Improving Reproductive Health: Assessing Determinants and Measuring Policy Impacts

Improving Reproductive Health:
Assessing Determinants and Measuring Policy Impacts

A Dissertation Presented

By

Slawa Rokicki

To

The Committee on Higher Degrees in Health Policy

In Partial Fulfilment of the Requirements

For the Degree of

Doctor of Philosophy

In the Subject of

Health Policy

Harvard University

Cambridge, Massachusetts

March 2016

Improving Reproductive Health:
Assessing Determinants and Measuring Policy Impacts

Abstract

In this thesis, I investigate policies and programs to improve reproductive health. My thesis makes a substantive contribution to reproductive health policy and a methodological contribution to quasi-experimental research.

In chapter 1, I evaluate the impact of a mobile phone intervention for adolescent girls. I design and implement a randomized controlled trial in Ghana to test whether sending information via mobile phones is an effective way to improve girls' knowledge of sexual health and to ultimately reduce teenage pregnancy. I find that mobile phone programs are effective not only in increasing knowledge, but also in decreasing risk of pregnancy among sexually active adolescents. I discuss the results in the context of sexual education policy in Ghana.

In chapter 2, I explore the complex interactions between migration and reproductive health. I reconstruct the complete migration and reproductive health histories of women residing in the urban slums of Accra, Ghana. Using individual fixed effects to reduce selection bias, I find an increased risk of pregnancy, miscarriage, and abortion in the 48 months after migration, with no significant increase in the chance of live birth during this time period. With half of abortions in Ghana classified as unsafe, these results suggest that policies which target the rapidly growing number of urban migrants by providing access to contraception and public hospital services may reduce unsafe abortion and improve maternal health outcomes.

Dissertation advisor: Professor Joshua A. Salomon Slawa Rokicki

In chapter 3, I investigate the bias of standard errors in difference-in-differences estimation, which typically evaluates the effect of a group-level intervention on individual data. Common modeling adjustments for grouped data, such as cluster-robust standard errors, are biased when the number of clusters is small. I run Monte Carlo simulations to investigate both the coverage and power of a wide variety of modeling solutions from the econometric and biostatistics fields, while varying the balance of cluster sizes, the degree of error correlation, and the proportion of treated clusters. I then apply my results to re-evaluate a recently published study on the effect of emergency contraception on adolescent sexual behavior. I find that the study's results claiming that emergency contraception increases risky sexual behavior may be spurious once proper adjustments for grouped data are applied.

Contents

Acknowledgements

I am very grateful to everyone who gave me support and guidance over the past four years to write this dissertation.

My advisor and mentor, Joshua Salomon, who I have relied on not only for exceptional academic advising but for a great deal of life coaching. Over the years, Josh has instilled in me a resiliency and self-confidence that I know will take me far in my career, not only because of his consistent praise for my work, but because he has allowed me to make (many) mistakes along the way without fear.

Günther Fink, who has supported every one of my ideas and interfered only to make them better. In the fall of 2012, I sent Günther a short email for the idea of what turned into Chapter 1 and the heart of this dissertation, and his immediate response was “definitely count me in as an advisor”. I have since counted on him an infinite number of times and he has come through for me on every one.

Jessica Cohen, who in addition to providing me with excellent advice on everything from RCTs to career moves, has been my role model and constant source of encouragement.

Mary Beth Landrum, who I instantly felt a connection to in my second year preparing for qualifying exams. She has added a much needed perspective to this work, both statistically and substantively, and taught me a great deal in the process.

In addition to my committee, several faculty have provided me with guidance and mentorship including Alan Zaslavsky, Kathy Swartz, David Cutler, Tom McGuire, Laura Hatfield, Maggie McConnell, and Joe Newhouse.

Debbie Whitney, Jessica Livingston, and Ayres Heller, who went above and beyond so many times to help me in my research endeavours and my academic needs, and who really care about my wellbeing and success as a graduate student.

My colleagues and my staff at Innovations for Poverty Action; I could not have done the SMART trial without their skills, support, and good humor. I am grateful to my first project manager, Comfort Bonney Arku, who worked with me side-by-side in the first year of the project. My other IPA staff who all worked tirelessly and always at 100%: project manager Grace Gletsu and team leaders Stephanie Adjovu, Carlina Apraku, Angela Adzogble, and Felicity Ansong. My IPA colleagues: Kwame Abrokwa, Richard Attibu, Maham Farhat, Loïc Watine, Aubrey Levine, Pace Phillips, Philip Amara. And above all, I’m very grateful to the participants who gave their time and energy over and over again during the course of this study.

My fellow students and good friends in the Health Policy Program: Prachi Sanghavi, Katherine Donato, Yusuke Tsugawa, and Natalie Carvalho.

My mom, Alina Rokicki, who taught me from a very young age the transformational power of education. I would not be where I am today without her.

My family, Konrad Rokicki, Ciocia Jola, Babcia, and Dave DiLillo, who are always there for me.

Laura Perovich, with whom I’ve had a million conversations about this thesis (and every other thing).

And Mark E. McGovern, who inspires me every day; who not only endured all of my fits of panic about so many aspects of this work, but who gave me the calm, clarity, and courage to go on; and who gives my life meaning and overwhelming happiness.

List of Tables

List of Figures

Synopsis

In this thesis, I investigate policies and programs to improve reproductive health. My thesis makes a substantive contribution to reproductive health policy and a methodological contribution to quasi-experimental research. I employ a variety of methodological techniques to obtain causal estimates, subject to the limitations that I discuss. In chapter one, I design a cluster-randomized controlled trial to investigate the effectiveness of a mobile phone program on adolescents' reproductive health knowledge and behavior. In chapter two, where a randomized trial is not possible, I use quasi-experimental methods on observational data to estimate the impact of mobility on the risk of unintended pregnancy. Finally, in chapter three, I simulate the potential bias in a common quasi-experimental analysis and demonstrate the implications of my findings for health policy research.

Sexual and reproductive health is central to human development, affecting a broad range of health, social, and economic outcomes. According to recent estimates by the Guttmacher Institute, 225 million women worldwide who want to avoid a pregnancy are not using an effective method of contraception [1]. Only half of the 125 million women who give birth each year get adequate care. Failing to meet women's contraceptive and maternal health care needs leads to an estimated 290,000 pregnancy-related deaths and 2.9 million infant deaths each year [1]. Adolescent pregnancies are particularly vulnerable to poor health outcomes – with teenage pregnancies more likely than those of older women to result in birth complications – and are also likely to have a damaging effect on girls' education and skill development [2]. Despite gains made over the past several decades in reducing maternal and newborn deaths, improving women's reproductive health outcomes remains a significant challenge. In my thesis, I explore policies and programs to improve reproductive health, in the United States and in developing countries.

I begin, with chapter one, by employing experimental methods to investigate the possibility of a mobile phone program to improve reproductive health among Ghanaian adolescent girls. As new mobile phone connections grow at 30% a year in sub-Saharan Africa [3], mobile health ("mHealth") programs have rapidly gained momentum in international development. In 2014, USAID listed more than 400 mHealth projects across 74 countries from over 100 different organizations [4]. However, very few of these programs have been rigorously evaluated with regards to their health impact, particularly in developing countries [5,6]. In chapter one, I fill a gap in the evidence on both the feasibility and the effectiveness of mHealth for adolescents. Using a cluster-randomized controlled trial, I show that mobile phone programs may be powerful tools not only for increasing adolescents' knowledge of sexual health issues, but also for reducing unintended pregnancy among sexually active adolescent girls. My research shows that interactive features in mHealth interventions are important for increasing knowledge, but that these features are not as important for impacting pregnancy and sexual behavior. The research contributes to our knowledge of what kinds of programs are effective in mHealth; moreover, as the first randomized trial of an mHealth program on adolescent health outcomes in a developing country, it opens the door for many other avenues of future mHealth research such as the role of information diffusion, the mediating influence of social networks, and generalizability of the results to other study populations.

Next, in chapter two, I use quasi-experimental methods to study the effect of female mobility on reproductive health outcomes. Over the coming decades, urbanization is expected to continue or accelerate in the developing world, with Africa and Asia urbanizing most rapidly [7]. One of the most significant recent trends in migration has been the entry of women into migration streams that in previous decades had been primarily male, with an increasing number of female migrants moving on their own [8]. Many rural-to-urban migrants settle in slums, contributing to a projection of a doubling of slum settlements over the next 30 years [9].

I examine how mobility impacts women's reproductive health outcomes. Although there is a comprehensive literature on migration and fertility, I provide a contribution in three key ways. First, I use quasi-experimental methods, in particular individual fixed effects models, to reduce selection bias that confounds the relationship between migration and sexual behaviour. Second, I focus on the context of migration to poor residential neighborhoods, the primary force underlying the rapid rates of urbanization observed in developing countries. Lastly, I examine the relationship between mobility and abortion; though previous studies have documented increases in HIV and risky sexual behaviour due to migration to slums, no previous study has investigated whether female migration has an impact on rates of induced abortion. I show that migration is associated with an increase in risk of pregnancy and abortion in the first 48 months after a move, even after accounting for time invariant unmeasurable attributes.

These findings have important policy implications. Africa is urbanizing rapidly: by 2035, 50% of sub-Saharan Africa will live in urban areas [7], and the concern over the welfare of migrants will become increasingly important to policy-makers. My research shows that policies that target recent migrants to slums may reduce risk of unintended pregnancy and improve health. Possible policies could include integrating quality family planning services into high volume clinic settings, minimizing stockouts and supply chain disruptions, and providing transportation and family planning service vouchers to new arrivals. Evaluating such policies provide further avenues of research.

Lastly, in chapter three, I examine the appropriateness of a common methodology used to make policy recommendations. In health policy research, difference-in-difference (DID) estimation is an increasingly popular way to evaluate the impact of a group-level policy using individual data, for example, the effect of a new state law or the adoption of a policy in some public hospitals but not in others. Because observations are grouped, modeling adjustments must be made to account for the correlation in outcomes. However, a large literature has shown that when the number of groups is small, common approaches for adjustment – such as the cluster-robust variance matrix – may lead to standard errors that are too small [10–12] and therefore inaccurate conclusions about whether policies are effective or not.

I contribute to this body of research in a number of ways. First, I compare the empirical performance of the most commonly used modeling solutions in DID estimation for panel data from the fields of both Econometrics and Statistics, including cluster-robust standard errors, wild cluster bootstrapping, random effects models, GEE with bias corrections, permutation tests, and aggregation.

Second, I present results from a Monte Carlo simulation study in which I test a wide range of scenarios, by varying the degree of error correlation, the balance of cluster sizes, and the proportion of treated clusters. Third, I directly compare empirical coverage rates to power for all models. Finally, using replication, I show the implications of the findings for health policy research. I replicate the results of a recent article that claims that over-the-counter emergency contraception increases risky sexual behaviour in teens. However, the proportion of treated groups (states in this case) in the original analysis is low. When I apply methodologies with adequate coverage rates, I find that the effects are no longer statistically significant. Overall, my research demonstrates the importance of appropriate methodology in the context of clustered data when the number of clusters is small. Without proper adjustment for serial correlation in DID estimation, spurious results may promote poor public policymaking by potentially wasting funding on ineffective programs or cutting effective ones.

In summary, this thesis has two main goals. First, I investigate the impact of two emerging issues in developing countries on reproductive health – the possibilities of mobile phone technology to improve the health of adolescent girls (chapter one) and the impact of migration & urbanization on women's reproductive health (chapter two). I discuss the policy implications of both papers. Second, I demonstrate the importance of appropriate methodology in a popular area of health policy research to avoid spurious results, and demonstrate the importance of my findings on U.S. health policy (chapter three). Together, this work contributes to an evidence base on the worldwide impacts of programs and policies to improve women's reproductive health. As next steps, my research will build on the work presented here. In the future, I aim to investigate how mobile phone programs may be combined with structural interventions focusing on educational and economic opportunities to improve girls' health outcomes. I will also explore the impact of social networks on information diffusion in the context of adolescent mHealth interventions. Finally, I will explore whether mHealth could play a role in identifying and reaching newly arrived migrants to urban areas, as a means of increasing female migrants' access to contraceptive products and health services.

References

1. Susheela Singh, Jacqueline E. Darrroch, Lori S. Ashford. Adding It Up: The Costs and Benefits of Investing in Sexual and Reproductive Health. New York: Guttmacher Institute and United Nations Population Fund; 2014.

2. Sawyer SM, Afifi RA, Bearinger LH, Blakemore S-J, Dick B, Ezeh AC, et al. Adolescence: a foundation for future health. The Lancet. 2012;379: 1630–1640.

3. UN-Habitat. State of the World's Cities 2012/2013. Prosperity of Cities. New York, NY: UN-Habitat; 2013.

4. USAID. mHealth Working Group Inventory of Projects [Internet]. 2014. Available: https://www.mhealthworkinggroup.org/projects/mhealth-working-group-inventory-projects

5. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? PLoS Med. 2013;10: e1001382.

6. Aranda-Jan CB, Mohutsiwa-Dibe N, Loukanova S. Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. BMC Public Health. 2014;14: 188. doi:10.1186/1471-2458-14-188

7. United Nations. World urbanization prospects: The 2011 revision (Economic & Social Affairs report). New York, NY: UN Population Division; 2012.

8. Zlotnik H. The global dimensions of female migration (UN International Migration Report: 2002). [Internet]. Washington D.C.: Migration Policy Institute; 2003. Available: http://www.migrationinformation.org/Feature/display.cfm?ID=109

9. UN-Habitat. Planning sustainable cities: Global report on human settlements 2009. London, UK: Earthscan; 2009.

10. Cameron AC, Miller DL. A Practitioner's Guide to Cluster-Robust Inference. J Hum Resour. 2015;50: 317–372. doi:10.3368/jhr.50.2.317

11. Donald SG, Lang K. Inference with Difference-in-Differences and Other Panel Data. Rev Econ Stat. 2007;89: 221–233. doi:10.1162/rest.89.2.221

12. Bertrand M, Duflo E, Mullainathan S. How Much Should We Trust Differences-In-Differences Estimates? Q J Econ. 2004;119: 249–275. doi:10.1162/003355304772839588

Chapter 1:

Impact of a Mobile Phone Program on Adolescent Reproductive Health: A Cluster-Randomized Trial

Abstract

Background: Knowledge of modern contraceptive methods remains limited among adolescent girls in developing countries, and rates of unintended pregnancy are high. We assess whether a mobile phone sexual education intervention improves reproductive health.

Methods: We conducted a cluster-randomized controlled trial among secondary school students of ages 14-23 in Accra, Ghana. We randomized 38 schools via computer-generated random numbers to Unidirectional intervention (12 schools), Interactive intervention (12), and control (14). The Unidirectional intervention sent participants short message service text messages (SMS) with reproductive health information. The Interactive intervention engaged adolescents in SMS quiz-games about reproductive health. Programs were run for 12 weeks. The main study outcomes were pregnancy at 15 months after program start and participants' reproductive health knowledge at 3 and 15 months. The primary analysis was by intent-to-treat.

Results: 247 students (12 schools) completed the Unidirectional intervention, 197 students (10 schools) completed the Interactive, and 277 students (12 schools) completed the trial in the control group. From baseline to 3 months, average knowledge scores increased from 25% to 33% in the Control, 30% to 44% in the Unidirectional (effect size 11, 95% CI (7 to 15), p-value<0·0001) and 30% to 57% in the Interactive (effect size 24, 95% CI (19 to 28), p-value<0·0001). Knowledge was retained at 15 months for the Unidirectional (45%) and Interactive (54%).There was no impact of the intervention on pregnancy in the past year for the full sample of participants. Among sexually active adolescents, Unidirectional and

Interactive interventions lowered the odds of pregnancy by 87% (5/63, Adjusted OR 0·13, 95% CI 0·03–0·68, p=0·02) and 84% (4/51, Adjusted OR 0·16, 95% CI 0·03–0·89, p=0.04), respectively, compared to the Control (9/58). Mixed results were found for intervention effects on contraceptive use.

Conclusions: Text messaging programs can lead to large and sustained improvements in reproductive health knowledge among adolescents in low-income settings. While effects on overall pregnancy rates are unclear, the programs may be effective in reducing pregnancy risk among sexually active girls.

Trial Registration: This trial is registered with ClinicalTrials.gov, NCT02031575, as the Study on mHealth and Reproductive Health in Teens (SMART).

Keywords: Reproductive Health • Sexual Health • Adolescent Health • Mobile Health • SMS

1.1 Introduction

More than 13 million adolescent girls give birth each year, and over 95% of these births occur in low- and middle-income countries (LMICs) [1]. Adolescent pregnancies are associated with an increased risk of unsafe abortion [2], maternal mortality [3,4], child mortality [5], impaired fetal growth [2,6], birth complications [7], stunting [8], and early school exit and social stigmatization [9]. Despite the large number of risk factors associated with teenage pregnancies, reproductive health knowledge and the adoption of modern contraception remain low in many developing countries [10–12], where more than 50% of unmarried, sexually active 15–19-year-olds have an unmet need for modern contraception [1].

Mobile phone-based programs offer a promising new platform to improve sexual and reproductive health, in particular among adolescents. In 2011, average mobile phone ownership was 40% among 15–18-year-olds in sub-Saharan Africa, with some countries reaching ownership rates of more than 80% in this age range [13]. The past decade has seen a rapid rise in short message service (SMS, or "text messaging") programs that aim to improve health (called "mHealth") [14–18]; however, systematic reviews have consistently found a dearth of high-quality peer-reviewed studies examining outcomes of mHealth programs [19–21], with no evidence on the effectiveness of these interventions among adolescent populations in developing countries.

To examine the potential of mobile phone sex education programs to improve adolescent reproductive health, we conducted a randomized controlled trial in Ghana, investigating the effectiveness of both unidirectional and interactive programs on knowledge and sexual behavior.

1.2 Methods

Study Setting

The cluster-randomized study was conducted in Accra, Ghana. Since 1996, Ghana's education policy has stipulated that secondary schools must provide family life education. However, more than 60% of adolescents say they receive their information about HIV/AIDS and contraception from mass media sources like radio, TV, and the internet [22]. A number of studies have exposed the misperceptions and lack of knowledge that young people have about reproductive health issues, indicating the gap between the policy and its implementation [11,23]. Particularly in religious countries such as Ghana, teachers are often uncomfortable teaching about sexuality and are usually inadequately prepared, leaving out basic information and key aspects of sexual health such as condoms and contraception [24]. Reproductive health knowledge among adolescents is low: 56% of Ghanaian adolescents consider washing after sex an option to prevent pregnancy and 62% are not aware a girl can get pregnant if she has sex standing up [22].

According to the 2014 Ghana Demographic and Health Survey, half of girls have sexual intercourse before the age of 18, but less than a third of sexually active unmarried 15-19-year-old females use any form of modern contraception [25]. From 2008 to 2014, birth rates increased from 66 to 76 births per 1000 girls aged 15-19; total unmet need for family planning is highest among women aged 15-19 compared to any other age group (51 percent) [25]. The prevalence of adolescent pregnancy remains high: 42% of sexually experienced 15–19-year-olds report prior pregnancies [26], with three in five births classified as unintended [11]. Abortion, though highly stigmatized, is common among adolescents: one study found that 40% of 15-19 year old girls who reported having ever been pregnant had obtained an abortion [27]. In the 2004 National Survey of Adolescents, 29% of 15-19 year olds girls said that a close friend had tried to end a pregnancy [26].

Human Subjects

IRB approval was granted by the Committee on the Use of Human Subjects in Research at Harvard University (IRB13-1647) as well as the Ghana Health Service Ethical Review Committee (GHS-ERC:05/09/13). The pre-specified statistical analysis plan was registered at AEAregistry.org (AEARCTR-0000180).

Participants

The sampling frame for the study was provided by the 2012–2013 Ghana Education Service Register of Secondary Schools in Greater Accra. The primary sampling unit for the study was secondary schools. Sampling was restricted to day schools (boarding schools were excluded). Within schools, sampling was restricted to girls aged 14–23. Participants gave written consent, with those under age 18 years obtaining parental consent, and were informed that they could exit from the study at any time.

Randomization

We randomized 38 schools to 12 Unidirectional intervention, 12 Interactive intervention, and 14 control schools. Randomization was done based on a computer-generated random number draw by the principal investigator. Randomization was stratified by school category (a measure of quality designated by the Ghana Education Service) and by whether the school had a home economics class. Study participants and data collection staff could not be masked because the intervention required overt participation. A cluster design was used to encourage communication about the intervention among participants in the same classroom.

1.3 Procedures

Schools were visited to secure agreement of the headmaster or headmistress and to select a specific class within the school. All chosen classes were in their second year of senior secondary school (similar to grade 11 in the USA). Classes were chosen with the objective of maximizing the number of girls with the following process. If a home economics class was offered at the school, it was chosen because most students studying home economics in Ghana are female; if a home economics class was not offered, the investigators worked with the school headmaster or headmistress to choose a class that had a large number of female students. Female students in the chosen class of each school were invited to participate in the study. Girls who refused consent and all boys were asked to step outside for the duration of study visit. Participants in all groups were told they would receive “health messages” on their phones, including such topics as reproductive health or malaria. Participants used their own mobile phones or could use a family member's phone. Participants without phones were eligible to be enrolled in the trial; however, phones were not provided. After enrollment, students in the Interactive intervention group received a brief training on how to respond to the quiz questions.

Interventions and Control

The study was designed to evaluate the effectiveness of two separate interventions. As part of the "Unidirectional" intervention, participants were sent one reproductive health message via SMS once a week. These messages focused on pregnancy prevention, and contained information on topics of reproductive anatomy, pregnancy, sexually transmitted infections (STIs), and contraception including male condoms, female condoms, birth control pills, and emergency contraception (see Appendix Table A.1 for complete message content). Message content was generated after extensive focus groups with young adults prior to the launch of the study, with the goal of understanding the most popular sexual health topics of interest, as well as guidance from the Ghana Health Service Health Promotion Unit, who edited wording and approved appropriateness of the content for this age group.

As part of the "Interactive" intervention, participants were not sent any information initially, but were instead sent one multiple-choice quiz question via SMS each week to which they were invited to respond free of charge. After responding, participants immediately received a confirmatory SMS informing them whether they answered correctly along with the correct answer and additional information, which corresponded to the information provided in the Unidirectional intervention. Participants who never responded were sent an SMS with the correct answer at the end of the week. Every two correct answers resulted in an airtime credit reward of 1GHS (0·38USD). The control group participants were sent placebo messages once a week with information about malaria. All programs ran for 12 weeks.

As part of the intervention, the Unidirectional and Interactive groups also received 4 extra tips about the effectiveness of condoms, the benefits of talking with their boyfriend about reproductive health, and the existence of a free public hotline number that they could call for reproductive health information (sent twice). This was done as a means of increasing access and communication of reproductive health information. After the 3-month follow-up, students in both intervention and control arms were offered a 30–45-minute lecture about reproductive health by a nurse recruited by the Alliance for Reproductive Health Rights, a Ghanaian NGO.

All messages were in English, the language of secondary school instruction in Ghana, and automatically sent to participants through a computerized SMS-messaging system. If a message was not delivered, it was resent. Interactive participants were sent up to two reminder messages encouraging them to respond if they had not yet responded. Airtime credit rewards were also sent at the end of the week, along with a message informing participants of how many questions they had correctly answered and encouraging them to continue participating. Study staff maintained a record of all incoming and outgoing SMS with participants.

To assess the interventions, students were administered a written baseline questionnaire, a follow-up questionnaire 3 months later, and a second follow-up questionnaire 15 months after baseline. Study staff proctored the questionnaires under test-taking conditions, and scores were calculated by a computer. Demographic information was recorded at baseline. Knowledge and secondary outcomes were recorded at baseline and both the 3-month and 15-month follow-up. Information on sexual activity and pregnancy was collected only at the 15-month follow-up. Questionnaires at baseline and the 3-month follow-up were self-administered on paper; at 15-months, the questionnaire was self-administered on tablet computers to maximize privacy for individual responses about sexual behavior [28].

Outcomes

The primary outcome was reproductive health knowledge. Participants completed a quiz with 24 true/false questions at both the 3-month and 15-month follow-ups (see Appendix Table A.2 for details).

At 15 months, we additionally evaluated the impact of the interventions on pregnancy, sexual activity, and contraceptive use (see Appendix Table A.3 for definitions of all outcome variables). As secondary outcomes, we evaluated the program impact on communication and attitudes about reproductive health. All outcomes are measured at the participant level.

Statistical analysis

The study was powered to detect an improvement of 15 percentage points in the knowledge score with power=0.9 and α=0.05 in pairwise comparisons between control and each of the two intervention arms. This calculation was based on an average of 30 participants in 12 schools in each arm, and an intraclass correlation coefficient of 0.05 (a DEFF of 2.5).

We used ordinary least squares models to estimate intent-to-treat effects on knowledge and multilevel logistic regression models for pregnancy and sexual behavior outcomes. For age at sexual debut, an ordinary least squares model was used. Multivariable regression models were estimated, adjusting for baseline individual- and school-level characteristics, including age, ethnicity, religion, mother's education, father's education, school size, and baseline knowledge.

We measured communication at the 3-month follow-up via 4 questions of the form "In the last 3 months, how often have you spoken to [X] about sex or reproductive health issues?" where [X] is replaced with "your close friends", "your parents", "a teacher, nurse, or any professional", and "your boyfriend". At the 15-month follow-up, this question was asked only about "friends" and "anyone". The responses provided by the participant were on a 5-point scale (i.e. "every day or almost every day", "at least once a week", "at least once a month", "less than once a month", "never"). We generated an indicator variable for speaking at least once a week about reproductive health with each contact and conducted multilevel logistic regression models as described above.

We measured attitudes via 18 questions on a 5-point Likert scale (from "Strongly Agree" to "Strongly Disagree") at the 3-month follow-up and via 7 questions on a 3-point Likert scale (from

"Agree" to "Disagree") at the 15-month follow-up (see Appendix Table A.4 for more details). To assess the effect on attitudes, we generated an indicator variable for agreeing or strongly agreeing with each statement and conducted multilevel logistic regression models as described above.

For linear regression models, standard errors were clustered at the school level to correct for within-school correlation of outcomes. For logistic regression models, we included school random effects. We used R (version 3.1.1) for all analyses. The study design was registered on ClinicalTrials.gov (NCT02031575).

The original trial registration described the protocol and primary and secondary outcomes for the trial only through the first follow-up, which was the period defined in the original project. After this follow-up was completed, we developed plans and secured additional funding to conduct a further follow-up round at 15-months to investigate whether knowledge gains were persistent and to measure the potential impact on reproductive health outcomes. The protocol was amended accordingly to accommodate the additional follow-up. None of the amendments were driven by specific data from the initial study.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

1.4 Results

Participants were recruited between January 15th and February 28th, 2014. A total of 38 schools were eligible for randomization (Figure 1.1). After randomization, three schools were found to be ineligible (they were boarding schools) and one refused on the basis of time constraint. The final sample included 34 schools with 12 schools assigned to the Unidirectional intervention, 10 schools assigned to the Interactive intervention, and 12 schools assigned to control. A total of 756 participants were enrolled in the study, of which 716 (94.7%) were successfully followed up at the 3-month follow-up and 721 (95.4%) were successfully followed up at the 15-month follow-up. Of those participants followed up at 3 months, 99% had provided a phone number at baseline and 83% claimed to have received at least one message. In the Interactive group, weekly response rates to the quiz questions remained relatively stable, ranging from 68 to 75% over the 12-week intervention duration. Table 1.1 shows baseline characteristics (age, ethnicity, religion, mother's education, father's education, and knowledge), which were evenly distributed between the groups. The observed ICCs for knowledge at 3 months, knowledge at 15 months, and pregnancy were 0.34, 0.16, and 0.00, respectively.

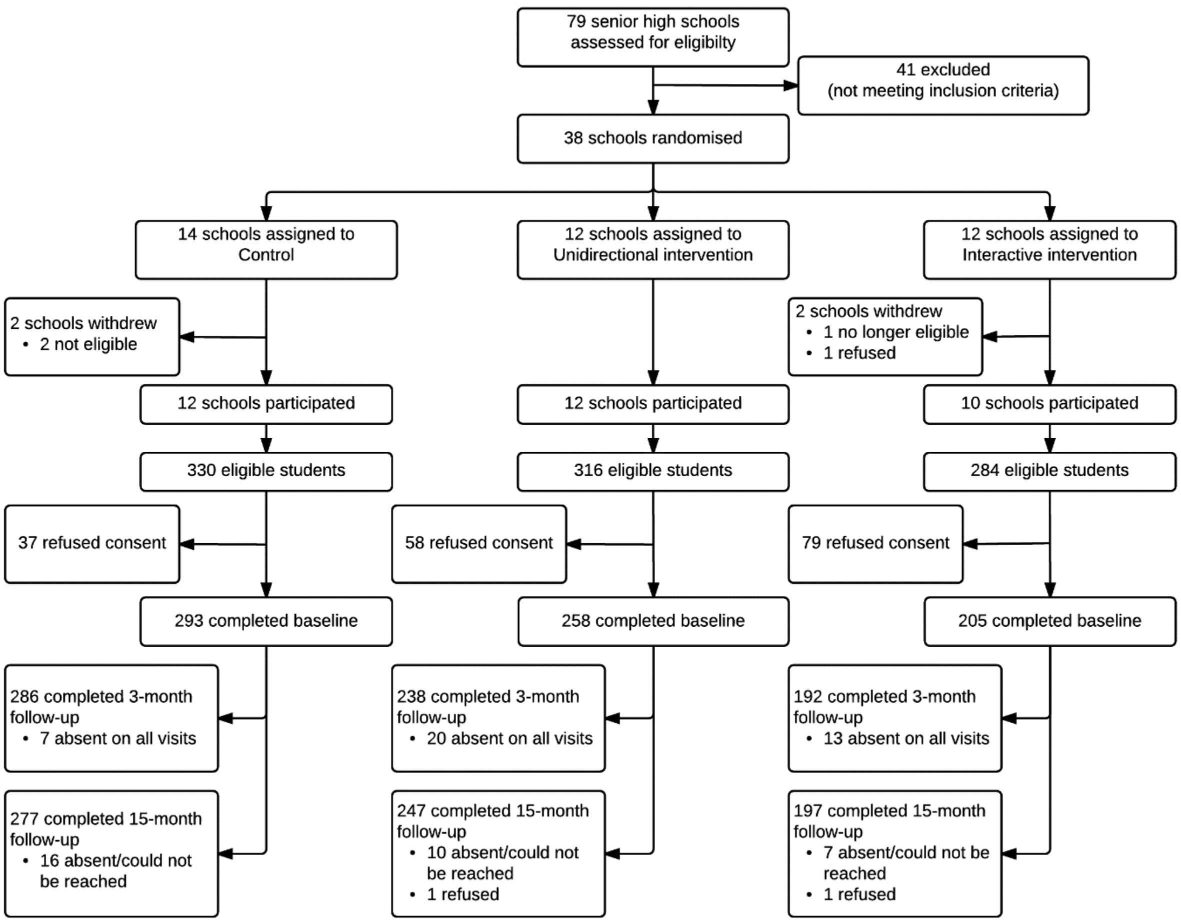


Figure 1.1: Trial Profile

Table 1.1: Baseline characteristics of the intention-to-treat population

	Control	**Unidirectional**	**Interactive**	**p-value**
Number of clusters	12	12	10	
Number of total participants	293	258	205	
Median participants per cluster	22.5 (6–47)	20.5 (2–42)	19.5 (1–39)	
Participated at 3 mo. follow-up	286 (98%)	238 (92%)	192 (94%)	
Participated at 15 mo. follow-up	277 (95%)	247 (96%)	197 (96%)	
Age (years)	17.8 (1.2)	17.6 (1.4)	17.6 (1.5)	0.45
Religion*:				
Muslim	52 (17.7%)	37 (14%)	24 (12%)	0.41
Catholic	21 (7%)	21 (8%)	18 (9%)	0.89
Spiritual/Pentecostal/Charismatic	128 (44%)	120 (47%)	93 (45%)	0.91
Protestant	61 (21%)	61 (24%)	54 (26%)	0.64
Other	26 (9%)	14 (5%)	12 (6%)	0.39
Mother's Education†:				
Don't know	72 (25%)	56 (22%)	47 (23%)	0.89
Less than Secondary	47 (16%)	46 (18%)	22 (11%)	0.15
At least Secondary	170 (58%)	154 (60%)	135 (66%)	0.48
Father's Education‡:				
Don't know	65 (22%)	42 (16%)	41 (20%)	0.52
Less than Secondary	119 (41%)	109 (42%)	77 (38%)	0.67
At least Secondary	105 (36%)	106 (41%)	86 (42%)	0.45
Ethnicity¶:				
Akan	112 (38%)	113 (44%)	70 (34%)	0.19
Ga	86 (29%)	61 (24%)	68 (33%)	0.32
Ewe	42 (14%)	49 (19%)	39 (19%)	0.26
Other	41 (14%)	23 (9%)	25 (12%)	0.39
Own phone**:				
Yes	247 (84%)	219 (85%)	177 (87%)	0.69
No, but have access	38 (13%)	29 (11%)	24 (12%)	0.75
No, no access	2 (0.7%)	5 (2%)	3 (2%)	0.43
Baseline knowledge score	0.26 (0.16)	0.3 (0.17)	0.31 (0.18)	0.10

***Notes*:** Data are number (%), mean (SD), median (range). *Data missing for 5 Control, 5 Unidirectional, and 4 Interactive participants. †Data missing for 4 Control, 2 Unidirectional, and 1 Interactive. ‡Data missing for 4 Control, 1 Unidirectional, and 1 Interactive. ¶Data missing for 12 Control, 11 Unidirectional, and 3 Interactive. **Data missing (although phone number was provided by all) for 7 Control, 9 Unidirectional, and 4 Interactive. p-value from F-statistic for differences across intervention groups.

Figure 1.2 shows the crude and adjusted means of the knowledge score for the Interactive, Unidirectional, and Control groups at 0 (baseline), 3, and 15 months (estimates are reported in Appendix Table A.5). From baseline to the 3-month follow-up, unadjusted average knowledge scores increased from 26% to 32% in the control group, 31% to 45% in the Unidirectional, and 31% to 60% in the Interactive groups. After adjusting for covariates, the increase in knowledge in the Unidirectional group was 11 percentage points (95% CI (7 to 15), p-value < 0.0001) greater than in the Control group, and the increase in the Interactive was 24 percentage points (95% CI (19 to 28), p-value < 0.0001) greater than in the Control. At 15 months, these gains were largely maintained for the Unidirectional (47%) and Interactive groups (56%), though the Control (42%) caught up in knowledge to the Unidirectional group. At 15-months, the increase in knowledge in the Interactive group was 11 percentage points (95% CI (8 to 15), p-value<0.0001) greater than in the Control group, while the Unidirectional intervention was no longer significantly different from the Control once adjusting for baseline characteristics (adjusted endline difference 3%, 95% CI (-1 to 7), p-value=0.17).

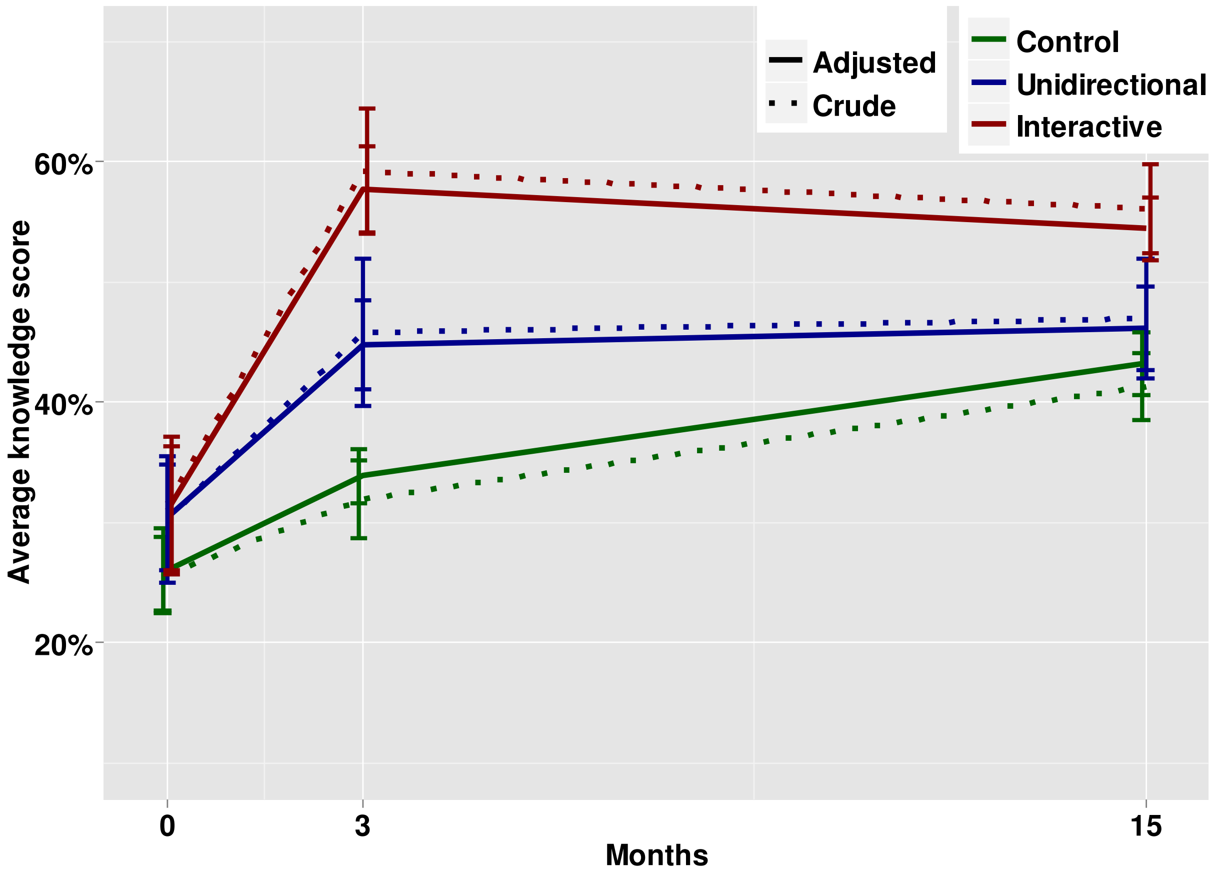


Figure 1.2 Crude and adjusted mean and 95% confidence intervals of knowledge score at 0 (Baseline), 3, and 15 months for Interactive, Unidirectional, and Control groups

Notes: Estimates come from a regression of knowledge score on intervention group. Crude model adjusted for blocking variables, that is, presence of home economics class and school category. Adjusted model additionally adjusted for age, religion, ethnicity, mother's education, father's education, school size, and baseline knowledge.

Table 1.2 shows the results for pregnancy and sexual behavior from both unadjusted and adjusted models. Though the direction of the effects found in both models stays the same, the point estimates vary and standard errors in the adjusted models are generally narrower as a result of the additional control variables. There was no significant impact of either intervention on *ever having sex*, on *having sex in the past year,* or on *pregnancy in the past year* for the full sample of participants.

Conditional on having sex in the past year, both the Unidirectional and the Interactive programs significantly lowered the odds of pregnancy by 86% in the adjusted models (OR 0.14, 95% CI 0.03–0.71, p=0.02) and 85% (OR 0.15, 95% CI 0.03–0.86, p=0.03), respectively, compared to the Control (Table

1.3, Panel A). The Interactive intervention increased odds of using the birth control pill in the past year (OR 13, 95% CI 1.1–160, p=0.04) although small sample sizes resulted in large confidence intervals. The Interactive intervention decreased odds of using emergency contraception (OR .22, 95%CI .06–.88, p=.03). The Interactive intervention appeared to increase risk of sex without a condom in the past year (OR 3.5, 95% CI 1.1-11, p=0.03). There was no impact on age of sexual debut for those who have ever had sex (Table 1.3, Panel B).

Table 1.2: Estimated intervention effects for pregnancy and sexual behavior for full sample

				Unidirectional – Control		**Interactive – Control**	
	Control n (%)	**Unidirectional n (%)**	**Interactive n (%)**	**Crude OR (95% CI)**	**Adj. OR (95% CI)**	**Crude OR (95% CI)**	**Adj. OR (95% CI)**
Ever had sex	88/273 (32%)	83/239 (35%)	64/196 (33%)	1.0 (0.71–1.5)	1.1 (0.71–1.6)	1.3 (0.85–2.0)	1.20 (0.8–1.9)
Sex in past year	58/273 (21%)	64/243 (26%)	51/196 (26%)	1.2 (0.8–1.8)	1.2 (0.8 –1.9)	1.5 (0.97–2.4)	1.6 (0.96–2.5)
Pregnant in past year	10/276 (4%)	5/243 (2%)	6/193 (3%)	.51 (.17-1.5)	.39 (.12-1.3)	.85 (.27-2.7)	.59 (.17-2.0)

Notes: Odds ratios from multilevel logistic regression model with school random effects. Crude model adjusted for blocking variables, that is, presence of home economics class and school category. Adjusted model additionally adjusted for age, religion, ethnicity, mother's education, father's education, school size, and baseline knowledge.*p<.05, **p<.01, ***p<.0001

Table 1.3: Estimated intervention effects for pregnancy and sexual behavior among adolescents sexually active in past year

				Unidirectional – Control		**Interactive – Control**	
Panel A: Binary Outcomes	**Control n (%)**	**Unidirectional n (%)**	**Interactive n (%)**	**Crude OR (95% CI)**	**Adj. OR (95% CI)**	**Crude OR (95% CI)**	**Adj. OR (95% CI)**
Pregnant in past year	9/58 (16%)	5/63 (8%)	4/51 (8%)	.40 (.12–1.4)	.14* (0.03–0.71)	.42 (.1–1.7)	.15* (0.03–0.86)
Used any contraception past year	26/56 (46%)	35/60 (58%)	25/46 (54%)	1.8 (.83–3.8)	1.5 (0.68–3.4)	1.3 (.56–2.9)	1.2 (0.48–2.9)
Used contraception last time had sex	27/54 (50%)	36/59 (61%)	27/50 (54%)	1.6 (.75–3.5)	1.4 (0.61–3.2)	1.3 (.57–2.9)	1.2 (0.48–2.8)
Used condom at sexual debut	30/54 (56%)	34/62 (55%)	27/49 (55%)	0.99 (0.46–2.1)	.83 (.36–1.9)	1.1 (0.5–2.6)	.97 (.39–2.4)
Had sex without condom past year	38/57 (67%)	48/62 (77%)	42/49 (86%)	1.5 (.65–3.5)	1.9 (0.73–4.7)	2.8* (1.0–7.7)	3.5* (1.1–11.0)
Used condom in past year	15/58 (26%)	17/64 (27%)	16/51 (31%)	1.2 (.51–2.7)	1.1 (0.47–2.8)	1.3 (.53–3.1)	1.2 (0.48–3.2)
Used birth control pill in past year	1/58 (2%)	5/64 (8%)	5/51 (10%)	4.9 (.55–43)	5.0 (0.5–50)	6.9 (.73–65)	13.0* (1.1–160)
Used emergency contraception past year	10/58 (17%)	11/64 (17%)	4/51 (8%)	1.0 (.37–2.7)	.81 (0.28–3.0)	0.38 (.1–1.4)	.22* (0.06–.88)
Panel B: Linear Outcomes	**Control mean**	**Unidirectional mean**	**Interactive mean**	**Crude Diff (95% CI)**	**Adj. Diff (95% CI)**	**Crude Diff (95% CI)**	**Adj. Diff (95% CI)**
Age at sexual debut	17.7 n=60	17.4 n=66	17.9 n=40	-0.25 (-0.88 to 0.38)	-0.25 (-0.88 to 0.37)	0.10 (-0.38 to 0.59)	0.17 (-0.40 to 0.74)

Notes: Panel A: Odds ratios from multilevel logistic regression model with school random effects, conditional on having sex in the past year. Crude model adjusted for blocking variables, that is, presence of home economics class and school category. Adjusted model additionally adjusted for age, religion, ethnicity, mother's education, father's education, school size, and baseline knowledge.
Panel B: Linear model with clustered standard errors at school level, conditional on ever having sex. Crude model adjusted as above. Adjusted model additionally adjusted for religion, ethnicity, mother's education, father's education, school size, and baseline knowledge.
1 participant in the Control and 2 in the Interactive reported being pregnant in the past year but not sex in the past year. Since this is physically possible, we did not recode them. Analysis including those participants in this sub-population does not change the direction or the significance of the results.
*p<.05, **p<.01, ***p<.0001

Figure 1.3 shows the odds ratios and 95% confidence interval for agreeing with each attitude statement asked at both the 3- and 15-month follow-ups for the Interactive and Unidirectional interventions as compared to the Control. Girls in the intervention arms appeared to have more confidence with use of birth control and condoms, although there was little sustained impact on attitudes of susceptibility to STIs, confidence to refuse sex, or comfort discussing reproductive health with friends. Full results for all attitudes items measured at both the 3- and 15-month follow-ups are shown in Appendix Tables A.6 and A.7.

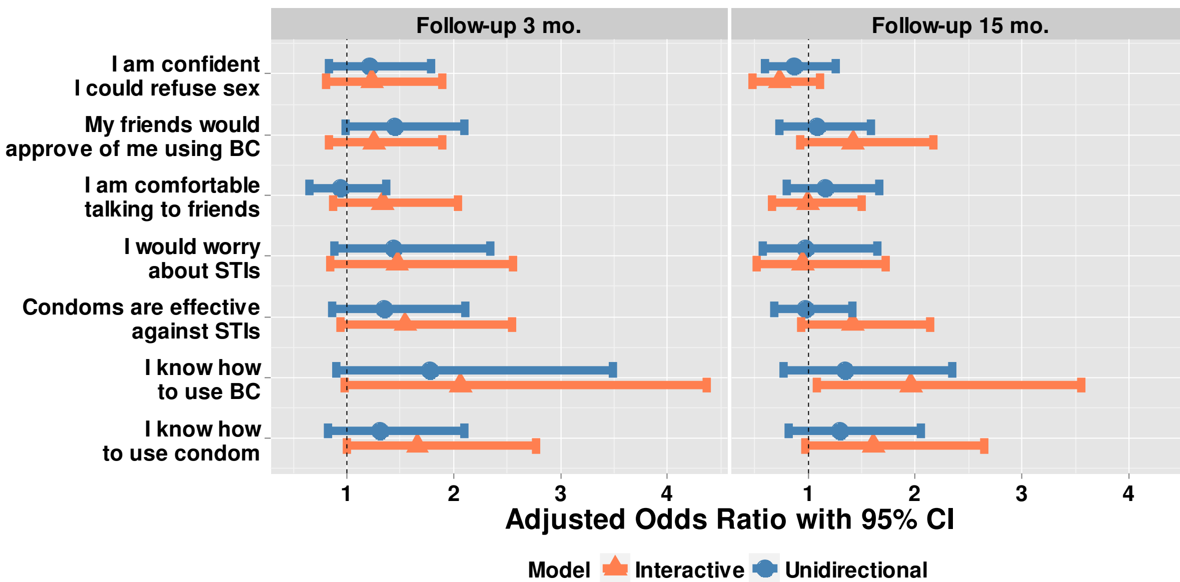


Figure 1.3: Adjusted odds ratio and 95% Confidence Interval of agreeing with attitude statement at 3-month and 15-month follow-ups

Notes: Odds ratios from multilevel logistic regression model with school random effects. Model adjusted for age, religion, ethnicity, mother's education, father's education, baseline attitude, school size, presence of home economics class, and school category. BC=birth control pill

Participants in the Interactive program increased communication about reproductive health among friends at 3 months. The results are shown in Appendix Table A.8.

1.5 Discussion

The results presented in this study suggest that mobile phone programs are effective tools to improve reproductive health knowledge among adolescents. Large improvements in knowledge were observed at 3 months and sustained after 15 months. For the sexual behavior outcomes, results were mixed. Among sexually active teenagers, we found both programs to be protective against pregnancies; however we found no significant impact on pregnancy in the full sample. We find no consistent evidence of a program impact on use of contraception; however, small sample sizes make it difficult to determine if there was no behavioral effect or if the study was underpowered. Larger impacts on reproductive health outcomes may be plausible once a majority of treated women become sexually active.

Somewhat surprisingly, we found that the Interactive interventions increased the likelihood of having sex without a condom among sexually active teenagers in the Interactive group. The main focus of the intervention content was on pregnancy prevention rather than STIs, which appears to have resulted in a move away from condoms and towards birth control pills. Birth control pills have the advantage of being completely in the control of women, rather than needing to be negotiated with men, and have a lower rate of failure than condoms in typical use. Other studies have found that fear of pregnancy, not of STIs, motivates Ghanaian adolescents to use contraceptives [22]. However, in other settings where HIV and other STI rates are high, these messages may not be appropriate. This study highlights the importance of carefully adjusting content and framing of mobile phone programs to local public health needs.

This study has several limitations. First, for reproductive health outcomes, the study exclusively relied on self-reported measures. It is possible that respondents in intervention arms may have felt more pressure to misreport their sexual behavior. Since they received messages that encouraged use of contraception to prevent unintended pregnancy, they may have consequently underreported pregnancy. This would lead to an overestimation of the effect of the intervention. This is not obvious, however, as the exposure to the programs may have increased familiarity and openness to sexual health questions, so that program participants may have been more likely to report undesired outcomes than the Control (such as sex without a condom), which would imply that we would underestimate the true impact of the program. To mitigate misreporting concerns, all questions at the 15-month follow-ups were asked using self-administered tablet computers, which have been shown to increase honesty in adolescent responses of sexual behavior [28]. A second limitation of the study is that we included only adolescent girls in secondary school in Accra; program impact may be different among high-risk girls, boys, and adolescents in rural areas. Recent evidence from 83 sexual education programs across the world evaluating the impact of sex education on knowledge, attitudes, and behaviors found that programs that had positive effects were equally effective in both rural and urban areas, among girls and boys, and among low- and middle-income youth, and that replication of effective studies in other settings yielded consistent results [12]. Third, neither the participants nor the study staff could be masked to assignment. However, staff were trained to provide the same description of the messages to all groups to prevent differential uptake. Similarity of baseline characteristics across groups indicates that the participants were comparable.

An important consideration is that of selective attrition. Although our follow-up rates at 15 months were very high across all intervention arms (95% in the Control, 95% in Unidirectional, and 96% in Interactive), one may be concerned about differential attrition driven by pregnancy. To ensure this was not the case, we asked the classmates and school administrations the whereabouts and pregnancy status of the 35 participants not followed up. 3 participants were not reachable for interview but personally confirmed they were not pregnant over the phone. 32 participants (16 in Control, 9 in Unidirectional, and 7 in Interactive) could not be reached by study staff. In the vast majority of cases, the classmates knew what had happened to the participant who was not found and confirmed the participant had not been pregnant. One participant lost to follow-up in the control arm was reported as pregnant by classmates. A total of 7 participants (2 in the control, 4 in the Unidirectional, and 1 in the Interactive) had an unknown status.

To our knowledge, this is the first randomized trial evaluating the effectiveness of mobile health communication to improve adolescent reproductive health in a developing country. For reproductive health knowledge, our results are contrary to a recent study in Uganda that found no impact of an automated self-directed SMS health information intervention on sexual health knowledge among adults [16]. One possible explanation of the divergent result is that the self-directed nature of the Ugandan program led to limited and highly heterogeneous use (most participants used the service fewer than 5 times) and a lack of clear learning objective. Our results show that a guided intervention led to sustained participation and large knowledge gains. It is also likely that adolescent users are more proficient using SMS technology than adults.

Generalizability of this program to rural areas may currently be limited. However, mobile technology availability continues to grow at a rapid pace – the number of cellphone connections has grown by 30% per year since 2001 in sub-Saharan Africa [29]. In Ghana, mobile phone penetration surpassed 116% in 2015 [30]. Although rural areas currently lag behind urban areas in mobile phone access, a priority of the government of Ghana is to narrow this gap in the next few years [31]. In addition, rapid urbanization in sub-Saharan Africa has led to more than 50% of people living in urban areas. These factors indicate that the use of mobile programs to improve health in both urban and rural areas will be an important frontier.

Our results complement findings of other programs that aim to reduce teen pregnancy, which show that providing teenagers with sex education does not change frequency of sex but can result in safer sex. Dupas (2011) found that providing information to teenage girls on the relative risk of HIV infection by type of partner led to a 61% decrease in the incidence of pregnancies with older partners [32]. A web-based sex education course in Colombia led to reductions in self-reported STIs [33]. A teacher-led in-school education program combined with youth-friendly health services, community activities, and condom distribution found an impact on knowledge and reduction in multiple partners, although no significant impact on reported pregnancies or other reported sexual behaviors [34].

Public Health Implications

School-based comprehensive sexual education in a study context has been found to be largely effective at increasing knowledge; behavioral impacts have been observed for some programs, though less consistently [35–39]. However, poor implementation of programs at-scale, including problems of curricula lacking basic information on condoms and contraception, poor teaching, and short program durations result in a lack of fidelity to the designed intervention and may therefore become largely ineffective [24]. Our study supports the idea that text messaging programs or other digital media-based interventions such as web programs may be effective ways to fill the gap, by providing high quality, accurate information over a long duration via a medium that adolescent are comfortable with. Moreover, they can be tailored to the audience both in terms of cultural and individual characteristics [40], and they can inexpensively reach a large and diverse population.

Young people are the most likely age group to use their phone to send text messages in LMICs [41], yet very few digital media interventions have been developed for and evaluated on adolescents in LMICs [19,20,40]. To our knowledge, this is the first randomized trial evaluating the effectiveness of mobile health communication to improve adolescent reproductive health in a low-resource country. More research on text messaging programs examining the full impact of such programs on objective measures of reproductive health and over the long-term is necessary.

1.6 References

1. Woog V, Singh S, Browne A, Philbin J. Adolescent Women's Need for and Use of Sexual and Reproductive Health Services in Developing Countries [Internet]. New York: Guttmacher Institute; 2015. Available from: www.guttmacher.org/pubs/AdolescentSRHS-Need-Developing-Countries.pdf

2. Sawyer SM, Afifi RA, Bearinger LH, Blakemore S-J, Dick B, Ezeh AC, et al. Adolescence: a foundation for future health. The Lancet. 2012;379(9826):1630–40.

3. Nove A, Matthews Z, Neal S, Camacho AV. Maternal mortality in adolescents compared with women of other ages: evidence from 144 countries. Lancet Glob Health. 2014 Mar;2(3):e155–64.

4. Cleland J, Conde-Agudelo A, Peterson H, Ross J, Tsui A. Contraception and health. The Lancet. 2012;380(9837):149–56.

5. Finlay JE, Özaltin E, Canning D. The association of maternal age with infant mortality, child anthropometric failure, diarrhoea and anaemia for first births: evidence from 55 low- and middle-income countries. BMJ Open. 2011 Jan 1;1(2):e000226.

6. Chen X-K, Wen SW, Fleming N, Demissie K, Rhoads GG, Walker M. Teenage pregnancy and adverse birth outcomes: a large population based retrospective cohort study. Int J Epidemiol. 2007 Apr 1;36(2):368–73.

7. Mayor S. Pregnancy and childbirth are leading causes of death in teenage girls in developing countries. BMJ. 2004 May 13;328(7449):1152.

8. Fink G, Sudfeld CR, Danaei G, Ezzati M, Fawzi WW. Scaling-Up Access to Family Planning May Improve Linear Growth and Child Development in Low and Middle Income Countries. PLoS ONE. 2014 Jul 14;9(7):e102391.

9. Hindin MJ, Fatusi AO. Adolescent sexual and reproductive health in developing countries: an overview of trends and interventions. Int Perspect Sex Reprod Health. 2009;58–62.

10. Singh S, Bankole A, Woog V. Evaluating the need for sex education in developing countries: sexual behaviour, knowledge of preventing sexually transmitted infections/HIV and unplanned pregnancy. Sex Educ. 2005 Nov 1;5(4):307–31.

11. Biddlecom A, Hessburg, Laura, Singh S, Bankole A, Darabi, Leila. Protecting the Next Generation In Sub-Saharan Africa: Learning from Adolescents to Prevent HIV and Unintended Pregnancy. New York: Guttmacher Institute; 2007.

12. Bearinger LH, Sieving RE, Ferguson J, Sharma V. Global perspectives on the sexual and reproductive health of adolescents: patterns, prevention, and potential. The Lancet. 2007 Apr 13;369(9568):1220–31.

13. Bob Tortora, Magali Rheault. Mobile Phone Access Varies Widely in Sub-Saharan Africa. Gallup Research Poll [Internet]. 2011 Sep 16; Available from: http://www.gallup.com/poll/149519/mobile-phone-access-varies-widely-sub-saharan-africa.aspx

14. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. The Lancet. 2010;376(9755):1838–45.

15. Raifman JRG, Lanthorn HE, Rokicki S, Fink G. The Impact of Text Message Reminders on Adherence to Antimalarial Treatment in Northern Ghana: A Randomized Trial. PLoS ONE. 2014 Oct 28;9(10):e109032.

16. Jamison JC, Karlan D, Raffler P. Mixed Method Evaluation of a Passive mHealth Sexual Information Texting Service in Uganda [Internet]. National Bureau of Economic Research; 2013 [cited 2014 Jul 1]. Available from: http://www.nber.org/papers/w19107

17. Ramachandran A, Snehalatha C, Ram J, Selvam S, Simon M, Nanditha A, et al. Effectiveness of mobile phone messaging in prevention of type 2 diabetes by lifestyle modification in men in India: a prospective, parallel-group, randomised controlled trial. Lancet Diabetes Endocrinol. 2013 Nov;1(3):191–8.

18. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, et al. The Effectiveness of Mobile-Health Technologies to Improve Health Care Service Delivery Processes: A Systematic Review and Meta-Analysis. PLoS Med. 2013 Jan 15;10(1):e1001363.

19. Gurman TA, Rubin SE, Roess AA. Effectiveness of mHealth Behavior Change Communication Interventions in Developing Countries: A Systematic Review of the Literature. J Health Commun. 2012 May 2;17(sup1):82–104.

20. Déglise C, Suggs LS, Odermatt P. Short Message Service (SMS) Applications for Disease Prevention in Developing Countries. J Med Internet Res. 2012 Jan 12;14(1):e3.

21. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? PLoS Med. 2013;10(2):e1001382.

22. Awusabo-Asare K, Biddlecom A, Kumi-Kyereme A, Patterson K. Adolescent Sexual and Reproductive Health in Ghana: Results from the 2004 National Survey of Adolescents. New York: Guttmacher Institute; 2006 Jun. Report No.: No. 22.

23. Hindin MJ, McGough LJ, Adanu RM. Misperceptions, misinformation and myths about modern contraceptive use in Ghana. J Fam Plann Reprod Health Care. 2014 Jan 1;40(1):30–5.

24. Chandra-Mouli V, Lane C, Wong S. What Does Not Work in Adolescent Sexual and Reproductive Health: A Review of Evidence on Interventions Commonly Accepted as Best Practices. Glob Health Sci Pract. 2015 Sep 10;3(3):333–40.

25. Ghana Statistical Service (GSS), Ghana Health Service (GHS), and ICF International. Ghana Demographic and Health Survey 2014 [Internet]. Rockville, Maryland, USA: GSS, GHS, and ICF International; 2015. Available from: http://dhsprogram.com/pubs/pdf/FR307/FR307.pdf

26. Hessburg L et al. Protecting the Next Generation in Ghana: New Evidence on Adolescent Sexual and Reproductive Health Needs. New York: Guttmacher Institute; 2007.

27. Agyei WK, Biritwum RB, Ashitey AG, Hill RB. Sexual behaviour and contraception among unmarried adolescents and young adults in Greater Accra and Eastern regions of Ghana. J Biosoc Sci. 2000;32(04):495–512.

28. Kelly CA, Soler-Hampejsek E, Mensch BS, Hewett PC. Social Desirability Bias in Sexual Behavior Reporting: Evidence from an Interview Mode Experiment in Rural Malawi [. Int Perspect Sex Reprod Health. 2013 Mar;39(01):014–21.

29. UN-Habitat. State of the World's Cities 2012/2013. Prosperity of Cities. New York, NY: UN-Habitat; 2013.

30. Ghana's mobile penetration now at 116%. GhanaWeb [Internet]. 2015 Jun 23; Available from: http://www.ghanaweb.com/GhanaHomePage/business/Ghana-s-mobile-penetration-now-at-116-363989

31. Ametorgoh M. Connecting rural Ghana to the world of opportunities. Graphic Online [Internet]. 2015 Sep 8; Available from: http://graphic.com.gh/news/general-news/49086-connecting-rural-ghana-to-the-world-of-opportunities.html

32. Dupas P. Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya. Am Econ J Appl Econ. 2011 Jan 1;3(1):1–34.

33. Chong A, Gonzalez-Navarro M, Karlan D, Valdivia M. Effectiveness and Spillovers of Online Sex Education: Evidence from a Randomized Evaluation in Colombian Public Schools [Internet]. National Bureau of Economic Research; 2013 Feb [cited 2015 Feb 3]. Report No.: 18776. Available from: http://www.nber.org/papers/w18776

34. Doyle AM, Ross DA, Maganja K, Baisley K, Masesa C, Andreasen A, et al. Long-Term Biological and Behavioural Impact of an Adolescent Sexual Health Intervention in Tanzania: Follow-up Survey of the Community-Based MEMA kwa Vijana Trial. PLoS Med. 2010 Jun 8;7(6):e1000287.

35. Fonner VA, Armstrong KS, Kennedy CE, O'Reilly KR, Sweat MD. School Based Sex Education and HIV Prevention in Low- and Middle-Income Countries: A Systematic Review and Meta-Analysis. PLoS ONE. 2014;9: e89692. doi:10.1371/journal.pone.0089692

36. Kirby DB, Laris BA, Rolleri LA. Sex and HIV Education Programs: Their Impact on Sexual Behaviors of Young People Throughout the World. J Adolesc Health. 2007;40: 206–217. doi:10.1016/j.jadohealth.2006.11.143

37. Napierala Mavedzenge SM, Doyle AM, Ross DA. HIV Prevention in Young People in Sub-Saharan Africa: A Systematic Review. J Adolesc Health. 2011;49: 568–586. doi:10.1016/j.jadohealth.2011.02.007

38. Speizer IS, Magnani RJ, Colvin CE. The effectiveness of adolescent reproductive health interventions in developing countries: a review of the evidence. J Adolesc Health. 2003;33: 324–348. doi:10.1016/S1054-139X(02)00535-9

39. Haberland N, Rogow D. Sexuality Education: Emerging Trends in Evidence and Practice. J Adolesc Health. 2015;56: S15–S21. doi:10.1016/j.jadohealth.2014.08.013

40. Guse K, Levine D, Martins S, Lira A, Gaarde J, Westmorland W, et al. Interventions Using New Digital Media to Improve Adolescent Sexual Health: A Systematic Review. J Adolesc Health. 2012;51: 535–543. doi:10.1016/j.jadohealth.2012.03.014

41. Pew Research Center. Cell Phones in Africa: Communication Lifeline. In: Pew Research Center's Global Attitudes Project [Internet]. 15 Apr 2015 [cited 4 Mar 2016]. Available: http://www.pewglobal.org/2015/04/15/cell-phones-in-africa-communication-lifeline/

Chapter 2:

Impact of Migration on Fertility and Abortion: Evidence from the Household and Welfare Study of Accra

Abstract

Over the last few decades, total fertility rates, child morbidity, and child mortality rates have declined in most parts of sub-Saharan Africa. Among the most striking trends observed are the rapid rate of urbanization and the often remarkably large gaps in fertility between rural and urban areas. Although a large literature has highlighted the importance of migration and urbanization within countries' demographic transitions, relatively little is known regarding the impact of migration on migrants' reproductive health outcomes in general and abortion in particular. In this article, we use detailed pregnancy and migration histories collected as part of the Household and Welfare Study of Accra (HAWS) to examine the association between migration and pregnancy outcomes among women residing in the urban slums of Accra, Ghana. We find that the completed fertility patterns of lifetime Accra residents are remarkably similar to those of residents who migrated. Our results suggest that recent migrants have an increased risk of pregnancy but not an increased risk of live birth in the first years post-move compared with those who had never moved. This gap seems to be largely explained by an increased risk of miscarriage or abortion among recent migrants. Increasing access to contraceptives for recent migrants has the potential to reduce the incidence of unwanted pregnancies, lower the prevalence of unsafe abortion, and contribute to improved maternal health outcomes.

Keywords: Migration • Abortion • Fertility • Reproductive health • sub-Saharan Africa

2.1 Introduction and Conceptualization

Over the coming decades, urbanization is expected to continue or accelerate in the developing world, with Africa and Asia urbanizing most rapidly [1]. Internal rural-to-urban migration accounts for more than one-half of the growth of cities in Africa [2]. One of the most significant recent trends in migration has been the entry of women into migration streams that in previous decades had been primarily male, with an increasing number of female migrants moving on their own [3,4]. Many rural-to-urban migrants settle in slums, contributing to a projection of a doubling of slum settlements over the next 30 years [5].

Over the past five decades, total fertility rates have declined across sub-Saharan Africa [6], with particularly rapid declines in urban areas. Although urban fertility rates have consistently been lower historically, the difference between urban and rural fertility rates has increased substantially from 0.3 children in 1960 to 1.9 children in 2000 [7]. Given that migrants from rural areas constitute an increasingly large fraction of the urban population, the increasing rural-urban gaps are rather remarkable. From an individual perspective, migration to urban centers constitutes a fundamental change in environment and lifestyle, which may be associated with increased risky sexual behavior, unintended pregnancies, and mistimed births [8–10].

Most research in the area of migration and fertility has relied on the theoretical framework proposed by Goldstein and Goldstein (1982). The framework is based on three mechanisms: disruption, adaptation, and selection. In the context of reproductive health, each of the three factors may increase or decrease sexual activity and risk of pregnancy [8]. “Disruption” can be interpreted as interruption in what otherwise would have been the anticipated fertility schedule of migrants. For example, separation of spouses or a desire to delay childbearing until after the move may reduce fertility in the short term [11,12]. Women who migrate to cities to marry or to join husbands are less likely to live with their spouses in the first few months, potentially lowering the probability of fertility in those years [13,14]. On the other hand, disruption may also cause an increase in conception, unintended pregnancies, and potentially abortion if spousal separation increases risk of extramarital sexual behavior [8,15]

“Adaptation” refers to the socialization of migrants: that is, the adjustment to the social, cultural, and sexual norms of the destination’s residential environment as well as to the economic constraints and opportunities that they face as a result of the move [12,13,16]. Female migrants who moved before their completed fertility may adjust their desired fertility to match the norms of the destination, thus reducing total fertility rates. Rural-urban migrants may also discover a change in lifestyle constraints in their new location. Newly arrived married couples may reduce or postpone having children until they can adapt to the new economic conditions [16,17]. Generational analysis of migrants in Ghana found evidence for the effects of adaptation in that migrants’ fertility begins to approach levels characteristic of the second generation, and differences between second-generation migrants and urban natives almost disappear [17].

On the other hand, rural-urban migration can also lead to a change of social networks and the removal of traditional village controls over sexual behavior [15]. Migrants may find themselves in an environment that is conducive to high-risk sexual behaviors, which may be especially predominant in slums where migrants tend to move [9]. Contraceptive use may also be low during the first two years after migrants’ arrival [13]. This increase in sexual behavior and reduced use of contraception may result in an increased risk of pregnancy.

Finally, the selection hypothesis captures the notion that mobile individuals differ from nonmigrating populations with respect to predisposed individual characteristics. These characteristics may be observable (such as level of education or employment status) or may be largely unobservable (e.g., ambition and openness to change) [12,18,19]. The decision to move, potentially over a long distance, to a new socioeconomic and cultural environment demonstrates a degree of risk-taking because consequences of the move are often uncertain and social networks at the destination are smaller or non-existent [8]. For example, the earnings of the immigrant population may be higher than the earnings of the native population because individuals with high earning potential are more likely to self-select into migration [18]. Previous research has found substantial support for the selection hypothesis among both rural-urban and urban-rural migrants in Ghana [12]. Another recent study on child mortality of rural-to-urban migrants found that migrants had lower child mortality before they migrated than rural nonmigrants and that their mortality levels dropped further after they arrived in urban areas [20]; these outcomes are evidence of both selection and adaptation.

Few studies have examined the effect of migration on abortion. Research on migration and sexual behavior has generally focused on HIV rates and condom use [8,9]. In Kenya, migrants were found to be more likely than nonmigrant counterparts to engage in sexual practices conducive to HIV infection, such as multiple partners and low condom use [8]. For Nairobi, migration to slums was also found to be associated with an increased likelihood of risky sexual behaviour [9]. In China, rural-to-urban migrant males were found to be significantly less likely to report condom use at first sex and consistent contraceptive use with the first partner compared with nonmigrants and urban-to-urban migrants [10]. However, to the best of our knowledge, no study has investigated whether female migration has an impact on rates of induced abortion.

In this article, we use the detailed data on migration and reproductive health collected as part of the Housing and Welfare Study of Accra (HAWS) to examine the relationship between mobility and reproductive health outcomes in the context of migration to poor residential neighborhoods, which has become the primary force underlying the rapid rates of urbanization observed in developing countries [5,21]. We take advantage of a unique data set that collected both full pregnancy histories and detailed migration histories in order to estimate the effect of migration on both completed fertility and pregnancy outcomes, including miscarriage, stillbirth, and induced abortion.

To disentangle the roles of disruption, adaptation, and selection in fertility and pregnancy outcomes, we start by comparing reproductive health outcomes of long-term residents with those of migrants. We find that completed fertility schedules of migrants are very similar to those of long-term Accra residents, suggesting both adaptation and selection effects. We then conduct an event-history analysis to evaluate the risk of pregnancy, stillbirth, and lost birth (abortion or miscarriage) of new arrivals compared with those who had never moved and those who are longer-term residents. We find that the probability of live birth is unchanged for new arrivals compared with those who had never moved, but the risk of pregnancy and lost birth is increased in the first two years after the move, suggesting that both selection and adaptation mechanisms are relevant in this context. To disentangle selection from adaption effects, we use individual fixed-effects models to compare the risk of adverse pregnancy outcomes among migrants before and immediately after the move. Our data allow us to apply a fixed-effects estimator to account for the unobservable individual-level factors affecting the decision to migrate as well as to have a child. We again find an increase in risk of pregnancy and lost birth in the years immediately following a move. The observed increases suggest a strong influence of the role of adaptation in the sexual behavior of migrant female populations: that is, female migrants appear to increase sexual activity after a move, but reduce their completed fertility via abortion or miscarriage.

The rest of the article is organized as follows. We provide background information on Ghana's fertility history and abortion laws in the following section. We then describe the data and the analytical methods. We present the summary statistics and analytical results, and then we conclude by discussing the policy implications of our findings.

2.2 Background

Reproductive Health in Ghana

In the last 20 years, fertility in Ghana has declined rapidly from a total fertility rate of 6.4 in 1988 to a rate of 4.0 in 2008 [22]. Infant mortality fell from 77 to 50 per 100,000 live births from 1988 to 2007, while contraceptive use among women aged 15–49 increased from 12% to 21% [22]. HIV prevalence is relatively low in Ghana compared with sub-Saharan Africa, estimated to be about 1.5% in 2011 [23]. Women's median age at first marriage is 18.7 in rural areas and 21.3 in urban areas. Women in the Greater Accra urban region marry five years later than women in the more rural region of the Upper East (22.9 years vs. 17.8 years); fertility varies substantially by region, mother's education, and wealth, with wealthier, more-educated urban women having the fewest children [24].

Abortion in some circumstances has been legal in Ghana since 1985. Abortion, by law, must be performed by a registered medical practitioner and is allowed when the physical or mental health of the pregnant woman is threatened, when the child is likely to be born with a serious physical abnormality, or when the pregnancy resulted from rape or incest. In all other situations, it is illegal (Morhee and Morhee 2006). Despite the long-term legality of abortion, unsafe abortion is the second-largest cause of maternal mortality in Ghana [25,26]. In 2010, Ghana's maternal mortality rate was estimated to be 350 maternal deaths per 100,000 live births (95% confidence interval 210–630), which is much higher than the average in the developing world of 210 per 100,000 live births [27]. The 2007 Ghana Maternal Health Survey estimated the ratio to be even higher, at 580 maternal deaths per 100,000 live births [22]. Of these maternal deaths, 11% are the result of unsafe abortion [22,26]. Stigma associated with abortion is high and prevents women from seeking medically safe abortions at a health facility, opting for clandestine abortions instead, which may lead to hemorrhaging, infection, or death. Additionally, a survey of health care facilities in 10 districts found that fewer than one in seven public health facilities reported offering legal abortion services [28]. Nearly one-half of Ghanaian women who recently obtained an abortion underwent the procedure unsafely [26]. Negative encounters with health care providers discourage women from seeking safe abortions or treating post-abortion complications safely with family planning services [25].

Women receive abortions for various reasons, the most common of which is not having the financial means to take care of a child [29]. Other reasons include the presence of relationship problems with the woman's partner, the desire to continue working or schooling, and the desire for spacing or limiting childbearing [29,30]. Women often do not disclose their abortion to their male partners because they fear the partner's reaction, including domestic violence or relationship dissolution [25].

Other studies have linked the legalization of abortion with lower fertility trends [31–33]. These studies have observed that the increase in modern contraceptive use in Ghana has not kept pace with the observed declines in fertility, suggesting that the empirical gap could be explained by increased induced abortion. Finlay and Fox (2013) used multivariate longitudinal regression to show that the timing of the liberalization of the abortion laws coincided with the onset of Ghana's fertility decline. Abortion as a method of birth control has thus been explored as a possible means for women to reduce their completed fertility in Ghana.

Migration in Ghana

Migration is very common in Ghana, with at least one migrant in more than 43 % of all households in 2005–2006 [34]. More than 80% of Ghanaian migrants stay in Ghana; and among them, 70% go to urban areas [34]. About 50.9% of the total population lives in an urban area [35]. The Greater Accra and Ashanti regions attract more than one-half of all internal migrants, and migrants make up a substantial share of the population in these regions [34]. Migration does not have to be permanent and can be two-directional; among households with migrants, 37% have at least one returned migrant.[1] However, differences in characteristics between migrants who return and those who do not have not been found to be significant or meaningfully large with respect to age, gender, and education [34].

The urbanization rate in Ghana is comparable with other sub-Saharan African countries. The average annual rate of change in the urban population of sub-Saharan Africa was 3.82% between 1970 and 2011 [1]. Accra's growth rate between 2005 and 2010 was 3.30%, comparable with other sub-Saharan African cities—such as Nairobi (4.50%), Lagos (3.76%), and Bamako (4.32%)—during the same period [36].

Data

The data used in this article come from the Housing and Welfare Study of Accra (HAWS), which was conducted between 2009 and 2010 in a collaborative effort between the Institute of Statistical, Social and Economic Research (ISSER) at the University of Ghana and the Harvard School of Public Health. The purpose of the HAWS survey was to assess the current health status and living standards of the population

[1] Returned migrants are defined as individuals who were away from the household for some time in the last five years but have since returned to the household (Ackah and Medvedev 2012).

in 37 enumeration areas classified as slums. The “slum” attribute was defined by the GSS, and was given to enumeration areas ranked in the bottom quartile on an index based on the housing and socioeconomic characteristics collected in the 2000 census [37]. The GSS index includes household-level dwelling characteristics, including lighting, water supply, toilet facilities, cooking fuel, cooking space, bathing facilities, and highest level of schooling and educational grade by any member of the household [37].

The HAWS survey consists of a household interview and individual interviews with all women aged 18 and older in the household. The individual woman’s questionnaire consists of sections on background characteristics, migration, health insurance, general health, mental health, nutrition, malaria, a full pregnancy history, prenatal and postnatal care, immunizations for children born in the last five years, marriage and sexual activity, reproductive health, family planning, and fertility preferences. A total of 2,095 women completed the individual interview, of which 1,488 had had at least one pregnancy.

The HAWS data set is unique in two ways. First, it focuses on urban dwellers in the poorest neighborhoods of Accra, where residential mobility is particularly common and health service provision may be more limited. Second, because the study collected full pregnancy histories in combination with detailed migration histories, we are able to identify reproductive health patterns before and after residential changes. The data set includes the outcome of each pregnancy, as well as the month and year of each pregnancy termination. It also includes the month and year of each residential move for the past four moves of each woman interviewed, the location she moved from, whether she knew anyone when moving, and the reason for the move. Information about residential duration in data sets such as the Demographic and Health Surveys (DHS) includes information only on duration in current residence, which makes it impossible to link birth outcomes to residential duration.

The HAWS survey interviewed only women in the slum areas of Accra, who may be systematically different from other Accra residents. The DHS in 2008 did not stratify based on slum areas; only 5 of 35 enumeration areas from the 2008 DHS overlapped with the HAWS study area [38]. Table 2.1 compares descriptive statistics for both the HAWS and DHS 2008 surveys for residents ages

18–49 in the Greater Accra region and shows *t* tests for the difference in means. Compared with DHS Accra residents, women in the HAWS data set are less educated and less likely to be Akan or Ewe ethnicity. They have a lower average age at first birth, are less likely to be married, and are more likely to have terminated a pregnancy.[2] However, they do not differ significantly in terms of average number of total children ever born or the length of time at their current residence (the only residential duration information available in the DHS).

Table 2.1: Sociodemographic characteristics of female residents of Accra aged 18–49 in the DHS (2008) and HAWS data sets

	DHS 2008[a] (N = 622[b])	HAWS (N = 1797[c])	2-Sample t Test p-value
Age (years)	29.996	28.91	0.003
No Education (%)	0.084	0.237	<.0001
Only Primary School (%)	0.144	0.166	0.113
At Least Middle School (%)	0.772	0.594	<.0001
Akan (%)	0.436	0.303	<.0001
Ewe (%)	0.164	0.112	0.001
Ga (%)	0.234	0.246	0.522
Other Ethnicity (%)	0.166	0.338	<.0001
Age at First Birth (years)	21.208	20.457	<.0001
Ever Married (%)	0.648	0.608	0.054
Ever Terminated Pregnancy (%)	0.248	0.351	<.0001
Number Children Ever Born	1.689	1.607	0.307
Up to 24 Months at Residence (%)	0.156	0.159	0.965
25-48 Months at Residence (%)	0.175	0.193	0.348
49-72 Months at Residence (%)	0.102	0.11	0.62
>72 Months at Residence (%)	0.36	0.334	0.141
Never Moved (%)	0.194	0.204	0.662

Notes: [a] Ghana Demographic and Health Survey 2008. Sample restricted to women in the Accra region.
[b] Summary statistics calculated using individual sample weights.
[c] Sample restricted to women aged 18–49.

[2] Measures of variables across data sets were not obtained in the same way for every variable. For example, the DHS asked, "Have you ever had a pregnancy that was terminated?" For the HAWS data, the participant was asked to list every pregnancy and identify the outcome as live, stillbirth, or lost. Reporting bias can act on these measures differently, which makes these rough rather than exact comparisons. Individual sample weights were used to calculate summary statistics of DHS variables.

About 75% of women in the complete HAWS data set moved at least once.[3] The average number of moves was 1.59. The most common age to move was in the late teens and early 20s, with the average age of any move at 22.8 years. About 55% of the sample moved either one or two times over a lifetime. We show the distribution of moves in the sample in Appendix Figure B.1.

Figure 2.1 shows a map of Ghana, with all cross-regional moves indicated by arrows from the origin to the destination. The map shows the density of all cross-regional origins and destinations of moves. Although women move to and from regions across the country and abroad, most moves in the sample are to the Greater Accra region. Moving from the Ashanti, Eastern, and Northern regions to the region of Greater Accra are the most popular cross-regional residential moves. This is partly a reflection of the data source in that all women were residing in Accra at the time of the interview, but reiterates the previously mentioned fact that 70% of moves in Ghana are to urban areas.

[3] For consistency in both descriptive statistics and analysis, we regard a move to be a "true" move only if it was out of the neighborhood in which the woman was residing. This constituted 85 % of all moves; see Appendix Figure B.3 for the distribution of all types of moves.

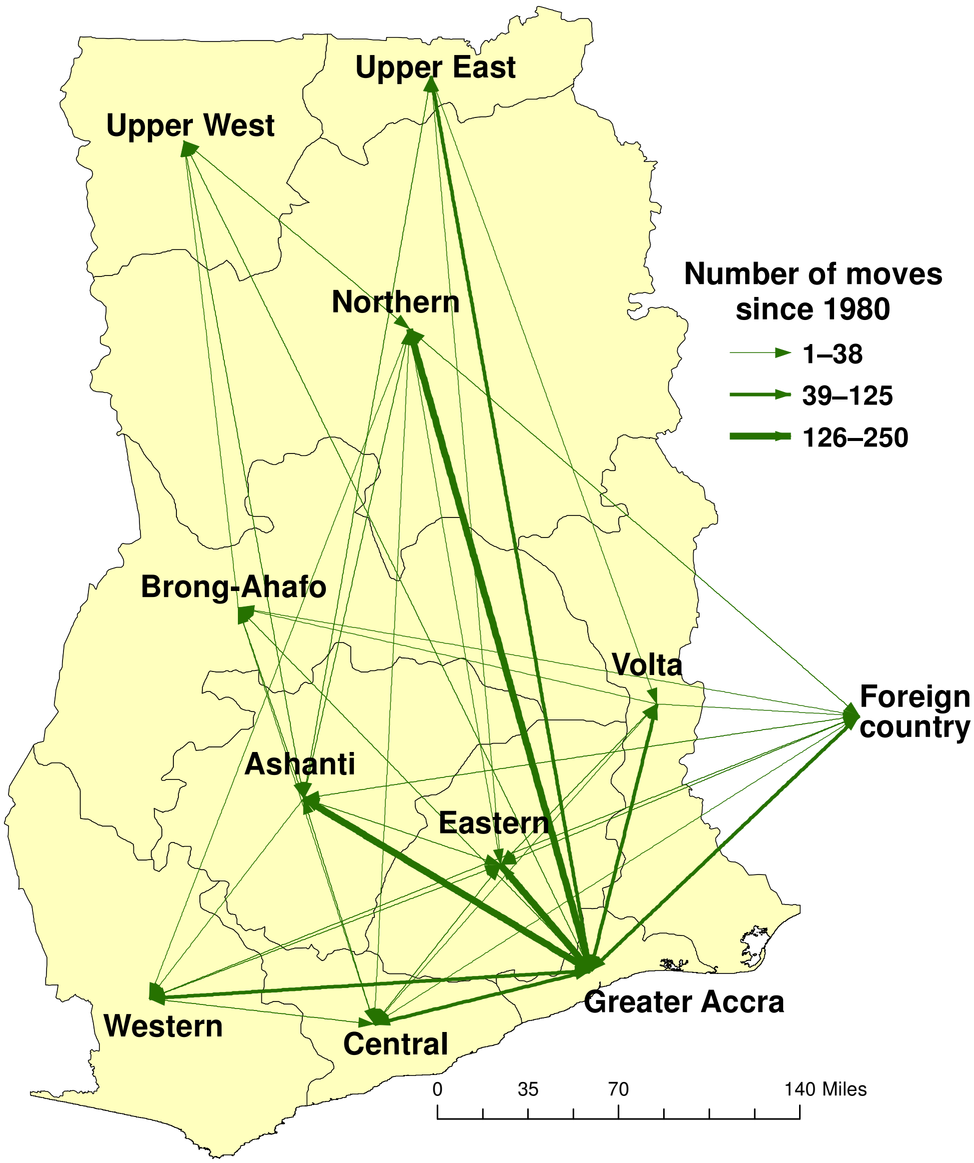


Figure 2.1: Frequency of moves by origin and destination

The migration history in the HAWS survey includes the location of the past three homes that a woman lived in prior to the current home where the study found her, the month and year of each move, the reason why she moved, and whether she knew someone at her destination when she moved. Figure 2.2 provides an overview of the migration patterns observed in the data and also the context for where and why women in this population are moving. Although about 55% of women moved from Greater Accra (which includes the urban center of Accra), residential moves were observed from all regions of Ghana.

The most common reason for moving was improved living conditions, followed closely by marriage. More than 70% of women knew someone when moving: most commonly, a spouse. However, in many cases, women reported moving to unknown neighborhoods, with almost 30% of women reporting not knowing anyone in the location to which they moved.

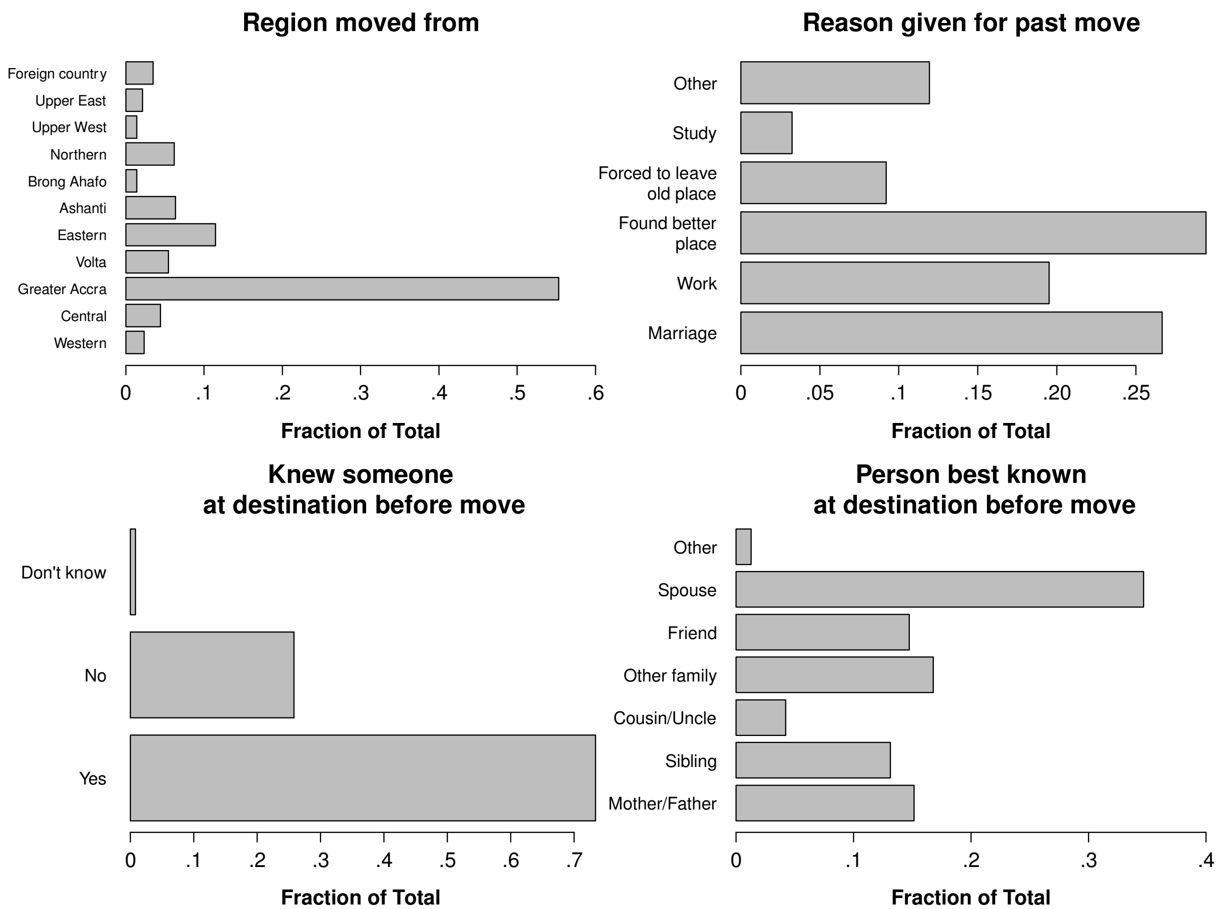


Figure 2.2: Distributions of move characteristics in HAWS sample, clockwise starting from upper left: the region women moved from, the reason given for a past move, the person best known by at the destination before the move, and whether women knew anyone at destination before the move

2.3 Analytical Methods

Total Fertility

The analytical work in this article is divided into three parts. In the first part, we investigate the effect of migration on the total number of children ever born. We use a Poisson model to compare the total fertility of those who had never moved with those who had moved within the area of Greater Accra and with those who had moved from another region. We also compare average cumulative children ever born by mother's age for our sample and the DHS data in order to compare migrants at destination with their nonmigrant counterparts at origin. Analyses were conducted in R (version 3.0.1) and Stata (version 12).

We conduct Poisson regressions with a log link to investigate whether having moved has an effect on total fertility:

$$Y_i \sim Poisson(\lambda_i)$$

Here, Y_i is one of three outcomes: the total number of children ever born, children ever born and still alive, or children born since 2005 and still alive. The incidence rate of birth, λ_i, is modeled by our explanatory variables of interest and individual covariates $\mathbf{X}_i$:

$$\lambda_i = \exp(\beta_0 + \beta_1 MovedWithinAccra_i + \beta_2 InmigratedFromOutsideAccra_i + X_i\gamma)$$

where $MovedWithinAccra_i$ is an indicator for whether the individual had moved but only within Accra, and $In\text{-}migratedFromOutsideAccra_i$ is an indicator for whether the individual moved from outside the Greater Accra region to inside the Greater Accra region. The parameters β_1 and β_2 are the parameters of interest—the effect of moving on completed fertility compared with those who had never moved. $\mathbf{X}_i$ is a vector of individual covariates including mother's age group, ethnicity, education (an indicator for completed at least middle school), and ever married. We also interact age group with education because the effect of age on fertility may differ across education groups. Move status was determined by whether an individual woman claimed to have ever moved outside the neighborhood that she was living in. Moves within the same neighborhood were not determined to be substantial enough to constitute a "true" move and thus were not counted as having moved. We calculate incidence rate ratios with robust standard errors.

Event-History Analysis for Pregnancy Outcomes

In the second part of the analysis, we conduct an event-history analysis using a person-year data structure. Each person-year between the ages of 15 and 47 and between the years of 1980 and 2009 constitutes an observation in the analysis. We chose these ages and years so that each pregnancy outcome would yield positive probabilities of occurring in our data [12]. Similar to Chattopadhyay et al. (2006), we chose a time interval of one year. Because we are interested in the effect of residential duration on reproductive health outcomes in a given year, we eliminate multiple pregnancy observations that were claimed to have happened in the same year.[4] We compare the risk of pregnancy, live birth, and pregnancy outcomes of those who had moved with a comparison group of never-movers.

The linear probability model is shown below:

$$Y_{it} = \alpha_1 + \rho_1 Residence(0-24mo.)_{it} + \rho_2 Residence(25-48mo.)_{it} \\ + \rho_3 Residence(49-72mo.)_{it} + \rho_4 Residence(>72mo.)_{it} + X_{it}\beta + Z_i\gamma + \varepsilon_{it}$$

where Y_{it} is a binary indicator variable for a pregnancy outcome for individual i in time t, $\mathbf{X}_{it}$ is a vector of individual time-varying controls, and $\mathbf{Z}_i$ is a vector of individual time-invariant controls.

Our parameters of interest are $\rho_1, \rho_2,\ \rho_3$, and ρ_4. $Residence(0–24mo.)_{it}$ is an indicator of whether individual i in year t had been living in their residence between 0 and 24 months, $Residence(25–48mo.)_{it}$ is an indicator of an individual i at time t living in their residence between 25 and 48 months, and similarly for the other residential duration–status indicators. $\mathbf{X}_{it}$ is a vector of time-varying covariates, including marital status, an indicator for whether the marriage occurred within the past year, mother's age group, an indicator for whether the woman already has a child, a dummy indicator for whether a previous child had died before time t, and the period of birth in five-year intervals to control for the time trend.[5] We include the dummy variable for "already had a child" because first and higher-order births belong to different biological and life processes and because first-order births are associated with risk of abortion [26]. These covariates were chosen based on the theoretical model and previous literature [12,17]. $\mathbf{Z}_i$ is a

[4] This could be possible if a woman has multiple stillbirths in the same year, for example. It could also be due to measurement error. However, whether the stillbirth happened once or twice in a person-year doesn't affect our analysis because the binary indicator of stillbirth for that person-year is 1, regardless. It is also rare, occurring in only 3.2% of observations.

[5] Results were robust to including year fixed effects instead of period fixed effects.

vector of time-invariant controls that includes both ethnicity and education (an indicator of having finished at least middle school). Again, we interact age group with education because the effect of age on a pregnancy outcome may differ across education groups.

Pregnancy and reproductive health indicators were obtained from detailed pregnancy histories of all women who had given birth. The year of the pregnancy was recorded for all pregnancies on the roster, as well as the outcome of the pregnancy. Induced abortion was differentiated from spontaneous abortion (lost birth or miscarriage) by a positive response to the question, “Did you or someone else put a hand to this pregnancy?” This question, which uses a common euphemism in Accra for induced abortion, was asked only if the outcome of the pregnancy was indicated as a lost birth. However, stigma of abortion is quite high in Ghana, resulting in potentially large measurement error owing to reporting bias. We thus also combine miscarriage and abortion for one estimate of lost birth from either cause. Separate results for miscarriage and abortion are shown in Appendix Table B.1.

Accounting for Selection Bias

Finally, in the third part of the analysis, we use individual fixed effects to account for the systematic differences between those who choose to move and those who do not. Using fixed effects accounts for all characteristics that are unique to that individual and constant over time, including unobserved characteristics such as fertility preferences, risk aversion, and general attitudes. Because we wish to analyze differences observed within each woman over time, our analysis is restricted to women who moved at least once.

The regression below describes our linear probability fixed-effects model:

$$Y_{it} = \alpha_i + \rho_1 Residence(0-24mo.)_{it} + \rho_2 Residence(25-48mo.)_{it}$$
$$+ \rho_3 Residence(49-72mo.)_{it} + X_{it}\beta + \varepsilon_{it}$$

where Y_{it} is a binary indicator of pregnancy outcome for individual i in year t; α_i is the individual fixed effect, which accounts for selection bias; $Residence(0–24mo.)_{it}$ is an indicator of whether individual i in year t had moved in the last 0–24 months (and similarly for $Residence(25–48mo.)_{it}$ and $Residence(49–72mo.)_{it}$); and $\mathbf{X}_{it}$ is the same matrix of time-varying covariates from the previous analyses. The reference category is a residential duration of more than 72 months.

2.4 Results

Total Fertility

We divide women in our sample into three migration status groups: those who never moved (N = 530), those who moved in their lifetime but only within Accra (N = 455), and those who moved in their lifetime across regions (N = 1,108). We show the descriptive statistics for the full data set in Table 2.2. Those who never moved were younger, were less likely to be married, and had fewer total children than those who had moved.

Table 2.2: Sociodemographic characteristics by respondent migration status

	Never Moved	Moved Within Accra	In-Migrated From Outside Accra
Sample Size	530	455	1,108
Ever Married, *N* (%)	250, 0.472	324, 0.712	807, 0.728
At Least Middle School, *N* (%)	336, 0.634	287, 0.631	557, 0.503
Ethnicity: Akan, *N* (%)	73, 0.138	124, 0.273	420, 0.379
Ethnicity: Ewe, *N* (%)	49, 0.092	47, 0.103	140, 0.126
Ethnicity: Ga, *N* (%)	291, 0.549	174, 0.382	89, 0.080
Ethnicity: Other, *N* (%)	117, 0.221	110, 0.242	459, 0.414
Age (years), Mean (SD)	31.02 (13.53)	35.55 (14.19)	33.63 (14.19)
Total Children Ever Born, Mean (SD)	1.68 (2.19)	2.44 (2.47)	2.20 (2.323)
Total Ever Born and Still Alive, Mean (SD)	1.52 (1.97)	2.21 (2.19)	1.96 (2.02)
Total Ever Born and Still Alive Since 2005, Mean (SD)	0.27 (0.56)	0.40 (0.71)	0.40 (0.64)

Notes: Standard deviation is shown in parentheses

Next, we examine the average cumulative children ever born for those in the HAWS data by migration status and compare with the DHS data by region (Figure 2.3). The curves labeled "Never moved," "In-migrated from outside Accra," and "Moved within Accra" come from the HAWS sample, and those labeled by region come from the DHS sample. We see remarkably similar profiles for the HAWS data profiles compared with those in the DHS who live in Greater Accra, regardless of move status. From age 15 to about 35, the observed profiles are directly atop of each other, but those of other regions are dramatically elevated, showing the contrast between rural and urban fertility patterns. A divergence occurs for the HAWS and Greater Accra DHS data following age 35, which may be due to selective, smaller sample sizes of women at those ages in the HAWS data.

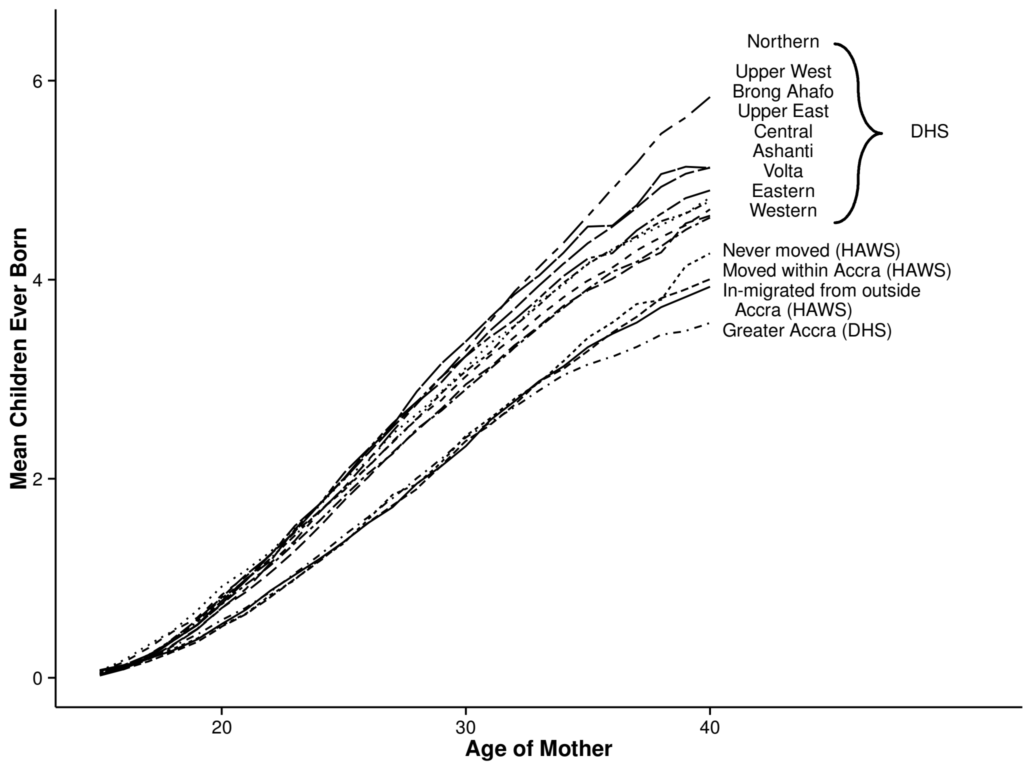


Figure 2.3: Average cumulative children ever born in HAWS sample and DHS 2008 samples

Table 2.3 shows the results from the Poisson regression models for three outcomes: total children ever born, total children born and still alive, and total children born and still alive since 2005.[6] We show

[6] The distribution of the outcome of total children ever born by each migration status group is shown in Appendix Figure B.2.

the incidence rate ratios for these outcomes, with robust standard errors. The reference category is the group of those who had never moved. None of the estimates for migration status group were significant at the .05 level after the model was adjusted for covariates, including marital status (ever married or not), age, education, an interaction of age and education, and ethnicity. Goodness-of-fit chi-squared tests for all models were statistically insignificant, indicating that the Poisson model is appropriate and fits the data.

Table 2.3: Completed fertility incidence rate ratios using Poisson regression of total children ever born, total alive children ever born, and total alive ever born since 2005

	Total Ever Born	Total Ever Born and Still Alive	Total Ever Born and Alive Since 2005
Moved Within Accra	1.026	1.025	$1.215^{\dagger}$
	(0.0486)	(0.0495)	(0.132)
In-Migrated From Outside	1.013	0.993	1.045
	(0.0479)	(0.0483)	(0.105)
Ever Married	4.054^{***}	4.263^{***}	5.835^{***}
	(0.403)	(0.397)	(0.732)
Age 25–29	1.651^{***}	1.619^{***}	$0.814^{\dagger}$
	(0.148)	(0.146)	(0.0998)
Age 30–40	2.739^{***}	2.566^{***}	0.610^{***}
	(0.218)	(0.208)	(0.0762)
Age >40	4.156^{***}	3.603^{***}	0.0754^{***}
	(0.323)	(0.278)	(0.0205)
At Least Middle School	0.616^{***}	0.599^{***}	0.757^{*}
	(0.0616)	(0.0616)	(0.0849)
At Least Middle × Age 25–29	1.171	1.189	$1.345^{\dagger}$
	(0.153)	(0.159)	(0.224)
At Least Middle × Age 30–40	1.259^{*}	1.352^{**}	1.569^{**}
	(0.142)	(0.158)	(0.262)
At Least Middle × Age >40	$1.196^{\dagger}$	1.328^{*}	1.103
	(0.130)	(0.148)	(0.477)
Ethnicity: Ewe	1.051	1.069	1.061
	(0.0551)	(0.0553)	(0.119)
Ethnicity: Ga	1.069	1.044	1.110
	(0.0503)	(0.0493)	(0.112)
Ethnicity: Other	1.003	0.949	1.028
	(0.0462)	(0.0427)	(0.0941)
N	2,093	2,093	2,093

Notes: Coefficients displayed are exponentiated to reflect incidence rate ratios (for example, 1.026 is a 2.6 % increase in the rate of children ever born). Reference categories are never-movers, age <25, and Akan ethnicity. Robust standard errors are shown in parentheses. $^{\dagger}p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

The predicted number of children ever born was 2.1 for all three migration status groups when all covariates were at their means. This prediction varied depending on women's characteristics. For example, an Akan woman over age 40 who has been married and did not finish primary school had predicted numbers of children of 5.1, 5.3, and 5.2 (respectively) for never-movers, moved within Accra, and in-migrated from outside Accra; and an Akan woman who is under 25, not married, and did finish primary school had a predicted number of children of 0.19, regardless of migration status.

Risk of Pregnancy and Pregnancy Outcome by Move Status

We create an event-history analysis to investigate the effect of moving on pregnancy in the first 0–5 years post-move. We compare those who moved in the past 0–24, 25–48, 49–72, and >72 months with the group of never-movers. We construct a panel data set, where one observation is a person-year between the ages of 15 to 47 and between the years 1980 and 2009. The final sample size for our data is 31,936 person-years, composed of 2,022 women.[7]

We generate our summary statistics and analysis based on this sample. Our sample contains a total of 3,989 pregnancies. Of these, 3,364 were live births, 520 were lost births (either miscarriage or induced abortions), and 105 were stillbirths. Of the 520 lost births, 206 were identified as induced abortions, and 314 were reported as miscarriages. In total, 350 (17.3%) women reported at least one lost birth, and the number of lost births per 100 pregnancies was 13. Another 80 women (4.0%) reported at least one stillbirth, and the number of stillbirths per 100 pregnancies was 2.6. A total of 147 women (7.3%) reported having at least one induced abortion, and the number of reported abortions per 100 pregnancies was 5.2. Having multiple induced abortions is not uncommon in Ghana [26]. A study of a hospital in Accra found that 37% of the women in the sample who presented with complications from induced abortions had obtained a previous induced abortion [30]. The measure of abortions per 100 pregnancies is low in our sample compared with other measures from recent urban surveys in Accra:

[7] The process for how we obtained the sample size is described in Appendix Figure B.4.

specifically, the Women's Health Study of Accra (11.2/100 pregnancies), and a clinic-based surveillance survey using preceding birth technique (14.0/100 pregnancies), although neither study focused on the slum population [39]. For this reason, for all of our analyses, we combine miscarriages and abortions. We show the separate analyses for miscarriages and abortions in Appendix Table B.1; results for both outcomes follow the same pattern as the combined measure.

Table 2.4 shows the descriptive statistics of person-years in the data set, by residential duration. For person-years with a shorter residential duration, women were younger, less likely to be married, and less likely to already have a child than those who had lived in the area longer. However, they were most likely to have married in the past year compared with any other duration and compared with those who had never moved. Women with a residential duration of 0–24 months or 25–48 months were the most likely to have a pregnancy, a live birth, an abortion, or a miscarriage.

Because we create the person-year data set, we inevitably have pregnancy years occurring in the same year as moves. For person-years in which a pregnancy occurs, we can distinguish whether the pregnancy or the move came first in the year if both birth month and month of move are not missing.[8] However, for person-years in which a birth did not occur, it does not make sense to infer which event came first. Thus, we do not want to induce bias by categorizing our explanatory variable by our outcome variable. Therefore, the coefficient on ρ_1 should be interpreted as an association between a migration and pregnancy outcome that occurred in the same year, not a causal effect of moving on pregnancy.

[8] Month of move is missing for 84% of all moves that happen in the same year as the current person-year. Birth month is missing for 25% of all births. Death month is missing for 56% of births who died since birth.

Table 2.4: Sociodemographic characteristics by duration of residence

	Duration 0–24 Months	Duration 25–48 Months	Duration 49–72 Months	Duration >72 Months	Never Moved	Total
Age in Years, Mean (SD)	24.58	25.89	26.83	27.09	25.03	26.29
	(6.987)	(7.299)	(7.589)	(9.143)	(8.403)	(8.564)
At Least Middle School, *N* (%)	2,199	1,867	1,381	9,240	3,424	18111
	(0.597)	(0.599)	(0.595)	(0.546)	(0.582)	(0.567)
Married, *N* (%)	1,979	1,856	1,445	9,551	2,401	17232
	(0.537)	(0.596)	(0.622)	(0.564)	(0.408)	(0.540)
Married in Past Year, *N* (%)	470	215	109	871	245	1910
	(0.128)	(0.069)	(0.047)	(0.051)	(0.042)	(0.060)
Previous Child Died, *N* (%)	73	65	66	494	201	899
	(0.020)	(0.021)	(0.028)	(0.029)	(0.034)	(0.028)
Already Have Child, *N* (%)	1,717	1,693	1,373	9,256	2,573	16612
	(0.466)	(0.543)	(0.591)	(0.547)	(0.438)	(0.520)
Ethnicity: Akan, *N* (%)	248	226	140	696	99	1409
	(0.067)	(0.073)	(0.060)	(0.041)	(0.017)	(0.044)
Ethnicity: Ewe, *N* (%)	69	74	42	278	77	540
	(0.019)	(0.024)	(0.018)	(0.016)	(0.013)	(0.017)
Ethnicity: Ga, *N* (%)	92	78	64	577	505	1316
	(0.025)	(0.025)	(0.028)	(0.034)	(0.086)	(0.041)
Ethnicity: Other, *N* (%)	185	184	126	682	144	1321
	(0.050)	(0.059)	(0.054)	(0.040)	(0.024)	(0.041)
Pregnancy, *N* (%)	546	439	315	2077	612	3989
	(0.148)	(0.141)	(0.136)	(0.123)	(0.104)	(0.125)
Live Birth, *N* (%)	432	364	269	1,768	531	3364
	(0.117)	(0.117)	(0.116)	(0.104)	(0.090)	(0.105)
Lost Birth, *N* (%)	99	66	36	248	71	520
	(0.027)	(0.021)	(0.016)	(0.015)	(0.012)	(0.016)
Abortion, *N* (%)	39	29	11	94	33	206
	(0.011)	(0.009)	(0.005)	(0.006)	(0.006)	(0.006)
Miscarriage, *N* (%)	60	37	25	154	38	314
	(0.016)	(0.012)	(0.011)	(0.009)	(0.006)	(0.010)
Stillbirth, *N* (%)	15	9	10	61	10	105
	(0.004)	(0.003)	(0.004)	(0.004)	(0.002)	(0.003)
Observations	3,684	3,115	2,322	16,935	5,880	31,936

The results from the linear probability multivariate models for all outcomes are shown in Table 2.5. Logistic models were substantively similar to linear probability models, and results are shown in Appendix Table B.2. The risk of pregnancy for women who had moved in the past 0–24 months and 25–

48 months (compared with those who had never moved) was elevated by 2.7 and 1.9 percentage points, respectively, with no significant change in risk of live birth. The risk of lost birth for women who had moved in the past 0–24 and 25–48 months was also highly significantly elevated—by 1.5 and 0.90 percentage points, respectively. There was no significant effect of any residential duration on stillbirth compared with never-movers. When all covariates were at their means, the change in risk of pregnancy represented an increase from 11.7% for never-movers to 13.6% for those with a residential duration of 25–48 months (risk ratio of 1.17), and the change in risk of lost birth represented an increase from 1.1% for never-movers to 2.0% for those with a residential duration of 25–48 months (risk ratio of 1.8). The increase in risk of lost birth was more than proportional to the increase in risk of pregnancy. By contrast, the increase in risk of live birth for the same groups was 10.3% to 11.3%—a risk ratio of 1.1, which is less than proportional to the increase in risk of pregnancy.

Consistent with previous literature, the mother's age of 30 and older was negatively associated with pregnancy and live birth compared with those younger than age 25, for those with only primary education or less. Having completed at least middle school was negatively associated with pregnancy and lost birth in the lowest age group. Being married was positively associated with pregnancy and live birth, but not with lost birth or miscarriage. Having been married within the past year was also positively and strongly significantly associated with pregnancy and live birth.

Table 2.5: Linear probability estimates for effect of residential duration on pregnancy outcome compared with those who had never moved

	Pregnancy	Live Birth	Lost Birth	Still Birth
Residence 0-24 Months	0.027***	0.01	0.015***	0.002
	(0.007)	(0.007)	(0.003)	(0.001)
Residence 25-48 Months	0.019*	0.01	0.009*	0.001
	(0.008)	(0.007)	(0.003)	(0.001)
Residence 49-72 Months	0.013	0.008	0.003	0.002
	(0.008)	(0.007)	(0.003)	(0.002)
Residence >72 Months	0.004	-0.001	0.004†	0.001
	(0.006)	(0.005)	(0.002)	(0.001)
Age 25-29	-0.003	-0.008	0.005	0
	(0.009)	(0.008)	(0.003)	(0.001)
Age 30-40	-0.056***	-0.055***	-0.002	0.001
	(0.009)	(0.008)	(0.003)	(0.001)
Age >40	-0.125***	-0.125***	-0.003	0.003
	(0.011)	(0.009)	(0.005)	(0.003)
At Least Middle School	-0.033***	-0.034***	0.001	-0.001
	(0.005)	(0.005)	(0.002)	(0.001)
At Least Middle x Age 25-29	0.030**	0.021*	0.006	0.002
	(0.011)	(0.01)	(0.005)	(0.002)
At Least Middle x Age 30-40	0.021*	0.023*	0	-0.001
	(0.01)	(0.009)	(0.004)	(0.002)
At Least Middle x Age >40	0.006	0.013	-0.005	-0.002
	(0.012)	(0.009)	(0.006)	(0.004)
Previous Child Had Died	0.023	0.022	0.002	-0.001
	(0.015)	(0.014)	(0.005)	(0.002)
Already Had Child	-0.008	-0.003	-0.006†	0.001
	(0.006)	(0.006)	(0.003)	(0.001)
Married	0.117***	0.111***	0.005†	0.002
	(0.007)	(0.006)	(0.003)	(0.001)
Married in Past Year	0.064***	0.062***	0.001	0.001
	(0.01)	(0.01)	(0.004)	(0.002)
1985-1989	-0.016	-0.016†	0.001	-0.001
	(0.01)	(0.009)	(0.004)	(0.002)
1990-1994	-0.040***	-0.048***	0.009*	0
	(0.01)	(0.009)	(0.004)	(0.002)
1995-1999	-0.071***	-0.074***	0.004	-0.002
	(0.009)	(0.009)	(0.004)	(0.002)
2000-2004	-0.067***	-0.074***	0.008*	-0.001
	(0.009)	(0.008)	(0.004)	(0.002)
2005-2009	-0.098***	-0.103***	0.006	0
	(0.009)	(0.008)	(0.004)	(0.002)

Table 2.5 (Continued)

	Pregnancy	Live Birth	Lost Birth	Still Birth
Ethnicity: Ewe	0.004	0.002	0.002	0
	(0.006)	(0.005)	(0.003)	(0.001)
Ethnicity: Ga	0.014**	0.014**	0.001	-0.001
	(0.006)	(0.005)	(0.003)	(0.001)
Ethnicity: Other	-0.010+	-0.003	-0.005*	-0.001
	(0.005)	(0.005)	(0.002)	(0.001)
Constant	0.153***	0.145***	0.006	0.002
	(0.011)	(0.01)	(0.004)	(0.002)
N	31,936	31,936	31,936	31,936
Number of clusters	2,022	2,022	2,022	2,022

Notes: Linear probability model. Reference categories: never-movers, age <25, 1980–1984, and Akan ethnicity. Clustered standard errors are shown in parentheses. $^{\dagger}p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

We show the results graphically in Figure 2.4 with parameter estimates and 95% confidence intervals for each residential duration compared with those who never moved; the dark horizontal line indicates no change in risk from those who never moved. As residential duration increased, risk of pregnancy, live birth, and lost birth decreased. There was an elevated risk of pregnancy and lost birth for those with a residential duration of 0–24 and 25–48 months compared with those who had never moved, with no elevated risk of live birth.

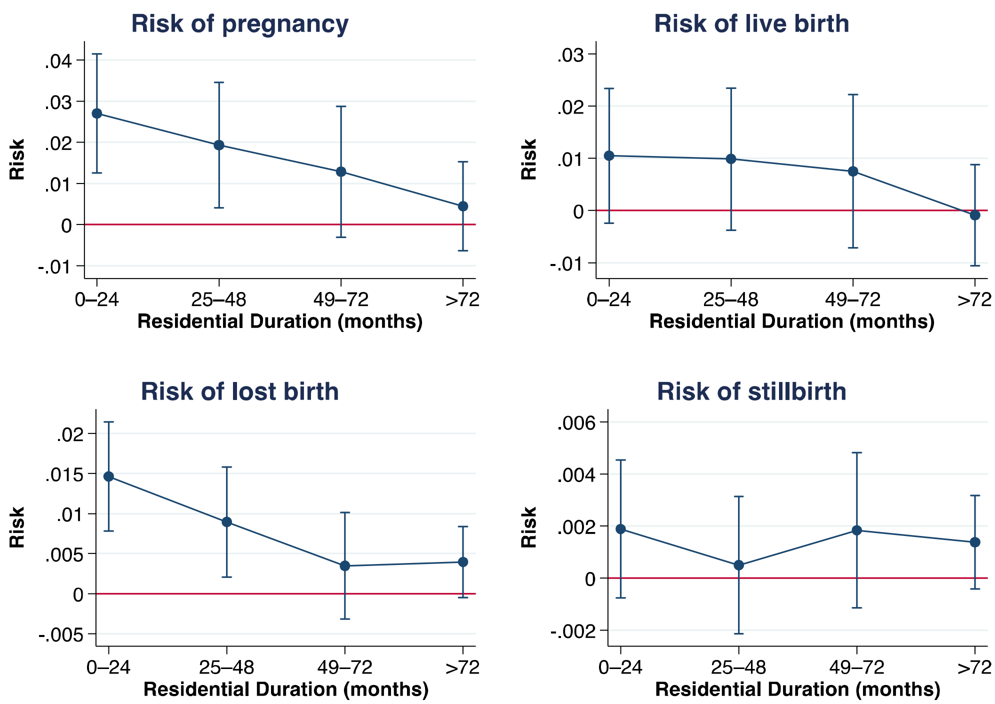


Figure 2.4: Risk of fertility outcome by residential duration as compared with those who never moved

Notes: Point estimates and 95% confidence intervals for risk of pregnancy or pregnancy outcome for movers compared with those who had never moved (linear probability models). The dark horizontal line is no change compared with never-movers

Fixed-Effects Analysis

The results from the fixed-effects analysis, which accounts for selection bias, are shown in Table 2.6. As stated earlier, all women who never moved are excluded from this analysis. We find that the first 24 months of residence are associated with a 2.2 percentage point increase (17.7%) in the likelihood of pregnancy. Subsequently, the likelihood decreases in magnitude when compared with a residential duration of more than 72 months.

Table 2.6: Linear probability estimates with individual fixed effects for effect of residential duration on pregnancy outcomes

	Pregnancy	Live Birth	Lost Birth	Stillbirth
Residence 0-24 Months	0.022**	0.011†	0.010**	0.001
	(0.007)	(0.006)	(0.003)	(0.001)
Residence 25-48 Months	0.014†	0.012†	0.004	-0.001
	(0.007)	(0.006)	(0.003)	(0.001)
Residence 49-72 Months	0.01	0.011	-0.001	0
	(0.008)	(0.007)	(0.003)	(0.002)
Age 25-29	-0.006	-0.008	0	0.001
	(0.009)	(0.008)	(0.004)	(0.002)
Age 30-40	-0.119***	-0.100***	-0.019***	0
	(0.013)	(0.012)	(0.005)	(0.002)
Age >40	-0.264***	-0.232***	-0.031***	-0.001
	(0.019)	(0.017)	(0.008)	(0.004)
Previous Child Died	-0.006	-0.01	0.003	0
	(0.023)	(0.022)	(0.007)	(0.003)
Already Had Child	-0.180***	-0.173***	-0.007*	0.001
	(0.008)	(0.008)	(0.003)	(0.001)
Married	0.222***	0.211***	0.009†	0.002
	(0.012)	(0.011)	(0.005)	(0.002)
Married in Past Year	-0.026*	-0.021†	-0.004	0
	(0.012)	(0.012)	(0.004)	(0.002)
Period 1985-1989	0.062***	0.057***	0.006	-0.001
	(0.012)	(0.012)	(0.004)	(0.002)
Period 1990-1994	0.073***	0.056***	0.017***	0
	(0.014)	(0.013)	(0.005)	(0.002)
Period 1995-1999	0.082***	0.066***	0.018**	-0.002
	(0.016)	(0.015)	(0.006)	(0.003)
Period 2000-2004	0.127***	0.100***	0.027***	0.001
	(0.019)	(0.017)	(0.008)	(0.003)
Period 2005-2009	0.132***	0.101***	0.031***	0.001
	(0.021)	(0.02)	(0.009)	(0.004)
N	26,056	26,056	26,056	26,056
Number of Clusters	1,582	1,582	1,582	1,582

***Notes*:** Clustered standard errors are shown in parentheses. Coefficients displayed represent parameter estimates based on a linear probability model. All models include individual fixed effects. Reference categories are residential duration >72 months, age <25, and 1980–1984. $^{\dagger}p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

The same model was applied to pregnancy outcomes, including live birth, stillbirth, and any lost birth (either abortion or miscarriage). There was no significant association between any residential duration with live birth or with stillbirth. The association between the first 24 months of residence and lost birth was positive and significant, at 1.0 percentage points (61%).

Consistent with previous research and theory, mother's age over 30 was negatively associated with all outcomes compared with age under 25, while being married was positively significantly associated with the likelihood of pregnancy and live birth [8,12]. Already having a child was negatively associated with the likelihood of having another pregnancy, live birth, or lost birth.

2.5 Discussion

This article investigates the relationship between migration and reproductive health outcomes in the modern urban sub-Saharan setting of Accra. We use a unique data set on detailed pregnancy and migration histories collected as part of the Household and Welfare Study of Accra (HAWS) to investigate the effect of migration on the likelihood of pregnancy and live birth, and on the risk of induced abortion, stillbirth, and miscarriage.

We find no difference in total children ever born for those who had never moved, those who had moved within Accra, and those who had in-migrated from outside Accra. Conceptually, these results are consistent with both selection and adaptation mechanisms as influential factors in the impact of migration on sexual behavior and fertility. Those who move to an urban environment may be different than their rural counterparts in their desired fertility. For example, they may seek easier access to modern contraception that can help them reduce their completed fertility. They may desire to invest in better educational opportunities for their children and thus desire a smaller number of children to invest in. Alternatively (and perhaps concurrently), migrants quickly adapt to their new surroundings and adjust their desired fertility and behaviors to match urban natives at destination.

However, we do find an elevated increase of risk of pregnancy and lost birth in the 48 months after migration, although there is no significant increase in live birth in this time period. The change in probability of lost birth represents an increase from 1.1% for never-movers to 2.0% for those that had moved 25–48 months before—almost an 80% increase. One of the concerns with observing pregnancy outcomes directly after the move is that women could already have been pregnant prior to the move. From this perspective, the results for the period 25–48 months after the move are interesting because the move had to have happened before the pregnancy began.

The findings presented in this article can be interpreted in a number of ways. First, women may increase sexual activity after a move because of their adaptation to the urban slum environment, subsequently find that they do not want the resulting pregnancy at that time, and terminate their pregnancy via induced abortion. Alternatively, sexual behavior may stay the same while desired fertility changes upon moving, resulting in more unwanted pregnancies. If contraception is not used and migration results in increased access to and knowledge of abortive measures, women may choose to use induced abortion to keep their fertility low. Third, the move may result in physical or occupational changes, such as an increase or decrease in weight gain or the carrying of heavy loads or bending, which are risk factors for miscarriage [40].

Although the results from the basic group comparisons are consistent with both adaption and selection theories, the same is not true for the fixed-effects models for which we are able to account for selection bias. We find that even after reducing the influence of selection, the likelihood of pregnancy is highest in the first 24 months of residence after a move. Subsequently, the likelihood decreases in magnitude and significance. After controlling for individual fixed effects as well as age, fertility characteristics, marital status, and a time trend, we find that the first 24 months of residence are associated with a 2.2 percentage point increase (17.7%) in the likelihood of pregnancy. The association between the first 24 months of residence and lost birth was positive and significant, at 1.0 percentage points (61%). These results are similar to those of the linear event-history model comparing movers with never-movers, which is evidence that selection is not the driving force in the effect of migration on pregnancy outcomes. The results are consistent with the theory of adaptation to the new environment as the cause of the increase in pregnancy and lost birth following a move.

This study has several limitations. The HAWS data are representative of women living in slums in Accra. As we show in this article, this group of women is highly migratory and differs from the Ghanaian average with respect to their education and assets. It is thus not clear whether the results presented would extend to the larger population of women in Ghana.

Additionally, although the level of detail of the HAWS data in regards to migration and pregnancy history is high in comparison with the DHS or other data sources, potential biases remain. First, the data collected in the survey represent the average slum population at a given point in time. By definition, this includes women who just moved into these areas, and women who move out of slums are not included. Thus, the results are representative only of women who stay long enough in the slum for observing completed fertility. If pregnancy or birth make women more likely to migrate out again, we may underestimate the true impact of migration; and if giving birth means that women become less mobile, the opposite would be true. Although the fixed-effects analysis accounts for selection bias, it does not account for women who out-migrate and are lost to follow-up. Because of the nature of the data, we do not have information on women who moved to Accra and subsequently moved away; and we are able to ascertain neither the frequency of such moves nor whether and in which direction this would bias our results.

Second, some women may not report abortion because of stigma, which can lead to reporting bias if the propensity to report is correlated with migrant status. Stigma of abortion is a significant problem in Ghana, and it is very likely that not all abortions were reported in the HAWS data [26].

Third, we may have residual confounding from omitted time-varying factors, such as health status. We also have no data on the reasons why women decided to obtain an abortion, whether this decision was based on health status or choice, or whether abortions were obtained in a clinical setting or in a clandestine setting.

Finally, because of the nature of the data, we can only make associations about pregnancy and pregnancy outcomes that occur for residential durations of 0–24 months. More research should be conducted to disentangle the temporal directionality of the two events among female migrants, for example, with in-depth qualitative interviews. However, for the estimates of residential duration of 25–48 months, we are sure that the move occurred before the pregnancy began. The pattern of high-to-low risk for pregnancy and lost birth outcomes also suggests that the estimates for 0–24 months after a move are indicative of the underlying trend that a move increases the risk of these fertility outcomes.

This study has important policy implications. Abortion has become more common in Ghana, especially among women aged 20–24. In the 2007 Ghana Maternal Health Survey, the number of abortions per 1000 women was 15 among those aged 15-49 and 25 among those aged 20-24. However, 30% of abortions occurred in the respondent's home, thus increasing the risk of injury and morbidity to the mother (GSS 2009b). Almost one-half of all abortions obtained in Ghana are unsafe [26]. Unsafe abortion is the second leading cause of maternal mortality in Ghana, at 350 maternal deaths per 100,000 live births (95% CI, 210–630), which is higher than the average in the developing world [27]. Thus, from a public health point of view, targeting recent migrants by providing both easy access to contraception and information on public hospital services may improve maternal health outcomes. Other studies have connected the legalization of abortion with lower fertility trends [31,32]. These studies have observed that the increase in modern contraception usage in Ghana has not kept pace with the observed declines in fertility, suggesting that the empirical gap could be explained by increased induced abortion. Abortion as a method of birth control has thus been explored as a possible means for women to reduce their completed fertility in Ghana. In this article, we show that a possible conclusion may be that recent migrants are at risk of such induced abortive measures, although more research should be conducted to fully understand the relationship between migration and induced abortion and miscarriage.

The Ghanaian experience may also inform the larger sub-Saharan African context. In sub-Saharan Africa, 14% of maternal deaths are due to unsafe abortion [41]. As sub-Saharan Africa's rapid urbanization continues, the concern over the welfare of migrants will become more and more important to policy-makers. Internal migration accounts for more than one-half the growth of cities in Africa [2]. One of the most significant trends in migration has been the entry of women into migration streams that had in previous decades been primarily male, with an increasing number of female migrants moving on their own [3,4]

This article explores the association between migration and reproductive health outcomes in a modern urban slum setting of Accra, Ghana. Our analysis complements other research in the field of migration and reproductive health by providing evidence of an increase in risk of pregnancy and abortion for recent migrants. This research highlights the importance of implementing policy to improve urban migrant women's access to reproductive health care services to reduce unwanted pregnancies and mistimed births.

Acknowledgments We are grateful to Allan Hill for the support of the project, as well as to Mark McGovern for his useful feedback and comments. This project was supported by Grant No. T32HS000055 from the Agency for Healthcare Research and Quality. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality. We would also like to thank Matthias Schündeln as well as the William and Flora Hewlett Foundation for the financial support for the HAWS project.

2.6 References

1. United Nations. World urbanization prospects: The 2011 revision (Economic & Social Affairs report). New York, NY: UN Population Division; 2012.

2. Barrios S, Bertinelli L, Strobl E. Climatic change and rural–urban migration: The case of sub-Saharan Africa. Journal of Urban Economics. 2006;60: 357–371.

3. Martin SF. Consultative meeting on migration and mobility and how this movement affects women [Internet]. Malmö Sweden; 2003 Dec. Report No.: Report of the consultative meeting. Available: https://repository.library.georgetown.edu/bitstream/handle/10822/551493/Consultative_meeting.pdf?sequence=1&isAllowed=y

4. Zlotnik H. The global dimensions of female migration (UN International Migration Report: 2002). [Internet]. Washington D.C.: Migration Policy Institute; 2003. Available: http://www.migrationinformation.org/Feature/display.cfm?ID=109

5. UN-Habitat. Planning sustainable cities: Global report on human settlements 2009. London, UK: Earthscan; 2009.

6. Sneeringer SE. Fertility transition in sub-Saharan Africa: A comparative analysis of cohort trends in 30 countries. [Internet]. Calverton, MD: ICF Macro; 2009. Report No.: DHS Comparative Reports No. 23. Available: http://www.popline.org/node/209470

7. Garenne MM. Fertility changes in Sub-Saharan Africa. [Internet]. Calverton, MD: Macro International Inc.; 2008. Report No.: DHS Comparative Reports No. 18. Available: http://www.popline.org/node/204204

8. Brockerhoff M, Biddlecom AE. Migration, sexual behavior and the risk of HIV in Kenya. International migration review. 1999;33: 833–856.

9. Greif MJ, Dodoo FN-A. Internal migration to Nairobi's slums: Linking migrant streams to sexual risk behavior. Health & place. 2011;17: 86–93.

10. Sudhinaraset M, Astone N, Blum RW. Migration and unprotected sex in Shanghai, China: correlates of condom use and contraceptive consistency across migrant and nonmigrant youth. Journal of Adolescent Health. 2012;50: S68–S74.

11. Brockerhoff M, Yang X. Impact of migration on fertility in sub-Saharan Africa. Social Biology. 1994;41: 19–43.

12. Chattopadhyay A, White MJ, Debpuur C. Migrant fertility in Ghana: Selection versus adaptation and disruption as causal mechanisms. Population studies. 2006;60: 189–203.

13. Brockerhoff M. Fertility and family planning in African cities: The impact of female migration. Journal of Biosocial Science. 1995;27: 347–358.

14. Bongaarts J, Frank O, Lesthaeghe R. The proximate determinants of fertility in sub-Saharan Africa. Population and Development Review. 1984; 511–537.

15. Anarfi JK. Sexuality, migration and AIDS in Ghana-A socio-behavioural study. Health Transition Review. 1993;3: 45–67.

16. Lindstrom DP. Rural-urban migration and reproductive behavior in Guatemala. Population Research and Policy Review. 2003;22: 351–372.

17. White MJ, Tagoe E, Stiff C, Adazu K, Smith DJ. Urbanization and the fertility transition in Ghana. Population Research and Policy Review. 2005;24: 59–83.

18. Borjas GJ. Self-Selection and the Earnings of Immigrants. The American Economic Review. 1987;77: 531–553.

19. Goldstein S, Goldstein A. Techniques for analysis of the interrelations between migration and fertility [Internet]. Santa Monica, CA: Rand Corporation; 1982. Report No.: Rand Paper No. P-6844. Available: http://pdf.usaid.gov/pdf_docs/pnaap927.pdf

20. Bocquier P, Madise NJ, Zulu EM. Is there an urban advantage in child survival in sub-Saharan Africa? Evidence from 18 countries in the 1990s. Demography. 2011;48: 531–558.

21. Bloom DE, Canning D, Fink G, Khanna T, Salyer P. Urban settlement: Data, measures, and trends [Internet]. Helsinki, Finland: World Institute for Development Economics Research; 2010. Report No.: Working Paper No. 2010, 12. Available: http://www.econstor.eu/handle/10419/54099

22. Ghana Statistical Service (GSS), Ghana Health Service (GHS), and ICF Macro. Ghana Maternal Health Survey 2007. Calverton, MD: Ghana Statistical Service (GSS), Ghana Health Service (GHS), and Macro International; 2009.

23. Ghana AIDS Commission. Ghana country AIDS progress report: Reporting period January 2010-December 2011 [Internet]. UNAIDS; 2012. Available: http://www.unaids.org/en/dataanalysis/knowyourresponse/countryprogressreports/2012countries/ce_GH_Narrative_Report[1].pdf

24. Ghana Statistical Service (GSS), Ghana Health Service (GHS), and ICF Macro. Ghana Demographic and Health Survey 2008: Key Findings. Calverton, Maryland, USA: GSS, GHS, and ICF Macro; 2009.

25. Schwandt HM, Creanga AA, Adanu RM, Danso KA, Agbenyega T, Hindin MJ. Pathways to unsafe abortion in Ghana: the role of male partners, women and health care providers. Contraception. 2013;88: 509–517.

26. Sundaram A, Juarez F, Bankole A, Singh S. Factors Associated with Abortion-Seeking and Obtaining a Safe Abortion in Ghana. Studies in family planning. 2012;43: 273–286.

27. WHO, UNICEF, UNFPA, and The World Bank. Trends in Maternal Mortality: 1990 to 2010; WHO, UNICEF, UNFPA and The World Bank estimates. [Internet]. Geneva, Switzerland: World Health Organization; 2012. Available: WHO, UNICEF, UNFPA and The World Bank estimates

28. Aboagye PK, Gebreselassie H, Asare GQ, Mitchell EM, Addy J. An assessment of the readiness to offer contraceptives and comprehensive abortion care in the Greater Accra Eastern and Ashanti regions of Ghana. [Internet]. Chapel Hill, NC: Ipas; 2007. Available: http://www.popline.org/node/207033

29. Sedgh G. Abortion in Ghana (Issues in Brief 2010 Series, No. 2). New York, NY: Guttmacher Institute; 2010.

30. Adanu RM, Ntumy MN, Tweneboah E. Profile of women with abortion complications in Ghana. Tropical Doctor. 2005;35: 139–142.

31. Agyei-Mensah S. Fertility transition in Ghana: looking back and looking forward. Population, Space and Place. 2006;12: 461–477.

32. Finlay JE, Fox AM. Reproductive health laws and fertility decline in Ghana. International Journal of Gynecology & Obstetrics. 2013;123: e24–e28. doi:10.1016/j.ijgo.2013.07.008

33. Geelhoed DW, Nayembil D, Asare K, Van Leeuwen JH, Van Roosmalen J. Contraception and induced abortion in rural Ghana. Tropical Medicine & International Health. 2002;7: 708–716.

34. Ackah C, Medvedev D. Internal migration in Ghana: determinants and welfare impacts. International Journal of Social Economics. 2012;39: 764–784.

35. Ghana Statistical Service (GSS). 2010 Population and Housing Census: Summary of Final Results [Internet]. 2012 May. Available: http://www.statsghana.gov.gh/docfiles/2010phc/Census2010_Summary_report_of_final_ results.pdf

36. UN-Habitat. State of the World's Cities 2012/2013. Prosperity of Cities. New York, NY: UN-Habitat; 2013.

37. Megill D. Recommendations for designing master sample for Ghana Intercensal Household Survey Program. Ghana and Washington DC: Demographic and Socioeconomic Research Services, International Programs Center, and U.S. Census Bureau; 2002.

38. Montana L. Estimating child mortality in resource poor settings with insufficient data (Unpublished doctoral dissertation). Cambridge, MA: Harvard School of Public Health Global Health and Population Dissertation; 2011.

39. Oliveras E, Ahiadeke C, Adanu RM, Hill AG. Clinic-based Surveillance of Adverse Pregnancy Outcomes to Identify Induced Abortions in Accra, Ghana. Studies in family planning. 2008;39: 133–140.

40. Florack EI, Zielhuis GA, Pellegrino JE, Rolland R. Occupational physical activity and the occurrence of spontaneous abortion. International journal of epidemiology. 1993;22: 878.

41. World Health Organization. Unsafe Abortion: Global and Regional Estimates of the Incidence of Unsafe Abortion and Associated Mortality in 2008 [Internet]. Geneva, Switzerland: WHO; 2011. Available: http://www.who.int/reproductivehealth/publications/unsafe_abortion/9789241501118/en/

Chapter 3:

Bias of standard errors in difference-in-differences analyses when number of clusters is small

Abstract

Difference-in-differences (DID) estimation is popular in public health and policy literature. Typically DID analyses evaluate group-level interventions using individual data. Many studies adjust standard errors to account for correlation using a cluster-robust variance matrix; however, when the number of clusters is small, variance may be biased downwards leading to confidence intervals that are too tight. We present a conceptual overview of DID estimation, the problems faced in small samples for commonly used analyses, and results from a simulation study examining the performance of alternative methods in which we vary the correlation structure, the balance of the data, and the proportion of treated groups. The results indicate that when the number clusters is less than 20, the cluster-robust variance matrix estimate is consistently biased downwards, particularly in settings with unbalanced cluster sizes and when proportion of treated groups is low, even with the inclusion of cluster-specific fixed effects. Aggregation, the wild cluster bootstrap, permutation tests, and bias-adjusted generalized estimating equations (GEE) generally provide accurate coverage rates for almost all scenarios, though GEE suffers from low power.

Key words: inference • clustering • difference-in-difference • Monte Carlo simulation • standard errors

3.1 Introduction

Difference-in-differences (DID) estimation has become increasingly popular in public health and policy literature. Most typically, DID analysis evaluates the effect of a group-level policy on individual-level outcomes. Because observations are grouped, errors are correlated across individuals within groups; models that do not account for this correlation will result in misleadingly small standard errors and incorrect inference [1,2].

DID estimation is often used to analyse data from a natural experiment and therefore the number of groups may be small; for example, when examining the effect of state-level policy changes in Medicare or Medicaid, the total number of possible groups (states in this case) is limited.

The most common approach to control for correlated errors is to use a cluster-robust variance estimate. However, when the number of clusters is small (generally less than 50), asymptotic properties necessary for correct inference may not apply. Simulation studies have shown that Wald tests from cluster-robust standard errors over-reject the null when the number of clusters is small, with variable type I error rates that can be more than twice that of targeted rates [3–5]. Tests from data that are imbalanced with respect to cluster size have been shown to be particularly prone to over-rejection, and similarly from data with a very low (or very high) proportion of groups experiencing policy changes [6–9]. Empirical work using longitudinal data is seldom balanced, particularly as the length of the time series lengthens.

A variety of approaches have been proposed to improve inference in correlated data with small number of groups including bias-adjusted generalized estimating equations (GEE) [10–13], bootstrapping methods [8,14,15], permutation tests [16–18] and aggregation [1,4]. However, not all methods have been applied in a multilevel DID setting where observations are clustered in groups and repeated over time. In addition, little prior work has compared the performance of more than one or two solutions [1,3,4,8,14]. Finally, few articles compare Type-1 error rates to Type-2 error rates across a wide range of approaches.

The contribution of our paper is four-fold. First, we compare the empirical performance of the most commonly used modelling solutions in DID estimation for panel data from the fields of both Econometrics and Statistics, including cluster-robust standard errors, wild cluster bootstrapping, random effects models, GEE with bias corrections, permutation tests, and aggregation. Second, we present results from a Monte Carlo simulation study in which we test a wide range of scenarios, by varying the degree of error correlation, the balance of cluster sizes, and the proportion of treated clusters. Third, we directly compare empirical coverage rates to power for all models. Finally, we apply our results to re-evaluate a recent article examining the effect of over-the-counter emergency contraception laws on teenage sexual behaviour using data from the National Longitudinal Survey of Youth 1997 to demonstrate the importance of our findings for health policy research.

3.2 Materials and Methods

Conceptual Review

The main idea of DID is to compare relative trends in treatment and control areas, before and after the imposition of a group-level treatment. Data used for DID may be either repeated cross sections over time or repeated observations on individual over time, and are grouped by the level of the treatment (for example, for state policy changes, individuals are grouped by states). Typically DID analyses use a regression framework to estimate differential changes in outcomes for the treated areas in the post-intervention period using changes in outcomes in the control areas to estimate background secular trends. The basic assumption for unbiased effect estimates is that of parallel trends; that is, the treatment areas would have had a trend parallel to the control areas in the post-treatment period, had they not been treated. In this article we assume these assumptions hold and we focus on the empirical challenge of serial correlation, in that grouped observations result in error correlation within groups and across time.

Models

The three most common methods to account for correlation in observations across groups in linear models are: (1) using post-hoc adjustments to the standard errors after ordinary least squares (OLS) estimation of the point estimates such as clustered standard errors, bootstrapping, or permutation tests (also known as randomization inference); (2) directly modelling the within cluster correlation matrix such as using random effects or GEE models; and (3) aggregating the data to the cluster level, thereby eliminating the correlation.

Post-hoc adjustments

We examine the performance of three post-hoc adjustments to the standard ordinary least squares model: clustered standard errors, wild cluster bootstrap, and permutation tests. Clustered standard errors (CSE) are a generalization of the White robust covariance sandwich estimator that allows for clustering in addition to heteroscedasticity [2,19]. Appendix Table C.1 shows the technical details for estimating the cluster-robust variance matrix. To account for serial correlation in outcomes, we cluster at the group, rather than group-time level, we include a finite sample adjustment that scales the residuals based on the number of groups, and we test our hypothesis using a Wald test with T(G-1) degrees of freedom [1,3]. These features are all standard to Stata's reg command, vce(cluster) option (StataCorp 2015). We test OLS models both with and without individual-level fixed effects. Fixed effects may reduce some, though not all, of the within-cluster correlation. However, we expect CSE adjustments will perform poorly in small number of clusters because the robust variance estimator is based on a sample variance estimate and residuals tend to underestimate the true error.

Next, we evaluate the wild-cluster bootstrap. In previous literature, cluster-specific bootstrapping (or block bootstrapping) has been found to result in estimation problems, particularly when the treatment variable of interest is binary and cluster invariant [14]. Resampling by cluster may result in some samples having no variation (or limited variation) in treatment, leading to replications that cannot be estimated (or near 0 standard errors). A variation on bootstrapping that does not produce these issues is the wild-cluster bootstrap [14]. For each iteration, this method resamples all observations, randomly transforms some of the residuals by multiplying them by a random variable, and re-estimates the treatment effect and Wald test statistic. Because the random variable multiplying the residuals is the same within each cluster, the within-cluster error correlation structure is preserved. The bootstrap p-value can then be calculated as the proportion of times that the original sample Wald statistic was as or more extreme than the bootstrapped statistics. Appendix Table C.1 provides details of the procedure. We use the Rademacher 2-point distribution for the random variable. This method was shown to dramatically reduce bias in standard errors in DID estimation of small samples in both Cameron and Miller (2015) and Cameron et al. (2008), when compared against the cluster-robust variance matrix.

Finally, we estimate exact p-values using permutation tests. Like bootstrapping, permutation tests (also called randomization inference) are nonparametric resampling methods [16–18,20]. They have been more recently applied to quasi-experimental settings [9,21–23]. The procedure reassigns entire groups to either treatment or control and recalculates the treatment effect in each reassigned sample, generating a randomization distribution. An exact p-value can be calculated as the probability of obtaining a test statistic as far or further from the observed [23].

Directly modelling error correlation

We directly model the error correlation in two ways. First, we estimate random effect (RE) models (also known as mixed models, hierarchical models, or varying intercept models), where we specify that the error is composed of three components: a group-specific component, an individual-specific component, and a random, uncorrelated component as follows:

$$u_{ig} = \alpha_g + \varphi_i + \varepsilon_{ig}$$

where $\alpha_g \sim N(0, \sigma_\alpha^2)$, $\varphi_i \sim N(0, \sigma_\varphi^2)$, and ε_{ig} is the random noise. The model assumes that all unobservable covariates are time-invariant, a strong assumption that is rarely true in grouped data (for example, when data are grouped by state, there are usually economic and social changes that vary over time within states). The RE models are estimated via maximum likelihood (MLE) methods, which are asymptotic-based (in G), so the sample size must be sufficiently large to produce consistent estimates. When G is small, the data supply little information about the distribution of group-level effects and thus the group-level variance may be poorly estimated [24].

Second, we estimate GEE models under various adjustments [2]. There are two main problems with the GEE in small samples. First, as with CSE, variance estimates are biased downward. This bias gets larger as G gets smaller, and can be estimated using a Taylor series approximation (Appendix Table C.1). Fay and Graubard (2001) use a first-order Taylor expansion to obtain a bias-corrected sandwich estimator. Second, the z-distribution is a poor approximation of the sampling distribution in small samples and leads to over-rejection of the null; a T-distribution has been shown to improve the accuracy of the test size [10–13,25,26]. We test the two modifications to the GEE by using both the bias adjustment and the T-distribution, as proposed by Fay and Graubard (2001) and implemented in the R package *saws*. All simulations use an exchangeable working correlation matrix.

Aggregation

Lastly, we collapse the data into group cells pre- and post-intervention, thus eliminating the error correlation. Similar to Bertrand et al. (2004), we first regress the outcome on individual controls (in the simulation, this is just the intercept) and form residuals. We then compute the difference of the mean of the residuals before and after the intervention for each group and estimate the OLS regression of the outcome on an indicator of whether the group was treated. We obtain the variance using OLS.

The additional problem of unbalanced data

Most data used in empirical analysis are unbalanced, meaning that the number of observations per cluster is unequal across clusters. Previous work has demonstrated that in unbalanced data, false rejection rates are higher than in balanced data for CSE [6–8] as well as for GEE [11]. Carter et al. (2013) provide the theory for this phenomenon; they demonstrate that the effective number of clusters is reduced when the cluster size varies across clusters. They provide a measure for calculating this effective number of clusters (G*) that scales down the true number of clusters (G). MacKinnon and Webb (2014) use this measure to produce critical values from the T(G*-1) distribution and compare rejection frequencies to those from the usual T(G-1) distribution. They find that the T(G*-1) distribution frequently (though not always) results in more accurate inferences, but that it can under-reject.

Additionally, Conley and Taber (2011) show that the proportion of treatment groups also impacts the standard errors in simulation studies [9]. They show that when this proportion is very small (or very large), the treatment effect, though unbiased, is no longer consistent (see full explanation and proof in Conley and Taber 2011).

Monte Carlo simulations

For each method above, we perform Monte Carlo simulation studies to obtain quantitative results of the empirical coverage and power. The data generating process is as follows:

$$Y_{igt} = \beta Trt_{igt} + u_g + v_i + w_{gt} + \varepsilon_{igt},$$

with $u_g \sim N(0, \sigma_u^2);\ v_i \sim N(0, \sigma_v^2);\ w_{gt} \sim AR(1)\ from\ N(0, \sigma_w^2);\ \varepsilon_{igt} \sim N(0, \sigma_{\varepsilon)}^2$

where Y_{igt} is the outcome for individual *i* in group *g* at time *t*. Trt_{igt} is an indicator for whether the intervention affected group *g* at time *t* and β is the DID estimate. Under the null, we set $\beta = 0$. Via this data generating process, the error is correlated within groups and within individuals as normally distributed disturbances, as well as within groups by an AR(1) process with normal disturbances and a correlation of $\rho = 0.8$. The AR(1) process allows data to be correlated across time within groups, as in the way state-specific economic or health conditions vary over time. Bertrand et al. (2004) show that this AR(1) process is too simple to be realistic in panel data; however, it is illustrative of the problems in serial correlation, and we follow the same process [1].

If σ_w^2 is 0 or near 0, then group-level fixed effects or random effects should account fully for the within cluster correlation as the correlation of errors is driven solely by a common shock process. However, previous research has shown that the inclusion of state fixed effects in state-year panel data does not eliminate the within-state correlation of the error [1,3]. Thus our data generating process induces correlation in the error even after accounting for group and time fixed effects.

Similar to Donald and Lang (2001), in the low correlation scenario, we set $\sigma_\varepsilon^2 = 10\sigma_v^2 = 100\sigma_u^2 = 100\sigma_w^2 = 1$. In the high correlation scenario, we change $\sigma_u^2 = \sigma_w^2 = 0.05$, and $\sigma_v^2 = 0.15$. Although our data generating process is unique, our intraclass correlations are similar to those of other studies [4,11].

The list of simulation scenarios is shown in Table 3.1. We begin our simulations with balanced data, where the number of individuals per group is always 30 for both low and high correlations and the proportion of treated groups is 0.5. We vary the time points per individual for each scenario, allowing for exactly 1 before and after the treatment (T=2), 2 before and 2 after (T=4), 5 before and 5 after (T=10), and 10 before and 10 after (T=20).

Next, we set the number of time points to T=20 and alter the scenario by generating unbalanced data. In the first unbalanced case, we allow the number of individuals per group to vary on a uniform distribution between 1 and 59 (for an average of 30). In the second unbalanced case, we test the case in which the proportion of treated groups is 0.2 (with balanced cluster sizes). Finally, we set $\sigma_w^2 = 0$ in balanced data to show how modelling solutions change when the correlation can be fully accounted for with the random effects model.

Table 3.1: Characteristics of simulation scenarios

Simulation Scenario	Correlation	Individuals per cluster	Time points per individual (T)	Proportion of treated clusters
Balanced data	Low	30	2, 4, 10, and 20	0.5
Balanced data	High	30	2, 4, 10, and 20	0.5
Unbalanced cluster size	High	1-59	20	0.5
Low proportion of treated clusters	High	30	20	0.2
Unbalanced cluster size, $\sigma_w^2 = 0$	High	30	20	0.5

For each simulation scenario, we simulate 2000 data sets under the null treatment effect and run all methods specified above. We estimate the coverage rate as the fraction of simulations in which the 95% confidence interval for β covers the null. (In the permutation test and wild cluster bootstrap, we test whether the p-value is greater than or equal to 0.05). Coverage rates lower than 0.95 indicate underestimation of standard errors, while higher coverage probabilities indicate overestimation.

Next, we impose a treatment effect of 0.6 standard deviation. We again simulate 2000 datasets for each scenario in Table 3.1, and we calculate power as the fraction of the data sets that resulted in a significant effect. All simulations are conducted using R, version 3.2.3.

3.3 Results

Simulation results

Figure 3.1 presents the results of our simulations on coverage rates in the high correlation scenario for 2, 4, 10, and 20 time points per individual. The red horizontal line is the nominal coverage of 0.95 and the red dotted lines indicate the Monte Carlo confidence interval. We show the results for the CSE, CSE with individual fixed effects, GEE with bias adjustment and F-distribution, aggregation, permutation, and the wild cluster bootstrap. We found that the GEE with a robust variance matrix and the GEE with only an F-distribution correction consistently underestimated coverage compared to the GEE with both a bias adjustment and the F-distribution; therefore, we exclude these models from our results (results are shown in Appendix Figure C.1). The random effects model performed extremely poorly in these scenarios because it could not adjust for the AR(1) process; we therefore do no show it on the graphs. However, we show the results of the RE model when $\sigma_w^2 = 0$ in Appendix Figure C.2; in this case, the RE model predictably provides accurate coverage even when the number of groups is small.

For each figure, as the number of groups increased, the models generally converged to the nominal coverage rate. The coverage rate of CSE was below acceptable below when G<10. When CSE included individual fixed effects, this produced more conservative and acceptable estimates when the number of time points per individual was 2 or 4; however, as serial correlation increased with more time points, the fixed effects were not able to account for all of the within-cluster serial correlation. The wild cluster bootstrap also had inadequate coverage when G<7, though performed better than CSE. MacKinnon and Webb (2014) find that a different distribution for the random variable (such as the Webb 6-point distribution rather than the Rademacher 2-point distribution used here) performs much better in small G [8].

Permutation was quite conservative and generally resulted in under-rejection; however, it performed better as the number of time points increased. Two models had consistently correct coverage within simulation 95% confidence intervals across both number of clusters and number of time points: aggregation of the data and GEE with the bias adjustment. Similar results were found for low correlation scenarios (Appendix Figure C.3).

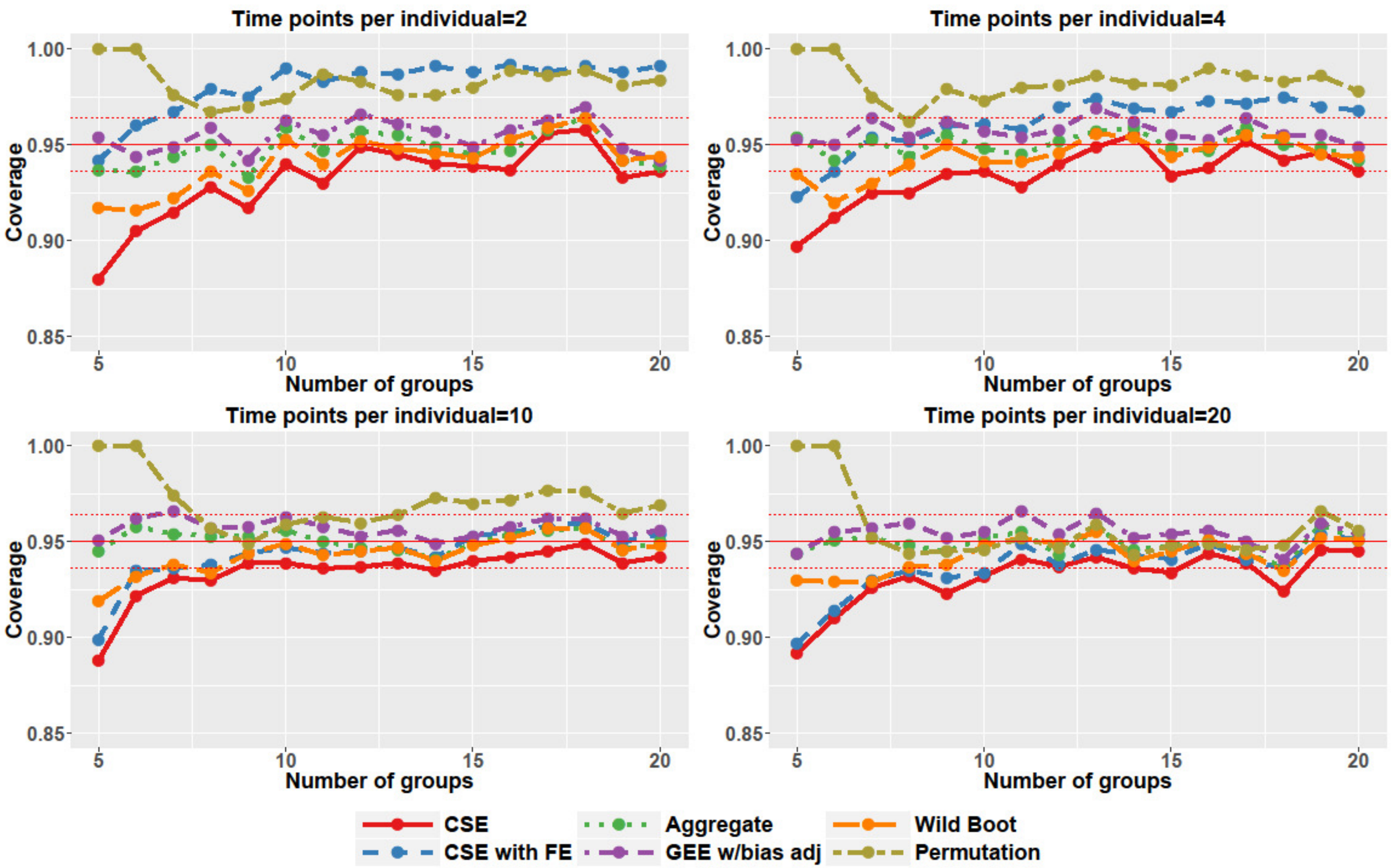


Figure 3.1: Coverage in high correlation and balanced data scenario as number of time points per individual increases

Notes: Horizontal thin red line shows 0.95, the nominal coverage. The Monte Carlo simulation standard error is 0.005 for $\hat{\alpha} = .05$ and 2000 iterations, therefore 95% of simulations are expected to yield estimated coverage in the range (0.9402, 0.9598).

Figure 3.2 shows the coverage rates for models in high correlation scenarios with unbalanced cluster sizes and a low proportion of treated groups, both with 20 time points per individual. The results are striking. In the unbalanced cluster size scenario (graph a) with G<15, CSE, whether with or without individual fixed effects, resulted in a coverage rate of between 85 – 92%, a severe underestimation. The wild cluster bootstrap also slightly over-rejected in small G<9 though not nearly as severely. Permutation performed similarly to the wild cluster bootstrap, though permutation resulted in coverage rate of 1 when G<7 due to limited number of permutations of the data resulting in p-values necessarily greater than 0.05. As before, aggregation and GEE with bias adjustment performed consistently well.

In the low proportion of treated groups scenario (graph b), CSE, both with and without individual fixed effects, led to severe over-rejection with coverage rates less than 90% even when G<16 and reached only 91% at G=20. CSE coverage was not monotonically increasing as G increased because although we aimed for a treated proportion of 0.2, the actual proportion of treated clusters was not constant; for example, when G was 7 the number of treated clusters was 2, resulting in a proportion of about .28 while when G was 10, the number of treated was still 2 and thus the proportion was 0.2. The dramatic swings as the treated proportion changes shows how sensitive CSE are to this issue when G is small.

Aggregation, GEE with bias adjustment, and permutation were most consistent, though the GEE severely over-rejected when G<7, most likely because there was only one treated cluster in those cases and the variance matrix of the GEE relies on averaging residuals across clusters. The wild cluster, on the other hand, was too conservative when G<12, a result of limited transformations of bootstrap residuals when very few clusters (or almost all clusters) are treated; again a different distribution for the transformation of the residuals may improve its performance.

Finally, we investigate the power of these models to detect a treatment effect (Figures 3.3 and 3.4). As expected, all methods resulted in unbiased treatment effects (Appendix Figure C.4); however, power varied widely. Figure 3.3 presents the results of coverage versus power when cluster sizes were imbalanced, with coverage on the x-axis and power on the y-axis; we present results for G<12. The wild cluster bootstrap, aggregation, and permutation provided the most power for adequate coverage when G< 11, although permutation had no power when G<7 because of the issue with limited number of permutations mentioned above. The GEE with bias adjustment provided excellent coverage but was consistently underpowered compared to aggregation and wild cluster bootstrap. In the low proportion of treated groups scenario (Figure 3.4), aggregation outperformed all models when G<7. Both aggregation and permutation performed well when G was between 7 and 12, while the wild cluster bootstrap lacked power. As shown previously, CSE did not provide adequate coverage.

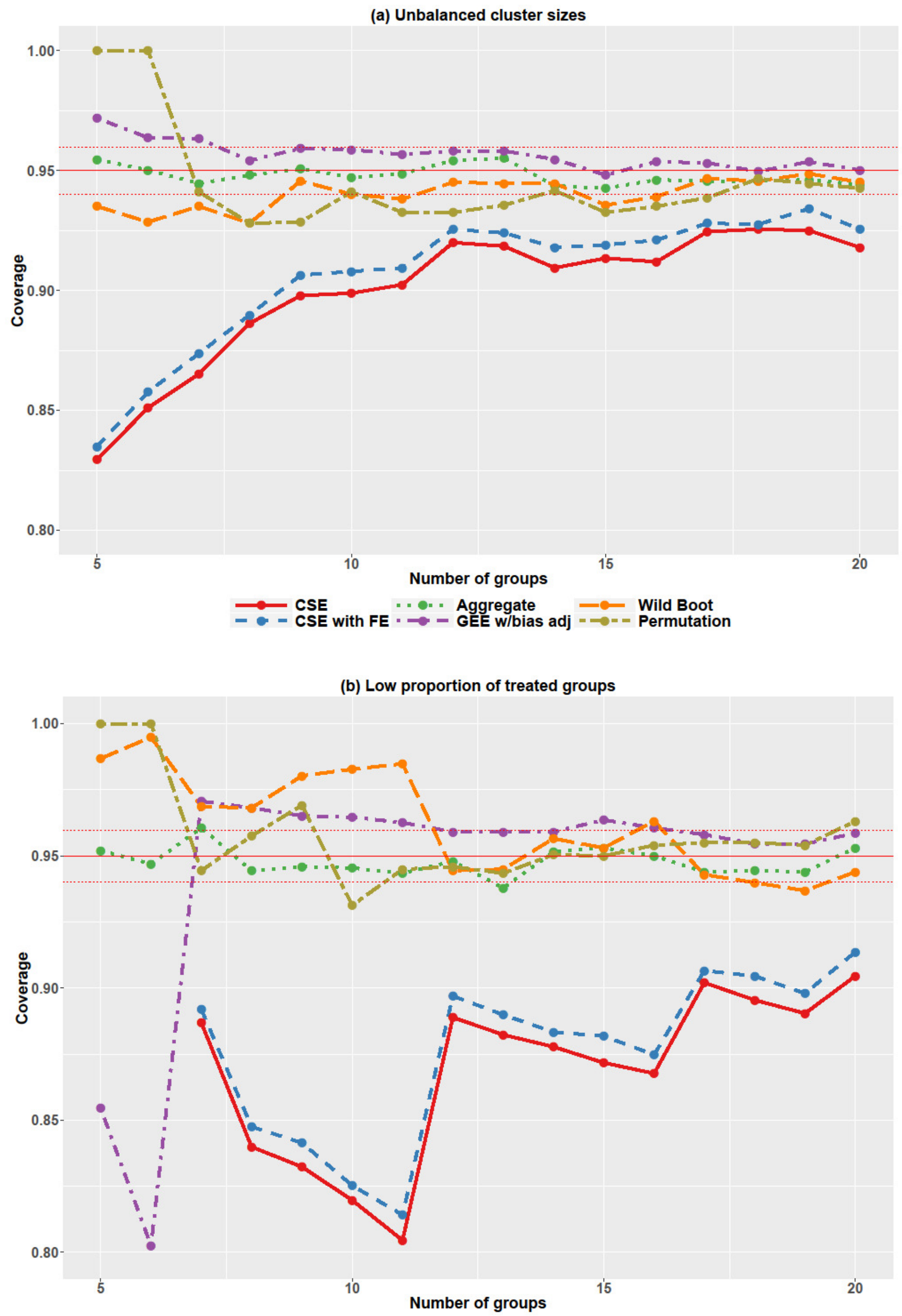


Figure 3.2: Coverage probabilities for (a) unbalanced cluster sizes and (b) a 0.2 proportion of treated groups

Notes: Horizontal red line shows 0.95, the nominal coverage and dotted lines show the Monte Carlo simulation 95% confidence intervals (0.9402, 0.9598). In graph (b), the coverage for CSE with individual fixed effects is off of the graph for G=5 and G=6, at 0.67 and 0.61 respectively, and for CSE at 0.69 and 0.64, respectively. For G=5 and G=6, GEE degrees of freedom using $\hat{d}_H$ rather than the usual $\tilde{d}_H$ because the latter has multicollinearity problems in the smoothing formula due to inclusion of only one treated cluster (see Fay and Graubard 2001). For G>6, $\tilde{d}_H$ is used, however, results were similar using $\hat{d}_H$

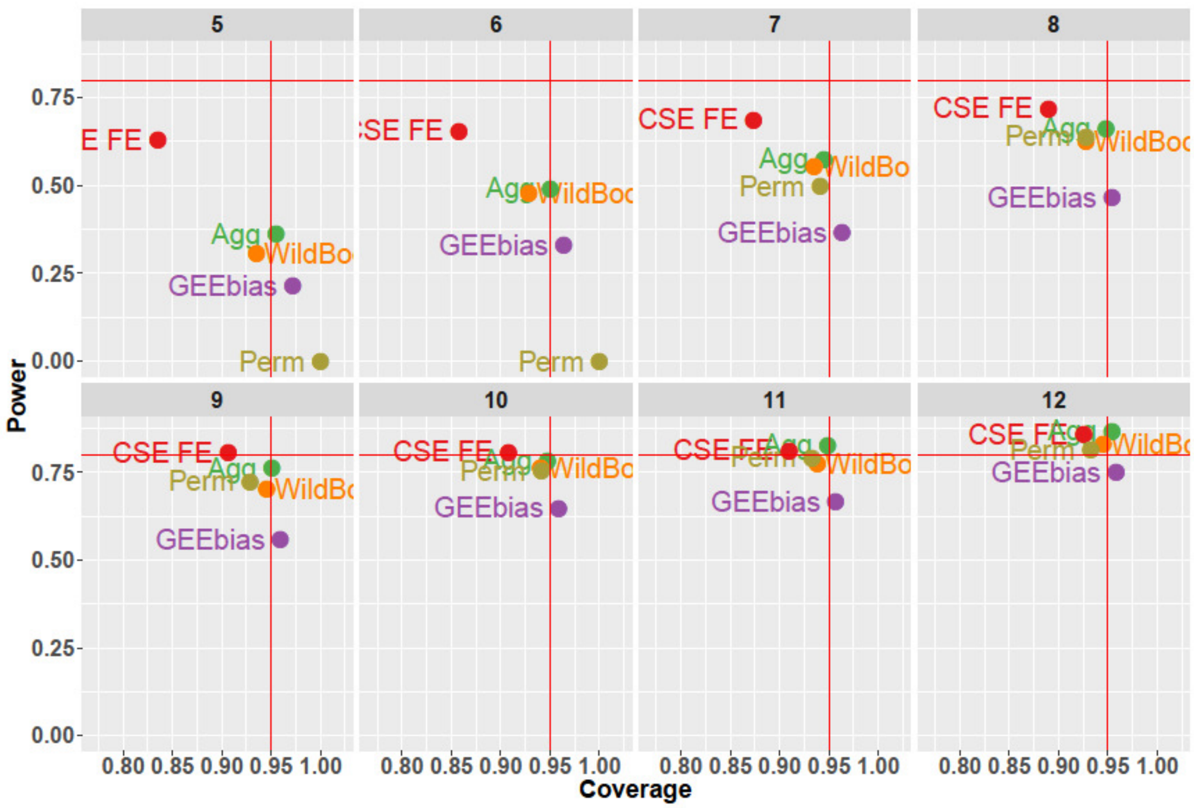


Figure 3.3: Coverage versus power in unbalanced cluster size as number of groups increases from 5 to 12.

Notes: Number of time points for each individual is 20. "CSE FE" is clustered standard errors with individual fixed effects, "Agg" is aggregation, "WildBoot" is wild cluster bootstrap, "GEEbias" is the GEE with bias adjustment and using the F-distribution, "Perm" is permutation.

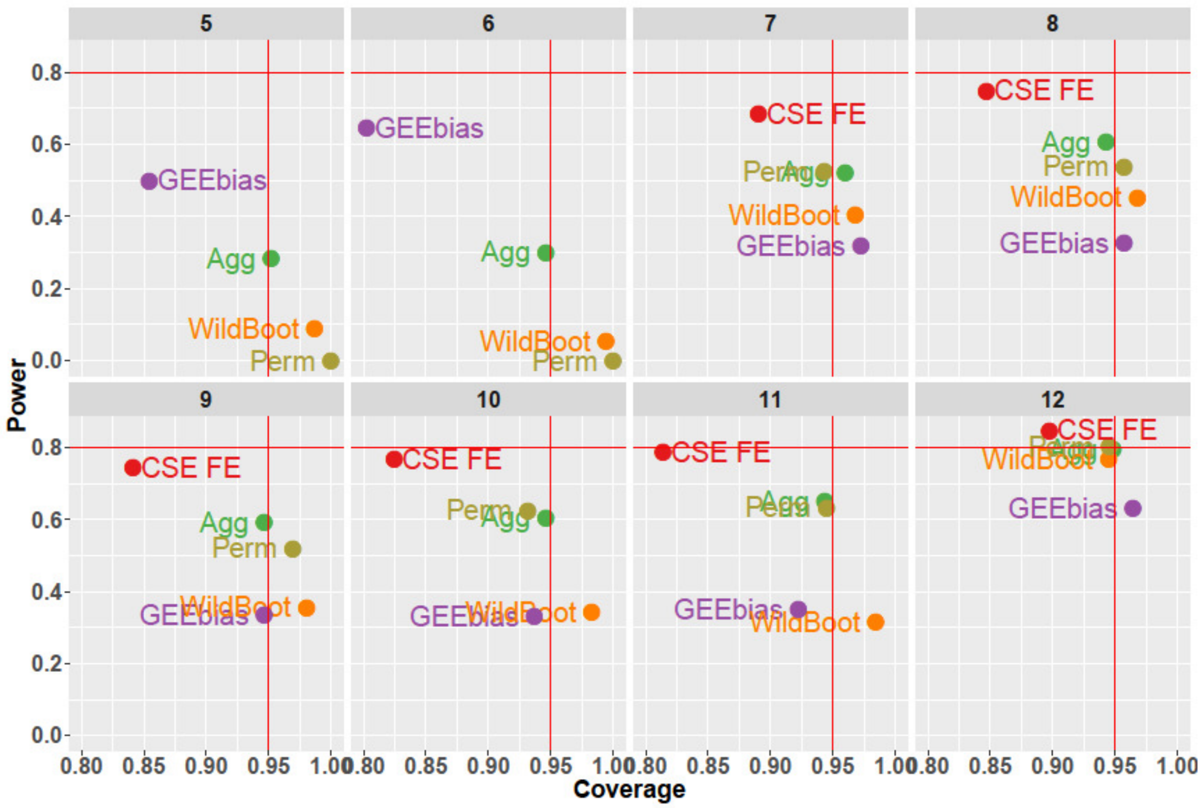


Figure 3.4: Coverage versus power in low proportion of treated groups scenario as number of groups increases from 5 to 12.

Notes: Number of time points for each individual is 20. "CSE FE" is clustered standard errors with individual fixed effects, "Agg" is aggregation, "WildBoot" is wild cluster bootstrap, "GEEbias" is the GEE with bias adjustment and using the F-distribution, "Perm" is permutation. For G=5 and G=6, CSE FEcoverage is below 0.8.

Replication results

To demonstrate the importance of these results, we conduct a replication of a recently published paper [27]. This paper exploits variation in state policy on over-the-counter (OTC) access to emergency birth control (EBC) to estimate the impact on sexual behaviour and finds that risky sexual behavior such as engaging in unprotected sex and number of sexual encounters increases as a result of OTC access to EBC.

The NLSY97 is designed to represent U.S. residents in 1997 that were born during the years 1980 through 1984, and who were between the ages of 12 and 16. The original sample includes 8,984 respondents. Respondents were re-interviewed annually until 2011 and biannually thereafter for a maximum of 17 periods between 1997 and 2013. The author restricts observations to females between the years 1997 to 2009.

The data are highly imbalanced: the number of observations per individual range from 1 to 13, with a median of 10. The number of individuals per state range from 4 to 662, with a median of 86. The author estimates linear probability models and estimates CSE after controlling for state and year fixed effects and individual-level covariates. The treatment variable in the estimation is binary; it is 1 when a state offered OTC EBC and 0 otherwise. Until the FDA ruled in favour of a national policy, only 9 states between 1998 and 2006 implemented such a policy. This produces a proportion of treatment groups of 9/50, or about 0.18.

Our aim was to replicate the results of the paper and compare them with the results from the models that provide adequate coverage. After replication of the point estimates and standard errors using CSE, we run the wild cluster bootstrap, permutation tests, and aggregation on the same data (since the GEE models were underpowered compared to these methods we do not estimate those results). Descriptive statistics are in shown in Appendix Table C.2. Regression results are in Table 3.2.

Table 3.2: Results of Mulligan analysis using various methods of accounting for correlation

	Mulligan (2015) Table IX, col 1		Replication OLS				Replication Aggregation	
	Est	p-value	Est	CSE p-value	Wild Boot p-value	Permute p-value	Est	p-value
Probability of having sex ever	0.02	.101	0.02	0.218	0.68	0.59	0.03	0.37
Probability of having sex in past 12 months	0.034**	.01	0.04	0.01**	0.36	0.61	0.04	0.11
Number of sexual encounters in past 12 months	13.8	.118	12.66	0.236	0.60	0.46	9.2	0.55
Number times used condom past 12 months	-5.43*	.101	-4.18	0.215	0.15	0.17	-9.2	0.07*
Probability of having risky sex	0.6***	<0.01	0.001	0.049**	0.24	0.44	-0.001	0.54

Notes: P-value in column 2 calculated based on reported point estimate and standard error. Risky sex defined in text as sex with a stranger or with an IV user (though most likely coded as both sex with a stranger and sex with an IV drug user, see Appendix Table C.2 for more details). Point estimates for probability of having sex ever, having sex in the past 12 months, and having risky sex indicate percentage point increases. *p<.1, **p<.05, ***p<.01

Most of the paper's results were replicable, though not perfectly. We could not replicate the point estimate for *had risky sex* that appears questionable considering its magnitude (the author's results would indicate that probability of risky sex (defined as sex with a stranger or IV drug user) increased 60 percentage points from a mean of 0.003); however, we were still able to find a significant result for this outcome using a cluster-robust variance matrix though with a much smaller point estimate.

Estimates for *sex in the past year* and *had risky sex*, which were significant under CSE, were no longer significant under models with correct coverage. None of the wild cluster bootstrap, permutation, or aggregation methods detected significant effects at the 0.05 level for these variables, and most p-values from these methods were quite large. These results indicate that the use of CSE to account for correlation in the data may have resulted in spurious results, leading to the inaccurate conclusion that OTC EBC led to women engaging in more and riskier sex, when in fact, it had no discernible impact on women's sexual behaviour or the data lacked power to detect policy-relevant effects.

3.4 Discussion

In this analysis, we compared the coverage of models in DID estimation with grouped data. We found that aggregation and GEE with bias adjustment provided accurate coverage under a wide variety of scenarios that modified the degree of correlation, number of time points, balance of data with respect to cluster sizes, and proportion of treated groups. The cluster-robust variance matrix estimate, even with the inclusion of individual fixed effects, was biased downwards when number of groups was less than 20, particularly in unbalanced data and in low proportion of treated groups, leading to severe over-rejection. GEE with bias adjustment, aggregation, and permutation tests performed adequately under the full range of scenarios. When G<9, the wild cluster bootstrap mildly over-rejected in scenarios with unbalanced cluster sizes and under-rejected when proportion of treated clusters was low. All methods with adequate coverage had low power to detect effects when G<10. The GEE with bias adjustment had consistently lower power to detect significant effects than the other methods.

This analysis has some limitations. First, our data generating process is unique and results may be specific to the process. For example, we could generate individual random effects that also follow an AR(1) process similar to the state random effects, which would induce serial correlation at the individual level. However, similarity of our results to other articles that use different data generating processes and different types of data give us confidence that our results are applicable to other settings [4,9,11,12,14]. Second, we do not explore models for binary data. However, again our results are comparable to those of other articles. Mancl & DeRouen (2001) compare the default robust standard errors of GEE with a logit link to bias corrected ones and calculate coverage probabilities to be 0.87 and 0.94 for GEE with robust variance and GEE with bias correction, respectively. We find those coverage probabilities to be 0.89 and 0.96, respectively (for similarly correlated scenarios). These results suggest that our findings on GEE models may be useful for researchers working with binary data; however, care must be taken to use the appropriate model for such data.

These results have important implications for epidemiologic research. First, when number of groups is less than 20, CSE should not be applied in DID estimation unless data are balanced and have a sufficient proportion of treated clusters. Reviews of articles that include small sample clustering should request that authors use appropriate methods, or at minimum compare their findings to either aggregation, permutation tests, GEE with bias adjustment, or the wild cluster bootstrap. Second, though the GEE with bias adjustment provides accurate coverage, it has particularly low power in DID estimation in small samples; researchers may consider permutation or aggregation as alternative methods. Third, since randomized controlled trials are increasingly analysed using DID, researchers can maximize power and avoid low coverage by designing cluster-randomized trials with equally sized clusters.

Lastly, these findings also have important implications for public policy. Understanding the extent of bias in DID estimation as well as how to properly adjust for correlated data is important for policy analysis and impact evaluations. Evaluations of policies that find a positive (or damaging) effect of a policy when one doesn't exist could promote poor public policymaking by potentially wasting funding on ineffective programs or cutting effective ones.

3.5 References

1. Bertrand M, Duflo E, Mullainathan S. How Much Should We Trust Differences-In-Differences Estimates? The Quarterly Journal of Economics. 2004;119: 249–275. doi:10.1162/003355304772839588

2. Liang K-Y, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika. 1986;73: 13–22. doi:10.1093/biomet/73.1.13

3. Cameron AC, Miller DL. A Practitioner's Guide to Cluster-Robust Inference. J Human Resources. 2015;50: 317–372. doi:10.3368/jhr.50.2.317

4. Donald SG, Lang K. Inference with Difference-in-Differences and Other Panel Data. Review of Economics and Statistics. 2007;89: 221–233. doi:10.1162/rest.89.2.221

5. McCaffrey DF, Bell RM. Improved hypothesis testing for coefficients in generalized estimating equations with small samples of clusters. Statist Med. 2006;25: 4081–4098. doi:10.1002/sim.2502

6. Carter AV, Schnepel KT, Steigerwald DG. Asymptotic Behavior of at Test Robust to Cluster Heterogeneity [Internet]. 2013. Available: http://works.bepress.com/douglas_steigerwald/26/?REMOTE_ADDR=66.249.74.195&

7. Imbens GW, Kolesar M. Robust Standard Errors in Small Samples: Some Practical Advice [Internet]. National Bureau of Economic Research; 2012 Oct. Report No.: 18478. Available: http://www.nber.org/papers/w18478

8. MacKinnon JG, Webb MD. Wild bootstrap inference for wildly different cluster sizes [Internet]. Queen's Economics Department Working Paper; 2014. Available: http://www.econstor.eu/handle/10419/97471

9. Conley TG, Taber CR. Inference with "difference in differences" with a small number of policy changes. The Review of Economics and Statistics. 2011;93: 113–125.

10. Morel J g., Bokossa M c., Neerchal N k. Small Sample Correction for the Variance of GEE Estimators. Biom J. 2003;45: 395–409. doi:10.1002/bimj.200390021

11. Mancl LA, DeRouen TA. A Covariance Estimator for GEE with Improved Small-Sample Properties. Biometrics. 2001;57: 126–134.

12. Fay MP, Graubard BI. Small-Sample Adjustments for Wald-Type Tests Using Sandwich Estimators. Biometrics. 2001;57: 1198–1206. doi:10.1111/j.0006-341X.2001.01198.x

13. Pan W, Wall MM. Small-sample adjustments in using the sandwich variance estimator in generalized estimating equations. Statist Med. 2002;21: 1429–1441. doi:10.1002/sim.1142

14. Cameron AC, Gelbach JB, Miller DL. Bootstrap-Based Improvements for Inference with Clustered Errors. Review of Economics and Statistics. 2008;90: 414–427. doi:10.1162/rest.90.3.414

15. Webb MD. Reworking Wild Bootstrap Based Inference for Clustered Errors [Internet]. Queen's Economics Department Working Paper; 2013. Report No.: 1315. Available: http://www.econstor.eu/handle/10419/97480

16. Fisher RA. The design of experiments. 1935. Oliver and Boyd, Edinburgh. 1935;

17. Rosenbaum PR. Observational studies [Internet]. Springer; 2002. Available: http://link.springer.com.ezp-prod1.hul.harvard.edu/chapter/10.1007/978-1-4757-3692-2_1

18. Ernst MD. Permutation methods: a basis for exact inference. Statistical Science. 2004;19: 676–685.

19. White H. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. Econometrica. 1980;48: 817–838. doi:10.2307/1912934

20. Cohen J, Dupas P. Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment. The Quarterly Journal of Economics. 2010;125: 1–45.

21. Bloom E, Bhushan I, Clingingsmith D, Hong R, King E, Kremer M, et al. Contracting for health: evidence from Cambodia [Internet]. Brookings Institution; 2006. Available: http://www.webprodserv.brookings.edu/~/media/Files/rc/papers/2006/07healthcare_kremer/20060720cambodia.pdf

22. Ho DE, Imai K. Randomization inference with natural experiments: An analysis of ballot effects in the 2003 California recall election. Journal of the American Statistical Association. 2006;101: 888–900.

23. Ryan AM, Burgess JF, Dimick JB. Why We Should Not Be Indifferent to Specification Choices for Difference-in-Differences. Health services research. 2014; Available: http://onlinelibrary.wiley.com.ezp-prod1.hul.harvard.edu/doi/10.1111/1475-6773.12270/epdf

24. Bryan ML, Jenkins SP. Multilevel Modelling of Country Effects: A Cautionary Tale. Eur Sociol Rev. 2015; jcv059. doi:10.1093/esr/jcv059

25. Bell RM, McCaffrey DF. Bias reduction in standard errors for linear regression with multi-stage samples. Survey Methodology. 2002;28: 169–182.

26. Gunsolley JC, Getchell C, Chinchilli VM. Small sample characteristics of generalized estimating equations. Communications in Statistics - Simulation and Computation. 1995;24: 869–878. doi:10.1080/03610919508813280

27. Mulligan K. Access to emergency contraception and its impact on fertility and sexual behavior. Health economics. 2015; Available: http://onlinelibrary.wiley.com.ezp-prod1.hul.harvard.edu/doi/10.1002/hec.3163/abstract

A. Appendix for Chapter 1

Table A.1: Content of SMS for Interactive, Unidirectional, and Control groups

Week	Interactive Group			Unidirectional Group Fact/Tip text	Control Group Fact text
	Quiz Question/Tip text	**Correct Answer**	**Response from SMART**		
1	SMART quiz:How many ovaries does a woman have? Reply SMT1 for 1 ovary or SMT2 for 2 ovaries. Reply to this number for free. Reply until you receive confirmation	SMT2	SMART:Right! A woman has 2 ovaries. This is where eggs are stored. She has a womb (uterus) where a fertilized egg implants and a pregnancy grows.Two fallopian tubes connect ovaries to the womb.The cervix connects the womb to the vagina. The vagina is a tube of muscle connecting cervix to outside of body	SMART fact: A woman has 2 ovaries. This is where eggs are stored. She has a womb (uterus) where a fertilized egg implants and a pregnancy grows.Two fallopian tubes connect ovaries to the womb.The cervix connects the womb to the vagina. The vagina is a tube of muscle connecting cervix to outside of body	SMART fact: In 2012, malaria killed over 483000 children under 5 years, or about 1 child every minute. Malaria kills over 45000 adolescents per year in Africa.
2	SMART quiz:When is the most likely time that a girl can get pregnant? Reply SMT1 for days 1-7 of her menses, reply SMT2 for days 8-19, or SMT3 for days 20-28.	SMT2	SMART answer: Correct! The menstrual cycle is usually 28 days. If day 1 is the first day of your menses, then days 8-19 are the most likely time that you can get pregnant. The egg is released from the ovaries between days 8-19. If sperms are present, then the egg may be fertilized, causing pregnancy.	SMART fact: The menstrual cycle is usually 28 days. If day 1 is the first day of your menses, then days 8-19 are the most likely time that you can get pregnant. The egg is released from the ovaries between days 8-19. If sperms are present, then the egg may be fertilized, causing pregnancy.	SMART fact:Malaria is caused by Plasmodium falciparum parasites.The only way the parasites are spread to people are thru bites of infected Anopheles mosquitoes.

3	SMART quiz: True or False: Standing up during sex can prevent a girl from getting pregnant. Reply SMT1 for true or SMT2 for false.	SMT2	SMART answer: Correct! Standing up during sex does NOT prevent pregnancy. When a man ejaculates (releases sperm), the sperms are deposited deep into the vagina immediately after ejaculation, allowing fertilization to take place. Bathing/washing will NOT prevent pregnancy either.	SMART fact: Standing up during sex does NOT prevent pregnancy. When a man ejaculates (releases sperm), the sperms are deposited deep into the vagina immediately after ejaculation, allowing fertilization to take place. Bathing/washing will NOT prevent pregnancy either.	SMART fact:Getting malaria while pregnant is very serious. About 9% of pregnant women in Ghana die of malaria. It can also result in low birth weight babies.
Tip 1: End of week 3	SMART tip: If you have any questions about your health, you can call 0302208585 or 080028585 (Toll free- Voda only) to speak to a nurse. It is confidential.			SMART tip: If you have any questions about your health, you can call 0302208585 or 080028585 (Toll free- Voda only) to speak to a nurse. It is confidential.	
4	SMART:Can you be a carrier of a Sexually Transmitted Infection (STI) and NOT be aware that you have it? Reply SMT1 for yes or SMT2 for no.	SMT1	SMART:Right!You can have STI without having any symptoms or knowing you are a carrier.It can take months to see symptoms like sores, itches and problems urinating.A partner may have a STI and it may be impossible for him or you to know that he has it.Condoms or abstinence are effective ways to prevent STI	SMART fact:You can be a carrier of a sexually transmitted infection (STI) without having any symptoms or knowing you are a carrier. It can take months to see symptoms like sores, itches and problems urinating. A partner may have a STI and it may be impossible for him or you to know that he has it.	SMART fact:The first symptoms of malaria are fever, headache, and chills. These occur 2-3 days after the mosquito bite.Other symptoms are body pain and nausea.

5	SMART quiz: True or False: A woman with an untreated gonorrhea may have severe lower abdominal pains. Reply SMT1 for true or SMT2 for false.	SMT1	SMART:Right! Untreated gonorrhea may lead to severe pains in lower abdomen called pelvic inflammatory disease. It can cause infertility.It also makes it easier to get HIV. It may take months to see signs of gonorrhea in females. In males it takes days. Its important to seek treatment from a health center.	SMART fact: Untreated gonorrhea may lead to severe pains in lower abdomen called pelvic inflammatory disease. It can cause infertility.It also makes it easier to get HIV. It may take months to see signs of gonorrhea in females. In males it takes days. Its important to seek treatment from a health center.	SMART:Malaria symptoms resemble diseases like pneumonia or typhoid.At health centers you can get rapid diagnostic test (just a few min) to identify the disease.
Tip 2: End of week 5	SMART Tip: Talking about reproductive health with friends, family, and a boyfriend/future boyfriend is smart. It can help you to be healthier and make good choices that are right for you. Be sure to talk to your friends about the SMART messages, and encourage them to participate! Win together!			SMART Tip: Talking about reproductive health with friends, family, and a boyfriend/future boyfriend is smart. It can help you to be healthier and make good choices that are right for you. Be sure to talk to your friends about the SMART messages and ask them about their opinions!	
6	SMART quiz: True or false: A woman can wear the female condom for up to 8 hours before she has sex. Reply SMT1 for true or SMT2 for false.	SMT1	SMART:Right! The female condom is made of a thin transparent and soft plastic that looks like a tube that is closed at one end.It is designed to fit into a woman's vagina. It can be	SMART fact: The female condom is made of a thin transparent and soft plastic that looks like a tube that is closed at one end.It is designed to fit into a woman's vagina. It can be worn up to 8 hours before a	SMART fact: You can cure malaria with drugs called ACTs like Artesunate-Amodiaquine. ACTs combine two drugs together into each pill.

			worn up to 8 hours before a woman has sex.It protects against both STIs and pregnancy.It is 95% effective if worn correctly.	woman has sex.It protects against both STIs and pregnancy.It is 95% effective if worn correctly.	They are 97% effective.
Tip 3: End of week 6	SMART Tip: Great job! Remember, if you don’t want to have sex, it’s ok to say no. Call 0302208585 or 080028585 (Toll free- Vodafone only) to speak to a nurse about strategies for saying no. It is completely confidential. You could also call this number if you have any questions bothering you.			SMART Tip: Great job! Remember, if you don’t want to have sex, it’s ok to say no. Call 0302208585 or 080028585 (Toll free- Vodafone only) to speak to a nurse about strategies for saying no. It is completely confidential. You could also call this number if you have any questions bothering you.	
7	SMART:When putting on a condom, should a man unroll it all the way first before putting it on the penis? Reply SMT1 for yes or SMT2 for no.	SMT2	SMART: Right! When putting on a condom, do NOT unroll the entire condom first. Open the package, hold the tip of the condom with one hand and roll it down the penis with the other hand. Leave space at the tip to collect semen. If there is no space at the tip the condom will burst open during ejaculation.	SMART: When putting on a condom, a man should NOT unroll the entire condom first.Open the package, hold the tip of the condom with one hand and roll it down the penis with the other hand. Leave space at the tip to collect semen.If there is no space at the tip the condom will burst open during ejaculation.	SMART fact:The malaria parasite has developed resistance to previous drugs like chloroquine. This means the drug no longer works to cure malaria.Only ACTs cure.

8	SMART:When using a condom, when should a man pull out of the vagina after ejaculation? Reply SMT1 for while penis is still stiff or SMT2 for when penis is soft.	SMT1	SMART answer: Right! When using a condom, it is important for the man to pull his penis out right after ejaculation, while it is still stiff. If the penis gets soft then the condom could fall off inside the woman's vagina. If this happens then it is possible that the woman will get pregnant.	SMART fact: When using a condom, it is important for the man to pull his penis out right after ejaculation, while it is still stiff. If the penis gets soft then the condom could fall off inside the woman's vagina. If this happens then it is possible that the woman will get pregnant.	SMART fact: If you take an ACT and don't finish all the pills, the malaria parasite will survive. This builds resistance to the medicine. Always finish ACTs.
Tip 4: End of Week 8	SMART Tip: Contraception means a method to prevent pregnancy.Birth control pills and condoms are types of contraception.Condoms are only effective if you use them correctly and use them every time you have sex. Then they are 98% effective against STDs and pregnancy.Condoms do NOT cause infertility in men.	NA	SMART Tip: Contraception means a method to prevent pregnancy.Birth control pills and condoms are types of contraception.Condoms are only effective if you use them correctly and use them every time you have sex. Then they are 98% effective against STDs and pregnancy.Condoms do NOT cause infertility in men.	SMART Tip: Contraception means a method to prevent pregnancy.Birth control pills and condoms are types of contraception.Condoms are only effective if you use them correctly and use them every time you have sex. Then they are 98% effective against STDs and pregnancy.Condoms do NOT cause infertility in men.	
9	SMART quiz:How often is the Pill taken (the birth control Pill)? Reply SMT1 for only after a woman has sex or reply	SMT2	SMART answer: Right! The Pill is taken once a day whether or not a woman has sex.If you choose to use the Pill as your contraceptive	SMART: The birth control Pill is taken once a day whether or not a woman has sex.If you choose to use the Pill as your contraceptive method then you	SMART fact: There are no vaccines against malaria. You can prevent malaria with treated mosquito

	SMT2 for once a day, everyday.		method then you must take it everyday or it is NOT effective. You can't just take it whenever you please! It contains low and safe doses of hormones and prevents pregnancy.	must take it everyday or it is NOT effective. You can't just take it whenever you please! It contains low and safe doses of hormones and prevents pregnancy.	nets.Traditional medicines are not effective in curing malaria.
10	SMART quiz: True or False: Birth control pills are effective even if a woman misses taking them for 2-3 days in a row. Reply SMT1 for true or SMT2 for false.	SMT2	SMART answer: Right! The Pill is NOT effective if a woman misses it for 2 or 3 days in a row. The Pill must be taken everyday and if a woman stops taking it then she may get pregnant after 2-3 days. It does NOT take 6 months to become pregnant after stopping birth control.	SMART fact:The Pill is NOT effective if a woman misses it for 2 or 3 days in a row. The Pill must be taken everyday and if a woman stops taking it then she may get pregnant after 2-3 days. It does NOT take 6 months to become pregnant after stopping birth control.	SMART fact:Children who survive episodes of severe malaria may develop learning problems, brain damage, or anemia (low iron in body which affects their growth).
11	SMART:True or False:A woman should take a rest from the Pill every year because the pills build up in the body over time.Reply SMT1 for true or SMT2 for false.	SMT2	SMART answer: Right! The Pill does NOT build up in the body so women do NOT need to take a rest from the Pill. If a woman has side effects like nausea, switching to another type or brand might help. The Pill protects against pregnancy but not STIs. The Pill does not cause infertility later in life.	SMART fact: The Pill does NOT build up in the body over time so women do NOT need to take a rest from the Pill. If a woman has side effects like nausea, switching to another type or brand might help. The Pill protects against pregnancy but not STIs. The Pill does not cause infertility later in life.	SMART fact: Common myths about how malaria is spread are that you can get infected from working too much in the sun or eating hot foods. These are NOT true.

12	SMART quiz: True or False. Emergency contraception must be taken within 1 hour of unprotected sex. Reply SMT1 for true, and SMT2 for false.	SMT2	SMART: Right! Emergency contraception (like Postinor-2) is a method to reduce chance of pregnancy after unprotected sex or when a condom breaks. The 2 pills must be taken within 5 DAYS of unprotected sex (that's 120 hours). It should only be used for emergencies, not as a regular method of contraception.	SMART fact: Emergency contraception (like Postinor-2) is a method to reduce chance of pregnancy after unprotected sex or when a condom breaks. The 2 pills must be taken within 5 DAYS of unprotected sex. It should only be used for emergencies, not as a regular method of contraception.	SMART fact:Increased prevention of malaria with nets and treatment with ACTs have led to more than 3million lives saved since 2010, mostly children under 5 yrs.

Table A.2: Knowledge Quiz and percent of participants responding correctly at baseline

Item	**% responding correctly at baseline**
Standing up during sex can help prevent pregnancy.	29 %
Condoms cause infertility in men.	37 %
To put on a condom, you should first unroll it all the way and then try to put it on the penis.	7 %
When putting on a condom, it is important to leave space at the tip.	28 %
When using a condom, it is important for the man to pull his penis out right after ejaculation, while it is still stiff.	18 %
Birth control pills (known as The Pill) are taken once every day, whether or not you have sex.	21 %
Birth control pills protect against sexually transmitted infections.	46 %
Birth control pills are effective even if a woman misses taking them for two or three days in a row.	17 %
It is important that women should "take a rest" from the pill every year because the pills build up in a woman's body over time.	7 %
If a woman is having side effects with one kind of pill, switching to another type or brand might help.	15 %
After a woman stops taking birth control pills, she is unable to get pregnant for at least six months.	19 %
The female condom can be worn up to 8 hours before having sex.	7 %
Emergency contraception must be taken within 1 hour of having unprotected sex.	8 %
Symptoms of gonorrhea in females will appear the day after becoming infected.	33 %
Gonorrhea infection makes it easier to get HIV and other STIs and pass them to sex partners.	52 %
If left untreated, sexually transmitted infections like gonorrhea can cause infertility in both men and women.	63 %
A woman with an untreated gonorrhea may have severe lower abdominal pains.	50 %
If day 1 is the first day of a woman's period, she has the greatest chance of becoming pregnant during days 8-19.	47 %
You can have a sexually transmitted infection without having any symptoms or knowing you are a carrier.	44 %
Every woman has 1 ovary where her eggs are stored.	30 %
STI symptoms can include sores, itches, and problems urinating.	Only asked at follow-up
Postinor-2 is a type of emergency contraception.	Only asked at follow-up
The female condom protects against both sexually transmitted infections and pregnancy.	Only asked at follow-up
Washing/bathing oneself after sex can prevent pregnancy.	Only asked at follow-up

Notes: Response choices for each item were "True", "False", and "Don't know". An incorrect answer, a "don't know", and a missing answer were counted as incorrect.

Table A.3. List of outcome variables and regression model

Outcome	Time Measured	Model
Full sample		
Knowledge of reproductive health	0, 3, 15 months	Linear
Ever had sex	15 months	Logit
Had sex in the past year	15 months	Logit
Pregnancy in the past year	15 months	Logit
Attitudes about reproductive health[a]	0, 3, 15 months	Logit
Subgroup who reported having sex in the past year		
Pregnancy in the past year	15 months	Logit
Used any contraception in past year	15 months	Logit
Used contraception last time had sex	15 months	Logit
Used condom at sexual debut	15 months	Logit
Had sex without a condom in past year	15 months	Logit
Used a condom in the past year	15 months	Logit
Used birth control pill in past year	15 months	Logit
Used emergency contraception in past year	15 months	Logit
Subgroup who reported ever having sex		
Age at sexual debut	15 months	Linear

Notes: [a]Secondary outcome

Table A.4: Attitude items at each follow-up and baseline proportion agreeing or strongly agreeing with each item

Item	Asked at 3-month Follow-up	Asked at 15-month Follow-up	% Participants agreeing or strongly agreeing at baseline
I know about the signs and symptoms of STDs.	x		74%
I know how to use a condom correctly.	x	x	20%
I know how to use the birth control Pill correctly.	x	x	13%
I am confident that I can use a condom every time I have sex.	x		33%
I could insist on using a condom during sex even if my boyfriend/girlfriend (or future boyfriend/girlfriend) does not want to use one.	x		53%
I am confident I could refuse to have sex if my boyfriend/girlfriend (or future boyfriend/girlfriend) does not want to use a condom.	x	x	62%
I would be embarrassed to buy condoms.	x		64%
It is too much of an inconvenience to use a condom every time you have sex.	x		40%
I would feel embarrassed to buy the birth control pill.	x		48%
Condoms are effective against sexually transmitted diseases.	x	x	56%
I would feel comfortable talking about avoiding or delaying sex with a boyfriend/girlfriend (or future boyfriend/girlfriend).	x		61%
I would be embarrassed to talk about using condoms with my boyfriend/girlfriend (or future boyfriend/girlfriend).	x		40%
I would be worried about getting an STI if I had sex without a condom at this time in my life.	x	x	69%
My friends think contraception should be used to prevent unwanted pregnancy.	x		62%
My friends think condoms should be used during sex before marriage.	x		55%

Table A.4 (Continued)

Item	Asked at 3-month Follow-up	Asked at 15-month Follow-up	% Participants agreeing or strongly agreeing at baseline
I feel comfortable talking to my friends about condoms and contraception.	x	x	49%
My friends would approve of me using contraception or condoms to avoid pregnancy.	x	x	53%
I feel comfortable talking to my parents about condoms and contraception.	x		31%

Table A.5: Estimated intervention effects for primary outcome

				Unidirectional – Control		Interactive – Control		Interactive – Unidirectional	
	Control mean	**Uni-directional** mean	**Inter-active** mean	**Crude Difference (95% CI)**	**Adjusted Difference (95% CI)**	**Crude Difference (95% CI)**	**Adjusted Difference (95% CI)**	**Crude Difference (95% CI)**	**Adjusted Difference (95% CI)**
Baseline	26% n=293	31% n=258	31% n=205	5 (-0.1 to 10)	4 (-1 to 10)	6 (-0.3 to 12)	5 (-1 to 11)	1 (-7 to 9)	0 (-7 to 8)
Follow-up – 3 months	32% n=286	45% n=238	60% n=192	14*** (7 to 21)	11*** (7 to 15)	27*** (21 to 33)	24*** (19 to 28)	13** (5 to 21)	13*** (8 to 18)
Follow-up – 15 months	42% n=277	47% n=247	56% n=197	6* (0.1 to 11)	3 (-1 to 7)	15*** (10 to 19)	11*** (8 to 15)	9** (3 to 15)	8*** (4 to 13)

Notes: Knowledge score is percentage correct of a 24-item index at follow-up (20-item index at baseline). Missing answer, "don't know' coded as incorrect answer. Crude model is adjusted for school category and presence of home economics class. Adjusted model is additionally adjusted for baseline knowledge, age, religion, ethnicity, mother completed at least secondary school, father completed at least secondary school, and school size. Clustered standard errors at school level in parentheses. *p<.05, **p<.01, ***p<.0001

Table A.6: Estimated intervention effects for attitudes measured at both 3 and at 15 months

	3-month Follow-up					**15-month Follow-up**				
	Control	**Unidirec -tional**	**Inter-active**	**Unidirec-Control Adj OR**	**Interactive-Control Adj OR**	**Control**	**Unidirec -tional**	**Inter-active**	**Unidirec – Control Adj OR**	**Interactive – Control Adj OR**
I know how to use condoms	55/286 (19%)	58/238 (24%)	53/192 (28%)	1.3 (.82–2.1)	1.7* (1.0–2.8)	50/277 (18%)	56/247 (23%)	56/197 (28%)	1.3 (0.82– 2.1)	1.6 (0.98– 2.7)
I know how to use birth control pills	31/286 (11%)	40/238 (17%)	37/192 (19%)	1.8 (.90–3.5)	2.0 (.98–4.4)	29/277 (10%)	36/247 (15%)	37/197 (19%)	1.3 (0.77– 2.3)	2.0* (1.1– 3.6)
Condoms effective against STI	150/286 (52%)	141/238 (59%)	129/192 (67%)	1.4 (0.86– 2.1)	1.6 (0.94– 2.5)	154/277 (56%)	139/247 (56%)	130/197 (66%)	0.98 (0.68–1.41)	1.4 (0.93– 2.1)
I would be worried about STI	195/286 (68%)	181/238 (76%)	152/192 (79%)	1.4 (0.88– 2.3)	1.5 (0.85– 2.6)	209/277 (75%)	185/247 (75%)	154/197 (78%)	0.97 (0.57– 1.6)	0.95 (0.52– 1.7)
I am comfortable talking with friends about contraception	152/286 (53%)	126/238 (53%)	121/192 (63%)	0.94 (0.65– 1.4)	1.3 (0.88–2.0)	149/277 (54%)	144/247 (58%)	115/197 58%)	1.2 (0.8– 1.7)	1.0 (0.66–1.5)
My friends would approve of me using contraception	131/286 (46%)	127/238 (53%)	110/192 (57%)	1.4 (1.0– 2.1)	1.3 (0.83– 1.9)	88/277 (32%)	85/247 (34%)	79/197 (40%)	1.1 (0.73– 1.6)	1.4 (0.93–2.2)
I am confident I could refuse sex	171/286 (60%)	154/238 (65%)	134/192 (70%)	1.2 (0.83– 1.8)	1.2 (0.81– 1.9)	180/277 (65%)	155/247 (63%)	122/197 (62%)	0.86 (0.59– 1.3)	0.73 (0.48–1.1)

Notes: Odds ratios from multilevel logistic regression model with school random effects. Model adjusted for age, religion, ethnicity, mother's education, father's education, baseline attitude, school size, presence of home economics class, and school category. *p<.05, **p<.01, ***p<.0001

Table A.7: Estimated intervention effects for attitudes measured only at 3 months

	Control	**Unidirec-tional**	**Inter-active**	**Unidirectional – Control Adj OR**	**Interactive – Control Adj OR**
I know signs of STIs	214/286 (75%)	174/238 (73%)	149/192 (78%)	0.96 (0.51–1.8)	1.1 (0.54–2.2)
I am confident I can use condoms	84/286 (29%)	75/238 (32%)	65/192 (34%)	1.2 (0.79–1.8)	1.2 (0.77–1.9)
I am embarrassed to talk about condoms with bf	104/286 (36%)	89/238 (37%)	83/192 (43%)	1.0 (0.70–1.5)	1.4 (0.92–2.1)
My friends think contraception should be used	167/286 (58%)	163/238 (68%)	144/192 (75%)	1.6* (1.1–2.4)	2.1** (1.3–3.4)
It is inconvenience to use condom every time	109/286 (38%)	95/238 (40%)	84/192 (44%)	1.2 (0.80–1.7)	1.4 (0.91–2.0)
I could insist on using condom even if bf didn't want	160/286 (56%)	133/238 (56%)	115/192 (60%)	0.98 (0.67–1.4)	0.93 (0.60–1.4)
My friends think condoms should be used	147/286 (51%)	126/116 (53%)	116/192 (60%)	1.1 (0.74–1.6)	1.2 (0.79–1.8)
Embarrassing to buy birth control pill	145/286 (51%)	134/238 (56%)	107/192 (56%)	1.4 (0.94–1.9)	1.1 (0.73–1.6)
I feel comfortable avoiding sex with bf	177/286 (62%)	146/238 (61%)	128/192 (67%)	0.91 (0.61–1.4)	0.98 (0.63–1.5)
I feel comfortable talking to parents about contraception	93/286 (33%)	63/238 (26%)	60/192 (31%)	0.77 (0.51–1.2)	0.83 (0.53–1.3)
Embarrassing to buy condoms	185/286 (65%)	160/238 (67%)	138/192 (72%)	1.2 (0.83–1.9)	1.3 (0.83–2.1)

Notes: Odds ratios from multilevel logistic regression model with school random effects. Model adjusted for age, religion, ethnicity, mother's education, father's education, baseline attitude, school size, presence of home economics class, and school category. *p<.05, **p<.01, ***p<.0001

Table A.8: Estimated intervention effects on communicating at least once a week with each contact

	Control	Unidirec-tional	Inter-active	Unidirectional – Control Adj. OR (95% CI)	Interactive – Control Adj. OR (95% CI)
3-month follow-up:					
Friends	103/286 (36%)	82/238 (34%)	92/192 (48%)	0.92 (0.61–1.4)	1.68* (1.1–2.6)
Professional	45/286 (17%)	26/238 (14%)	31/192 (16%)	0.58 (0.30–1.13)	0.84 (0.41–1.7)
Boyfriend	53/286 (19%)	47/238 (20%)	45/192 (23%)	1.2 (0.72–1.9)	1.4 (0.86–2.4)
Parents	50/286 (17%)	34/238 (20%)	37/192 (19%)	0.80 (0.48–1.3)	0.94 (0.55–1.6)
15-month follow-up:					
Friends	119/277 (43%)	107/247 (43%)	84/197 (43%)	0.85 (.59–1.2)	0.91 (.60–1.4)
Anyone	104/277 (38%)	100/247 (40%)	87/197 (44%)	1.0 (.71–1.5)	1.1 (.75–1.7)

Notes: Odds ratios from multilevel logistic regression model with school random effects. Crude model is adjusted for school category and presence of home economics class. Adjusted model is additionally adjusted for religion, ethnicity, mother's education, father's education, school size, and baseline communication. Missing values were replaced with overall median for each outcome, which was 0 in every case. *p<.05, **p<.01, ***p<.0001

B. Appendix for Chapter 2

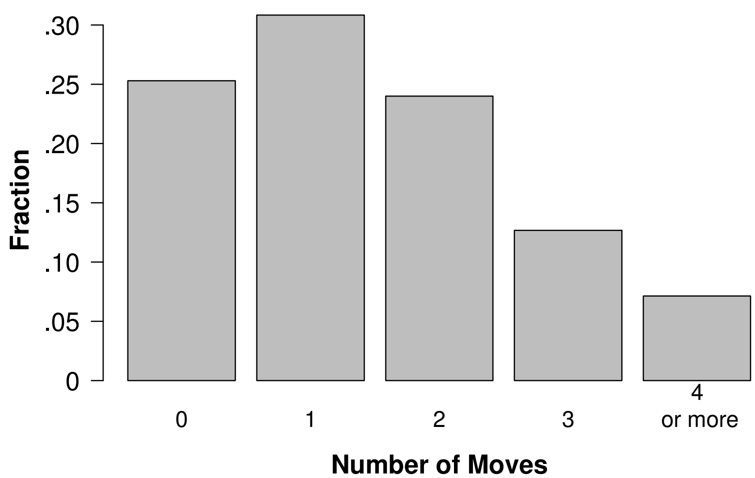


Figure B.1: Distribution of number of total lifetime moves among women in HAWS sample, $N = 1,488$

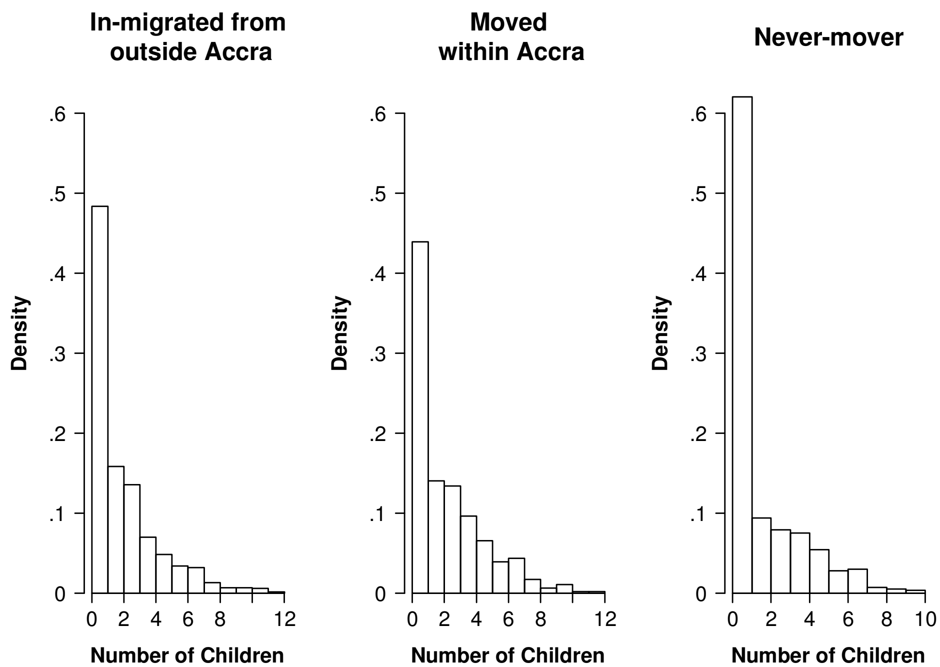


Figure B.2: Distribution of total number children alive by lifetime mover status

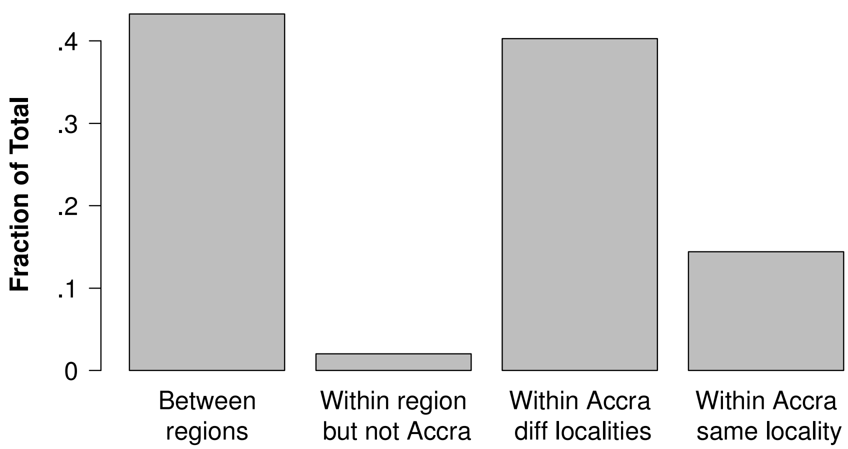
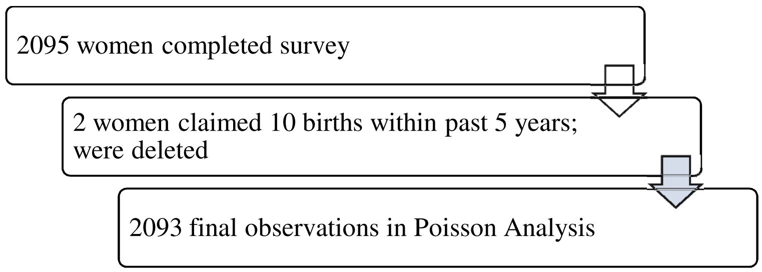


Figure B.3: Distribution of distance of moves identified in HAWS sample

Poisson Analysis Data Set:



Event History Analysis Data Set:

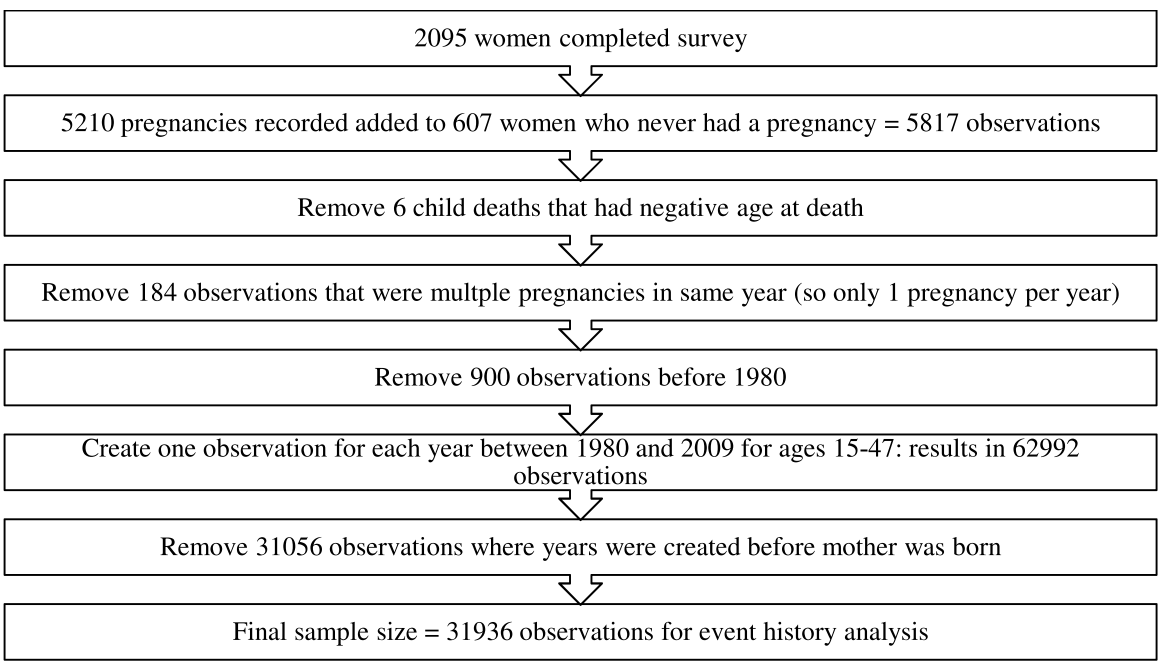


Figure B.4: Process for obtaining the final sample size for all analyses

Table B.1: Linear probability estimates for effect of residential duration on lost birth, miscarriage, and abortion outcomes compared with those who had never moved

	Lost Birth	Miscarriage	Abortion
Residence 0–24 Months	0.0146***	0.0107***	0.00390†
	(0.00348)	(0.00278)	(0.00210)
Residence 25–48 Months	0.00895*	0.00602*	0.00293
	(0.00350)	(0.00272)	(0.00221)
Residence 49–72 Months	0.00348	0.00475†	–0.00128
	(0.00339)	(0.00282)	(0.00195)
Residence >72 Months	0.00396†	0.00371*	0.000248
	(0.00226)	(0.00178)	(0.00143)
Age 25–29	0.00484	0.00671*	–0.00187
	(0.00334)	(0.00274)	(0.00187)
Age 30–40	–0.00186	0.000748	–0.00261
	(0.00298)	(0.00239)	(0.00171)
Age >40	–0.00305	0.00252	–0.00556***
	(0.00517)	(0.00494)	(0.00150)
At Least Middle School	0.00114	0.00224	–0.00110
	(0.00248)	(0.00168)	(0.00186)
At Least Middle × Age 25–29	0.00626	0.00223	0.00403
	(0.00479)	(0.00392)	(0.00265)
At Least Middle × Age 30–40	–0.000422	–0.00153	0.00111
	(0.00389)	(0.00308)	(0.00241)
At Least Middle × Age >40	–0.00486	–0.00597	0.00111
	(0.00578)	(0.00541)	(0.00209)
Previous Child Had Died	0.00238	0.00220	0.000178
	(0.00454)	(0.00400)	(0.00217)
Already Had Child	–0.00551†	–0.00554*	0.0000268
	(0.00293)	(0.00246)	(0.00161)
Married	0.00475†	0.00786***	–0.00310†
	(0.00283)	(0.00231)	(0.00161)

Table B.1 (Continued)

	Lost Birth	Miscarriage	Abortion
Married in Past Year	0.000908	0.00133	–0.000422
	(0.00391)	(0.00343)	(0.00192)
1985–1989	0.00140	–0.00195	0.00335†
	(0.00360)	(0.00295)	(0.00202)
1990–1994	0.00861*	0.00463	0.00398*
	(0.00376)	(0.00312)	(0.00196)
1995–1999	0.00420	0.00135	0.00286+
	(0.00353)	(0.00308)	(0.00159)
2000–2004	0.00773*	0.00279	0.00494**
	(0.00366)	(0.00315)	(0.00173)
2005–2009	0.00575	0.000259	0.00549**
	(0.00352)	(0.00302)	(0.00171)
Ethnicity: Ewe	0.00175	0.000697	0.00105
	(0.00311)	(0.00210)	(0.00247)
Ethnicity: Ga	0.00106	0.00482*	–0.00375**
	(0.00276)	(0.00232)	(0.00143)
Ethnicity: Other	–0.00516*	–0.000227	–0.00494**
	(0.00244)	(0.00177)	(0.00173)
Constant	0.00591	–0.000974	0.00688**
	(0.00416)	(0.00320)	(0.00259)
N	31,936	31,936	31,936

Notes: Clustered standard errors are shown in parentheses. Reference categories are never-movers, age <25, 1980–1984, and Akan ethnicity. $^{\dagger}p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table B.2: Logistic regression estimates for effect of residential duration on pregnancy outcome compared with those who had never moved

	Pregnancy	Live Birth	Lost Birth	Stillbirth
Residence 0–24 months	1.358***	1.199*	2.240***	2.131
	(0.0992)	(0.0926)	(0.424)	(1.050)
Residence 25–48 months	1.268**	1.188*	1.754**	1.412
	(0.0974)	(0.0956)	(0.366)	(0.789)
Residence 49–72 months	1.183*	1.143	1.291	2.042
	(0.0955)	(0.0982)	(0.309)	(1.041)
Residence >72 months	1.099	1.050	1.329†	1.827
	(0.0665)	(0.0659)	(0.229)	(0.795)
Age 25–29	0.934	0.883†	1.338	1.068
	(0.0601)	(0.0600)	(0.261)	(0.466)
Age 30–40	0.628***	0.598***	0.848	1.249
	(0.0462)	(0.0460)	(0.195)	(0.471)
Age >40	0.304***	0.251***	0.769	1.798
	(0.0425)	(0.0343)	(0.345)	(0.997)
At Least Middle School	0.718***	0.680***	1.073	0.793
	(0.0372)	(0.0366)	(0.175)	(0.277)
At Least Middle × Age 25–29	1.351***	1.322**	1.234	1.787
	(0.114)	(0.118)	(0.289)	(0.914)
At Least Middle × Age 30–40	1.231*	1.292**	1.020	0.796
	(0.116)	(0.125)	(0.289)	(0.413)
At Least Middle × Age >40	0.695	0.619†	0.688	0.733
	(0.164)	(0.155)	(0.380)	(0.551)
Previous Child Had Died	1.214	1.232†	1.148	0.836
	(0.144)	(0.152)	(0.300)	(0.487)
Already Had Child	0.936	0.977	0.721†	1.294
	(0.0524)	(0.0576)	(0.126)	(0.389)
Married	3.050***	3.461***	1.337†	1.699
	(0.216)	(0.267)	(0.229)	(0.677)

Table B.2 (Continued)

	Pregnancy	Live Birth	Lost Birth	Stillbirth
Married in Past Year	1.414***	1.460***	1.012	1.264
	(0.0913)	(0.0968)	(0.196)	(0.508)
1985–1989	0.916	0.907	1.094	0.724
	(0.0614)	(0.0623)	(0.318)	(0.316)
1990–1994	0.771***	0.700***	1.734*	0.962
	(0.0541)	(0.0502)	(0.468)	(0.425)
1995–1999	0.589***	0.548***	1.337	0.539
	(0.0400)	(0.0389)	(0.372)	(0.264)
2000–2004	0.615***	0.546***	1.658†	0.819
	(0.0409)	(0.0380)	(0.459)	(0.359)
2005–2009	0.438***	0.371***	1.478	0.833
	(0.0284)	(0.0257)	(0.403)	(0.353)
Ethnicity: Ewe	1.046	1.028	1.102	1.068
	(0.0618)	(0.0625)	(0.182)	(0.395)
Ethnicity: Ga	1.145**	1.166**	1.076	0.793
	(0.0598)	(0.0658)	(0.174)	(0.274)
Ethnicity: Other	0.890*	0.943	0.707*	0.631
	(0.0469)	(0.0516)	(0.114)	(0.227)
N	31,936	31,936	31,936	31,936

Notes: Odds ratios standard errors are shown in parentheses. Reference categories are never-movers, age <25, 1980–1984, and Akan ethnicity. †$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

C. Appendix for Chapter 3

Table C.1: Details of estimation methods

Method	Estimation
Cluster-robust SEs	$\hat{V}_{clu}[\hat{\beta}] = (X'X)^{-1}\hat{B}_{clu}(X'X)^{-1}$ $\hat{B}_{clu} = \sum_{g=1}^{G} X'_g \hat{u}_g \hat{u}'_g X_g$ $\hat{u}_g = y_g - X_g\hat{\beta}$ Finite sample adjustment: $\sqrt{c}\hat{u}_g$ instead of $\hat{u}_g$, where $c = \frac{G}{G-1}\frac{N-1}{N-k}$
Cluster Boot-strapping	1. Resamples data by cluster with replacement 2. Calculates $\hat{\beta}_b$ for each b^th sample for b=1,…B and B=400 3. $\hat{V}_{boot}[\hat{B}] = \frac{1}{B-1}\sum_{b=1}^{B}(\hat{B}_b - \hat{\bar{B}})(\hat{B}_b - \hat{\bar{B}})'$
Wild Cluster Bootstrapping	1. Re-estimate OLS subject to the restriction that $\beta = 0$. 2. Estimate the b^th resample by randomly assigning each cluster with the weight v_g where v_g is a random variable that takes on 1 with probability 0.5 and -1 with probability 0.5. 3. With the new residuals, generate a new y-vector, re-estimate OLS with the new y-vector, and calculate the Wald-statistic, w_b*. 4. Conduct this procedure B=400 times. 5. The p-value for the test is then the proportion $\lvert w\rvert > \lvert w_b^*\rvert$, b=1,…,B.
GEE	GEE robust variance matrix: $V_s = V_m\left(\sum_{i=1}^{K}\hat{U}_i\hat{U}_i^T\right)V_m$ Bias adjusted variance sandwich estimator (see Fay and Graubard 2001 for full proof): $V_a = V_m\left(\sum_{i=1}^{k}H_i\hat{U}_i\hat{U}_i^T H_i\right)V_m$ Where H_i is a pxp diagonal matrix with jj^th element equal to $\left\{1 - \min(b, \{\hat{\Omega}_i V_m\}_{jj})\right\}^{-1/2}$ Degree of freedom adjustment using $\tilde{d}_H$ (d3 and d5 options in *saws* function): $\hat{\psi}_i = H_i\hat{U}_i\hat{U}_i^T H_i^T$ $\tilde{d}_H = \frac{\{trace(\tilde{\psi}B_1\}^2}{trace(\tilde{\psi}B_1\tilde{\psi}B_1)}$

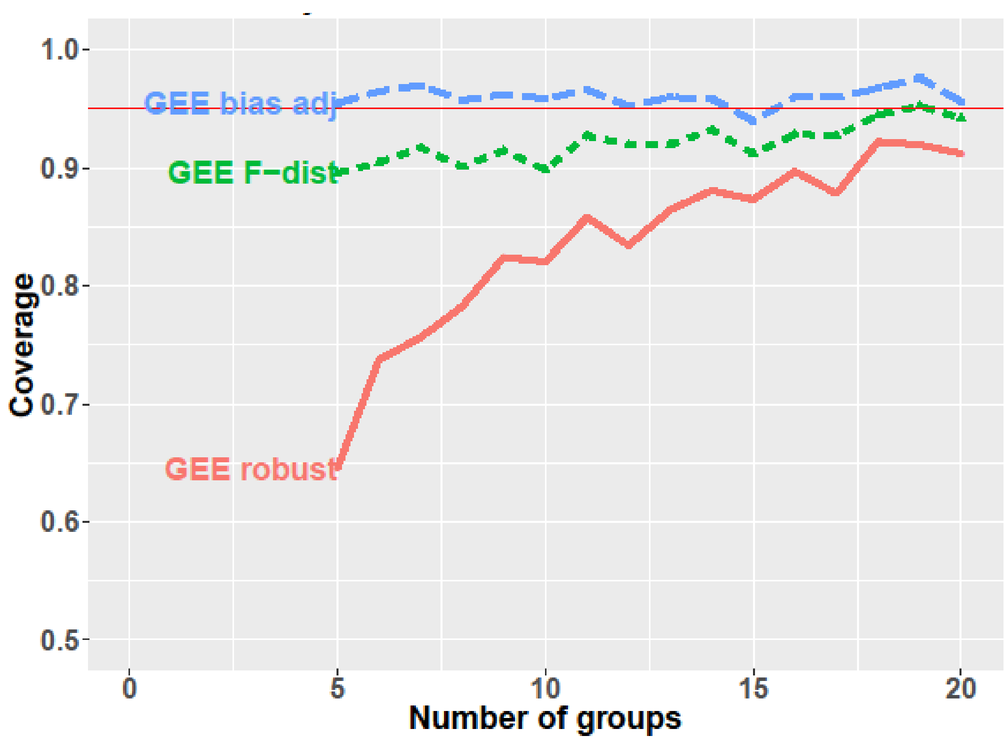


Figure C.1: Coverage for adjustments on GEE model for unbalanced cluster sizes

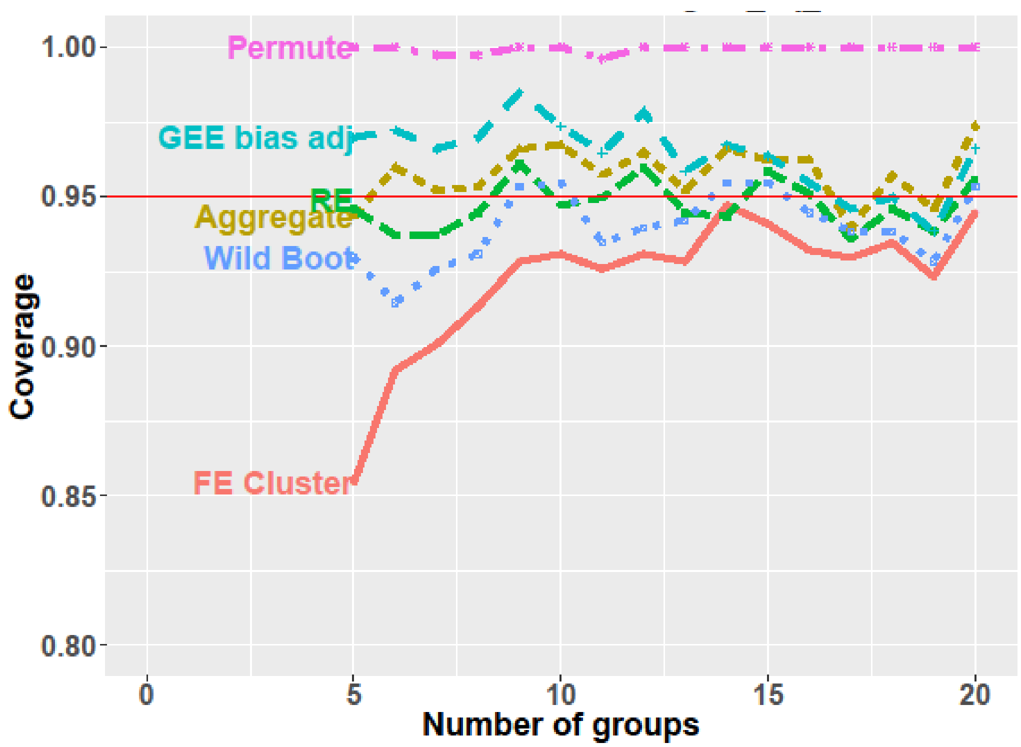


Figure C.2: Coverage when $\sigma_w^2 = 0$ with unbalanced cluster sizes

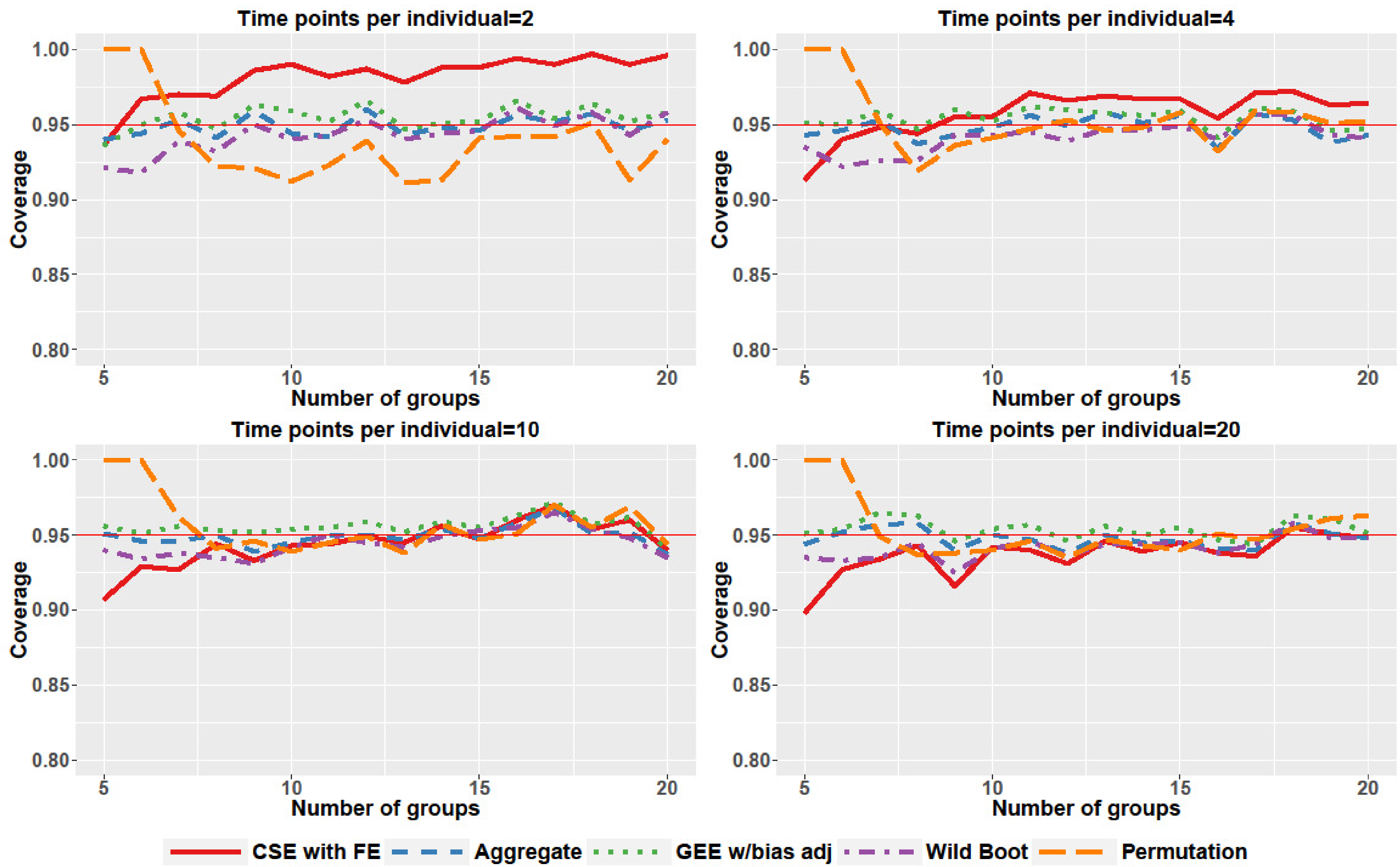


Figure C.3: Coverage in low correlation and balanced data scenario as number of time points per individual increases

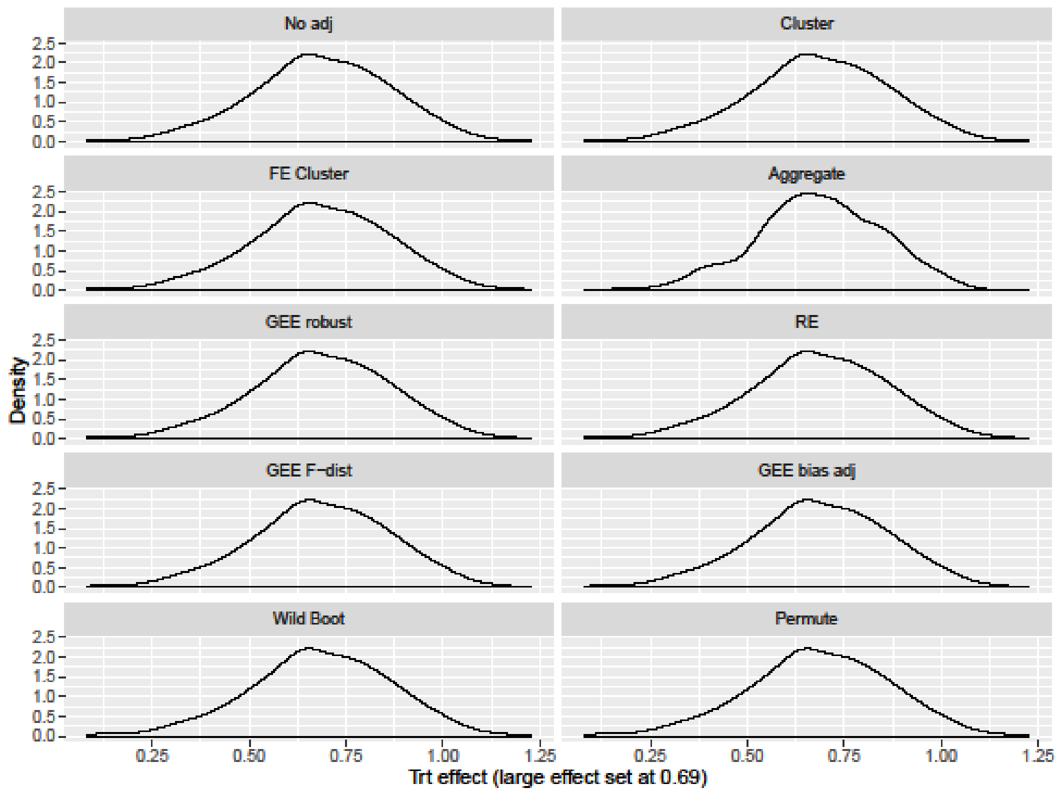


Figure C.4: Histogram of treatment effect for all models

Table C.2**:** Replication of Mulligan (2015) descriptive statistics

	Mulligan's Table III, panel 1			**Replication**		
	Mean	SD	N	Mean	SD	N
Had sex ever	0.645	0.479	38,567	0.680	0.467	43,014
Had sex in past year	0.728	0.445	31,890	0.719	0.449	31,888
Number of sexual encounters in past year	128.7	208	25,431	124.9	209.6	25,430
Number times used condom past year	33	98.5	24,627	32.9	100	24,626
Had risky sex	0.003	0.054	28,151	0.002	0.043	26,728

Notes**:** Had risky sex is defined in the text as sex with a stranger or sex with an IV drug user in the past year. However, for the mean of this variable to be 0.003, Mulligan must have coded the variable as both had sex with a stranger AND an IV drug user.