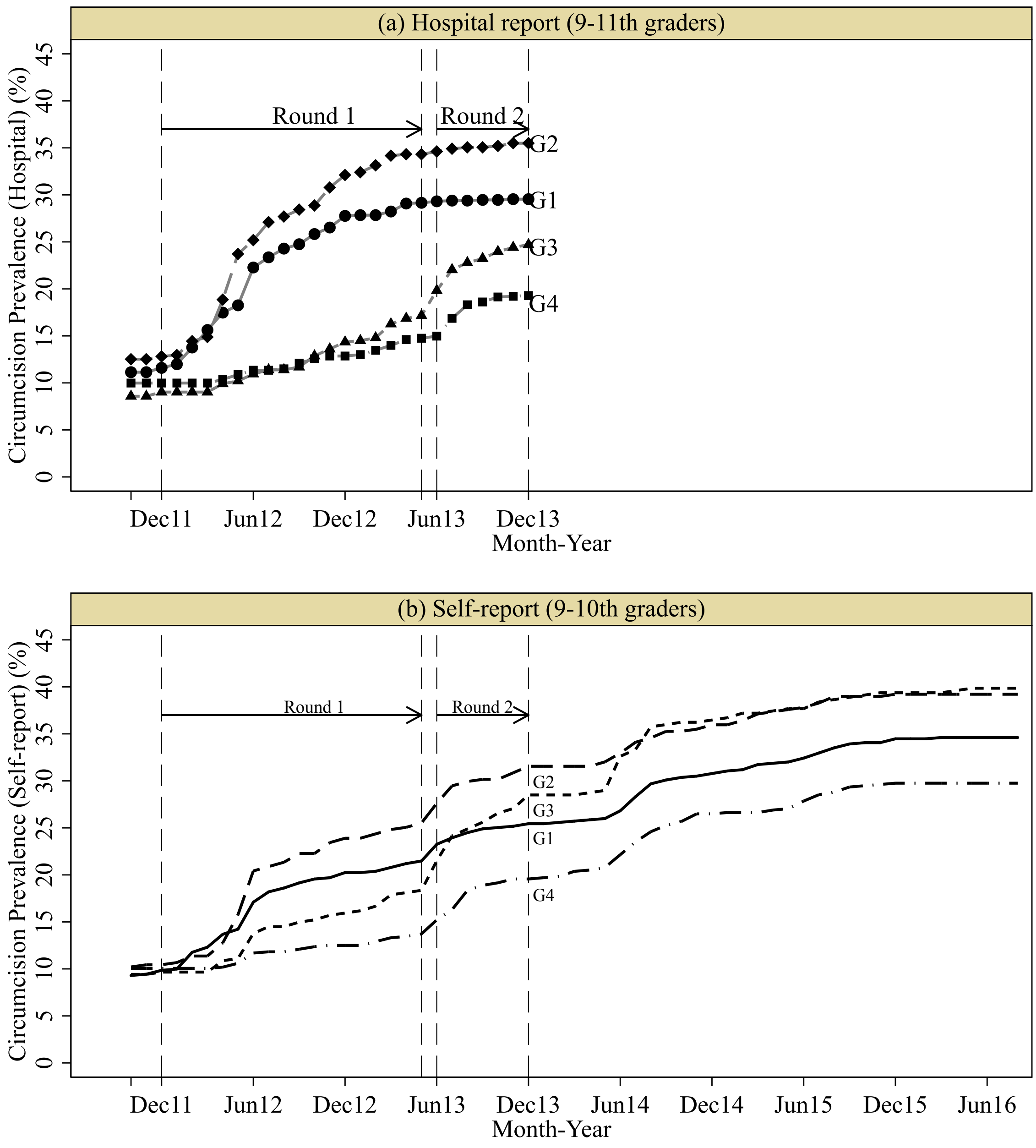
This study has some limitations. The lack of baseline data for HIV and HSV-2 makes comparison of incidences between the treatment and control group impossible. However, there are several reasons to support our finding are robust. First, the treatment and control group had similar baseline characteristics. So it is unlikely that our biomarker results are caused by differences of baseline prevalence of HIV or HSV-2 or baseline variables that are associated with HIV or HSV-2 infection. Second, we have similar results in IgM test for HSV-2 which detects only recent infection so that helps us to prevent bias that might be driven by difference in baseline prevalence. Lastly, the difference in biomarker outcomes are supported by the changes in sexual behaviors measured by item count technique and self-reported.

Our findings have a number of implications for public policies related to the scale-up of male circumcision. First, while a lack of accessibility to male circumcision is major barrier, our results suggest that free male circumcision with well-designed incentives such as transportation support can increase demand for male circumcision substantially. Second, this study sheds light on the important role that peer effects play in the decision to get circumcised, but it also suggests that those who are taking up an intervention as a result of their peers might display very different behavioral responses.

This study also shows that male circumcision scale-up project might not be successful to prevent HIV and HSV-2 infection when those circumcised engage in risk compensation. Our results suggest that male circumcision scale-up projects should be combined with programs to address risk compensation among circumcised men such as intensive public health campaign and education.

## 3.7 Figures

Figure 3.1: Cumulative prevalence of male circumcision over time




*Notes*: These figures present cumulative male circumcision prevalence rate over time based on the sample of 3,970 9th-11th graders (Full sample) in Panel (a) and 2,312 9th-10th graders who completed the 2nd follow-up survey in Panel (b). Panel (a) and (b) show the prevalence based on hospital administration and self-reported data, respectively.

## 3.8 Tables

Table 3.1: Experimental Design

| | Group | Assignment | Classrooms (Full Sample) | Students (Full Sample) | Classrooms (9th-10th grade) | Students (9th-10th grade) |
|---|---|---|---|---|---|---|
| 100% Treatment | G1 | Treated in Round 1 | 41 | 1,293 | 27 | 861 |
| 50% Treatment | G2 | Treated in Round 1 | 41 | 679 | 28 | 481 |
| 50% Treatment | G3 | Treated in Round 2 | 41 | 676 | 28 | 470 |
| No Treatment | G4 | Treated in Round 2 | 42 | 1,322 | 28 | 851 |
| Total | | | 124 | 3,970 | 83 | 2,663 |

*Notes*: This table present two stage randomization design. First, 124 available classrooms within the 33 schools were stratified by grade and randomly assigned into three groups: 100% treatment, 50% treatment, and No treatment. Second, within 50% treatment classrooms, only half of the students were randomly assigned to treatment in the first round at individual level. Full sample includes all male students from 9th, 10th, and 11th grade at the baseline. The 2nd follow-up survey was conducted only for 9th and 10th grade students at baseline.

Table 3.2: Baseline Statistics and Randomization Balance

| | Mean | Difference in Mean (Full Sample) | | | Mean | Difference in Mean (Baseline 9th-10th grade) | | |
|---|---|---|---|---|---|---|---|---|
| | G4 | (G1vs.G4) | (G2vs.G4) | (G3vs.G4) | G4 | (G1vs.G4) | (G2vs.G4) | (G3 vs.G4) |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| ***Panel A. Socio-demographic Characteristics*** | | | | | | | | |
| Age (Year) | 16.809 | -0.219 | -0.278 | -0.229 | 16.163 | 0.070 | -0.091 | 0.026 |
| Circumcising ethnicity | 0.158 | 0.033 | 0.020 | -0.011 | 0.168 | 0.022 | 0.013 | -0.019 |
| Muslim | 0.050 | 0.021* | 0.021 | 0.000 | 0.054 | 0.017 | 0.019 | -0.005 |
| Orphan | 0.068 | -0.006 | -0.025** | -0.021** | 0.060 | -0.001 | -0.027** | -0.013 |
| Father's tertiary education | 0.172 | 0.007 | 0.020 | 0.014 | 0.166 | 0.020 | 0.034 | 0.020 |
| Mother's tertiary education | 0.068 | -0.004 | 0.005 | 0.003 | 0.069 | -0.009 | -0.000 | 0.005 |
| Father's white-collar job | 0.223 | 0.028 | 0.029 | 0.011 | 0.241 | 0.006 | 0.000 | -0.002 |
| Mother's white-collar job | 0.096 | -0.004 | 0.008 | 0.005 | 0.105 | -0.008 | 0.002 | -0.003 |
| Household asset count (0-16) | 7.313 | 0.382 | -0.192 | -0.184 | 7.424 | 0.181 | -0.408 | -0.437 |
| Conventional schools | 0.186 | -0.010 | 0.179* | 0.169 | 0.095 | 0.070 | 0.337*** | 0.337*** |
| ***Panel B. HIV/AIDS Knowledge and Sexual Behavior*** | | | | | | | | |
| HIV/AIDS knowledge (0-20) | 17.371 | -0.132 | -0.025 | -0.052 | 17.398 | -0.272** | -0.059 | -0.049 |
| Belief in the efficacy of MC | 0.671 | -0.045* | -0.001 | -0.038 | 0.680 | -0.065* | 0.004 | -0.025 |
| MC is painful | 0.358 | 0.049** | 0.041 | 0.048* | 0.345 | 0.069** | 0.037 | 0.047 |
| MC is only for Muslim | 0.149 | 0.001 | 0.005 | 0.017 | 0.147 | 0.010 | -0.003 | 0.012 |
| Ever had sex | 0.308 | -0.012 | 0.010 | 0.002 | 0.254 | 0.024 | 0.038 | 0.038 |
| Sexually active | 0.094 | -0.021 | 0.008 | 0.007 | 0.059 | 0.002 | 0.043* | 0.037** |
| Multiple partners | 0.012 | 0.004 | 0.004 | 0.003 | 0.011 | 0.003 | 0.002 | 0.002 |
| Inconsistent use of condoms | 0.045 | -0.014* | -0.005 | -0.002 | 0.041 | -0.012 | -0.004 | 0.002 |
| Number of condoms purchased | 0.919 | -0.193* | 0.004 | 0.051 | 0.784 | -0.029 | 0.175 | 0.225* |
| Already circumcised | 0.100 | 0.011 | 0.025 | -0.014 | 0.106 | -0.001 | 0.013 | -0.012 |
| Observations | 1,322 | 2,615 | 2,001 | 1,998 | 851 | 1,712 | 1,332 | 1,321 |

*Notes*: This table reports means of selected baseline variables and shows tests for balance between treatment arms. Panel A summarizes demographic and socioeconomic information, and Panel B summarizes HIV/AIDS knowledge and individual sexual behaviors. Columns 1 and 5 show summary statistics for those initially assigned to G4. Columns 2-4 and 6-8 report mean differences (and significance levels for difference of mean tests) between groups having different treatment status. Circumcising ethnicity refers to a tribe of which more than 20% population reported being circumcised in 2010 MDHS. HIV/AIDS knowledge is constructed by counting the correct answers from 20 HIV/AIDS related questions. Inconsistent use of condoms is an indicator variable which becomes one if the respondent did not use condoms at least once during the last sexual intercourse with three most recent partners in the past 12 months. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 3.3: Impacts on Male Circumcision Take-up

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Sample | Full sample (9th-11th graders) | | | | 9th and 10th graders | | | |
| Timing | Round 1 | | Round 2 | | Round 1 | | Round 2 | |
| Data | Hospital Administration Data | | | | | | | |
| ***Panel A*** | | | | | | | | |
| G1 (100% Treatment) | 0.137*** | 0.139*** | 0.099*** | 0.099*** | 0.136*** | 0.141*** | 0.105*** | 0.108*** |
| | (0.023) | (0.022) | (0.024) | (0.023) | (0.017) | (0.017) | (0.019) | (0.019) |
| G2 (50% Treatment) | 0.172*** | 0.174*** | 0.133*** | 0.133*** | 0.165*** | 0.169*** | 0.113*** | 0.116*** |
| | (0.027) | (0.025) | (0.026) | (0.025) | (0.024) | (0.024) | (0.023) | (0.024) |
| G3 (50% No Treatment) | 0.040* | 0.042* | 0.064** | 0.064** | 0.032 | 0.040* | 0.062*** | 0.068*** |
| | (0.022) | (0.022) | (0.025) | (0.025) | (0.021) | (0.022) | (0.023) | (0.023) |
| $F$ test (Prob.$>F$) | | | | | | | | |
| G1=G2 | 0.258 | 0.243 | 0.287 | 0.256 | 0.258 | 0.291 | 0.764 | 0.762 |
| G1=G3 | 0.000 | 0.000 | 0.221 | 0.214 | 0.000 | 0.000 | 0.066 | 0.094 |
| G2=G3 | 0.000 | 0.000 | 0.005 | 0.005 | 0.000 | 0.001 | 0.044 | 0.065 |
| R-Squared | 0.093 | 0.103 | 0.080 | 0.090 | 0.099 | 0.107 | 0.096 | 0.105 |
| ***Panel B*** | | | | | | | | |
| G1 & G2 combined | 0.150*** | 0.152*** | 0.112*** | 0.112*** | 0.147*** | 0.151*** | 0.108*** | 0.111*** |
| | (0.019) | (0.018) | (0.019) | (0.019) | (0.015) | (0.015) | (0.017) | (0.017) |
| G3 | 0.036 | 0.039* | 0.061** | 0.061** | 0.023 | 0.032 | 0.060** | 0.066*** |
| | (0.022) | (0.022) | (0.024) | (0.024) | (0.024) | (0.026) | (0.023) | (0.024) |
| $F$ test (Prob.$>F$) | | | | | | | | |
| G1& G2=G3 | 0.000 | 0.000 | 0.018 | 0.018 | 0.000 | 0.000 | 0.025 | 0.042 |
| R-Squared | 0.092 | 0.102 | 0.079 | 0.089 | 0.099 | 0.107 | 0.096 | 0.104 |
| Observations | 3,970 | 3,937 | 3,970 | 3,937 | 2,663 | 2,643 | 2,663 | 2,643 |
| Mean of Dep. Var. (G4) | 0.048 | | 0.093 | | 0.035 | | 0.075 | |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes |

*Notes*: The circumcision status is based on hospital administration data. All the specifications include school fixed effects. Controls include standard control variables described in Section 4. Robust standard errors clustered by classroom are in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 3.4: Externalities on Male Circumcision Take-up (50% Treatment Classroom)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Sample | Full sample (9th-11th graders) | | | | 9th and 10th graders | | | |
| Timing | Round 1 | | Round 2 | | Round 1 | | Round 2 | |
| Data | Hospital Administration Data | | | | | | | |
| ***Panel A*** | | | | | | | | |
| MC offer | 0.132*** | 0.132*** | 0.068*** | 0.069** | 0.133*** | 0.132*** | 0.050* | 0.054* |
| | (0.028) | (0.029) | (0.025) | (0.025) | (0.035) | (0.037) | (0.025) | (0.027) |
| R-squared | 0.100 | 0.113 | 0.098 | 0.114 | 0.099 | 0.111 | 0.105 | 0.124 |
| ***Panel B*** | | | | | | | | |
| Rate of close friends who got MC offer | 0.033 | 0.026 | 0.055 | 0.050 | 0.057 | 0.047 | 0.071 | 0.067 |
| | (0.038) | (0.038) | (0.043) | (0.045) | (0.039) | (0.038) | (0.046) | (0.046) |
| R-squared | 0.066 | 0.080 | 0.092 | 0.108 | 0.064 | 0.078 | 0.104 | 0.121 |
| ***Panel C*** | | | | | | | | |
| MC offer | 0.133*** | 0.132*** | 0.069*** | 0.070*** | 0.134*** | 0.132*** | 0.050* | 0.055** |
| | (0.028) | (0.030) | (0.025) | (0.025) | (0.035) | (0.037) | (0.025) | (0.026) |
| Rate of close friends who got MC offer | 0.040 | 0.035 | 0.059 | 0.055 | 0.059 | 0.050 | 0.072 | 0.068 |
| | (0.039) | (0.040) | (0.044) | (0.046) | (0.039) | (0.038) | (0.046) | (0.047) |
| R-squared | 0.101 | 0.113 | 0.099 | 0.115 | 0.101 | 0.113 | 0.108 | 0.126 |
| ***Panel D*** | | | | | | | | |
| MC offer | 0.080** | 0.088** | 0.028 | 0.037 | 0.074* | 0.081* | -0.001 | 0.010 |
| | (0.033) | (0.033) | (0.033) | (0.033) | (0.040) | (0.041) | (0.033) | (0.033) |
| Rate of close friends who got MC offer | -0.033 | -0.026 | 0.002 | 0.010 | -0.020 | -0.018 | 0.003 | 0.009 |
| | (0.040) | (0.043) | (0.055) | (0.060) | (0.039) | (0.041) | (0.062) | (0.064) |
| MC offer× | 0.142** | 0.119* | 0.112 | 0.087 | 0.153** | 0.133* | 0.133 | 0.115 |
| Rate of close friends who got MC offer | (0.066) | (0.066) | (0.077) | (0.080) | (0.073) | (0.071) | (0.087) | (0.087) |
| R-squared | 0.104 | 0.115 | 0.101 | 0.116 | 0.104 | 0.116 | 0.110 | 0.127 |
| Observations | 1,355 | 1,339 | 1,355 | 1,339 | 951 | 942 | 951 | 942 |
| Mean of Dep. Var. (G4) | 0.152 | | 0.196 | | 0.142 | | 0.198 | |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes |

*Notes*: This analysis includes only the 50% Treatment classroom sample. Robust standard errors are shown in parentheses clustered at the classroom level. "Rate of close friends who got MC offer" is a variable for male circumcision offer to friends defined as the proportion (rate) of friends who are offered male circumcision. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 3.5: My take-up related to my friend's take-up decision timing (50% Treatment Classroom)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Dependent vars. | My uptake | | My uptake× No friend uptake | | My uptake×Friend's uptake with me | | My uptake×Friend's uptake before/after me | |
| Timing | Round 1 | | | | | | | |
| Data | Hospital Administration Data | | | | | | | |
| ***Panel A. Overall*** | | | | | | | | |
| My MC offer | 0.136*** | 0.134*** | 0.093*** | 0.091*** | 0.020** | 0.019** | 0.023* | 0.023* |
| | (0.031) | (0.032) | (0.017) | (0.017) | (0.009) | (0.009) | (0.012) | (0.012) |
| Observations | 2,937 | 2,903 | 2,937 | 2,903 | 2,937 | 2,903 | 2,937 | 2,903 |
| R-squared | 0.108 | 0.120 | 0.061 | 0.070 | 0.027 | 0.032 | 0.036 | 0.039 |
| ***Panel B1. When Friend got MC offer*** | | | | | | | | |
| My MC offer | 0.164*** | 0.160*** | 0.085*** | 0.083*** | 0.043** | 0.041** | 0.036*** | 0.036*** |
| | (0.033) | (0.034) | (0.014) | (0.016) | (0.017) | (0.016) | (0.013) | (0.013) |
| Observations | 1,501 | 1,482 | 1,501 | 1,482 | 1,501 | 1,482 | 1,501 | 1,482 |
| R-squared | 0.134 | 0.146 | 0.065 | 0.073 | 0.055 | 0.063 | 0.056 | 0.063 |
| ***Panel B2. When Friend got MC offer & He thinks MC is painful*** | | | | | | | | |
| MC offer | 0.133*** | 0.129*** | 0.090*** | 0.086*** | 0.023 | 0.022 | 0.020 | 0.021 |
| | (0.035) | (0.035) | (0.026) | (0.026) | (0.016) | (0.016) | (0.017) | (0.018) |
| Observations | 633 | 624 | 633 | 624 | 633 | 624 | 633 | 624 |
| R-squared | 0.125 | 0.140 | 0.082 | 0.094 | 0.056 | 0.070 | 0.061 | 0.071 |
| ***Panel B3. When Friend got MC offer & He doesn't think MC is painful*** | | | | | | | | |
| MC offer | 0.183*** | 0.178*** | 0.080*** | 0.078*** | 0.057** | 0.053** | 0.047** | 0.047** |
| | (0.041) | (0.044) | (0.018) | (0.020) | (0.021) | (0.020) | (0.019) | (0.020) |
| Observations | 868 | 858 | 868 | 858 | 868 | 858 | 868 | 858 |
| R-squared | 0.162 | 0.174 | 0.078 | 0.086 | 0.070 | 0.083 | 0.081 | 0.096 |
| ***Panel C1. When Friend didn't get MC offer*** | | | | | | | | |
| MC offer | 0.109*** | 0.109*** | 0.104*** | 0.103*** | -0.003 | -0.002 | 0.008 | 0.009 |
| | (0.033) | (0.034) | (0.027) | (0.027) | (0.006) | (0.006) | (0.018) | (0.018) |
| Observations | 1,436 | 1,421 | 1,436 | 1,421 | 1,436 | 1,421 | 1,436 | 1,421 |
| R-squared | 0.097 | 0.113 | 0.077 | 0.095 | 0.020 | 0.023 | 0.045 | 0.048 |
| ***Panel C2. When Friend didn't get MC offer & He thinks MC is painful*** | | | | | | | | |
| MC offer | 0.066** | 0.068** | 0.070** | 0.075** | -0.001 | -0.001 | -0.003 | -0.005 |
| | (0.032) | (0.033) | (0.030) | (0.030) | (0.004) | (0.004) | (0.022) | (0.023) |
| Observations | 584 | 576 | 584 | 576 | 584 | 576 | 584 | 576 |
| R-squared | 0.118 | 0.130 | 0.087 | 0.100 | 0.063 | 0.068 | 0.063 | 0.073 |
| ***Panel C3. When Friend didn't get MC offer & He doesn't think MC is painful*** | | | | | | | | |
| MC offer | 0.134*** | 0.128*** | 0.121*** | 0.116*** | -0.004 | -0.003 | 0.017 | 0.016 |
| | (0.041) | (0.042) | (0.031) | (0.031) | (0.009) | (0.008) | (0.020) | (0.020) |
| Observations | 852 | 845 | 852 | 845 | 852 | 845 | 852 | 845 |
| R-squared | 0.118 | 0.148 | 0.110 | 0.142 | 0.022 | 0.027 | 0.064 | 0.075 |
| Mean of Dep. Var. | 0.152 | | 0.112 | | 0.013 | | 0.027 | |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes |

*Notes*: This analysis uses a stacked 50% treatment classroom sample where unit of observation is single friendship relationship. Robust standard errors are shown in parentheses clustered at the classroom level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 3.6: Impacts on HSV2 and HIV infections (9th and 10th Grade)

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent vars. | HSV2 IgG Positive | | HSV2 IgM Positive | | HIV Positive | |
| Sample | 9th and 10th graders | | | | | |
| Data | Hospital Administration Data | | | | | |
| ***Panel A*** | | | | | | |
| G1 (100% Treatment) | 0.038*** | 0.039*** | 0.017** | 0.019*** | -0.002 | -0.002 |
| | (0.013) | (0.013) | (0.007) | (0.007) | (0.003) | (0.003) |
| G2 (50% Treatment) | 0.028 | 0.035 | 0.020*** | 0.022*** | -0.001 | -0.002 |
| | (0.021) | (0.022) | (0.006) | (0.007) | (0.003) | (0.003) |
| G3 (50% No Treatment) | -0.000 | -0.001 | 0.002 | 0.003 | -0.004 | -0.004 |
| | (0.019) | (0.019) | (0.007) | (0.007) | (0.004) | (0.004) |
| $F$ test (Prob.$>F$) | | | | | | |
| G1=G2 | 0.590 | 0.819 | 0.720 | 0.620 | 0.878 | 0.885 |
| G1=G3 | 0.067 | 0.037 | 0.118 | 0.128 | 0.698 | 0.606 |
| G2=G3 | 0.354 | 0.216 | 0.062 | 0.047 | 0.587 | 0.692 |
| R-Squared | 0.052 | 0.062 | 0.035 | 0.038 | 0.017 | 0.021 |
| ***Panel B*** | | | | | | |
| G1 & G2 combined | 0.034** | 0.037*** | 0.018*** | 0.020*** | -0.002 | -0.002 |
| | (0.014) | (0.014) | (0.005) | (0.006) | (0.003) | (0.003) |
| G3 | 0.003 | -0.000 | 0.001 | 0.002 | -0.004 | -0.004 |
| | (0.022) | (0.021) | (0.008) | (0.007) | (0.004) | (0.004) |
| $F$ test (Prob.$>F$) | | | | | | |
| G1& G2=G3 | 0.230 | 0.130 | 0.060 | 0.051 | 0.603 | 0.642 |
| R-Squared | 0.052 | 0.062 | 0.035 | 0.038 | 0.017 | 0.021 |
| Observations | 2,074 | 2,058 | 2,074 | 2,058 | 2,074 | 2,058 |
| Mean of Dep. Var. (G4) | 0.084 | | 0.012 | | 0.003 | |
| Controls | No | Yes | No | Yes | No | Yes |

*Notes*: This analysis includes only 9th and 10th graders who were surveyed in the 2nd follow-up. Dependent variables are the probability of HSV-2 infection measured by IgG and IgM, and HIV infection. Robust standard errors are shown in parentheses clustered at the classroom level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 3.7: Impacts on Sexual Behaviors (ICT)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent vars. | Using condoms for casual sex (ICT) | | Multiple partners (ICT) | |
| Sample | 9th and 10th graders | | | |
| Data | Hospital Administration Data | | | |
| ***Panel A*** | | | | |
| G1 (100% Treatment) | 0.154 | 0.174* | 0.033 | 0.047 |
| | (0.099) | (0.099) | (0.084) | (0.085) |
| G2 (50% Treatment) | 0.248*** | 0.257*** | 0.014 | 0.017 |
| | (0.094) | (0.097) | (0.099) | (0.101) |
| G3 (50% No Treatment) | 0.163 | 0.176 | 0.014 | 0.010 |
| | (0.112) | (0.119) | (0.120) | (0.123) |
| Long | 0.610*** | 0.619*** | 0.144* | 0.151* |
| | (0.081) | (0.082) | (0.085) | (0.085) |
| G1 (100% Treatment)×Long | -0.220 | -0.235 | -0.074 | -0.087 |
| | (0.142) | (0.143) | (0.102) | (0.103) |
| G2 (50% Treatment)×Long | -0.234* | -0.233* | -0.073 | -0.065 |
| | (0.128) | (0.130) | (0.139) | (0.139) |
| G3 (50% No Treatment) ×Long | -0.092 | -0.083 | -0.108 | -0.088 |
| | (0.153) | (0.157) | (0.189) | (0.190) |
| R-Squared | 0.059 | 0.062 | 0.026 | 0.031 |
| ***Panel B*** | | | | |
| G1 & G2 combined | 0.187** | 0.203** | 0.027 | 0.036 |
| | (0.080) | (0.081) | (0.079) | (0.080) |
| G3 | 0.135 | 0.149 | 0.020 | 0.016 |
| | (0.102) | (0.108) | (0.118) | (0.120) |
| Long | 0.607*** | 0.616*** | 0.145* | 0.152* |
| | (0.081) | (0.082) | (0.085) | (0.085) |
| G1 & G2 combined×Long | -0.221* | -0.231* | -0.074 | -0.079 |
| | (0.115) | (0.116) | (0.101) | (0.101) |
| G3 (50% No Treatment)×Long | -0.088 | -0.078 | -0.109 | -0.089 |
| | (0.152) | (0.156) | (0.188) | (0.189) |
| R-Squared | 0.058 | 0.062 | 0.026 | 0.031 |
| Observations | 2,311 | 2,294 | 2,312 | 2,295 |
| Mean of Dep. Var. (G4 & Short) | 1.760 | | 2.172 | |
| Controls | No | Yes | No | Yes |

*Notes*: This analysis includes only 9th and 10th graders who were surveyed in the 2nd follow-up. "Using condoms for casual sex" shows the ICT result for the question of which an extra item states. Actual statement for Columns 1 and 2 is "I think I have to use a condom in case of sex with somebody that I do not know well.", and that for Columns 3 and 4 is "I had sex with more than two people in last 12 months." *Long* refers to a set of questions that include a sensitive item. Robust standard errors are shown in parentheses clustered at the classroom level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

## Table 3.8: Impacts on Self-reported Sexual Behaviors

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dependent vars. | No. of condoms purchased | | Ever had sex | | Age at sexual debut | | Sexually active | | Multiple partners in past 12 mon. | | Multiple partners in lifetime | | Inconsistent condom use | | Unprotected sex with last partner | |
| Sample | 9th and 10th graders | | | | | | | | | | | | | | | |
| ***Panel A1. 1st Follow-up (9-10th grade sample)*** | | | | | | | | | | | | | | | | |
| G1 (100% Treatment) | 0.057 | 0.016 | 0.002 | -0.009 | -0.196 | -0.190 | -0.008 | -0.017 | -0.005 | -0.004 | -0.030 | -0.036 | -0.019 | -0.025* | -0.021 | -0.026* |
| | (0.124) | (0.125) | (0.039) | (0.036) | (0.351) | (0.255) | (0.025) | (0.024) | (0.004) | (0.005) | (0.026) | (0.025) | (0.016) | (0.014) | (0.016) | (0.014) |
| G2 (50% Treatment) | 0.165 | 0.179 | 0.019 | 0.025 | -1.090** | -1.038** | 0.008 | 0.006 | -0.010 | -0.010 | 0.019 | 0.023 | -0.020 | -0.026* | -0.019 | -0.023 |
| | (0.180) | (0.165) | (0.053) | (0.041) | (0.518) | (0.439) | (0.033) | (0.028) | (0.007) | (0.008) | (0.037) | (0.031) | (0.017) | (0.015) | (0.017) | (0.015) |
| G3 (50% No Treatment) | -0.117 | -0.177 | 0.033 | 0.037 | -0.438 | -0.466 | 0.004 | 0.002 | 0.007 | 0.010 | 0.011 | 0.019 | -0.009 | -0.015 | -0.008 | -0.014 |
| | (0.139) | (0.143) | (0.050) | (0.038) | (0.437) | (0.368) | (0.036) | (0.031) | (0.006) | (0.007) | (0.034) | (0.031) | (0.019) | (0.017) | (0.019) | (0.017) |
| R-squared | 0.052 | 0.065 | 0.084 | 0.150 | 0.193 | 0.327 | 0.046 | 0.084 | 0.083 | 0.094 | 0.077 | 0.117 | 0.047 | 0.056 | 0.050 | 0.060 |
| ***Panel A2. 1st Follow-up (9-10th grade sample)*** | | | | | | | | | | | | | | | | |
| G1 & G2 combined | 0.098 | 0.079 | 0.009 | 0.004 | -0.470 | -0.442 | -0.002 | -0.008 | -0.007* | -0.006 | -0.011 | -0.013 | -0.020 | -0.025** | -0.020 | -0.025** |
| | (0.109) | (0.108) | (0.037) | (0.030) | (0.374) | (0.286) | (0.024) | (0.021) | (0.004) | (0.004) | (0.026) | (0.023) | (0.014) | (0.012) | (0.013) | (0.012) |
| G3 (50% No Treatment) | -0.146 | -0.219 | 0.028 | 0.029 | -0.084 | -0.121 | 0.000 | -0.004 | 0.008 | 0.011 | -0.002 | 0.004 | -0.009 | -0.015 | -0.008 | -0.014 |
| | (0.135) | (0.150) | (0.046) | (0.036) | (0.471) | (0.410) | (0.038) | (0.035) | (0.008) | (0.009) | (0.034) | (0.032) | (0.018) | (0.017) | (0.018) | (0.017) |
| R-squared | 0.052 | 0.064 | 0.084 | 0.149 | 0.184 | 0.320 | 0.045 | 0.084 | 0.083 | 0.094 | 0.076 | 0.116 | 0.047 | 0.056 | 0.050 | 0.060 |
| Observations | 1,851 | 1,836 | 1,843 | 1,828 | 529 | 527 | 1,854 | 1,839 | 1,844 | 1,829 | 1,843 | 1,828 | 1,844 | 1,829 | 1,844 | 1,829 |
| Mean of Dep. Variable (G4) | 1.013 | | 0.271 | | 15.8 | | 0.107 | | 0.008 | | 0.134 | | 0.037 | | 0.034 | |
| ***Panel B1. 2nd Follow-up (9-10th grade sample)*** | | | | | | | | | | | | | | | | |
| G1 (100% Treatment) | 0.124 | 0.088 | 0.022 | 0.018 | -0.110 | -0.160 | 0.011 | 0.003 | 0.002 | 0.003 | -0.010 | -0.007 | 0.012 | 0.011 | 0.010 | 0.010 |
| | (0.143) | (0.139) | (0.027) | (0.021) | (0.144) | (0.117) | (0.026) | (0.022) | (0.005) | (0.005) | (0.026) | (0.022) | (0.012) | (0.008) | (0.012) | (0.009) |
| G2 (50% Treatment) | 0.377* | 0.348 | -0.008 | 0.003 | -0.214 | -0.114 | -0.006 | 0.002 | 0.029*** | 0.030*** | -0.021 | -0.017 | 0.017 | 0.021 | 0.011 | 0.015 |
| | (0.224) | (0.240) | (0.038) | (0.032) | (0.183) | (0.178) | (0.037) | (0.032) | (0.010) | (0.009) | (0.036) | (0.034) | (0.019) | (0.017) | (0.018) | (0.017) |
| G3 (50% No Treatment) | 0.190 | 0.148 | -0.024 | -0.021 | -0.010 | 0.026 | -0.019 | -0.018 | 0.011 | 0.012 | 0.025 | 0.032 | -0.011 | -0.010 | -0.013 | -0.011 |
| | (0.243) | (0.242) | (0.039) | (0.031) | (0.198) | (0.176) | (0.038) | (0.033) | (0.013) | (0.013) | (0.035) | (0.031) | (0.018) | (0.016) | (0.017) | (0.016) |
| R-Squared | 0.022 | 0.024 | 0.054 | 0.097 | 0.063 | 0.139 | 0.068 | 0.101 | 0.019 | 0.026 | 0.045 | 0.070 | 0.046 | 0.072 | 0.047 | 0.073 |
| ***Panel B2. 2nd Follow-up (9-10th grade sample)*** | | | | | | | | | | | | | | | | |
| G1 & G2 combined | 0.218 | 0.184 | 0.011 | 0.012 | -0.150 | -0.142 | 0.005 | 0.002 | 0.012** | 0.013** | -0.014 | -0.011 | 0.014 | 0.015 | 0.010 | 0.012 |
| | (0.137) | (0.140) | (0.027) | (0.021) | (0.144) | (0.127) | (0.027) | (0.022) | (0.006) | (0.006) | (0.024) | (0.021) | (0.011) | (0.009) | (0.011) | (0.009) |
| G3 (50% No Treatment) | 0.116 | 0.073 | -0.015 | -0.016 | 0.023 | 0.012 | -0.014 | -0.018 | 0.003 | 0.004 | 0.028 | 0.035 | -0.013 | -0.012 | -0.013 | -0.012 |
| | (0.225) | (0.220) | (0.037) | (0.030) | (0.183) | (0.160) | (0.036) | (0.031) | (0.013) | (0.014) | (0.036) | (0.032) | (0.018) | (0.017) | (0.018) | (0.017) |
| R-squared | 0.021 | 0.024 | 0.053 | 0.096 | 0.063 | 0.139 | 0.068 | 0.101 | 0.017 | 0.025 | 0.045 | 0.070 | 0.046 | 0.072 | 0.047 | 0.073 |
| Observations | 2,312 | 2,295 | 2,306 | 2,289 | 1,454 | 1,443 | 2,312 | 2,295 | 2,303 | 2,286 | 2,306 | 2,289 | 2,303 | 2,286 | 2,303 | 2,286 |
| Mean of Dep. Variable (G4) | 1.397 | | 0.607 | | 17.8 | | 0.549 | | 0.026 | | 0.329 | | 0.071 | | 0.069 | |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |

*Notes*: In Panel A1 and A2, we ran a weighted regression because 15 percent of students in the attrition sample were randomly selected for intensive home-visit survey in the 1st follow-up. Column (1)-(2) provide the results from an experiment we conducted during survey by giving students 10 kwachas and selling two condoms at subsidized price 5 kwachas (Thornton, 2008). Robust standard errors are shown in parentheses clustered at the classroom level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

# Appendix to Chapter 3

# 3.A Appendix. Figures

Figure 3.A.1: Study Design

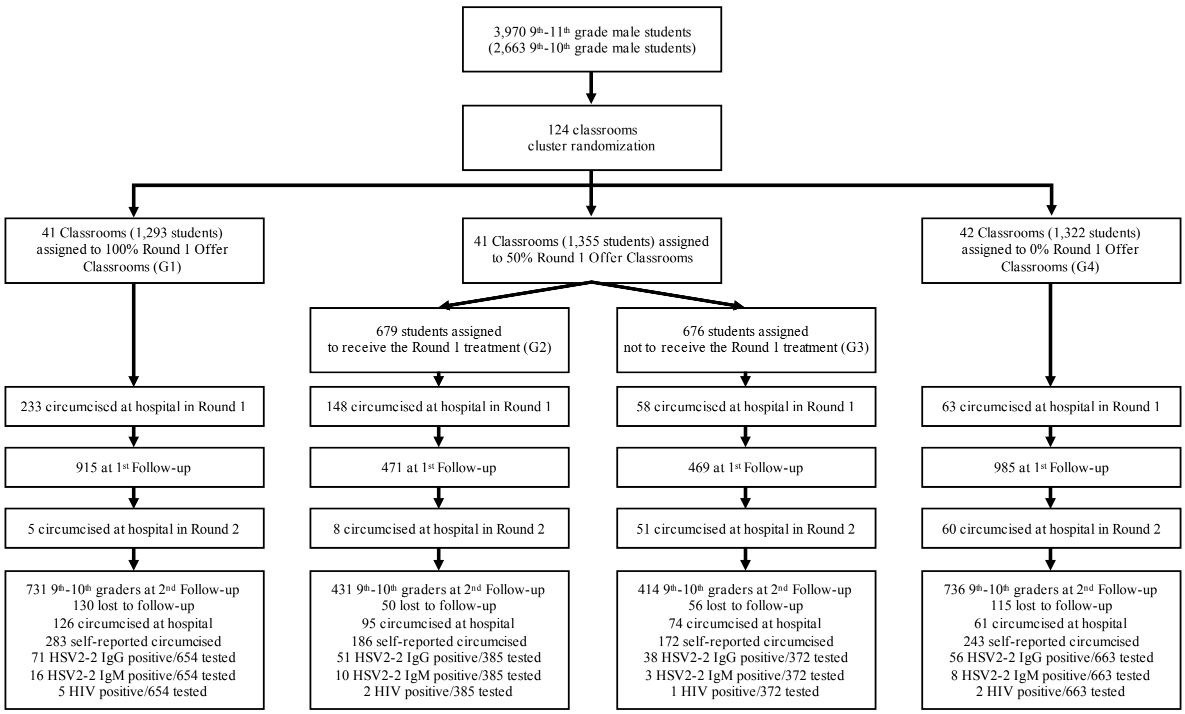


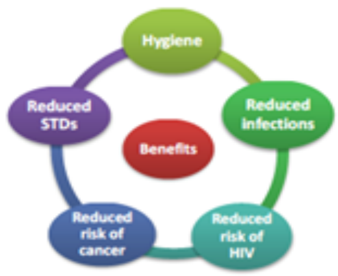
Figure 3.A.2: Male Circumcision Program Brochure

## 7 things you want to know about male circumcision

### Q1 What is male circumcision?

**A:** Male circumcision (MC) is the surgical removal of the foreskin; the skin that covers the tip of the penis

### Q2 What are the benefits of MC?




### Q3 Is it safe to get circumcised?

**A:** Likewise any other surgeries, there are some risks to male circumcision. But, the complications are rare and are easily resolved. MC provided at Daeyang Luke Hospital is performed by professionally trained clinical officers. It is performed at an adequately equipped and sterile health facility.

### Q4 Is MC only for the Muslim?

**A:** ***No***. Although MC is important for some traditions, MC is being performed regardless of faiths and cultures around the world.

MC provided at DLH is different from traditional MC found in some parts of Malawi. It is performed to follow the guideline of the World Health Organization (WHO).
**It's clean, sterile, and safe.**

### Q5 How does MC protect HIV and STIs?

**A:** The inside of the foreskin is soft and mist and is more likely to get a tidy tear or sore that allows HIV to enter the body more easily.

The circumcised male genital is dry providing protection against entry of the virus. Male circumcision modifies acidity of the penis, which is unfavorable for HIV to survive. Hence, MC reduces chances of HIV and STI transmission.

### Q6 How long will it take to recover?

**A:** You will be sore for a few days after surgery but you can continue with your normal routine after 2-3 days of the surgery.

**Full healing will take 4-6 weeks.**
Follow-up visits which ensure proper healing are provided after 1 - 3 days at your school.

### Q7 MUST DOs after surgery

- Don't get your dressing wet
- Return to the hospital if you experience serious pain, bleeding or discharge
- Avoid disturbing the sutures through physical activity or bicycle riding
- ***Avoid sex for 6-weeks*** **until healing is complete!**

  : Resuming sex before full healing can cause damage to your penis and put you and your partner at risk for HIV infection.

## MUST REMEMBER

**Though it is proved that MC is very effective to reduce HIV infection risk, it doesn't offer full protection against HIV infection. It's important to always practice safe sex**


source: Banja la Mtsogolo

It takes a team approach to defend yourself

Male Circumcision **(MC)** and ABC are one team altogether against HIV

- **A**bstaining from sex
- **B**eing faithful to one partner
- **C**ondom use

## WHY MALE CIRCUMCISION?

**1** It can help to prevent HIV infection.
It is estimated that a man who is circumcised appears to be **60% less likely** to get HIV

**2** It can reduce the risk of STI's.

**3** It keeps your partner safe from HIV infection indirectly as well as yourself.

**4** It is an investment for your invaluable Health & Future.

**Daeyang Luke Hospital**, in the capital of Malawi, Lilongwe was opened in March 2008. The hospital continues to build its reputation for excellent service and superior medical treatment throughout the nation. The hospital serves people who are suffering without means to receive proper medical care.

Contacts: Daeyang Luke Hospital. P.O.BOX 30330, Capital City, Lilongwe 3, Malawi

Tel: Chimwemwe: 0995586711
Gerald: 0999288740
Bright: 0995703699

## Daeyang Luke Hospital

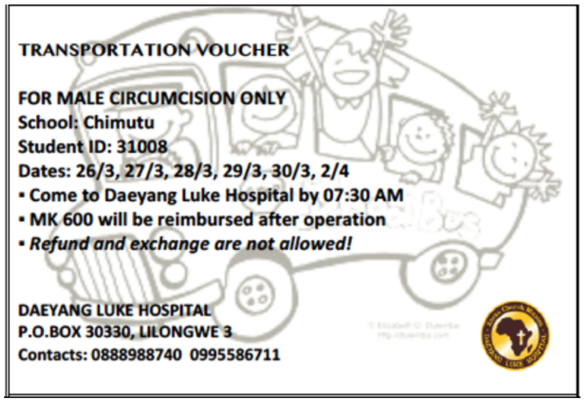
**MALE CIRCUMCISION (MC)**
**MDULIDWE WA ABAMBO**


Source: www.malecircumcision.org

*Notes*: This brochure was translated in Chichewa and distributed to students and used for counseling.

Figure 3.A.3: Transportation Vouchers

(a) 1st transportation voucher for Round 1

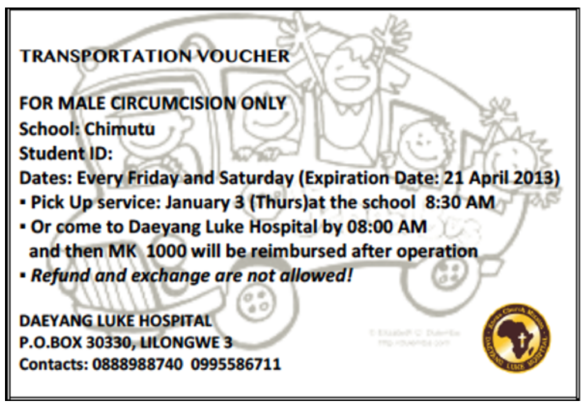


TRANSPORTATION VOUCHER

FOR MALE CIRCUMCISION ONLY
School: Chimutu
Student ID: 31008
Dates: 26/3, 27/3, 28/3, 29/3, 30/3, 2/4
▪ Come to Daeyang Luke Hospital by 07:30 AM
▪ MK 600 will be reimbursed after operation
▪ *Refund and exchange are not allowed!*

DAEYANG LUKE HOSPITAL
P.O.BOX 30330, LILONGWE 3
Contacts: 0888988740 0995586711

(b) 2nd transportation voucher for Round 1

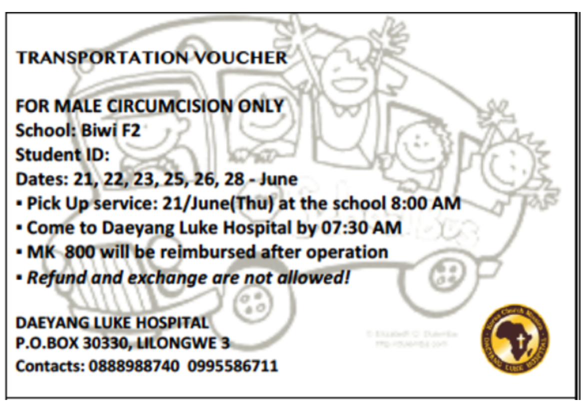


TRANSPORTATION VOUCHER

FOR MALE CIRCUMCISION ONLY
School: Chimutu
Student ID:
Dates: Every Friday and Saturday (Expiration Date: 21 April 2013)
▪ Pick Up service: January 3 (Thurs)at the school 8:30 AM
▪ Or come to Daeyang Luke Hospital by 08:00 AM
and then MK 1000 will be reimbursed after operation
▪ *Refund and exchange are not allowed!*

DAEYANG LUKE HOSPITAL
P.O.BOX 30330, LILONGWE 3
Contacts: 0888988740 0995586711

(c) Transportation voucher for Round 2



TRANSPORTATION VOUCHER

FOR MALE CIRCUMCISION ONLY
School: Biwi F2
Student ID:
Dates: 21, 22, 23, 25, 26, 28 - June
▪ Pick Up service: 21/June(Thu) at the school 8:00 AM
▪ Come to Daeyang Luke Hospital by 07:30 AM
▪ MK 800 will be reimbursed after operation
▪ *Refund and exchange are not allowed!*

DAEYANG LUKE HOSPITAL
P.O.BOX 30330, LILONGWE 3
Contacts: 0888988740 0995586711

Figure 3.A.4: Item Count Technique Questions

(a) Long-form with a sensitive item

Q231.How many of the statements below apply to you?

(A) I like to play with animals.
(B) I have moved house in last three years.
(C) I would consider myself a sports fan.
(D) I know the full name of the president.
(E) I think I have to use condom in case of a sex with somebody that you do not know well.

0 1 2 3 4 5

Q232.How many of the statements below apply to you?

(A) I sleep less than 7 hours per night
(B) I have few photos of myself
(C) I go to the church or mosque almost every week.
(D) I have never been in a traffic accident.
(E) I had sex with more than two people in last 12 months.

0 1 2 3 4 5

(b) Short-form without a sensitive item

Q231.How many of the statements below apply to you?

(A) I like to play with animals.
(B) I have moved house in last three years.
(C) I would consider myself a sports fan.
(D) I know the full name of the president.

0 1 2 3 4

Q232.How many of the statements below apply to you?

(A) I sleep less than 7 hours per night
(B) I have few photos of myself
(C) I go to the church or mosque almost every week.
(D) I have never been in a traffic accident.

0 1 2 3 4

## 3.B Appendix. Tables

Table 3.B.1: Summary Statistics of In-class Friendship Networks

| ***Panel A: Friendship Reconstruction*** | | | |
|---|---|---|---|
| | Raw count | Eligible male | Reordered eligible male |
| | (1) | (2) | (3) |
| First-best friend | 3,844 | 3,102 | 3,832 |
| Second-best friend | 3,839 | 2,818 | 3,135 |
| Third-best friend | 3,860 | 2,668 | 1,621 |

| ***Panel B: Friendship link treatment status*** | | |
|---|---|---|
| | Case | Percentage |
| No friend treated | 1,697 | 42.75% |
| One friend treated | 833 | 20.98% |
| Two friends treated | 823 | 20.73% |
| Three friends treated | 617 | 15.54% |

*Notes*: Panel A Column (1) includes raw friendship data including friends without baseline survey and female friends. Column (2) excludes friends without baseline survey and further excludes female friends from column (1). Finally, we reorder the remaining friendship data from column (2) as first, second, or third best male friends if available. Panel B present friendship statistics based on reordered eligible male best friend data (Panel A, column (3)).

Table 3.B.2: Baseline Statistics and Randomization Balance by Fraction of Treated Friends

| | Mean | Difference in Mean (Full Sample) | | | Mean | Difference in Mean (Baseline 9th-10th grade) | | |
|---|---|---|---|---|---|---|---|---|
| | No friend treated (1) | (1 vs. 0 treated) (2) | (2 vs. 0 treated) (3) | (3 vs. 0 treated) (4) | No friend treated (5) | (1 vs. 0 treated) (6) | (2 vs. 0 treated) (7) | (3 vs. 0 treated) (8) |
| ***Panel A. Socio-demographic Characteristics*** | | | | | | | | |
| Age (Year) | 16.760 | -0.114 | -0.198 | -0.285 | 16.151 | 0.005 | 0.051 | 0.069 |
| Circumcising ethnicity | 0.165 | 0.002 | 0.005 | 0.019 | 0.172 | 0.012 | 0.000 | -0.005 |
| Muslim | 0.056 | 0.005 | 0.010 | 0.009 | 0.058 | 0.005 | 0.005 | 0.010 |
| Orphan | 0.065 | -0.014 | -0.003 | -0.023** | 0.060 | -0.024** | -0.002 | -0.013 |
| Father's tertiary education | 0.176 | 0.004 | 0.008 | 0.009 | 0.170 | 0.023 | 0.032 | -0.003 |
| Mother's tertiary education | 0.070 | 0.012 | -0.012 | -0.009 | 0.070 | 0.011 | -0.012 | -0.016 |
| Father's white-collar job | 0.226 | 0.022 | 0.016 | 0.033 | 0.237 | 0.017 | 0.021 | -0.015 |
| Mother's white-collar job | 0.098 | 0.015 | -0.009 | -0.018 | 0.104 | 0.014 | -0.005 | -0.026 |
| Household asset count (0-16) | 7.338 | 0.064 | 0.049 | 0.073 | 7.361 | -0.072 | 0.091 | -0.212 |
| Conventional schools | 0.207 | 0.118* | 0.060 | -0.015 | 0.158 | 0.214*** | 0.138** | 0.034 |
| ***Panel B. HIV/AIDS Knowledge and Sexual Behavior*** | | | | | | | | |
| HIV/AIDS knowledge (0-20) | 17.354 | 0.010 | -0.118 | -0.104 | 17.375 | -0.022 | -0.204* | -0.219** |
| Belief in the efficacy of MC | 0.674 | -0.039 | -0.036 | -0.057** | 0.686 | -0.034 | -0.055* | -0.069** |
| MC is painful | 0.364 | 0.025 | 0.054** | 0.060** | 0.349 | 0.027 | 0.054** | 0.099*** |
| MC is only for Muslim | 0.152 | 0.007 | 0.008 | -0.014 | 0.148 | 0.012 | 0.005 | 0.004 |
| Ever had sex | 0.313 | -0.003 | -0.006 | -0.033 | 0.267 | 0.008 | 0.031 | -0.002 |
| Sexually active | 0.100 | -0.020 | -0.006 | -0.036** | 0.071 | -0.003 | 0.018 | -0.005 |
| Multiple partners | 0.015 | -0.001 | -0.001 | -0.002 | 0.012 | -0.003 | 0.002 | 0.005 |
| Inconsistent use of condoms | 0.047 | -0.008 | -0.009 | -0.026*** | 0.043 | -0.010 | -0.003 | -0.019** |
| Number of condoms purchased | 0.916 | -0.013 | -0.113 | -0.154 | 0.813 | 0.131 | 0.043 | -0.028 |
| Already circumcised | 0.108 | -0.010 | 0.006 | -0.007 | 0.106 | -0.009 | 0.011 | -0.004 |
| Observations | 1,697 | 2,530 | 2,520 | 2,314 | 1,098 | 1,655 | 1,683 | 1,521 |

*Notes*: This table reports means of selected baseline variables and mean differences (and significance levels for difference of mean tests) between groups having friendship link treatment status as presented in Table A1 Panel B. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 3.B.3: Attrition

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Dependent var. | Surveyed in 1st follow-up (full sample) | | Surveyed in 1st follow-up (9-10th grade | | Surveyed in 2nd follow-up (9-10th grade | | Biomaker Testing HSV-2 & HIV | |
| ***Panel A*** | | | | | | | | |
| G1 (100% Treatment) | -0.014 | -0.010 | -0.029 | -0.021 | -0.007 | -0.006 | 0.003 | 0.004 |
| | (0.020) | (0.020) | (0.026) | (0.025) | (0.018) | (0.020) | (0.022) | (0.024) |
| G2 (50% Treatment) | -0.015 | -0.010 | -0.048 | -0.042 | 0.025 | 0.023 | 0.028 | 0.031 |
| | (0.026) | (0.026) | (0.031) | (0.031) | (0.022) | (0.021) | (0.028) | (0.028) |
| G3 (50% No Treatment) | 0.011 | 0.013 | 0.005 | 0.011 | 0.010 | 0.012 | 0.018 | 0.022 |
| | (0.022) | (0.022) | (0.030) | (0.031) | (0.022) | (0.022) | (0.024) | (0.024) |
| $F$ test (Prob.$>F$) | | | | | | | | |
| G1=G2 | 0.973 | 1.000 | 0.613 | 0.558 | 0.133 | 0.184 | 0.341 | 0.301 |
| G1=G3 | 0.376 | 0.421 | 0.366 | 0.394 | 0.488 | 0.477 | 0.530 | 0.459 |
| G2=G3 | 0.258 | 0.337 | 0.068 | 0.069 | 0.436 | 0.571 | 0.749 | 0.777 |
| R-Squared | 0.045 | 0.051 | 0.074 | 0.082 | 0.024 | 0.033 | 0.026 | 0.039 |
| Observations | 3,970 | 3,937 | 2,663 | 2,643 | 2,663 | 2,643 | 2,663 | 2,643 |
| Mean of Dep. Var. (G4) | 0.947 | | 0.940 | | 0.865 | | 0.779 | |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes |

*Notes*: This table is to test for systemic attrition by regressing a dummy of the survey or testing completion on a set of indicators for treatment arms. Two mean dependent variables from G4 in Columns (1)-(4) refer to the effective survey rate. The effective survey rate (ESR) a function of the regular school follow-up rate (RFR) and intensive homevisit follow-up rate (HFR) after we random selected 15% students from the attrition sample (Baird et al. 2012): ESR = RFR + (1-RFR)$\times$HFR. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 3.B.4: Heterogeneous effects by prior beliefs

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent var. | MC Take-up | | | | | |
| Sample | Full Sample (9th-11th graders) | | | 9th and 10th graders | | |
| | (G1 vs. G4) | (G2 vs. G4) | (G3 vs. G4) | (G1 vs. G4) | (G2 vs. G4) | (G3 vs. G4) |
| Timing | Round 1 | | | | | |
| Data | Hospital Administration Data | | | | | |
| MC offer | 0.173*** | 0.191*** | 0.029 | 0.165*** | 0.186*** | 0.021 |
| | (0.029) | (0.037) | (0.029) | (0.031) | (0.041) | (0.025) |
| Knowing MC benefit | -0.004 | 0.005 | -0.001 | 0.006 | 0.004 | 0.003 |
| | (0.013) | (0.013) | (0.013) | (0.013) | (0.014) | (0.013) |
| MC offer×Knowing MC benefit | 0.008 | -0.031 | -0.004 | 0.002 | -0.032 | -0.022 |
| | (0.023) | (0.031) | (0.027) | (0.028) | (0.034) | (0.027) |
| Think that MC is very painful | -0.017 | -0.019 | -0.016 | -0.016 | -0.024* | -0.014 |
| | (0.012) | (0.012) | (0.012) | (0.013) | (0.012) | (0.012) |
| MC offer×Think that MC is very painful | -0.034 | -0.060* | 0.002 | -0.042 | -0.054 | 0.005 |
| | (0.022) | (0.035) | (0.023) | (0.028) | (0.033) | (0.026) |
| Think that MC is only for Muslim | -0.028** | -0.031** | -0.031** | -0.014 | -0.017 | -0.014 |
| | (0.013) | (0.013) | (0.014) | (0.017) | (0.016) | (0.018) |
| MC offer×Think that MC is only for Muslim | -0.022 | -0.013 | 0.054 | -0.029 | -0.058 | 0.045 |
| | (0.028) | (0.059) | (0.034) | (0.034) | (0.052) | (0.046) |
| R-squared | 0.136 | 0.165 | 0.080 | 0.120 | 0.184 | 0.072 |
| Observations | 2,590 | 1,975 | 1,977 | 1,697 | 1,315 | 1,305 |
| Mean of Dep. Var. (G4) | | 0.048 | | | 0.035 | |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

*Notes*: This table shows the heterogeneous effects on take-up of male circumcision. MC offer variable equals 1 when students get MC offer either from 100% Treatment classrooms or from 50% Treatment classrooms. All columns use school fixed effects and robust standard errors clustered by classroom are in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 3.B.5: When the most popular kid got MC offer in the 50% treatment classrooms

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent vars. | My MC Take-up | | | |
| Sample | 9th-11th graders | | | |
| Data | Hospital Administration Data | | | |
| ***Panel A: Overall*** | | | | |
| My MC offer | 0.154*** | | 0.154*** | 0.133*** |
| | (0.019) | | (0.019) | (0.031) |
| Most popular kid got MC offer | | -0.268* | -0.264* | -0.282* |
| | | (0.148) | (0.150) | (0.151) |
| My MC offer×Most popular kid got MC offer | | | | 0.035 |
| | | | | (0.039) |
| R-Squared | 0.152 | 0.108 | 0.152 | 0.153 |
| Observations | 1,350 | 1,350 | 1,350 | 1,350 |
| ***Panel B: When most popular kid thinks that MC is painful*** | | | | |
| My MC offer | 0.181*** | | 0.181*** | 0.123*** |
| | (0.026) | | (0.026) | (0.047) |
| Most popular kid got MC offer | | -0.294* | -0.284* | -0.331** |
| | | (0.153) | (0.155) | (0.156) |
| My MC offer×Most popular kid got MC offer | | | | 0.087 |
| | | | | (0.056) |
| R-Squared | 0.185 | 0.113 | 0.185 | 0.189 |
| Observations | 613 | 613 | 613 | 613 |
| ***Panel C: When most popular kid thinks that MC is not painful*** | | | | |
| My MC offer | 0.134*** | | 0.134*** | 0.143*** |
| | (0.027) | | (0.027) | (0.040) |
| Most popular kid got MC offer | | -0.103 | -0.094 | -0.085 |
| | | (0.078) | (0.075) | (0.076) |
| My MC offer×Most popular kid got MC offer | | | | -0.018 |
| | | | | (0.054) |
| R-Squared | 0.140 | 0.110 | 0.140 | 0.140 |
| Observations | 737 | 737 | 737 | 737 |
| Controls | Yes | Yes | Yes | Yes |

*Notes*: This analysis uses a stacked 50% treatment classroom sample where unit of observation is single friendship relationship. Robust standard errors are shown in parentheses clustered at the classroom level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 3.B.6: Compliers Characteristics

| Sample | MC takers in G1, G2, G3 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *N* | Coefficient on G12 | Std. Error | Constant | Std. Error | R-squared | Mean of dep. var. | Std. Dev. |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| ***Panel A. Baseline Characteristics*** | | | | | | | | |
| Age (Year) | 439 | -0.668* | (0.359) | 17.414*** | (0.368) | 0.013 | 16.834 | (1.955) |
| Circumcising ethnicity | 438 | 0.061 | (0.055) | 0.121** | (0.051) | 0.003 | 0.174 | (0.379) |
| Muslim | 438 | 0.017 | (0.020) | 0.017 | (0.018) | 0.001 | 0.032 | (0.176) |
| Orphan | 439 | 0.035 | (0.033) | 0.052* | (0.029) | 0.002 | 0.082 | (0.275) |
| Father's tertiary education | 438 | 0.037 | (0.058) | 0.155** | (0.059) | 0.001 | 0.187 | (0.391) |
| Mother's tertiary education | 436 | 0.001 | (0.035) | 0.052 | (0.034) | 0.000 | 0.053 | (0.224) |
| Father's white-collar job | 439 | 0.089* | (0.046) | 0.155*** | (0.048) | 0.005 | 0.232 | (0.423) |
| Mother's white-collar job | 437 | -0.024 | (0.050) | 0.103* | (0.053) | 0.001 | 0.082 | (0.275) |
| Household asset count (0-16) | 439 | 0.861 | (0.581) | 6.397*** | (0.649) | 0.008 | 7.144 | (3.292) |
| Conventional schools | 439 | -0.084 | (0.127) | 0.362** | (0.148) | 0.004 | 0.289 | (0.454) |
| Ever had sex | 438 | -0.065 | (0.081) | 0.397*** | (0.080) | 0.002 | 0.340 | (0.474) |
| Sexually active | 439 | -0.033 | (0.043) | 0.103** | (0.041) | 0.002 | 0.075 | (0.264) |
| Multiple partners | 438 | 0.004 | (0.016) | 0.017 | (0.016) | 0.000 | 0.021 | (0.142) |
| Inconsistent use of condoms | 437 | 0.002 | (0.022) | 0.034 | (0.024) | 0.000 | 0.037 | (0.188) |
| Number of condoms purchased | 439 | -0.316 | (0.224) | 1.103*** | (0.222) | 0.005 | 0.829 | (1.578) |
| Think that MC is painful | 438 | -0.033 | (0.059) | 0.362*** | (0.054) | 0.001 | 0.333 | (0.472) |
| Think that MC is only for Muslim | 438 | -0.088 | (0.062) | 0.207*** | (0.063) | 0.008 | 0.130 | (0.337) |
| ***Panel B. 1st Follow-up Characteristics*** | | | | | | | | |
| Ever had sex | 350 | 0.009 | (0.076) | 0.326*** | (0.075) | 0.000 | 0.334 | (0.472) |
| Sexually active | 351 | -0.034 | (0.062) | 0.152*** | (0.054) | 0.001 | 0.123 | (0.328) |
| Multiple partners | 350 | 0.013** | (0.006) | -0.000 | (0.000) | 0.002 | 0.011 | (0.106) |
| Inconsistent use of condoms | 350 | 0.014 | (0.025) | 0.022 | (0.022) | 0.001 | 0.034 | (0.182) |
| Number of condoms purchased | 353 | 0.071 | (0.276) | 0.913*** | (0.255) | 0.000 | 0.975 | (1.693) |
| Think that MC is painful | 352 | 0.065 | (0.054) | 0.109** | (0.047) | 0.003 | 0.165 | (0.372) |
| Think that MC is only for Muslim | 352 | 0.021 | (0.021) | 0.022 | (0.019) | 0.001 | 0.040 | (0.196) |

*Notes*: This table compares compliers characteristics of students in G1 and G2 (G12) and those in G3 by restricting the sample to circumcision takers in G1, G2 and G3. Standard errors clustered by classroom are in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 3.B.7: Externalities on HSV2 and HIV Infections (50% Treatment class only)

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent vars. | HSV2 IgG Positive | | HSV2 IgM Positive | | HIV Positive | |
| Sample | 9th and 10th graders | | | | | |
| Data | Hospital Administration Data | | | | | |
| ***Panel A*** | | | | | | |
| MC offer | 0.028 | 0.036 | 0.018* | 0.016 | 0.003 | 0.001 |
| | (0.031) | (0.029) | (0.010) | (0.010) | (0.005) | (0.005) |
| R-squared | 0.060 | 0.073 | 0.033 | 0.045 | 0.015 | 0.030 |
| ***Panel B*** | | | | | | |
| Rate of close friends who got MC offer | 0.002 | 0.007 | -0.012 | -0.013 | -0.001 | -0.001 |
| | (0.038) | (0.037) | (0.015) | (0.014) | (0.001) | (0.001) |
| R-squared | 0.058 | 0.070 | 0.029 | 0.042 | 0.015 | 0.030 |
| ***Panel C*** | | | | | | |
| MC offer | 0.028 | 0.036 | 0.018* | 0.016 | 0.002 | 0.001 |
| | (0.031) | (0.029) | (0.010) | (0.010) | (0.005) | (0.005) |
| Rate of close friends who got MC offer | 0.003 | 0.008 | -0.011 | -0.012 | -0.001 | -0.001 |
| | (0.037) | (0.036) | (0.015) | (0.014) | (0.001) | (0.001) |
| R-squared | 0.060 | 0.073 | 0.034 | 0.046 | 0.015 | 0.030 |
| ***Panel D*** | | | | | | |
| MC offer | 0.069 | 0.078 | 0.010 | 0.008 | 0.002 | 0.001 |
| | (0.048) | (0.046) | (0.019) | (0.020) | (0.005) | (0.005) |
| Rate of close friends who got MC offer | 0.059 | 0.066 | -0.022 | -0.024 | -0.001 | -0.001 |
| | (0.068) | (0.064) | (0.018) | (0.019) | (0.002) | (0.003) |
| MC offer× | -0.106 | -0.110 | 0.021 | 0.022 | 0.000 | 0.000 |
| Rate of close friends who got MC offer | (0.092) | (0.090) | (0.035) | (0.036) | (0.002) | (0.004) |
| R-squared | 0.062 | 0.075 | 0.034 | 0.046 | 0.015 | 0.030 |
| Observations | 757 | 750 | 757 | 750 | 757 | 750 |
| Mean of Dep. Var. (G4) | 0.118 | | 0.017 | | 0.004 | |
| Controls | No | Yes | No | Yes | No | Yes |

*Notes*: This analysis includes only the 50% Treatment classroom sample. Robust standard errors are shown in parentheses clustered at the classroom level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

# Bibliography

**Almond, Douglas, Hongbin Li, and Shuang Zhang**, "Land Reform and Sex Selection in China," *Journal of Political Economy*, November 2017, *Forthcoming*.

**_ , Lena Edlund, and Marten Palme**, "Chernobyl's Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden," *The Quarterly Journal of Economics*, November 2009, *124* (4), 1729–1772.

**_ , Yuyu Chen, Michael Greenstone, and Hongbin Li**, "Winter Heating or Clean Air? Unintended Impacts of China's Huai River Policy," *American Economic Review*, April 2009, *99* (2), 184–190.

**Ananat, Elizabeth Oltmans, Jonathan Gruber, Phillip B Levine, and Douglas Staiger**, "Abortion and Selection," *Review of Economics and Statistics*, February 2009, *91* (1), 124–136.

**Angrist, Joshua David and Joern-Steffen Pischke**, *Mostly harmless econometrics: an empiricist's companion*, Princeton: Princeton University Press, 2009.

**Anttila-Hughes, Jesse Keith and Solomon M. Hsiang**, "Destruction, Disinvestment, and Death: Economic and Human Losses Following Environmental Disaster," *SSRN Electronic Journal*, 2013.

**Artadi, Elsa V.**, "Going into Labor: Earnings vs. Infant Survival in Rural Africa." Working Paper, Harvard University November 2005.

**Auffhammer, Maximilian and Ryan Kellogg**, "Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality," *American Economic Review*, October 2011, *101* (6), 2687–2722.

**_ , V. Ramanathan, and Jeffrey R. Vincent**, "Climate change, the monsoon, and rice yield in India," *Climatic Change*, March 2012, *111* (2), 411–424.

**Auvert, Bertran, Dirk Taljaard, Emmanuel Lagarde, Joelle Sobngwi-Tambekou, Remi Sitta, and Adrian Puren**, "Randomized, Controlled Intervention Trial of Male Circumcision for Reduction of HIV Infection Risk: The ANRS 1265 Trial," *PLoS Medicine*, October 2005, *2* (11), e298.

**Bailey, Robert C, Stephen Moses, Corette B Parker, Kawango Agot, Ian Maclean, John N Krieger, Carolyn FM Williams, Richard T Campbell, and Jeckoniah O Ndinya-Achola**, "Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial," *The Lancet*, February 2007, *369* (9562), 643–656.

**Baird, Sarah J, Richard S Garfein, Craig T McIntosh, and Berk Ozler**, "Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomised trial," *The Lancet*, April 2012, *379* (9823), 1320–1329.

**Bandiera, Oriana, Iwan Barankay, and Imran Rasul**, "Social Incentives in the Workplace," *Review of Economic Studies*, April 2010, *77* (2), 417–458.

**Banzhaf, H. Spencer and Randall P Walsh**, "Do People Vote with Their Feet? An Empirical Test of Tiebout's Mechanism," *American Economic Review*, May 2008, *98* (3), 843–863.

**Becker, Gary S.**, *An Economic Analysis of Fertility* number no. c2387. In 'NBER book chapter series.', Cambridge, Mass: National Bureau of Economic Research, 1960.

**Behrman, Jere R.**, "Intrahousehold Allocation of Nutrients in Rural India: Are Boys Favored? Do Parents Exhibit Inequality Aversion?," *Oxford Economic Papers*, 1988, *40* (1), 32–54.

**Belanger, Daniele**, "Son Preference in a Rural Village in North Vietnam," *Studies in Family Planning*, December 2002, *33* (4), 321–334.

**Bengo, Joseph Mfutso, Kondwani Chalulu, Jobiba Chinkhumba, Lawrence Kazembe, Kenneth M Maleta, Francis Masiye, and Don Mathanga**, "Situation Analysis of Male Circumcision in Malawi," Technical Report, College of Medicine, University of Malawi, University of Malawi April 2010.

**Bhalotra, Sonia and Rudi Rocha**, "The Timing of Fertility in Response to Income Shocks." Unpublished Manuscript, University of Bristol 2012.

_ **and Tom Cochrane**, "Where Have All the Young Girls Gone? Identification of Sex Selection in India," Technical Report IZA Discussion Paper No. 5381 December 2010.

**Bharadwaj, Prashant and Leah K. Lakdawala**, "Discrimination Begins in the Womb: Evidence of Sex-Selective Prenatal Investments," *Journal of Human Resources*, 2013, *48* (1), 71–113.

**Bhattacharya, Jayanta, Thomas DeLeire, Steven Haider, and Janet Currie**, "Heat or Eat? Cold-Weather Shocks and Nutrition in Poor American Families," *American Journal of Public Health*, July 2003, *93* (7), 1149–1154.

**Bingenheimer, Jeffrey B. and Arline T. Geronimus**, "Behavioral Mechanisms in HIV Epidemiology and Prevention: Past, Present, and Future Roles," *Studies in Family Planning*, 2009, *40* (3), 187–204.

**Black, Dan A., Natalia Kolesnikova, Seth G. Sanders, and Lowell J. Taylor**, "Are Children "Normal"?," *Review of Economics and Statistics*, March 2013, *95* (1), 21–33.

**Bongaarts, John and Christophe Z. Guilmoto**, "How Many More Missing Women? Excess Female Mortality and Prenatal Sex Selection, 1970-2050," *Population and Development Review*, June 2015, *41* (2), 241–269.

**Brook, R. D.**, "Inhalation of Fine Particulate Air Pollution and Ozone Causes Acute Arterial Vasoconstriction in Healthy Adults," *Circulation*, April 2002, *105* (13), 1534–1536.

**Brooks, Ronald A., Mark Etzel, Lee E. Klosinski, Arleen A. Leibowitz, Sharif Sawires, Greg Szekeres, Mark Weston, and Thomas J. Coates**, "Male Circumcision and HIV Prevention: Looking to the Future," *AIDS and Behavior*, October 2010, *14* (5), 1203–1206.

**Burke, Marshall, Erick Gong, and Kelly Jones**, "Income Shocks and HIV in Africa," *The Economic Journal*, June 2015, *125* (585), 1157–1189.

**Cameron, Elissa Z.**, "Facultative adjustment of mammalian sex ratios in support of the Trivers-Willard hypothesis: evidence for a mechanism," *Proceedings of the Royal Society B: Biological Sciences*, August 2004, *271* (1549), 1723–1728.

**Cassell, Michael M, Daniel T Halperin, James D Shelton, and David Stanton**, "Risk compensation: the Achilles' heel of innovations in HIV prevention?," *BMJ : British Medical Journal*, March 2006, *332* (7541), 605–607.

**Chatterjee, Shoumitro and Tom Vogl**, "Escaping Malthus: Economic Growth and Fertility Change in the Developing World." Working Paper, Princeton University May 2017.

**Chay, Kenneth Y. and Michael Greenstone**, "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession," *The Quarterly Journal of Economics*, August 2003, *118* (3), 1121–1167.

**_ and _** , "Does Air Quality Matter? Evidence from the Housing Market," *Journal of Political Economy*, April 2005, *113* (2), 376–424.

**Chen, Yuyu, Ginger Zhe Jin, Naresh Kumar, and Guang Shi**, "Gaming in Air Pollution Data? Lessons from China," *The B.E. Journal of Economic Analysis & Policy*, January 2012, *12* (3).

**_ , Hongbin Li, and Lingsheng Meng**, "Prenatal Sex Selection and Missing Girls in China: Evidence from the Diffusion of Diagnostic Ultrasound," *Journal of Human Resources*, January 2013, *48* (1), 36–70.

**Chinkhumba, Jobiba, Susan Godlonton, and Rebecca Thornton**, "The Demand for Medical Male Circumcision," *American Economic Journal: Applied Economics*, April 2014, *6* (2), 152–177.

**Chong, Alberto, Marco Gonzalez-Navarro, Dean Karlan, and Martin Valdivia**, "Effectiveness and Spillovers of Online Sex Education: Evidence from a Randomized Evaluation in Colombian Public Schools," Technical Report w18776, National Bureau of Economic Research, Cambridge, MA February 2013.

**Cialdini, Robert B.**, *Influence: science and practice*, 5th ed ed., Boston: Pearson Education, 2009.

**Coffman, Katherine, Lucas Coffman, and Keith M. Marzilli Ericson**, "The Size of the LGBT Population and the Magnitude of Anti-Gay Sentiment are Substantially Underestimated," Technical Report w19508, National Bureau of Economic Research, Cambridge, MA October 2013.

**Committee for Population, Family and Children [Vietnam] and ORC Macro**, "Vietnam Demographic and Health Survey 2002," Technical Report, Committee for Population, Family and Children and ORC Macro, Calverton, Maryland, USA September 2003.

**Corno, Lucia, Nicole Hildebrandt, and Alessandra Voena**, "Age of Marriage, Weather Shocks, and the Direction of Marriage Payments." NBER Working Paper No. 23604, National Bureau of Economic Research, Cambridge, MA July 2017.

**Currie, Janet and Hannes Schwandt**, "Short- and long-term effects of unemployment on fertility," *Proceedings of the National Academy of Sciences*, October 2014, *111* (41), 14734–14739.

_ **and Matthew Neidell**, "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?," *The Quarterly Journal of Economics*, August 2005, *120* (3), 1003–1030.

_ **and Reed Walker**, "Traffic Congestion and Infant Health: Evidence from E-ZPass," *American Economic Journal: Applied Economics*, January 2011, *3* (1), 65–90.

_ **, Joshua Graff Zivin, Jamie Mullins, and Matthew Neidell**, "What Do We Know About Short- and Long-Term Effects of Early-Life Exposure to Pollution?," *Annual Review of Resource Economics*, November 2014, *6* (1), 217–247.

_ **, Lucas Davis, Michael Greenstone, and Reed Walker**, "Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings," *American Economic Review*, February 2015, *105* (2), 678–709.

_ **, Matthew Neidell, and Johannes F. Schmieder**, "Air pollution and infant health: Lessons from New Jersey," *Journal of Health Economics*, May 2009, *28* (3), 688–703.

**DeFranco, Emily, Eric Hall, Monir Hossain, Aimin Chen, Erin N. Haynes, David Jones, Sheng Ren, Long Lu, and Louis Muglia**, "Air Pollution and Stillbirth Risk: Exposure to Airborne Particulate Matter during Pregnancy Is Associated with Fetal Death," *PLOS ONE*, March 2015, *10* (3), e0120594.

**Dehejia, Rajeev and Adriana Lleras-Muney**, "Booms, Busts, and Babies' Health," *The Quarterly Journal of Economics*, August 2004, *119* (3), 1091–1130.

**Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken**, "What Do We Learn from the Weather? The New Climate-Economy Literature," *Journal of Economic Literature*, September 2014, *52* (3), 740–798.

**Dercon, Stefan**, "Income Risk, Coping Strategies, and Safety Nets," *The World Bank Research Observer*, 2002, *17* (2), 141–166.

**Deryugina, Tatyana, Garth Heutel, Nolan Miller, David Molitor, and Julian Reif**, "The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction," Technical Report w22796, National Bureau of Economic Research, Cambridge, MA November 2016.

**Deschenes, Olivier and Enrico Moretti**, "Extreme Weather Events, Mortality, and Migration," *Review of Economics and Statistics*, November 2009, *91* (4), 659–681.

**_ , Michael Greenstone, and Jonathan Guryan**, "Climate Change and Birth Weight," *American Economic Review*, April 2009, *99* (2), 211–217.

**Dettling, Lisa J. and Melissa S. Kearney**, "House prices and birth rates: The impact of the real estate market on the decision to have a baby," *Journal of Public Economics*, February 2014, *110*, 82–100.

**Dockery, Douglas W., C. Arden Pope, Xiping Xu, John D. Spengler, James H. Ware, Martha E. Fay, Benjamin G. Ferris, and Frank E. Speizer**, "An Association between Air Pollution and Mortality in Six U.S. Cities," *New England Journal of Medicine*, December 1993, *329* (24), 1753–1759.

**Dominici, Francesca, Roger D. Peng, Michelle L. Bell, Luu Pham, Aidan McDermott, Scott L. Zeger, and Jonathan M. Samet**, "Fine Particulate Air Pollution and Hospital Admission for Cardiovascular and Respiratory Diseases," *JAMA*, March 2006, *295* (10), 1127.

**Downs, Jennifer A, Agrey H Mwakisole, Alphonce B Chandika, Shibide Lugoba, Rehema Kassim, Evarist Laizer, Kinanga A Magambo, Myung Hee Lee, Samuel E Kalluvya, David J Downs, and Daniel W Fitzgerald**, "Educating religious leaders to promote uptake of male circumcision in Tanzania: a cluster randomised trial," *The Lancet*, March 2017, *389* (10074), 1124–1132.

**Duflo, E. and E. Saez**, "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *The Quarterly Journal of Economics*, August 2003, *118* (3), 815–842.

**Duflo, Esther, Pascaline Dupas, and Michael Kremer**, "Education, HIV, and Early Fertility: Experimental Evidence from Kenya," *American Economic Review*, September 2015, *105* (9), 2757–2797.

**Dyson, Tim**, "Demographic Responses To Famines In South Asia," *IDS Bulletin*, October 1993, *24* (4), 17–26.

**Eaton, Lisa and Seth C Kalichman**, "Behavioral Aspects of Male Circumcision for the Prevention of HIV Infection," *Current HIV/AIDS reports*, November 2009, *6* (4), 187–193.

**Ebenstein, Avraham**, "The "Missing Girls" of China and the Unintended Consequences of the One Child Policy," *Journal of Human Resources*, 2010, *45* (1).

**Edlund, Lena, Hongbin Li, Junjian Yi, and Junsen Zhang**, "Sex Ratios and Crime: Evidence from China," *Review of Economics and Statistics*, December 2013, *95* (5), 1520–1534.

**EPA, U.S.**, "Integrated Science Assessment for Particulate Matter (Final Report)," 2009.

**Erickson, Anders C. and Laura Arbour**, "The Shared Pathoetiological Effects of Particulate Air Pollution and the Social Environment on Fetal-Placental Development," *Journal of Environmental and Public Health*, 2014, *2014*, 1–20.

**Evans, William N. and John D. Graham**, "Risk Reduction or Risk Compensation? The Case of Mandatory Safety-Belt Use Laws," *Journal of Risk and Uncertainty*, 1991, *4* (1), 61–73.

**Faiz, A. S., G. G. Rhoads, K. Demissie, L. Kruse, Y. Lin, and D. Q. Rich**, "Ambient Air Pollution and the Risk of Stillbirth," *American Journal of Epidemiology*, August 2012, *176* (4), 308–316.

**Faiz, Ambarina S., George G. Rhoads, Kitaw Demissie, Yong Lin, Lakota Kruse, and David Q. Rich**, "Does Ambient Air Pollution Trigger Stillbirth?:," *Epidemiology*, July 2013, *24* (4), 538–544.

**Foster, Gigi**, "It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism," *Journal of Public Economics*, September 2006, *90* (8-9), 1455–1475.

**Funk, Chris, Pete Peterson, Martin Landsfeld, Diego Pedreros, James Verdin, Shraddhanand Shukla, Gregory Husak, James Rowland, Laura Harrison, Andrew Hoell, and Joel Michaelsen**, "The climate hazards infrared precipitation with stations-a new environmental record for monitoring extremes," *Scientific Data*, December 2015, *2*, 150066.

**Gammeltoft, Tine and Hanh Thi Thuy Nguyen**, "The Commodification of Obstetric Ultrasound Scanning in Hanoi, Viet Nam," *Reproductive Health Matters*, January 2007, *15* (29), 163–171.

**Gendron-Carrier, Nicolas, Marco Gonzalez-Navarro, Stefano Polloni, and Matthew Turner**, "Subways and Urban Air Pollution," Technical Report w24183, National Bureau of Economic Research, Cambridge, MA January 2018.

**Ghanem, Dalia and Junjie Zhang**, "'Effortless Perfection:' Do Chinese cities manipulate air pollution data?," *Journal of Environmental Economics and Management*, September 2014, *68* (2), 203–225.

**Gibson, John and Bonggeun Kim**, "Quality, Quantity, and Nutritional Impacts of Rice Price Changes in Vietnam," *World Development*, March 2013, *43*, 329–340.

**Glewwe, Paul, Nisha Agrawal, and David Dollar, eds**, *Economic growth, poverty, and household welfare in Vietnam* World Bank regional and sectoral studies, Washington, DC: World Bank, 2004.

**Godlonton, Susan and Rebecca Thornton**, "Peer effects in learning HIV results," *Journal of Development Economics*, January 2012, *97* (1), 118–129.

**Goodkind, Daniel**, "On Substituting Sex Preference Strategies in East Asia: Does Prenatal Sex Selection Reduce Postnatal Discrimination?," *Population and Development Review*, 1996, *22* (1), 111–125.

**Gray, Ron, Godfrey Kigozi, Xiangrong Kong, Victor Ssempiija, Frederick Makumbi, Stephen Wattya, David Serwadda, Fred Nalugoda, Nelson K. Sewenkambo, and Maria J. Wawer**, "The effectiveness of male circumcision for HIV prevention and effects on risk behaviors in a posttrial follow-up study:," *AIDS*, March 2012, *26* (5), 609–615.

**Gray, Ronald H, Godfrey Kigozi, David Serwadda, Frederick Makumbi, Stephen Watya, Fred Nalugoda, Noah Kiwanuka, Lawrence H Moulton, Mohammad A Chaudhary, Michael Z Chen, Nelson K Sewankambo, Fred Wabwire-Mangen, Melanie C Bacon, Carolyn FM Williams, Pius Opendi, Steven J Reynolds, Oliver Laeyendecker, Thomas C Quinn, and Maria J Wawer**, "Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial," *The Lancet*, February 2007, *369* (9562), 657–666.

**Groeger, Andre and Yanos Zylberberg**, "Internal Labor Migration as a Shock Coping Strategy: Evidence from a Typhoon," *American Economic Journal: Applied Economics*, April 2016, *8* (2), 123–153.

**Guilmoto, Christophe Z., Xuyen Hoang, and Toan Ngo Van**, "Recent Increase in Sex Ratio at Birth in Viet Nam," *PLoS ONE*, February 2009, *4* (2), e4624.

**Guiteras, Raymond, Amir Jina, and A. Mushfiq Mobarak**, "Satellites, Self-reports, and Submersion: Exposure to Floods in Bangladesh," *American Economic Review*, May 2015, *105* (5), 232–236.

**Gupta, Monica Das, Jiang Zhenghua, Li Bohua, Xie Zhenming, Woojin Chung, and Bae Hwa-Ok**, "Why is Son preference so persistent in East and South Asia? a cross-country study of China, India and the Republic of Korea," *Journal of Development Studies*, December 2003, *40* (2), 153–187.

**Ha, Sandie, Hui Hu, Dikea Roussos-Ross, Kan Haidong, Jeffrey Roth, and Xiaohui Xu**, "The effects of air pollution on adverse birth outcomes," *Linking Exposure and Health in Environmental Public Health Tracking*, October 2014, *134* (0), 198–204.

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

**HEI International Scientific Oversight Committee**, "Outdoor Air Pollution and Health in the Developing Countries of Asia: A Comprehensive Review." Special Report 18, Health Effects Institute, Boston, MA November 2010.

**Hoek, Gerard, Ranjini M Krishnan, Rob Beelen, Annette Peters, Bart Ostro, Bert Brunekreef, and Joel D Kaufman**, "Long-term air pollution exposure and cardio- respiratory mortality: a review," *Environmental Health*, 2013, *12* (1), 43.

**Hou, Xiaohui**, "Can Drought Increase Total Calorie Availability? The Impact of Drought on Food Consumption and the Mitigating Effects of a Conditional Cash Transfer Program," *Economic Development and Cultural Change*, July 2010, *58* (4), 713–737.

**IPCC**, *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Barros, V.R., C.B. Field, D.J. Dokken, M.D. Mastrandrea, K.J. Mach, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]*, Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2014.

**Isen, Adam, Maya Rossin-Slater, and W. Reed Walker**, "Every Breath You TakeEvery Dollar Youll Make: The Long-Term Consequences of the Clean Air Act of 1970," *Journal of Political Economy*, June 2017, *125* (3), 848–902.

**Ito, Koichiro and Shuang Zhang**, "Willingness to Pay for Clean Air: Evidence from Air Purifier Markets in China," Technical Report w22367, National Bureau of Economic Research, Cambridge, MA June 2016.

**Jaffee, Steven M., Son Kim Dang, Do Anh Tuan Nguyen, Emilie Cassou, Trang Thi Thu Truong, Nguyen Thi Thuy, Mateo Ambrosio Albala, Donald F. Larson, and Bernardita Ledesma**, "Vietnam-Agricultural modernization transforming Vietnamese agriculture: Gaining more for less," Technical Report AUS15856, World Bank Group, Washington, D.C. April 2016.

**Jayachandran, Seema**, "Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries," *Journal of Political Economy*, 2006, *114* (3), 538–575.

_ , "Air Quality and Early-Life Mortality: Evidence from Indonesia's Wildfires," *Journal of Human Resources*, September 2009, *44* (4), 916–954.

_ , "The Roots of Gender Inequality in Developing Countries," *Annual Review of Economics*, August 2015, *7* (1), 63–88.

_ , "Fertility Decline and Missing Women," *American Economic Journal: Applied Economics*, January 2017, *9* (1), 118–139.

**Jia, Ruixue and Hyejin Ku**, "Is China's Pollution the Culprit for the Choking of South Korea? Evidence from the Asian Dust," *Working Paper*, January 2018.

**Kalichman, Seth, Lisa Eaton, and Steven Pinkerton**, "Circumcision for HIV Prevention: Failure to Fully Account for Behavioral Risk Compensation," *PLoS Medicine*, March 2007, *4* (3), e138.

**Kaur, Supreet**, "Nominal Wage Rigidity in Village Labor Markets." NBER Working Paper No. 20770, National Bureau of Economic Research, Cambridge, MA December 2014.

**Khandker, Shahidur R.**, "Seasonality of income and poverty in Bangladesh," *Journal of Development Economics*, March 2012, *97* (2), 244–256.

**Kim, Hyuncheol Bryant and Sun mi Lee**, "When public health intervention is not successful: Cost sharing, crowd-out, and selection in Korea's National Cancer Screening Program," *Journal of Health Economics*, May 2017, *53*, 100–116.

**Kim, Jin-Bae, Changsoo Kim, Eunmi Choi, Sanghoon Park, Hyelim Park, Hui-Nam Pak, Moon-Hyoung Lee, Dong Chun Shin, Ki-Chul Hwang, and Boyoung Joung**, "Particulate air pollution induces arrhythmia via oxidative stress and calcium calmodulin kinase II activation," *Toxicology and Applied Pharmacology*, February 2012, *259* (1), 66–73.

**Knittel, Christopher, Douglas Miller, and Nicholas Sanders**, "Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health," Technical Report w17222, National Bureau of Economic Research, Cambridge, MA July 2011.

**Kong, X., G. Kigozi, F. Nalugoda, R. Musoke, J. Kagaayi, C. Latkin, R. Ssekubugu, T. Lutalo, B. Nantume, I. Boaz, M. Wawer, D. Serwadda, and R. Gray**, "Assessment of Changes in Risk Behaviors During 3 Years of Posttrial Follow-up of Male Circumcision Trial Participants Uncircumcised at Trial Closure in Rakai, Uganda," *American Journal of Epidemiology*, November 2012, *176* (10), 875–885.

**Kremer, Michael and Dan Levy**, "Peer Effects and Alcohol Use among College Students," *Journal of Economic Perspectives*, July 2008, *22* (3), 189–206.

**Lavaine, Emmanuelle and Matthew Neidell**, "Energy Production and Health Externalities: Evidence from Oil Refinery Strikes in France," *Journal of the Association of Environmental and Resource Economists*, June 2017, *4* (2), 447–477.

**Lee, Seungmin, Chang-Hoi Ho, and Yong-Sang Choi**, "High-PM10 concentration episodes in Seoul, Korea: Background sources and related meteorological conditions," *Atmospheric Environment*, December 2011, *45* (39), 7240–7247.

**_ , _ , Yun Gon Lee, Hyoung-Jin Choi, and Chang-Keun Song**, "Influence of transboundary air pollutants from China on the high-PM10 episode in Seoul, Korea for the period October 1620, 2008," *Atmospheric Environment*, October 2013, *77*, 430–439.

**Lin, Jintai, Da Pan, Steven J. Davis, Qiang Zhang, Kebin He, Can Wang, David G. Streets, Donald J. Wuebbles, and Dabo Guan**, "China's international trade and air pollution in the United States," *Proceedings of the National Academy of Sciences*, February 2014, *111* (5), 1736–1741.

**Lindo, Jason M.**, "Are Children Really Inferior Goods?: Evidence from Displacement-Driven Income Shocks," *Journal of Human Resources*, March 2010, *45* (2), 301–327.

**Lleras-Muney, Adriana**, "The Needs of the Army: Using Compulsory Relocation in the Military to Estimate the Effect of Air Pollutants on Children's Health," *Journal of Human Resources*, 2010, *45* (3), 549–590.

**Lobell, David B and Christopher B Field**, "Global scale climate-crop yield relationships and the impacts of recent warming," *Environmental Research Letters*, March 2007, *2* (1), 014002.

**Lohmann, Steffen and Tobias Lechtenfeld**, "The Effect of Drought on Health Outcomes and Health Expenditures in Rural Vietnam," *World Development*, August 2015, *72*, 432–448.

**Lovenheim, Michael F. and Kevin J. Mumford**, "Do Family Wealth Shocks Affect Fertility Choices? Evidence from the Housing Market," *Review of Economics and Statistics*, May 2013, *95* (2), 464–475.

**Lyle, David S**, "Estimating and Interpreting Peer and Role Model Effects from Randomly Assigned Social Groups at West Point," *Review of Economics and Statistics*, May 2007, *89* (2), 289–299.

**Maccini, Sharon and Dean Yang**, "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall," *American Economic Review*, May 2009, *99* (3), 1006–1026.

**MacDorman, Marian F. and Sharon Kirmeyer**, "Fetal and perinatal mortality, United States, 2005," *National vital statistics reports*, January 2009, *57* (8).

**Mas, Alexandre and Enrico Moretti**, "Peers at Work," *American Economic Review*, February 2009, *99* (1), 112–145.

**Mathews, F., P. J Johnson, and A. Neil**, "You are what your mother eats: evidence for maternal preconception diet influencing foetal sex in humans," *Proceedings of the Royal Society B: Biological Sciences*, July 2008, *275* (1643), 1661–1668.

**Mattson, Christine L., Richard T. Campbell, Robert C. Bailey, Kawango Agot, J. O. Ndinya-Achola, and Stephen Moses**, "Risk Compensation Is Not Associated with Male Circumcision in Kisumu, Kenya: A Multi-Faceted Assessment of Men Enrolled in a Randomized Controlled Trial," *PLoS ONE*, June 2008, *3* (6), e2443.

**MDHS**, "Malawi Demographic and Health Survey 2010," Technical Report, NSO and ICF Macro, Zomba, Malawi and Calverton, Maryland, USA 2011.

**Miguel, Edward and Michael Kremer**, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, January 2004, *72* (1), 159–217.

**Miller, J. D.**, “A New Survey Technique for Studying Deviant Behavior,” Technical Report Ph.D. thesis, The George Washington University 1984.

**Minnis, A. M., M. J. Steiner, M. F. Gallo, L. Warner, M. M. Hobbs, A. van der Straten, T. Chipato, M. Macaluso, and N. S. Padian**, “Biomarker Validation of Reports of Recent Sexual Activity: Results of a Randomized Controlled Study in Zimbabwe,” *American Journal of Epidemiology*, October 2009, *170* (7), 918–924.

**NAC**, “Malawi Voluntary Medical Male Circumcision Communication Strategy 2012-2016,” Technical Report, National AIDS Commission (NAC) Malawi 2012.

**National Research Council, ed.**, *Global sources of local pollution: an assessment of long-range transport of key air pollutants to and from the United States*, Washington, DC: National Acad. Press, 2010.

**National Statistical Office Malawi and ICF Macro**, “Malawi Demographic and Health Survey 2010,” Technical Report, NSO/Malawi and ICF Macro, Zomba, Malawi 2011.

**Obasi, Angela, Frank Mosha, Maria Quigley, Zebedayo Sekirassa, Tom Gibbs, Katua Munguti, James Todd, Heiner Grosskurth, Philippe Mayaud, John Changalucha, David Brown, David Mabey, and Richard Hayes**, “Antibody to Herpes Simplex Virus Type 2 as a Marker of Sexual Risk Behavior in Rural Tanzania,” *The Journal of Infectious Diseases*, January 1999, *179* (1), 16–24.

**Oster, Emily and Rebecca Thornton**, “Determinants of Technology Adoption: Peer Effects in Menstrual Cup Take-up,” *Journal of the European Economic Association*, December 2012, *10* (6), 1263–1293.

**Padian, Nancy S, Anne Buve, Jennifer Balkus, David Serwadda, and Ward Cates**, “Biomedical interventions to prevent HIV infection: evidence, challenges, and way forward,” *The Lancet*, August 2008, *372* (9638), 585–599.

**Palen, Lori-Ann, Edward A. Smith, Linda L. Caldwell, Alan J. Flisher, Lisa Wegner, and Tania Vergnani**, “Inconsistent Reports of Sexual Intercourse Among South African High School Students,” *Journal of Adolescent Health*, March 2008, *42* (3), 221–227.

**PATH and Reproductive Health Department**, “Examining the Cost of Providing Medical Abortion in Vietnam,” Technical Report, Ministry of Health, Hanoi, Vietnam August 2006.

**Paxson, Christina H.**, “Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand,” *The American Economic Review*, 1992, *82* (1), 15–33.

_ , “Consumption and Income Seasonality in Thailand,” *Journal of Political Economy*, February 1993, *101* (1), 39–72.

**Peltzman, Sam**, “The Effects of Automobile Safety Regulation,” *Journal of Political Economy*, 1975, *83* (4), 677–725.

**Pereira, L A, D Loomis, G M Conceicao, A L Braga, R M Arcas, H S Kishi, J M Singer, G M Boehm, and P H Saldiva**, "Association between air pollution and intrauterine mortality in Sao Paulo, Brazil.," *Environmental Health Perspectives*, June 1998, *106* (6), 325–329.

**Pham, Bang Nguyen, Wayne Hall, and Peter Stewart Hill**, "Indirect evidence of the contribution of prenatal sex selection practices to the high sex ratio at birth in Vietnam," *Journal of Population Research*, 2011, *28* (4), 293–299.

**_ , _ , Peter S Hill, and Chalapati Rao**, "Analysis of socio-political and health practices influencing sex ratio at birth in Viet Nam," *Reproductive Health Matters*, January 2008, *16* (32), 176–184.

**Pitt, Mark M. and Wendy Sigle**, "Seasonality, Weather Shocks and the Timing of Births and Child Mortality in Senegal." Working Paper, Brown University April 1997.

**Pop-Eleches, Cristian**, "The Impact of an Abortion Ban on Socioeconomic Outcomes of Children: Evidence from Romania," *Journal of Political Economy*, August 2006, *114* (4), 744–773.

**Pope, C. Arden, Michael J. Thun, Mohan M. Namboodiri, Douglas W. Dockery, John S. Evans, Frank E. Speizer, and Clark W. Heath**, "Particulate Air Pollution as a Predictor of Mortality in a Prospective Study of U.S. Adults," *American Journal of Respiratory and Critical Care Medicine*, March 1995, *151* (3_pt_1), 669–674.

**Pope, C Arden, Richard T Burnett, Michael J Thun, Eugenia E Calle, Daniel Krewski, Kazuhiko Ito, and George D Thurston**, "Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution," *JAMA : the journal of the American Medical Association*, March 2002, *287* (9), 1132–1141.

**Rasmussen, Svein, Susanne Albrechtsen, Lorentz M. Irgens, Knut Dalaker, Helga Maartmann-Moe, Ljiljana Vlatkovic, and Trond Markestad**, "Unexplained antepartum fetal death in Norway, 1985-97: diagnostic validation and some epidemiologic aspects," *Acta Obstetricia et Gynecologica Scandinavica*, February 2003, *82* (2), 109–115.

**Ritz, B and F Yu**, "The effect of ambient carbon monoxide on low birth weight among children born in southern California between 1989 and 1993.," *Environmental Health Perspectives*, January 1999, *107* (1), 17–25.

**Ritz, B., M. Wilhelm, K. J. Hoggatt, and J. K. C. Ghosh**, "Ambient Air Pollution and Preterm Birth in the Environment and Pregnancy Outcomes Study at the University of California, Los Angeles," *American Journal of Epidemiology*, August 2007, *166* (9), 1045–1052.

**Roncati, Luca, Teresa Pusiol, Francesco Piscioli, Giuseppe Barbolini, Antonio Maiorana, and Anna Lavezzi**, "The First 5-Year-Long Survey on Intrauterine Unexplained Sudden Deaths from the Northeast Italy," *Fetal and Pediatric Pathology*, September 2016, *35* (5), 315–326.

**Rose, Elaina**, "Consumption Smoothing and Excess Female Mortality in Rural India," *The Review of Economics and Statistics*, 1999, *81* (1), 41–49.

**Sacerdote, B.**, "Peer Effects with Random Assignment: Results for Dartmouth Roommates," *The Quarterly Journal of Economics*, May 2001, *116* (2), 681–704.

**Samet, Jonathan M., Francesca Dominici, Frank C. Curriero, Ivan Coursac, and Scott L. Zeger**, "Fine Particulate Air Pollution and Mortality in 20 U.S. Cities, 1987-1994," *New England Journal of Medicine*, December 2000, *343* (24), 1742–1749.

**Sanders, Nicholas J. and Charles Stoecker**, "Where have all the young men gone? Using sex ratios to measure fetal death rates," *Journal of Health Economics*, May 2015, *41* (0), 30–45.

**Sawires, Sharif R, Shari L Dworkin, Agnes Fiamma, Dean Peacock, Greg Szekeres, and Thomas J Coates**, "Male circumcision and HIV/AIDS: challenges and opportunities," *The Lancet*, February 2007, *369* (9562), 708–713.

**Schlenker, Wolfram and W. Reed Walker**, "Airports, Air Pollution, and Contemporaneous Health," *The Review of Economic Studies*, April 2016, *83* (2), 768–809.

**Schultz, T. Paul**, "Changing World Prices, Women's Wages, and the Fertility Transition: Sweden, 1860-1910," *Journal of Political Economy*, December 1985, *93* (6), 1126–1154.

**Sedgh, Gilda, Jonathan Bearak, Susheela Singh, Akinrinola Bankole, Anna Popinchalk, Bela Ganatra, Clementine Rossier, Caitlin Gerdts, Ozge Tuncalp, Brooke Ronald Johnson, Heidi Bart Johnston, and Leontine Alkema**, "Abortion incidence between 1990 and 2014: global, regional, and subregional levels and trends," *The Lancet*, July 2016, *388* (10041), 258–267.

_ **, Stanley K. Henshaw, Susheela Singh, Akinrinola Bankole, and Joanna Drescher**, "Legal Abortion Worldwide: Incidence and Recent Trends," *Perspectives on Sexual and Reproductive Health*, 2007, *39* (4), 216–225.

**Shah, Manisha and Bryce Millett Steinberg**, "Drought of Opportunities: Contemporaneous and Long-Term Impacts of Rainfall Shocks on Human Capital," *Journal of Political Economy*, March 2017, *125* (2), 527–561.

**Sobotka, Tomas, Vegard Skirbekk, and Dimiter Philipov**, "Economic Recession and Fertility in the Developed World," *Population and Development Review*, June 2011, *37* (2), 267–306.

**Teerawichitchainan, Bussarawan and Sajeda Amin**, "The Role of Abortion in the Last Stage of Fertility Decline in Vietnam," *International Perspectives on Sexual and Reproductive Health*, June 2010, *36* (2), 80–82.

**Thirumurthy, Harsha, Samuel H. Masters, Samwel Rao, Megan A. Bronson, Michele Lanham, Eunice Omanga, Emily Evens, and Kawango Agot**, "Effect of Providing Conditional Economic Compensation on Uptake of Voluntary Medical Male Circumcision in Kenya: A Randomized Clinical Trial," *JAMA*, August 2014, *312* (7), 703.

**Thompson, D C, R S Thompson, and F P Rivara**, "Risk compensation theory should be subject to systematic reviews of the scientific evidence," *Injury Prevention*, June 2001, *7* (2), 86.

**Thornton, Rebecca L**, "The Demand for, and Impact of, Learning HIV Status," *American Economic Review*, November 2008, *98* (5), 1829–1863.

**Tobian, Aaron A.R., David Serwadda, Thomas C. Quinn, Godfrey Kigozi, Patti E. Gravitt, Oliver Laeyendecker, Blake Charvat, Victor Ssempijja, Melissa Riedesel, Amy E. Oliver, Rebecca G. Nowak, Lawrence H. Moulton, Michael Z. Chen, Steven J. Reynolds, Maria J. Wawer, and Ronald H. Gray**, "Male Circumcision for the Prevention of HSV-2 and HPV Infections and Syphilis," *New England Journal of Medicine*, March 2009, *360* (13), 1298–1309.

**Townsend, Robert M.**, "Risk and Insurance in Village India," *Econometrica*, May 1994, *62* (3), 539.

**UNAIDS**, "Global report: UNAIDS report on the global AIDS epidemic 2013," Technical Report, UNAIDS, UNAIDS November 2013.

**United Nations Population Fund**, *Ageing in the twenty-first century: a celebration and a challenge.* 2012.

**Vu, Linh and Paul Glewwe**, "Impacts of Rising Food Prices on Poverty and Welfare in Vietnam," *Journal of Agricultural and Resource Economics*, 2011, *36* (1), 14–27.

**Wainwright, Fiona and Carol Newman**, "Income shocks and household risk-coping strategies: evidence from rural Vietnam," *Institute for International Integration Studies Discussion paper*, 2011, (358).

**Weiss, Helen A, Kim E Dickson, Kawango Agot, and Catherine A Hankins**, "Male circumcision for HIV prevention: Current research and programmatic issues," *AIDS (London, England)*, October 2010, *24* (0 4), S61–S69.

**Whittaker, Andrea, ed.**, *Abortion in Asia: local dilemmas, global politics* number v. 20. In 'Fertility, reproduction, and sexuality.', New York: Berghahn Books, 2010.

**WHO/UNAIDS**, "New Data on Male Circumcision and HIV Prevention: Policy and Programme Implications," Technical Report, Montreux, Switzerland 2007.

**Wilson, Nicholas L., Wentao Xiong, and Christine L. Mattson**, "Is sex like driving? HIV prevention and risk compensation," *Journal of Development Economics*, January 2014, *106*, 78–91.

**Woods, Robert**, "Long-term trends in fetal mortality: implications for developing countries," *Bulletin of the World Health Organization*, 2008, *86* (6), 460–466.

**Workowski, Kimberly A. and Gail A. Bolan**, "Sexually Transmitted Diseases Treatment Guidelines, 2015," Technical Report MMWR Recomm Rep 2015;64(No. RR-3):1-137, Center for Disease Control and Prevention June 2015.

**World Health Orgarnization (WHO)**, "Progress in scaling up voluntary medical male circumcision for HIV prevention in East and Southern Africa, January-December 2012," Technical Report, World Health Orgarnization, World Health Orgarnization 2012.

**Xiao, Qingyang, Zongwei Ma, Shenshen Li, and Yang Liu**, "The Impact of Winter Heating on Air Pollution in China," *PLOS ONE*, January 2015, *10* (1), e0117311.

**Zimmerman, David J.**, "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment," *Review of Economics and Statistics*, February 2003, *85* (1), 9–23.

**Zivin, Joshua Graff and Matthew Neidell**, "Environment, Health, and Human Capital," *Journal of Economic Literature*, September 2013, *51* (3), 689–730.