



LEGAL BANS AND TRADITIONAL PRACTICES IN AFRICA: THE EFFECT OF CRIMINALIZING FEMALE GENITAL CUTTING ON THE PRACTICE AND WELL-BEING

A Dissertation

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

in Partial Fulfillment of the Requirements for the Degree of

Ph.D. in Development Economics

by

Hailegabriel Abebe Fenta

October 2021

ABSTRACT

Female genital cutting (FGC), the act of removing some part of the female external reproductive organs for non-medical reasons, has raised concerns about public health and human rights. Consequently, the traditional practice has been subject to eradication through policy. However, systematic evaluation for the effectiveness of these eradication policies has been lacking. Moreover, it remains an important policy question of how much of the recent decline in FGC can be due to these reforms. Using data from Demographic and Health Surveys and Multiple Indicator Cluster Surveys, this dissertation empirically estimates the effect of criminalizing FGC on the traditional practice.

The difference-in-differences estimated results suggest a substantial effect of criminalization on the practice, exploiting the variation across cohorts and within-country region-ethnicity groups. The reduction in the likelihood of ever-experiencing FGC is more pronounced for youngest cohorts at reform and those cohorts in the highest pre-reform FGC prevalence region-ethnicity groups. The effectiveness of the criminalization policy might be due to the perception of the society towards the law as potentially enforceable and fear of possible punishment if being caught practicing FGC. Moreover, the law can improve the bargaining and decision-making power of those females who are willing to abandon the practice.

Although female genital cutting (FGC) has long been a prerequisite for marriage in many African families, almost no empirical evidence shows the impact of criminalization on females' well-being. Empirically examining the criminalization effect on several female well-being outcomes, this dissertation finds that the reform increases the risk of early marriage and fertility. Moreover, the reform lowers educational attainment and reduces spousal quality. However, the reform does not have significant effects on reproductive health outcomes.

In the presence of criminalization, if parents do not substitute education for FGC as a pre-marital investment in their daughters, then they may alternatively substitute early marriage for FGC because young girls are more valuable. The adverse effect of the reform on education could be directly related to stigmatization associated with FGC abandoning,

and indirectly through early marriage leading them to drop out of school. Additionally, the lower spousal quality associated with the reform might be related to a shorter searching period for a better match as parents hurry to marry off their daughter.

ACKNOWLEDGMENT

I would like to express my deepest appreciation to my main supervisor, Professor Stacey H. Chen for her guidance and mentorship throughout my dissertation writing. Since the time she has learned my academic and research aspirations from our first meeting, she has been encouraging and supporting me in every stage of the thesis work with patience. Her invaluable expertise and insightful feedback enabled me to progress on my work with passion. Every email exchange with her has been a lesson for me.

I am also extremely grateful to Professors Chikako Yamauchi, Dainn Wie, and Yoshito Takasaki who have been other members of my dissertation committee, for their constructive comments to improve my thesis. Also, my gratitude extends to my program director, Professor Alistair Munro, for his guidance and encouragement throughout the Ph.D. program and constructive comments during the initial presentations of my thesis chapters.

I must also thank the Japanese Government for providing me a full scholarship for five years of study in Japan. Without this financial support, the completion of my Ph.D. study would not have been possible.

I would like to extend my sincere thanks to Professors Akifumi Ishihara, Alistair Munro, Chikako Yamauchi, Dainn Wie, Fumio Hayashi, Hady George Kahy, Izumi Ohno, Jonna P. Estudillo, Junichi Fujimoto, Keijiro Otsuka, Makoto Hasegawa, Minchung Hsu, Ponpoje Porapakkarm, Yoichi Arai, Roberto Leon-Gonzalez, Shinichi Takeuchi, Shinsuke Ikeda, Stacey Chen, Stephan Litschig, Takasi Fukushima, Tetsushi Sonobe, Tomoya Matsumoto, and Yosuke Yasuda who taught me several crucial courses at GRIPS enabling to strengthen my academic background. I am also grateful to the Academic Support Team, General Affairs Team, International Programs Team, and Student Office of Academic and Student Affairs Division, Research Support and International Affairs Division, and all other supporting staff at GRIPS, including the I.T. Support Center, Health Service Center, and Security Guards.

I also had the great pleasure of working as a research and teaching assistant with Professors and researchers, including Daiju Aiba, Dainn Wie, Makoto Tanaka, Masako Kurosawa, and Stacey H. Chen. Also, many thanks to my colleagues in my batch Chie Aoyagi, Joseph Upile Matola, Md Arifur Rahman, Raphael E. Ayibor, Sharif Mosharraf Hossain, and Vu Thanh Trung whom I studied together with all of them and we helped each other.

I would like to acknowledge the support and great love of my family, my wife, Merkeza, who stayed with me throughout my stay in Japan; and my father, Abebe; my mother, Dinkayehu; and my brothers and sisters in my home country. They kept me progressing, and this long academic journey would not have been possible without their encouragement. Further, I would like to recognize Misato City and Saitama Prefecture offices for their effort to ease my stay in Japan.

TABLE OF CONTENTS

| | |
|---|-----|
| ABSTRACT | i |
| ACKNOWLEDGMENT | iii |
| TABLE OF CONTENTS | v |
| LIST OF FIGURES | ix |
| LIST OF TABLES | xii |
| CHAPTER 1 | 1 |
| INTRODUCTION | 1 |
| 1.1 Background and motivation | 1 |
| 1.2 Objective and methodology | 2 |
| 1.3 Main findings | 3 |
| 1.4 Contributions..... | 4 |
| 1.5 Organization of the dissertation | 5 |
| CHAPTER 2 | 7 |
| LITERATURE REVIEW | 7 |
| 2.1 Traditional institutions and female well-beings..... | 7 |
| 2.2 Enforcement and effectiveness of the laws against female genital cutting..... | 9 |
| 2.3 Female genital cutting and well-being | 10 |
| 2.3.1 Women’s well-being and economic development | 10 |
| 2.3.2 Female genital cutting and education | 11 |
| 2.3.3 Female genital cutting and health..... | 11 |

| | |
|---|----|
| 2.3.4 Female genital cutting and family formation | 13 |
| CHAPTER 3 | 14 |
| INSTITUTIONAL DETAILS AND DATA SOURCES | 14 |
| 3.1 The practice of female genital cutting..... | 14 |
| 3.2 Female genital cutting bans..... | 15 |
| 3.2.1 A long way toward banning female genital cutting | 15 |
| 3.2.2 The regulation of female genital cutting in Africa | 16 |
| 3.2.3 Possible determinants of timing of ban adoption and reform type..... | 18 |
| 3.2.4 Cohort trends in female genital cutting prevalence by reform type | 20 |
| 3.3 Data sources and related issues | 21 |
| 3.3.1 Data sources | 21 |
| 3.3.2 Potential problems with measurement errors in self-reporting | 22 |
| 3.3.3 Age at cutting and non-reporting of age at cutting..... | 23 |
| 3.3.4 Pre-reform FGC prevalence rate across region-ethnicity groups..... | 24 |
| 3.3.5 Managing the issue with the inconsistent region and ethnicity information | 25 |
| CHAPTER 4 | 27 |
| THE EFFECT OF CRIMINALIZING FEMALE GENITAL CUTTING ON THE PRACTICE..... | 27 |
| 4.1 Defining cohorts of interest and policy variables | 27 |
| 4.2 Identification strategy and empirical specification | 27 |
| 4.2.1 Difference-in-differences event-study approach | 28 |
| 4.2.2 Difference-in-differences approach comparing young and old cohorts | 30 |
| 4.2.4 Threats to identification..... | 30 |

| | |
|---|----|
| 4.3 Results for the effect of criminalizing female genital cutting on the practice | 31 |
| 4.3.1 Results for the difference-in-differences event-study estimated effects | 31 |
| 4.3.2 Results for the difference-in-differences comparing young and old cohorts | 33 |
| 4.3.3 Heterogeneity in the effectiveness of criminalization | 34 |
| 4.3.4 Sensitivity analysis to the results for the effect of criminalization | 35 |
| 4.4 Potential mechanisms for criminalization to affect female genital cutting..... | 37 |
| CHAPTER 5 | 39 |
| THE IMPACT OF CRIMINALIZING FEMALE GENITAL CUTTING ON WELL-BEING | 39 |
| 5.1 Descriptive analysis for the female well-being outcome variables..... | 39 |
| 5.2 Empirical specification for the reduced-form analysis of criminalization effect on female well-being..... | 40 |
| 5.2.1 Difference-in-difference event-study approach..... | 40 |
| 5.2.2 Difference-in-difference approach comparing young and old cohorts at reform..... | 41 |
| 5.3 Results for the reduced-form estimated effects of criminalization on female well-being | 42 |
| 5.3.1 Impact of criminalization on health outcomes | 42 |
| 5.3.2 Impact of criminalization on marriage and fertility outcomes | 43 |
| 5.3.3 Impact of criminalization on own educational outcomes..... | 43 |
| 5.3.4 Impact of criminalization on spousal attributes and family wealth..... | 44 |
| 5.4 Potential mechanisms for criminalization to affect female well-being..... | 45 |
| CHAPTER 6 | 47 |
| CONCLUSION..... | 47 |
| 6.1 Summary | 47 |
| 6.2 Policy implications..... | 47 |

6.3 Limitations and future research..... 48

REFERENCES 51

APPENDIX A..... 101

APPENDIX B 116

LIST OF FIGURES

| | |
|--|----|
| Figure 3.1: Rollout of criminal and non-criminal ban on female genital cutting, 1994-2011..... | 60 |
| Figure 3.2: Prevalence of FGC and FGC before age 15 by cohort. | 61 |
| Figure 3.3: Pre-reform age at cutting distribution. | 62 |
| Figure 3.4: Distributions for pre-reform FGC prevalence rate by region-ethnicity groups. | 63 |
| Figure 4.1: Event-study estimated effect of criminalization on ever-experiencing FGC by cohort and pre-reform FGC prevalence rate region-ethnicity groups | 64 |
| Figure 4.2: Event-study estimated effect of criminalization on ever-experiencing FGC..... | 65 |
| Figure 4.3: Event-study estimated effect of criminalization on ever-experiencing FGC by residence (rural/urban). | 66 |
| Figure 4.4: Event-study estimated effect of criminalization on ever-experiencing FGC by religion. | 67 |
| Figure 4.5: Event-study estimated effect of criminalization on ever-experiencing FGC by age at survey..... | 68 |
| Figure 4.6: Event-study estimated effect of criminalization on ever-experiencing FGC by degree of enforcement. | 69 |
| Figure 4.7: Event-study estimated effect of criminalization on ever-experiencing FGC by excluding countries with less within-country variation in the pre-reform FGC prevalence rate. | 70 |
| Figure 4.8: Event-study estimated effect of criminalization on ever-experiencing FGC by different clustering. | 71 |
| Figure 4.9: Event-study estimated effect of criminalization on ever-experiencing FGC by data type. | 72 |
| Figure 5.1: Reduced-form estimated effect of FGC criminalization on genital problems and child death among females aged 15 to 49 at the interview..... | 73 |
| Figure 5.2: Reduced-form estimated effect of FGC criminalization on marriage and fertility among females aged 15 to 49 at the interview..... | 74 |

| | |
|---|-----|
| Figure 5.3: Reduced-form estimated effect of FGC criminalization on early marriage and fertility among females aged 15 to 49 at the interview..... | 75 |
| Figure 5.4: Reduced-form estimated effect of FGC criminalization on own education and work status among females aged 15 to 49 at the interview..... | 76 |
| Figure 5.5: Reduced-form estimated effect of FGC criminalization on own general health among females aged 15 to 49 at the interview..... | 77 |
| Figure 5.6: Reduced-form estimated effect of FGC criminalization on spousal education among females aged 30 to 49 at the interview..... | 78 |
| Figure 5.7: Reduced-form estimated effect of FGC criminalization on spousal education among females aged 15 to 49 at the interview..... | 79 |
| Figure 5.8: Reduced-form estimated effect of FGC criminalization on spousal occupation among females aged 30 to 49 at the interview..... | 80 |
| Figure 5.9: Reduced-form estimated effect of FGC criminalization on spousal occupation among females aged 15 to 49 at the interview..... | 81 |
| Figure 5.10: Reduced-form estimated effect of FGC criminalization on family wealth among females aged 30 to 49 at the interview..... | 82 |
| Figure 5.11: Reduced-form estimated effect of FGC criminalization on family wealth among females aged 15 to 49 at the interview..... | 83 |
| Figure A3.1: Pre-reform distribution of age at female genital cutting for criminalizing countries..... | 101 |
| Figure A3.1: Pre-reform distribution of age at female genital cutting for criminalizing countries (continued). | 102 |
| Figure A3.2: Pre-reform distribution of age at female genital cutting for non-criminalizing countries. | 103 |
| Figure A3.3: Pre-reform FGC prevalence rate distribution across region-ethnicity groups for criminalizing countries..... | 104 |
| Figure A3.3: Pre-reform FGC prevalence rate distribution across region-ethnicity groups for criminalizing countries (continued). | 105 |
| Figure A3.4: Pre-reform FGC prevalence rate distribution across region-ethnicity groups for non-criminalizing countries. | 106 |
| Figure A4.1: Event-study estimated effect of criminalization on ever-experiencing FGC for young age females at the survey..... | 107 |

| | |
|--|-----|
| Figure A4.2: Event-study estimated effect of criminalization on ever-experiencing FGC for old age females at the survey..... | 108 |
| Figure A4.3: Event-study estimated effect of criminalization on ever-experiencing FGC including very old cohorts at reform | 109 |
| Figure A4.4: Event-study estimated effect of criminalization on ever-experiencing FGC using all the available cohorts at reform..... | 110 |
| Figure A4.5: Event-study estimated effect of non-criminal laws on ever-experiencing FGC using all the available cohorts at reform..... | 111 |
| Figure A5.1: Reduced-form estimated effect of FGC criminalization on marriage and fertility among females aged 15 to 29 at the interview..... | 112 |
| Figure A5.2: Reduced-form estimated effect of FGC criminalization on own education among females aged 15 to 29 at the interview..... | 113 |
| Figure A5.3: Reduced-form estimated effect of FGC criminalization on genital problems and child health among females aged 15 to 29 at the interview..... | 114 |
| Figure A5.4: Reduced-form estimated effect of FGC criminalization on own general health among females aged 15 to 29 at the interview..... | 115 |

LIST OF TABLES

| | |
|--|-----|
| Table 3.1: Data availability of African countries with an FGC ban as of 2013, year and content of reform, pre-reform FGC prevalence rate and post-reform surveys. | 84 |
| Table 3.1: Continued. | 85 |
| Table 3.2: Can the pre-reform country and region-ethnicity characteristics explain the timing of reform adoption and type of reform? | 86 |
| Table 3.3: Summary statistics for country attributes by reform type. | 88 |
| Table 3.4: Correlation between FGC experience and non-reporting the age at cutting. | 89 |
| Table 4.1: African countries with FGC criminalization as of 2013 and had DHS/MICS respondents affected fully or partially by the reform. | 90 |
| Table 4.2: The effect of FGC criminalization on the probability of ever-experiencing FGC for each cohort. | 91 |
| Table 4.2: Continued. | 92 |
| Table 4.3: Unconditional reform effect in a difference-in-difference design: Correlation between FGC criminalization and the probability of ever-experiencing FGC. | 93 |
| Table 4.4: The Effect of FGC criminalization on the probability of ever-experiencing FGC. | 94 |
| Table 4.5: The heterogeneous effects of FGC criminalization on the probability of ever-experiencing FGC by residence and religion. | 95 |
| Table 5.1: Summary statistics for outcome variables. | 96 |
| Table 5.2: The reduced-form estimated effect of criminalizing FGC on education, marriage, and health outcomes.. | 97 |
| Table 5.3: The reduced-form estimated effect of criminalizing FGC on spousal attributes and family wealth. | 98 |
| Table 5.4: Falsification test for the reduced-form estimated effect of criminalizing FGC on own education, marriage, and health. | 99 |
| Table 5.5: Falsification test for the reduced-form estimated effect of criminalizing FGC on spousal attributes and family wealth. | 100 |
| Table B3.1: African countries with an FGC ban as of 2013, pre-reform surveys and FGC prevalence rate at country level, and construction of pre-reform FGC prevalence rate by region-ethnicity groups. | 116 |

| | |
|--|-----|
| Table B3.1: Continued..... | 117 |
| Table B4.1: African countries with non-criminal FGC bans as of 2013 and had DHS/MICS respondents affected fully or partially by the reform. | 118 |

CHAPTER 1

INTRODUCTION

1.1 Background and motivation

How do policies against traditional practices affect the practices and the well-being of those practicing them have been important research considerations for scholars. On one hand, such policies could be effective, at least in part, in helping in the eradication efforts. Examining policies against child marriage in Ethiopia, for example, Chow and Vivalt (2021), and McGavock (2021) suggest the effectiveness of the interventions in reducing the likelihood of child marriage. On the other hand, the policies may have a backfiring effect rather than directly affecting it in the initially intended way (Acemoglu and Jackson 2017).¹

Gender-based discriminatory traditions are becoming the subject of deliberate changes or eradication by international organizations in the name of human rights. Researchers and policymakers have viewed FGC, the act of removing some part of the female external reproductive organs for non-medical reasons, as one of the significant social customs that discriminatorily hurts women (OHCHR 1995; Platteau, Camilotti, and Auriol 2018; WHO 2016).² The United Nations (UN), for example, adopted the convention on the elimination of all forms of discrimination against women in 1979. Moreover, the 2015 Sustainable Development Goals set up by the UN General Assembly (Goal 5.3) recognize the need to eliminate all forms of harmful practices. Although several scholars and international organizations (e.g., Mackie 1996; WHO 2016) consider FGC against human rights and subject to eradication through policy or social changes, it remains prevalent.³

¹ For more discussion on the effectiveness of policies against traditional practice, see section 2.2 of Chapter 2.

² Other customs listed in OHCHR (1995) include early marriage and pregnancy, dowry price, son preference, female infanticide, nutritional taboos, practices related to childbirth, and violence against women.

³ Mackie (1996) compares female genital cutting with the custom of female foot-binding in China. The paper considers both as social norms, self-enforcing conventions, practiced mainly to ensure marriageability. Though the eradication effort of foot binding was successful by forming

Despite decades of anti-FGC movements in Africa, little empirical analysis (e.g., Camilotti 2016; Diop and Askew 2009) systematically evaluates the effectiveness of the eradication policies. Several African countries have started regulating the practice of FGC, particularly since 1994. Out of the 27 African countries where FGC is prevalent, 22 have banned FGC practices as of 2013 (Shell-Duncan et al. 2013; UNICEF 2013). On the other hand, FGC has been declining although persistent in many regions (UNICEF 2013). Thus, it remains an important policy question of how much of the recent decline in FGC is due to the national bans. Few previous studies (e.g., Camilotti 2016) have evaluated the contribution of such policies to the declining trend in FGC. However, these previous analyses focus only on case studies and do not address the issue that concurrent policies other than FGC bans might confound the estimated impact of the reforms.

Moreover, although FGC has long been a prerequisite for marriage in many African households (Mackie 1996), almost no empirical evidence shows the impact of criminalization (or FGC) on females' well-being. Surprisingly, few large-scale quantitative studies focus on the FGC impact (e.g., Wagner 2015) despite its importance on child protection, public health, and human rights (Cook, Dickens, and Fathalla 2002; Shell-Duncan 2008). Thus, this dissertation attempts to fill in these research gaps.

1.2 Objective and methodology

How criminalization affected the traditional practice and how it affected the well-being of the victims, including health, education, and marriage are the research questions in this dissertation. To address these questions, this dissertation aims to evaluate the policy reform of FGC criminalization empirically.

This dissertation employs a difference-in-difference approach exploiting the within-country variation in reform effect across region-ethnicity groups and cohorts to address these research questions. In this approach, the study undertakes event-study analyses to examine the impact of criminalization per cohort and compare young cohorts at reform with

associations of parents who committed not to foot bind their daughters nor let their sons marry foot-bound women, female genital cutting still exists.

those old. This dissertation uses all the available FGC survey data for entire Africa from the demographic and health surveys (DHS) and the multiple indicator cluster surveys (MICS) to undertake these empirical analyses.

1.3 Main findings

This dissertation first discovers that criminalization substantially reduces the likelihood of ever-experiencing FGC. The policy change appears most effective among the younger females in the region-ethnicity group on the top tercile of the pre-reform FGC prevalence. One potential mechanism behind the effectiveness of the reform could be people's perception of the law as potentially enforceable and fear of possible punishment if they practice FGC even without actual enforcement. Another one might be through the effect of the law in increasing the bargaining and decision-making power of females (most likely the mothers) who are willing to abandon the practice. Moreover, this dissertation shows effect heterogeneity by religion, urbanization, and degree of enforcement. Interestingly, the results also suggest the estimated impact of the reform varies with data sources.

The reduced-form results from estimating the effects of criminalization on female well-being (measured by health, education, and marriage) show that the reform *increases* the risk of early marriage and fertility at the expense of quality marriage. Moreover, the reform *lowers* literacy and educational attainment, mainly the secondary level or above, suggesting that education is not the mechanism for the reform's effects (or FGC) on early marriage. Facing FGC criminalization, parents do not substitute education for FGC as a pre-marital investment to improve the value of their daughters. Instead, they replace FGC with early marriage because young-age girls are more valuable in marriage markets. Criminalizing FGC can lower female education potentially because uncut girls are likely to face stigmatization, discrimination, and peer isolation. Alternatively, the decreased educational achievement by criminalization could come from an increased risk of early marriage, leading them to drop out of school.

Furthermore, criminalization lowers husband quality, measured by educational achievement and engagement in a skilled type of occupation. This result can be related to a shorter searching time for a quality partner because parents

hurry to accept marriage proposals for an uncut daughter. However, the reform does not affect the reproductive health outcomes significantly. Also, the results do not suggest adverse effects of the reform on the general health well-being of females, except leading them to be overweight. The mechanisms could be related to the stress and tension associated with the feeling of stigma in abandoning FGC and their lower educational achievement.

These findings pass several falsification tests. Also, the event-study results support the validity of the research design in this dissertation. Major limitations of this study are potential issues with self-reported FGC status and lack of data on parental backgrounds from the same data sources, particularly on maternal conformity to the FGC practice. Because questions related to FGC are sensitive, women may misreport their actual status. However, understanding the direction of bias may be challenging as both under-reporting and over-reporting are possible. Women may under-report their true status of cutting (or deny their cutting) if they might fear punishment in the presence of criminalization. Moreover, over-reporting might be possible due to the feeling of the stigma associated with FGC abandoning. Further, as maternal conformity to the FGC practice is unobservable in the data, the observed relationship between criminalization and women's FGC status might be spurious even given the same region, ethnicity, and birth cohort.

1.4 Contributions

Few large-scale multi-country studies focus on the effect of FGC (or reform on FGC) despite its importance on child protection, public health, and human rights.⁴ Wagner (2015) is the only previous study that I am aware of that has attempted to cover multiple countries quantitatively. This dissertation fills this research gap and addresses the external validity concerns of past research. It is the first to pool all the available FGC survey data for the entire Africa, including the demographic and health surveys (DHS) and the multiple indicator cluster surveys (MICS). Also, it reviews historical and institutional details for entire Africa on FGC regulations and policy interventions against FGC.

These discussions and intensive data efforts are necessary for the identification strategy (the difference-in-differences approach), which uses variations in the reform timing, type, and enforcement across birth cohorts, countries, and *region-*

⁴ Cook et al. (2002) and Shell-Duncan (2008) show how FGC is a concern of reproductive health and human rights.

ethnicity groups within countries. This study's key novelty is to exploit further the variation in the exposure to criminalization across *region-ethnicity groups within countries*, in addition to the variation across birth cohorts and countries considered in the previous studies (Camilotti (2016) and Wagner (2015), to be discussed in detail in sections 2.2 and 2.3). This new design makes it possible to allow the reform's impact to be *nonlinear* in FGC prevalence among region-ethnicity groups. Allowing for nonlinearity is crucial in this empirical task because the standard linear models may not capture well the women's complex behavior and attitude towards the custom in the presence of laws.

This novel strategy and the intensive data work substantially improve the reliability and precision of the previous research, and contribute to policy. The findings related to the effectiveness of FGC criminalization have important policy implications aiming to eradicate FGC, such as promoting national bans on the practice. This dissertation provides empirical evidence for the role of laws against deeply enrooted traditional practices, even with limited actual enforcement. The law can possibly be effective by changing people's perception of the reform as potentially enforceable, leading to the fear of punishment. Also, the reform might work through enhancing the bargaining and decision-making power of those women against FGC. As a result, policymakers may have to consider such possible mechanisms while trying to improve the effectiveness of the interventions. Further, the results related to the variation in the effectiveness of the reform by cohort and *region-ethnicity groups* help the policymakers on their decision on whom and where to target and give priority.

Moreover, the findings related to the effect of the reform on well-being show that FGC is interlinked with another custom, early marriage, suggesting that policymakers should think over the risk of early marriage while trying to reform FGC. Another finding related to education shows that the reform lowers educational achievement, suggesting the requirement of additional policy promoting female education by addressing the feeling of stigma for those abandoning FGC.

1.5 Organization of the dissertation

The remainder of this dissertation proceeds as follows: Chapter 2 reviews related literature, including the relationship between traditional institutions and female well-being, the enforcement and effectiveness of laws against conventional practices, including female genital cutting (FGC), and the effects of FGC on well-being, measured by education, health, and marriage formation.

Chapter 3 describes institutional details and data and reports the related analyses. The institutional details include additional literature reviews on FGC practice, previous efforts to ban FGC, and the rollout and content of bans on FGC in Africa. Moreover, this chapter presents and discusses the potential determinants of the timing of FGC reform adoption and the reform content. Further, the chapter discusses data sources and management and some issues concerning the nature of the data.

Chapter 4 describes the identification strategy and presents the model specification to estimate criminalization's effect on the practice empirically. The next sub-sections of this chapter present and interpret the results for the effect of criminalizing FGC on the practice and discuss the possible mechanisms behind the findings.

Chapter 5 first discusses the empirical strategy and then reports the results for the reduced-form estimated effects of criminalization on several outcomes of female well-being. Then, this chapter proposes possible channels for the effect of the reform on the well-being outcomes. Chapter 6 concludes the dissertation, discussing policy implications related to the main chapters, limitations, and future research ideas.

CHAPTER 2

LITERATURE REVIEW

2.1 Traditional institutions and female well-beings

Intensive evidence has indicated that social norms can substantially affect female well-beings, measured by health, education, marriage, and family formation choices.⁵ Among the traditional practices, marriage customs play a crucial role in women's future well-being because many women do not work outside their homes in developing countries (Jayachandran 2015). Ashraf et al. (2020) show that *bride price* customs influence the success of development programs. They find that policies aiming at boosting female education are more effective in societies with a marriage custom of bride price using data from Indonesia and Zambia with large-scale school construction projects. The mechanism is that more educated girls have a higher bride price. Using data from a region in the Democratic Republic of the Congo, Lowes and Nunn (2018) find a positive association between bride price and marriage quality (measured by the self-reported wife happiness, opinion on domestic violence, and quality time as a couple) but no evidence of an association between large bride prices and early marriage or high fertility.

Jacoby and Mansuri (2010) show that Pakistan's bride exchange tradition (which involves the simultaneous marriage of a brother-sister pair from two households) can lower a husband's chance of mistreating his wife. Vogl (2013) suggests that the tradition of marrying daughters in birth order reduces older sisters' marriage quality in the presence of a younger sister. Chen, Chen, and Liu (2019) show that the son-preferring fertility stopping rule helps Taiwanese females with younger brothers get more education by reducing family size and canceling son preference's rivalry effect.

In a study on Kyrgyzstan's bride kidnapping tradition (abducting a girl to forcibly marrying), Bazarkulova and Compton (2021) show that the custom negatively affects education. However, the effect may vary with the degree of

⁵ We use the terminologies; social norm, culture, traditional practice or custom, and informal institution interchangeably; all refer to the "patterns of behavior that are self-enforcing within a group: Everyone conforms, everyone is expected to conform, and everyone wants to conform when they expect everyone else to conform" (Young 2015, pp.359).

traditionalism across regions. Furthermore, they suggest that an intervention to eradicate the bride kidnapping custom will not adversely affect girls' education.

Moreover, Tertilt (2005) investigates polygyny, the tradition of marrying multiple wives, in Sub-Saharan Africa. The result suggests that banning polygyny contributes to poverty alleviation by substantially reducing fertility and increasing savings and productivity. Because if polygyny is forbidden, then the demand for wives decreases, which decreases the equilibrium bride price. As a result, men (fathers) invest more in physical assets than wives (daughters), raising the aggregate capital stock. A high capital stock combined with low fertility due to a lack of incentive to have many children increases output (Tertilt 2005). However, enforcement is difficult due to cultural and economic factors. Economically, the initial generations of men lose from the reform, although output increases in the long run (Schoellman and Tertilt 2006).

Because of the difficulty of enforcing the law against polygyny, Tertilt (2006) suggests an alternative policy of allowing girls to decide on their marriage by themselves instead of parental involvement. However, both interventions have similar economic effects, although more pronounced for the ban. To enable girls to make their own marriage decisions, the author suggests improving gender equality (Tertilt 2006).

In addition to marriage customs, traditions related to social class, such as the caste system, affect women's well-being. Examining the role of religion and social class in the business activity of Indian women, Field, Jayachandran, and Pande (2010) find that financial literacy training is more effective for those women facing more cultural restrictions. More specifically, the training enhances the amount of borrowing and income for non-Muslim women with more restrictions in mobility and social interactions, the upper-caste women.

Moreover, Munshi and Rosenzweig (2006) examine the role of the *caste system* in shaping career choices by gender using data from India. They find that lower-caste girls are more likely to join schools leading to office work or professional jobs (non-traditional occupations) than boys in similar social classes. The mechanism is that lower-caste girls, unlike boys, have a low labor market experience and poor network ties, and as a result, are unlikely to enter into schools leading to the traditional occupations.

Using data from the program for international student assessment, Nollenberger, Rodríguez-Planas, and Sevilla (2016) show a significant effect of traditional beliefs about the role of women on gender differences in math. Moreover,

they suggest that the mechanisms for this effect include not only the math-specific generalizations against women but also other gender-related stereotypes.

2.2 Enforcement and effectiveness of the laws against female genital cutting

Aldashev et al. (2012a, 2012b) and Acemoglu and Jackson (2017) theoretically examine the effectiveness of the laws against the traditional practices and the conditions of potential backfire. The law can be effective by serving as an external helping agent for those who are willing to abandon the practice. In other words, the law creates an exit option by improving the bargaining power of the disadvantaged groups (Aldashev et al. 2012a, 2012b).

Investigating the effects of two interventions against child marriage in Ethiopia; community discussions about the tradition and economic incentives to stay in school, Chow and Vivalta (2021) find that both interventions are effective in reducing the likelihood of child marriage. Moreover, their findings show an increase in the women's decision-making power within the household due to the interventions, a possible mechanism for the decrease in the probability of child marriage. However, they do not find a significant effect of the interventions on education, suggesting that the mechanism for the reduction in child marriage is not through education. Moreover, exploiting the rollout of policies against child marriage across Ethiopian regions, McGavock (2021) suggest the effectiveness of the reforms in delaying women's marriage and fertility, particularly in areas with higher pre-reform early marriage prevalence rate.

However, laws may have a backfiring effect for those who are not willing to abandon the tradition and lead to the evolution of a custom instead of directly affecting it in the initially intended way (Acemoglu and Jackson 2017). The unintended consequences work mainly for those laws which are in intense conflict with the existing deeply enrooted social norms such as female genital cutting practices (Acemoglu and Jackson 2017; Belloc and Bowles 2013). Anti-FGC legislation in Senegal, for example, leads to a decrease in age at cutting rather than reducing the prevalence of FGC among daughters (Camilotti 2016). According to Acemoglu and Jackson (2017), the situation of snitching on those violating the law may be less likely when almost everybody breaks the law. Enforcement of laws may then become difficult in the absence of exposing the lawbreakers since it is costly to identify and convict the offender. In addition to

the cost of finding who violates the law, the type of punishments and responses of the lawbreakers also matter for the optimal level of law enforcement (Becker 1968).

Theoretically examining the effectiveness of criminalizing levirate marriage (that is, a marriage tradition forcing the brother of a deceased man to marry his brother's widow), Kudo (2021) suggests that the reform affects women's welfare negatively because of the community's higher cost of offering the levirate marriage by convincing the woman. However, systematic empirical evidence is lacking regarding the impact of policy changes against discriminatory customs, particularly FGC.

Few studies examine the effectiveness of interventions against FGC. Camilotti (2016) is the closest to our research and finds an insignificant effect of anti-FGC legislation on the cut status but a significant and negative effect on age at cutting using data for Senegal. Using data from the same country, García-Hombrados and Salgado (2019) suggest the effectiveness of the law in reducing FGC prevalence. Moreover, they find that the law improves girls' education, suggesting that education is substituting FGC in the marriage market given both education and FGC increase the quality of girls. However, whether the estimates for Senegal and West African countries apply to the rest of Africa remains an open question.

Studies on the impacts of several other interventions rather than the national bans against FGC show that their effectiveness and success are limited. These interventions include training health personnel, education of female students, multifaceted community activities, and village empowerment. Diop and Askew (2009) show that a community-based education program for community members in rural Senegal reduces the prevalence of FGC among daughters of age 10 or below. However, it is difficult to rely on the result due to concerns about the research design and methodological quality, such as the unrepresentativeness of samples (Berg and Denison 2012).

2.3 Female genital cutting and well-being

2.3.1 Women's well-being and economic development

Well-being is multi-dimensional by its concept and can be measured using different indicators, including income, health, education, and other social measures (Jordá and Sarabia 2015; Konu and Rimpela 2002; Saith and Harriss-White 1999). This dissertation measures women's well-being using education, health, and family formation. Education and health are human capital associated with labor market performance and intergenerational transmission. Recent evidence shows that better female education and health contribute substantially to economic growth and development (Bloom, Kuhn, and Prettnner 2020). Moreover, marriage and spousal attributes matter for women's financial security in developing countries because many women do not work outside their homes (Jayachandran 2015).

2.3.2 Female genital cutting and education

Arguably, FGC affects female educational achievement positively or negatively (Pesambili and Mkumbo 2018). FGC protects those cut against stigma and isolation associated with the abandonment of the practice, and those uncut against the risk of early marriage, enabling both to continue their education better. By contrast, FGC restricts the opportunities for education for those cut when it serves as a rite of passage and a signal to marriage. This is because most girls drop out of school when they are ready to marry. Moreover, FGC may lead to a feeling of stigma and isolation for those uncut. Pesambili (2013) discusses isolation as social, family, and peer isolation, and stigma, including the use of offensive and provoking words, verbal insults, and discrimination.

Exploiting variation in exposure to FGC legislation across ethnic groups in the same country, García-Hombrados and Salgado (2019) suggest the effectiveness of the law in reducing FGC prevalence and improving girls' education. The mechanism is that education is substituting FGC in the marriage market given both education and FGC increase the quality of girls. However, they could not observe the marriage outcomes directly from the data, making the analysis incomplete.

2.3.3 Female genital cutting and health

Views on the health impacts of FGC vary widely; see Berg et al. (2014), Wagner 2015, and WHO (2016) for reviews. Several medical and observational studies find FGC causes a list of physical (e.g., Kaplan et al. 2013) and mental (e.g., Mulongo, Martin, and McAndrew 2014) health problems, ranging from immediate and short-term complications to long-term risks. While the physical problems associated with FGC mostly include obstetrical and gynecological complications (Berg et al. 2014; Berg and Underland 2013; Kaplan et al. 2013; Klein et al. 2018; Wagner 2015), the psychological impacts of FGC include stress, anxiety, and depression, among others (Grose et al. 2019; Mulongo et al. 2014; Pesambili and Mkumbo 2018).

By contrast, other studies suggest no health impairments induced by FGC (e.g., Balachandran et al. 2018 and Morison et al. 2001). Balachandran et al. (2018) show insignificant FGC effects on obstetric and neonatal outcomes using observational data from 242 females in the UK. Moreover, Morison et al. (2001) do not find apparent FGC effects on morbidity using data from rural Gambia.

Wagner (2015) categorizes the health indicators into general and reproductive health measures, suggesting different results. Her findings suggest that FGC is significantly associated with reproductive health complications, mainly sexually transmitted diseases and genital problems. However, she finds no negative health impacts using general health measures, including body mass index and hemoglobin levels.

However, most studies on the impacts of female genital cutting suffer from data and methodological concerns. The results are inconclusive and mostly rely on observational data with a small sample size or lack of representation (Berg et al. 2014; Wagner 2015). As a result, policymakers recently shifted to the human rights issue to justify their intervention against FGC as they lack reliable evidence on the health consequences of the practice (Shell-Duncan 2008).

Deep-rooted traditional practices, particularly gender norms such as FGC, disproportionally affect women and hinder the implementation of their human rights (Giuliano 2020; Ssenyonjo 2007). As a result, the human rights advocates and international community demonize local culture as an obstacle to the reforms on human rights protection without a proper understanding of the practice. Instead of recognizing cultural diversity and even building upon culture, if the human rights movements merely resist culture, they face accusations of having different motives, including discriminatory, imperialist, and colonizing attitudes (Krivenko 2015; Merry 2003).

2.3.4 Female genital cutting and family formation

Wagner (2015) is the only exception providing evidence on the impact of FGC on marriage and fertility, in addition to health. Using cross-sectional data from 13 African countries, she finds that the cut women are more exposed to the risk of marrying and child-raising sooner, leading to a large family size. However, the identifying condition for the study is that the decision to cut is independent of any unobservable factor (e.g., family tradition or parental education) given the observed covariates. As a result, FGC may not be exogenous and it is difficult to consider the findings as causal effects of FGC, although she tries to minimize the concern by comparing cut and uncut women within the same cluster. Moreover, such cross-sectional studies are more susceptible to biases related to reporting and interviewing.

This dissertation attempts to narrow this research gap in estimating the effect of FGC bans (or FGC) on the practice and women's well-being by applying a difference-in-difference approach. The identification strategy exploits variations in the effectiveness of FGC policy across cohorts and region-ethnicity groups within countries. Further, we undertake event-study analyses showing cohort-to-cohort contrast across pre-reform FGC prevalence region-ethnicity groups to allow nonlinearity.

CHAPTER 3

INSTITUTIONAL DETAILS AND DATA SOURCES

3.1 The practice of female genital cutting

WHO (2016) classifies FGC practices into four categories; all involve an act of removing part of female external reproductive organs for non-medical reasons. The least severe and most popular is the first category, referred to as the partial or total cutting of the clitoris or the prepuce (*clitoridectomy*). The second category covers partial or complete removal of the clitoris and the labia minora, either with or without excision of the labia majora. The third type is the most severe because it involves reducing the vaginal opening by cutting and stitching together with the labia minora or the labia majora, either with or without excision of the clitoris. Finally, the fourth type covers the rest of the FGC practices, such as piercing, pricking, and scraping.

The first place for FGC and the initial way of cutting is unknown with certainty. Historians claim that the practice dates back from the second to fifth century B.C. (Kouba and Muasher 1985; Mackie 1996; Ruderman 2013). The Greek historian, Herodotus, witnessed the prevalence of female genital cutting in Egypt around the middle of the fifth century B.C. (Kouba and Muasher 1985). The ancient Phoenicians, the Hittites, and the Ethiopians had supposedly practiced the cutting around this time, too (Ruderman 2013; Taba 1980). The Greek geographer, Strabo, and the German traveler, Niebuhr, also reported the prevalence of such practice in Egypt later (Kouba and Muasher 1985; Taba 1980).

FGC is more prevalent in Africa than in the rest of the world, although Asia, the Middle East, and the immigrant communities of Europe and North America have some form of it. FGC is also prevalent among certain ethnic groups in Central and South America (WHO 2016). According to UNICEF (2013), FGC predominantly exists in 29 countries, of which 27 are in Africa. More than 125 million females in those countries in Africa have undergone some form of FGC.

As Antonazzo (2003) has noted, FGC has long been a widely acceptable custom in Africa for several reasons, mainly related to a common belief that FGC can increase a daughter's marriageability. Most parents use FGC as a rite of

purification and evidence for a daughter's tolerance for pain and respect for culture and tradition. Many of those parents consider FGC as a way to fulfill religious requirements, although no religion has required it.⁶ Additionally, those social factors persist over time due to peer pressures.

FGC has many adverse health consequences. These health impairments can be immediate, medium-term, and long-term risks. The short to medium-term risks may include pain, fever, shock, blood poisoning, tetanus, hemorrhage, keloid scar, cysts, dysmenorrhea, infections, genital tissue swelling, urination, and wound healing problems, and even death. The long-term consequences include genital tissue damage, vaginal discharge, menstrual problems, reproductive tract, and other infections. Moreover, it can lead to obstetric risks (e.g., cesarean section, prolonged labor, postpartum hemorrhage, and stillbirth and early neonatal death), sexual functioning problems (e.g., decreased sexual satisfaction and reduced sexual desire), and psychological risks such as anxiety disorders, depression, and post-traumatic stress disorder. However, no health benefits of FGC have been discovered yet (Berg and Denison 2013; Mackie 1996; WHO 2016).

3.2 Female genital cutting bans

3.2.1 A long way toward banning female genital cutting

Early opposition efforts against FGC began in Africa by Christian Missionaries in the first half of the 20th century before independence from colonization. Missionaries of Scotland Church in Kikuyu of Kenya were moving against FGC in 1906. There were education campaigns by Britain in Sudan in 1946 and even efforts to ban infibulation, the severe type. However, all these efforts were not generally successful and led instead to the politicization of the issue. As a result, colonial governments avoided interfering in domestic matters in order not to aggravate the tensions (Boyle and Preves 2000). The unsuccessful local opposition efforts weakened in the late 1950s and were then followed by the rise of international interest in the FGC (Boyle and Preves 2000; Shell-Duncan et al. 2013).

⁶ Although it is not mentioned explicitly in the Quran, some Muslim respondents believe cutting is a religious requirement.

The international community had been hesitating to oppose the practice openly and formally due to sovereign autonomy issues. It was in the 1964 conference that the U.N., for the first time, put a formal statement against FGC, considering it as a violation of human dignity and right to health. Other critical global milestones include the declaration of 1975 as the International Year of Women and 1975 to 1985 as the International Decade for Women (Boyle and Preves 2000). In 1979, the U.N. adopted the Convention on the Elimination of All Forms of Discrimination against Women (CEDAW), which considers FGC as a violation of human rights. The Organization of African Unity (now the African Union) adopted the African Charter on the Rights and Welfare of the Child in 1990 (UNICEF 2013). The U.N. further put a strong opposition against FGC in the 1994 International Conference on Population and Development (Boyle and Preves 2000). The eradication of harmful traditional practices, particularly FGC, was emphasized at the Fourth World Conference on Women in Beijing in 1995. Then, WHO, UNICEF, and UNFPA issued a joint statement opposing the FGC in 1997 (UNICEF 2013).

Western countries began to adopt national bans on FGC despite the very low prevalence in the small number of immigrant communities. France and Sweden adopted a national ban in 1982, while Switzerland took it in 1983. The United Kingdom instituted a federal ban in 1985, then Belgium and Netherlands passed later in 1988. Since France and United Kingdom were the colonizers of most of the countries, and there have been continued colonial ties even after independence, their stand on the FGC might matter on the timing of FGC ban adoption for African countries. United States adopted a national ban in 1996 and linked its provision of foreign aid to policies against FGC (Boyle and Preves 2000).

3.2.2 The regulation of female genital cutting in Africa

Out of the 54 African countries, the prevalence of FGC appears in 27 of them, exactly about half of the total number of countries (UNICEF 2013).⁷ Out of the 27 African countries where FGC is prevalent, 22 have banned FGC practices

⁷ The countries with FGC prevalence in Africa include Benin, Burkina Faso, Cameroon, Central African Republic, Chad, Cote d'Ivoire, Djibouti, Egypt, Eritrea, Ethiopia, Gambia, Ghana, Guinea, Guinea Bissau, Kenya, Liberia, Mali, Mauritania, Niger, Nigeria, Senegal, Sierra Leone, Somalia, Sudan, Tanzania, Togo, and Uganda.

as of 2013. Among the 22 countries, we exclude Eritrea and Somalia because of the unavailability of post-reform surveys. Moreover, although already criminalized FGC since 2007, Eritrea has no public-use DHS available for analysis. Table 3.1 presents anti-FGC laws in Africa varying by the timing of adoption, type of law, and degree of enforcement.⁸

The period of FGC ban adoption for the African countries in our sample ranges from as early as 1994 for Ghana to as recent as 2011 for Guinea Bissau. Although the Central African Republic and Guinea initially introduced some form of bans on FGC earlier in the 1960s, we focus on their later reforms to observe the younger cohorts comparable with those in other countries. For Nigeria and Sudan, only a few states have rolled out FGC bans instead of national bans throughout the countries. The states in Nigeria include Edo, Cross River, Ebonyi, Enugu and Bayelsa, and Rivers, having FGC bans since 1999, 2000, 2001, 2004, and 2009. Only two states in Sudan, South Kordofan, and Gadaref introduced an FGC ban in 2008 and 2009.

We categorize the type of the reforms into two: criminal and non-criminal bans, as shown in panels A and B of Table 3.1. While the criminal bans include criminal laws and legislations, the non-criminal ones include reproductive health laws, the children's Act, the prevention of FGC Act, provisional constitutional decree, and other laws preventing FGC. Figure 3.1 presents the rollout of reform on FGC for 14 African countries with criminal laws and 6 countries with non-criminal laws. The solid and dashed line in the figure indicates the cumulative percentage of countries with criminal and non-criminal bans on FGC, respectively. To compute the cumulative percentages, we consider the number of African countries with criminal or non-criminal laws against FGC as the numerator and all the 27 countries in Africa where FGC is prevalent as a denominator. The figure shows that more and more countries have been adopting either criminal or non-criminal laws against FGC recently.

⁸ Although we lack complete data, variations in FGC bans are also available concerning restrictions, subjects to be punished, and the severity of punishments by those bans. Although most African countries adopted FGC bans applicable throughout every group of societies, in some countries (e.g., Tanzania), FGC is considered illegal if the victims are minors only. In other countries (e.g., Ethiopia, Nigeria, Senegal, and Togo), those who participate in the cutting, although they do not perform the cutting, will be criminally liable. Other countries (e.g., Benin, Burkina Faso, Cote d'Ivoire Ethiopia, and Kenya) impose fines on both the practitioners of FGC and those who know about the practice but fail to report it (Shell-Duncan et al. 2013).

3.2.3 Possible determinants of timing of ban adoption and reform type

In this sub-section, we investigate whether the macro-characteristics of the countries in the years before the introduction of the first anti-FGC policy in 1994 explain the timing of reform adoption. Because, in addition to the laws against FGC, other macro-level socio-economic and political factors may affect the prevalence of FGC (Engelsma, Mackie, and Merrell 2020). Several pre-reform country characteristics --- such as urbanization, colonial ties,⁹ foreign aid dependency and population structure might induce anti-FGC legislation at a certain timing. Also, we examine whether within-country characteristics, such as the pre-reform FGC prevalence rate at region-ethnicity groups, matter in the reform adoption timing. For this purpose, we regress the year of reform on the list of pre-reform country characteristics, including the pre-reform FGC prevalence, square of pre-reform FGC prevalence, and the 1989-1993 average for total fertility rate, infant mortality rate, urban population percentage, population size, foreign aid, colonial ties, and the interactions between the colonizer dummies and the foreign aid indicator.¹⁰

Table 3.2 presents the potential determinants of the timing of ban adoption and type of reform, criminal or non-criminal. Considering all the 20 countries in column (4), the timing of the reform adoption has a positive and significant association with the fertility rate. Specifically, every 1 additional birth per woman is associated with approximately 5 years of later reform adoption.

For criminalizing countries in column (5), the timing of the reform adoption has a significant and negative correlation with the pre-reform FGC prevalence rate in region-ethnicity groups and a positive association with the square of the pre-reform FGC rate, although not significant. Every 1 percentage point increase in the pre-reform FGC rate correlates with approximately 5 years early adoption of criminalization.

⁹ Several scholars discuss the role of colonial history in shaping the current institutions of the previous colonies (e.g., Acemoglu, Johnson, and Robinson 2001; Banerjee and Iyer 2000; and La Porta et al. 1998, 1999).

¹⁰ The fertility rate (or total fertility rate) is an indicator of the total number of births per woman. It is defined as the number of children that would be born to a woman during her entire childbearing period. The infant mortality rate indicates the number of deaths per 1000 live births of children under one year of age. Foreign aid indicates the net official development assistance (ODA) received and is given as a percent of Gross National Income (GNI). OECD (2020) defines ODA as government aid, including grants, soft loans (at least 25 percent grant element), and technical assistance, aiming at the economic development and welfare of developing countries.

Moreover, higher population size is positively associated with late reform adoption. A 1 percent increase in the country's population size is correlated with approximately 8 years of late reform adoption, possibly due to the higher cost of implementation as the population increases. Further, a higher share of foreign aid is associated with late reform adoption, although not sure about the causality. Specifically, every 1 percentage point increase in the share of foreign aid is associated with approximately 0.8 years (or 10 months) late reform adoption. Those countries adopting reform late might not give priority to the reform initially by themselves. Later, however, donors might provide them with incentives or put pressure on them with conditional grants or loans. However, the positive association between the share of foreign aid and the timing of the reform is less pronounced for British colonies.

For non-criminalizing countries in column (6), the share of foreign aid is associated with the timing of reform adoption. Every 1 percentage point increase in the share of foreign aid is associated with approximately a 1.2-year late adoption of a reform. However, the association is less pronounced for former French colonies. Further, while countries with higher fertility rates adopt non-criminal bans later, those with higher infant mortality rates adopt it sooner. Specifically, every 1 additional birth per woman is associated with approximately 12 years of later reform adoption. Also, every 1 additional infant death per 1000 live births is associated with approximately 0.1 years of earlier reform adoption. Additionally, urbanization seems to be associated with late adoption of the reform, possibly due to low initial prevalence even before the reform, as people may access education better. A 1 percentage point increase in the urban population share is associated with approximately 1-year late reform adoption.

Moreover, to investigate the possible determinants of the reform type, we regress the dummy for criminalizing countries on the same macro-level factors as shown in column (7) of Table 3.2. A larger population size is positively associated with a higher probability of adopting criminal laws. Specifically, every 1 percent increase in the country's population size is correlated with an approximately 0.3 percentage point increase in the likelihood of criminalization. Moreover, being a former British colony is negatively associated with the probability of adopting criminal laws by approximately 1.7 percentage points. Considering these preliminary findings, we specify our empirical model in the next chapters by controlling for country-specific effects capturing the unobservable changes and the pre-reform FGC prevalence rate in region-ethnicity groups.

Further, Table 3.3 summarizes the average country-specific pre-reform characteristics and region-by-ethnicity-specific pre-reform FGC prevalence rate for criminalizing and non-criminalizing countries and their difference. The country-specific pre-reform characteristics include the total fertility and infant mortality rates, foreign aid, share of the urban population, population size, and former colonies. Overall, criminalizing countries are not significantly different from non-criminalizing countries in terms of pre-reform characteristics, except for the share of the urban population. Criminalizing countries are approximately 11 percentage points more likely to have an urban population.

3.2.4 Cohort trends in female genital cutting prevalence by reform type

In this sub-section, we compare the cohort trends in FGC prevalence rate by reform type, including criminal and non-criminal laws. FGC prevalence rate refers to the proportion of women who practiced FGC from the total sample of women in a given country or region-ethnicity group. Unlike the women in countries with non-criminal laws, those in criminalizing countries show a decreasing trend in FGC prevalence as we move from old to young cohorts at reform.

Figure 3.2 presents the average prevalence of FGC by age at reform. In the top panel of the figure, the solid and dashed lines indicate the mean FGC prevalence in criminalizing and non-criminalizing countries/states, respectively. The solid and dashed lines in the bottom panel indicate the mean prevalence of FGC below age 15 for criminalizing and non-criminalizing countries/states, respectively. Notably, FGC is less prevalent among cohorts with lower age at reform in criminalizing countries, although such a trend is not clear for non-criminalizing countries. In other words, for criminalizing countries, younger cohorts at reform exhibit a lower FGC rate because they are more likely to be affected by the reform.

Further, the difference between the top and bottom panels of the figure suggests a large number of women still experience FGC even after age 15. Because the data do not show a lower FGC prevalence among younger cohorts at reform in countries with non-criminal laws, we focus only on those countries with criminal laws to empirically examine the reform effects in chapters 4 and 5.

3.3 Data sources and related issues

3.3.1 Data sources

Our master data files include the demographic and health surveys (DHS) and the multiple indicator cluster surveys (MICS). The DHS and MICS, assisted by the United States Agency for International Development (USAID) and United Nations International Children's Emergency Fund (UNICEF), respectively, include several indicators on the wellbeing of children and women. Both datasets include internationally comparable and nationally representative household surveys. The DHS are nationally representative household surveys concerning population, health, and nutrition. The *standard* DHS surveys cover detailed information on child health, education, FGC, fertility, infant and child mortality, maternal health, maternal mortality, nutrition, and wealth, among others, which are comparable across countries. The sample sizes are large and usually conducted every five years for the sake of comparison over time.¹¹ Since the central issue of the study concerns FGC, we use the woman questionnaire of the DHS in which the respondents are girls and women aged 15 to 49.

We combine DHS and MICS for our analysis because the agents for both data sets work together for the harmonization and comparability of their surveys. As a result, we can pool them for a combined study in developing countries. Checking for their comparability in terms of stratification and clustering techniques may be important because stratification improves the precision of sampling estimates, and the clustering of the sample reduces it (Deaton 2018). For both of them, the sampling frame is limited to the population residing in fixed households despite their recommended sample size per cluster. While DHS considers about 30 to 40 women per rural cluster and 20 to 25 women per urban cluster, MICS takes 15 to 30 households per cluster (Hancioglu and Arnold 2013).¹²

Table 3.1 presents the list of post-reform surveys in both DHS and MICS that we use for our analysis. To investigate the policy impact, we rely on the post-reform surveys from both data sets. For several countries, however, the post-

¹¹ The sample size is commonly between 5000 and 30,000 for each country. See <https://dhsprogram.com/What-We-Do/Survey-Types/DHS.cfm>

¹² To consider any difference in sampling method between DHS and MICS, we keep an indicator for the data source in the data for checking robustness using DHS and MICS, separately.

reform surveys come from only DHS or MICS.¹³ To test the sensitivity of our results by data type, we further extend our analysis using each of the data separately.

3.3.2 Potential problems with measurement errors in self-reporting

One concern in using the DHS and MICS data for FGC analysis is a potential measurement error leading to the reform effect bias because FGC is a self-reported variable. Specific questions include whether the respondent ever-heard about FGC practice and then whether she, herself, ever-experienced FGC. Such questions related to FGC might be sensitive, leading the respondents to misreport their actual status (De Cao and Lutz 2018; Gibson et al. 2018). Particularly in the developing countries where institutions are weak, trust and individual freedoms are limited, and corruption is high, this problem of measurement error might be worse (Chuang et al. 2021).

One can argue that there might be cases of both underreporting and overreporting of the actual FGC status. As a result, understanding the direction of a possible bias of the estimated effect of the reform becomes difficult. Underreporting might exist if the respondents might be afraid of reporting their actual FGC status after the introduction of the ban if they had been cut (Yoder and Wang 2013). Moreover, the undertaking of the cutting a long time ago may lead to recall bias since they may not remember their cutting status well to give exact information (UNICEF 2013).¹⁴ Or they might not want to remember the issue due to the associated trauma associated with the practice. In such cases, a possible decrease in the prevalence rate of FGC after the introduction of FGC bans may not indicate the actual impact of anti-FGC policy interventions. On the other hand, over-reporting might occur if uncut women report as if they cut due to the feeling of the stigma associated with FGC abandoning (Wagner 2015).

However, several studies (e.g., Elmusharaf, Elhadi, and Almroth 2006; Morison et al. 2001) suggest that the respondents are less likely to forget and neglect their FGC status, and the self-reported and medical data are more likely to be comparable (Wagner 2015). Yoder and Wang (2013) compare the prevalence rates of cutting across cohorts over

¹³ Countries with post-reform surveys from only DHS include Egypt, Ethiopia, Kenya, Niger, Tanzania, and Uganda. We find several other countries whereby only MICS data is available after reform. These include the Central African Republic, Djibouti, Guinea Bissau, Mauritania, and Sudan. The rest 9 countries have post-reform surveys from both DHS and MICS data.

¹⁴ To address this, UNICEF (2013) suggests the use of data on women in the younger age cohort to minimize recall bias and censoring.

the years to examine the issues of under-reporting. They suggest that the prevalence rate of FGC for age cohorts 15 to 19, for example, in one survey, must be the same on average to the FGC rate for age cohort 20 to 24 in a similar study after 5 years if the respondents tell the truth. Using this technique, they find such evidence of under-reporting in only a few countries.

Unlike Camilotti (2016), who uses the sample of daughters below age fifteen, this study focuses on respondents aged between 15 and 49. The average age at cutting was about 6, and the cutting period would be 9 to 43 years before the interview time on average. Since the practice was undertaken during their childhood, without their consent a long time ago, the respondents are unlikely afraid of lawbreaking in truth-telling regarding their actual status of cutting. Because they could blame their parents and the previous generation who decided on their FGC status. At the same time, they might know that their parents and previous generation will not be accountable as the practice was during their childhood a long time ago.

3.3.3 Age at cutting and non-reporting of age at cutting

Most women undertake FGC during childhood before marriage; however, data for age at cutting is missing in several surveys. Figure 3.3 presents the percent distributions for the age at FGC in criminalizing and non-criminalizing countries. In this analysis, we exclude two countries without age at cutting information, one from criminalizing and another one from non-criminalizing countries. Those countries without age at cutting data in any of their surveys include Djibouti and Uganda. The rest of the countries have the age at cutting data in at least one of their surveys. The histogram shows that most women in criminalizing countries experience FGC before age 19. Observing the jumps in FGC prevalence across the age at cutting, we consider several thresholds for analyzing the exposure to an FGC ban. Ages 0, 4 to 5, 6 to 7, 9 to 10, and 11 to 12 are crucial for girls in criminalizing countries. However, we also check if trend change occurs at these thresholds using event-study graphical analysis later.

In non-criminalizing countries, however, most women undertake FGC from age 2 to 15, more precisely from age 4 to 15. For these countries, the jumps appear in ages 3 to 4, 5 to 10, 11 to 12, and 14 to 15. Appendix Figures A2.1 and A2.2 show the percent distributions for the age at cutting separately for each country.

Table 3.4 presents the association between the pre-reform FGC prevalence rate in the region-ethnicity groups and the non-reporting of the age at cutting. Considering all the 20 countries with FGC ban in columns (1) to (3), the results suggest that higher pre-reform FGC prevalence is associated with more chance of reporting the age at cutting (that is, less non-reporting). Seemingly, demographic controls capture the non-linearity for the 20 countries. Columns (4) and (5) report the results of similar analysis, separately for criminalizing and non-criminalizing countries. For these analyses, the specifications include the full set of controls. For criminalizing countries, higher pre-reform FGC prevalence is associated with less chance of reporting the age at cutting, suggesting non-linearity. For non-criminalizing countries, however, the negative effect of pre-reform FGC prevalence on non-reporting becomes even stronger in the region-ethnicity groups with higher pre-reform FGC prevalence. As a result, one who wants to examine the reform effect by age at cutting should be cautious of non-reporting bias of the age at cutting.

3.3.4 Pre-reform FGC prevalence rate across region-ethnicity groups

We use the pre-reform surveys to compute the pre-reform FGC prevalence across countries, regions, and ethnicities. However, we lack pre-reform surveys for several countries. As shown from columns (4) to (7) of Appendix Table B3.1, for the countries with pre-reform surveys, we use the latest pre-reform survey to compute the average pre-reform FGC prevalence rate. We adjust the cohorts to create identical young cohorts with those potentially affected by reforms for 22 years. Considering Benin, for example, we use the cohorts aged 15 to 37 in DHS 2001 survey to calculate the mean FGC prevalence.

For those countries without pre-reform surveys, we determine the average pre-reform FGC prevalence using the surveys just after the reforms and adjusting cohorts. For Burkina Faso, for example, we use the cohorts aged 17 to 39 in 1998-99 DHS survey to approximate for FGC prevalence in 1995, a year before the reform. We skip those aged 15 to 16 to adjust for the two years difference between the 1996 year of reform and the 1998-99 DHS survey year for Burkina Faso. Moreover, we exclude those aged 37 or above at reform because they become too old to compare with the potentially partially affected cohorts aged 22 or below at reform.

Columns (2) and (3) of Appendix Table B3.1 present the mean and standard deviations for the pre-reform FGC prevalence by the country for both criminalizing and non-criminalizing countries. Among the criminalizing countries, Egypt and Djibouti have the highest pre-reform FGC prevalence rate, exceeding 90 percent for each of them. On the other extreme, criminalizing countries with the lowest pre-reform rates include Ghana, Niger, and Congo, with less than 10 percent rate for each of them. For the countries with a non-criminal ban, Guinea and Sudan appear to have the highest FGC prevalence prior to the reform, about 98 and 92 percent, respectively. Uganda is the country with the least pre-reform FGC prevalence among the non-criminalizing countries, with a 1 percent prevalence rate.

Appendix Figure A3.2 reports the percent distributions for the pre-reform FGC prevalence across regions and ethnicities for each country.¹⁵ We observe significant variation in the pre-reform FGC prevalence for most of the countries except few countries with skewed distribution.

Moreover, Figure 3.4 presents the percent distributions for the pre-reform FGC prevalence rate in the ethnicity-by-region group, separately for criminalizing and non-criminalizing countries. The figure illustrates a considerable variation in FGC prevalence rate across regions by ethnicities before the reform for both criminalizing and non-criminalizing countries. We leverage the varying impact of the reform to identify the effect of FGC ban exposure.

3.3.5 Managing the issue with the inconsistent region and ethnicity information

We construct dummies for 130 regions and 88 ethnic groups, although not presented in the table for the sake of brevity.¹⁶ Previous studies suggest that these factors are essential determinants of FGC (Camilotti 2016; Howard and Gibson 2017; UNICEF 2013; Wagner 2015; Yoder and Wang 2013). We take the missing values in these dummies as reference groups due to the lack of this information in some countries or specific surveys.¹⁷ The information on region, religion, ethnicity, and residence is for the current household rather than the parental background. However, we control

¹⁵ For Nigeria and Sudan, we compute the pre-reform FGC prevalence across regions and ethnicities for each state with reforms in the countries.

¹⁶ We compile relatively smaller regions and ethnic groups together into one region and one ethnic group as other regions and other ethnic groups, respectively.

¹⁷ For instance, ethnicity information is missing for Egypt and Tanzania.

for these variables considering that they reflect the respective information on the respondents' family of origin since marriage usually happens within the same class (Burdett and Coles 1997; Wagner 2015).

The major challenge is to maintain consistent region and ethnicity coding for all the 20 countries across data sets and surveys. We prepare an excel table for this task to check for the list of consistent regions and ethnicities for each survey in both the DHS and MICS. Then, we label those inconsistent regions and ethnicities as "other groups." For those countries having a long list of regions or ethnicities, we recode them consistently by regrouping several regions and ethnicities with a small sample size into "other groups."

CHAPTER 4

THE EFFECT OF CRIMINALIZING FEMALE GENITAL CUTTING ON THE PRACTICE

4.1 Defining cohorts of interest and policy variables

Based on the period of exposure to an FGC reform, Table 4.1 classifies cohorts into three; the least affected, partially affected and fully affected cohorts. The "least affected" cohorts include those females aged 23 or above at reform year, and as a result, they would be too old to be affected by the reform. Because most women often undertake FGC before age 19, as shown in Figure 3.3, we expect the reform is more effective for those females aged 19 or below at the reform year. To be more precise on the threshold, however, we check for a trend break around this threshold using formal event-study analysis in the next sub-sections. As a result, we find that the trend becomes almost flat and approaches zero once it arrives at age 23 at reform, as shown in Figure 4.1. The partially affected cohorts include those aged 1 to 22 at reform, and the fully affected are those born at the reform year or after. For our analysis, we focus on the comparison between the least and partially affected cohorts and exclude the fully affected cohorts because they come only from 7 countries and cover less than 2 percent of our total sample.

4.2 Identification strategy and empirical specification

In this chapter, the dissertation aims to empirically estimate the effect of criminalizing FGC on African women's FGC experience. To investigate the effectiveness of the criminalization policy, first in sub-section 4.2.1, we use a difference-in-differences event-study approach exploiting variation in each cohort and pre-reform FGC prevalence rate region-ethnicity groups. The identification strategy focuses on comparing each similar cohort across pre-reform FGC

prevalence region-ethnicity groups. Then, in sub-section 4.2.2, we focus on a difference-in-differences method comparing young and old cohorts for identification. Sub-section 4.2.3 discusses possible threats to the identification.

4.2.1 Difference-in-differences event-study approach

In this sub-section, we undertake an event study analysis to examine whether the timing of FGC reform matters and to check for a trend break at a certain age at reform. In this model, we allow for the effect of criminalization to vary with the individual's age at reform and pre-reform FGC prevalence rate region-ethnicity groups. For this purpose, we construct our policy variables as a series of dummies indicating the age at reforms. Moreover, we divide the pre-reform FGC prevalence rate into terciles; top, middle, and bottom indicating high, medium, and low prevalence. Then, we include all the interactions of top, middle, and bottom prevalence rate terciles with cohort dummies, except the last dummy (age 36 at reform) that we omit as a reference and the full set of fixed effects and other controls to estimate the event study analysis. Following Duflo (2001), we undertake the following interaction terms analysis to estimate cohort-by-cohort contrasts. The equation to be estimated is therefore

$$F_{acgit} = \sum_{a=1}^{35} (d_a \times T_g) \alpha_{1a} + \sum_{a=1}^{35} (d_a \times M_g) \alpha_{2a} + \sum_{a=1}^{35} (d_a \times B_g) \alpha_{3a} + \alpha_4 P_g + \gamma_{ac} + X'_{acgit} \pi + \varepsilon_{acgit} \quad (4.1)$$

Here F is the probability of ever experiencing FGC, and a , c , g , i , and t indicate the age at the reform, country, region-ethnicity group, individual, and survey year.

The variables T , M , and B are *dummy* indicators if females are in the top, middle, and bottom terciles of the pre-reform FGC prevalence rate region-ethnicity groups. d is a dummy indicating whether the female is aged a at reform. We include 35 age-at-reform dummies interacted with top, middle, and bottom tercile dummies of the pre-reform FGC prevalence region-ethnicity groups. The control group we omit from the regression includes the females aged 36 at reform. We consider each coefficient α_{1a} as an estimate of the criminalization effect on a given cohort in the top tercile, α_{2a} in the middle tercile, and α_{3a} in the bottom tercile.

P represents an indicator for pre-reform FGC prevalence rate in region-ethnicity groups. The coefficient α_4 indicates the effect of pre-reform differential prevalence across regions-by-ethnicities on the current prevalence rate. The vector X' includes indicators for the region, religion, residence (rural/urban), and the fixed effect for a given survey/interview year. ε is the error term. The Greek letters are coefficients; in particular, γ is a cohort-by-country fixed effect, capturing the cohort-specific effects of pre-reform country characteristics. Our analysis limits the sample to the partially- and least-affected cohorts while excluding the fully-affected ones (that is, the females born at or after the reform) because most FGC bans are too recent to let researchers observe the fully-affected females in the surveys.

To improve the precision of the criminalization effect, we impose restrictions on the control group. The strategy is to contrast females across terciles of pre-reform FGC prevalence; top versus bottom and middle versus bottom. We compare those females aged a at reform in the top tercile, for example, with those females age a in the bottom tercile. Similarly, we compare those females aged a at reform in the middle tercile with those females aged a in the bottom tercile. We keep this contrast for each cohort, focusing on young cohorts aged 1 to 25 at reform.¹⁸ In this analysis, we omit all those females in the bottom tercile as a reference group. The revised equation to be estimated is then,

$$F_{acgit} = \sum_{a=1}^{25} (d_a \times T_g) \alpha_{1a} + \sum_{a=1}^{25} (d_a \times M_g) \alpha_{2a} + \alpha_4 P_g + \gamma_{ac} + X'_{acgit} \pi + \varepsilon_{acgit} \quad (4.2)$$

To consider possible shocks that are the same across all observations in a given country and cohort, we cluster the standard errors at the country-cohort level. However, we report additional results as sensitivity tests by clustering at cluster/village level following Wagner (2015) focusing on the within-cluster comparison of females for FGC analysis. The need to cluster at the cluster/neighborhood level is due to the sampling design in both DHS and MICS data, a clustered sampling.¹⁹ Also, we report results from clustering at the region-by-ethnicity level.

¹⁸ Based on the event-study graph in Figure 4.1, the "young" cohorts include those females aged 1 to 22 at reform. However, in this specific analysis, we include some more "old" cohorts, aged 23 to 25, to check if the top to bottom or middle to the bottom difference among these cohorts becomes almost insignificant.

¹⁹ Cameron and Miller (2015) discuss the loss of estimator precision due to cluster sampling, *design effect*, and the need to cluster at the primary sampling unit level at the minimum. Cluster is the primary sampling unit, an enumeration area whereby households are randomly drawn (Wagner 2015).

4.2.2 Difference-in-differences approach comparing young and old cohorts

In this analysis, we compare young cohorts aged 1 to 22 with those old cohorts aged 23 to 36 at reform, focusing on cohorts in the top and middle terciles of the pre-reform FGC prevalence groups. We take the age 22/23 threshold to classify the cohorts into young and old based on the results in Figures 4.1 and 4.2 and Table 3.2. The coefficients become almost zero for the cohorts aged 23 and above at reform in all terciles. As a result, we omit all those old females aged 23 to 36 as a reference group. The revised equation to be estimated is then,

$$F_{acgit} = \alpha_1 Young_a \times T_g + \alpha_2 Young_a \times M_g + \alpha_3 P_g + \gamma_{ac} + X'_{acgit} \pi + \varepsilon_{acgit} \quad (4.3)$$

Here *Young* is a dummy indicating whether the individual belongs to the young cohort, aged 1 to 22 at reform in a criminalizing country *c*. Our parameters of interest include the coefficients α_1 and α_2 , measuring the causal effects of the policy reform on FGC practice. To interpret the coefficients as causal impacts of FGC policies, we assume no differential trends in FGC prevalence between partially-affected and least-affected females by the reform. In addition to the common trends assumption, the timing of reform adoption should be exogenous for the validity of our research design. We discuss the validity of these assumptions in the following section.

4.2.4 Threats to identification

One concern of identification is the endogeneity of the FGC reform adoption timing and reform type for each country. In chapter 3, we find that several pre-reform country and region-ethnicity characteristics --- such as pre-reform FGC prevalence rate, population size, foreign aid dependency, and colonial ties induce FGC reform at certain timing and criminal law adoption. Consequently, the estimated impact of the policy using variation over time and across within-country region-ethnicity groups cannot be causal. Thus, the validity of our research design relies on the exogeneity of

the timing and type of the reform adoption. For this purpose, we control for the pre-reform FGC prevalence at the region-ethnicity group and country-by-cohort fixed effect to capture the possible differential trend in unobservable country characteristics.

Our strategy to include country-by-cohort fixed effect addresses another concern of potential spillover effects from neighboring countries. Specifically, the concern is the possible endogeneity of reform adoption timing if reform adoption in neighboring countries affects the timing of reform adoption for each country.

An additional concern of identification is the possible coincident expansion of other programs. The confounding programs include several anti-FGC interventions such as training of health personnel, female students' education, multifaceted community activities, and village empowerment. The interventions are mostly by NGOs and target specific regions with high FGC prevalence within countries. For instance, the multifaceted community programs target the Afar region in Ethiopia and the Enugu state in Nigeria. Moreover, an empowerment program through education, called Tostan Program, targets the Kati region in Mali, the Kolda region in Senegal, and the Central Mossi region in Burkina Faso (Berg and Denison 2012; Diop and Askew 2009). To address such possible confounders at regional levels, we present a specification controlling for the region-specific cohort trend.

4.3 Results for the effect of criminalizing female genital cutting on the practice

4.3.1 Results for the difference-in-differences event-study estimated effects

Figure 4.1 presents the criminalization effect for each cohort and terciles of the pre-reform FGC prevalence on ever-experiencing FGC. The impact of the reform on FGC practice increases with pre-reform FGC prevalence, moving from bottom to top terciles. Notably, the effect of criminalization is totally insignificant among females in the bottom tercile. As a result, cohorts in the bottom terciles can be a good control group. Moreover, the impacts become smaller if the exposure begins in late childhood and even eventually diminishes if it starts later in life, implying that the reform may not matter for the very old cohorts as we expect. The age at reform trend becomes flat for the least-affected cohorts,

particularly after age 22 at reform. Also, the coefficients become almost zero for the least-affected cohorts, particularly in the middle and bottom terciles.

Figure 4.2 and Table 4.2 present similar results for the estimated effect of criminalization on ever-experiencing FGC by cohort. In this analysis, we restrict the control group to those females in the bottom tercile, as shown in equation (4.2). Then, we compare each similar cohort in the top and middle tercile of the pre-reform FGC prevalence region-ethnicity groups with their counterparts in the bottom tercile. Notably, as shown in Figure 4.2, the stronger impacts of the reform on FGC practice appear among the youngest cohorts in both top and middle terciles, particularly for those aged below 14 at reform. Moreover, the policy impact on each cohort is stronger among the females in the top tercile than the middle tercile.

Table 4.2 shows the criminalization effect by cohort with several specifications. Columns (1) to (4) present the coefficients of the interactions between age at reform and the pre-reform FGC prevalence rate across region-by-ethnicity groups in four specifications. Including the year of interview and country-by-cohort fixed effects and controlling for pre-reform FGC prevalence rate, column (1) reports the coefficients of the interactions between cohort and top and middle terciles of pre-reform FGC rate from a single regression. Additionally, controlling for residence and religion, column (2) presents the coefficients from a separate regression. Also, we include region fixed-effect in the third specification shown in column (3). Finally, we control for region-specific linear cohort trends in the last column, column (4).

The estimated effects are negative for all columns before age 23 at reform, except an insignificant result for age 22 in the middle tercile in column (3). Focusing on column (4) with the full set of controls, all the estimates of the top tercile are significant before age 23 at reform. Also, all the middle tercile coefficients are significant before age 16 at reform, except at age 10. Overall, we observe a substantial impact of criminalization for those aged 1 to 8 at reform and a moderate impact for those aged 9 to 13 in both the top and middle terciles of the pre-reform FGC prevalence rate groups. Since age 14 at reform, we still observe a lower effect of the criminalization before age 23 at reform in the top tercile. However, the impact almost vanishes after age 15 at reform in the middle tercile, except at age 18.

4.3.2 Results for the difference-in-differences comparing young and old cohorts

In this sub-section, we report the criminalization effect focusing on the comparison of young and old cohorts. Table 4.3 illustrates the basic idea behind exploiting the variation in pre-reform FGC prevalence across region-ethnicity groups as an identification strategy. Columns (1) to (3) report the mean FGC among females in the top, middle, and bottom terciles for young and old cohorts. Columns (4) and (5) show the unconditional difference-in-difference results.

Assuming no systematic difference in the FGC decrease across region-ethnicity groups in the absence of the reform, we can consider the differences in these columns as causal effects of the reform. In panel A, a young female in the top tercile of the pre-reform FGC prevalence region-ethnicity group undertakes FGC 11 percentage points less likely than her counterpart (old female). These findings provide preliminary evidence that the reform differentially influences the likelihood of FGC experience by the timing of reform across region-ethnicity groups.

To test the identification assumption that old cohorts are least affected, we undertake a falsification experiment in panel B of Table 4.3. For this experiment, we compare cohorts age 23 to 29 at reform with those aged 30 to 36 at reform. As expected, the estimated difference-in-differences are almost zero, suggesting the validity of our design. However, we provide more convincing evidence by estimating the effect conditional on several covariates.

Table 4.4 presents the FGC criminalization effect on the probability of ever-experiencing FGC. In column (1), the specification includes the year of interview and country-by-cohort fixed effects and pre-reform FGC prevalence rate at the region-ethnicity group. We control for residence (rural/urban) and religion in column (2). Further, in column (3), we control for region-specific effects possibly correlated with the reform. Finally, column (4) presents the results capturing the region-specific cohort trend.

In panel A, we compare females aged 1 to 22 at reform in the top-and middle-terciles of pre-reform FGC rate with females aged 23 to 36 at reform in the same terciles, and with females aged 1 to 22 at reform in the bottom-tercile. In the first cell of column (4), the result suggests that those females aged 1 to 22 at reform in the top-tercile are 12.7 percentage points, about 37 percent of the mean FGC in criminalizing countries, less likely to experience FGC than old females. The estimate in the second cell of column (4) shows that those females aged 1 to 22 at reform in the middle-

tercile are 5.2 percentage points, about 15 percent of the sample mean, less likely to experience FGC than their counterparts.

Panel B of Table 4.4 reports the results of the falsification experiment using only old cohorts. In this panel, we compare females aged 23 to 29 at reform with those females aged 30 to 36 at reform. However, the criminalization effect becomes very small and insignificant for females in both the top and middle-terciles.

4.3.3 Heterogeneity in the effectiveness of criminalization

This sub-section compares the FGC criminalization effects across residence and religion, as it may shed some light on the mechanisms behind the basic results. Table 4.5 divides the whole sample by residence into rural or urban areas and religion as Muslims or Christians. All the columns report the difference-in-difference estimated effects of FGC criminalization, including the full set of controls. Columns (1) and (2) of Table 4.5 show that criminalization is highly effective in reducing the probability of ever-experiencing FGC among those females living in both rural and urban areas in the top tercile of the pre-reform FGC prevalence region-ethnicity groups. In the middle tercile, the effect of criminalization is less effective among females in both rural and urban areas, particularly for those residing in urban areas. If we consider the effect in the top tercile, for example, criminalization reduces the likelihood of experiencing FGC by 11 percentage points (=23 percent) and 16 percentage points (= 42 percent) among females residing in rural and urban areas, respectively. However, in the middle tercile, the reduction becomes by 6 percentage points (14 percent) and 4 percentage points (12 percent) for rural and urban areas, respectively. Moreover, Figure 4.3 presents the results of the criminalization effect by cohort, separately for rural and urban areas. Notably, the reform is effective for all cohorts before age 23 at reform in the rural areas. In urban areas, however, criminalizing FGC is effective for youngest cohorts, particularly before about age 14 at reform.

Christians seem to be more sensitive to criminalization than Muslims, as shown from columns (3) and (4) of Table 4.5 and Figure 4.4. The first cell in column (3) of Table 4.5, for example, suggests those Christian females aged 1 to 22 at reform in the top tercile are approximately 16 percentage points, about 46 percent of the Christian sample in

criminalizing countries, less likely to experience FGC than old females. For the case of Muslim females in the first cell of column (4), the reform reduces the likelihood of experiencing FGC among Muslims by approximately 13 percentage points, about 38 percent of the sample mean. The reason may be because many Muslims in Africa consider the practice as a requirement for their religion, although religious books do not explicitly mention it (Antonazzo 2003). Both of the heterogeneity analyses pass the falsification test, at least for the cohorts in the top tercile, as shown in panel B of Table 4.5.

4.3.4 Sensitivity analysis to the results for the effect of criminalization

We test the sensitivity of our results by estimating the effect of FGC criminalization on the practice for different age groups at survey/interview. Figure 4.5 presents the criminalization effect per cohort, separately for those females aged 15 to 29 and those aged 30 to 49 at survey. Notably, criminalization is more effective in reducing the likelihood of ever-experiencing FGC for those aged 15 to 29 at survey. For those females aged 30 to 49, the reform becomes less effective because they become too old to be affected by the reform. We provide further results for such analysis by age at survey in Appendix Figures A4.1 and A4.2. All these results show the reform is less effective for older cohorts at the interview because they are more likely to be in the least-affected category.

Detailed information on law enforcement cases is not widely available. Relying on the available data, we first categorize the degree of enforcement for each country as weak and strong, as shown in Table 3.1 of Chapter 3. For this purpose, we compute the median number of arrests and prosecutions in the countries concerned as a threshold to consider strong or weak enforcement. Then, we estimate the criminalization effect for each cohort, separately for countries with strong and weak enforcement cases. Figure 4.6 presents the estimated effects of criminalization on the probability of ever-experiencing FGC, considering the degree of law enforcement. As expected, our results suggest that criminalization is more effective in countries with strong enforcement.

Further, we provide results for the estimated effects of criminalization by excluding several countries with less within-country variation in pre-reform FGC prevalence. We identify five countries with less variation in pre-reform FGC prevalence, including Djibouti and Egypt having more than 90 percent prevalence rates, and Ghana, Niger, and Togo,

with less than 10 percent prevalence rates. Figure 4.7 reports the results, excluding separately Djibouti and Egypt, and Ghana, Niger, and Togo. However, both results are almost similar to the result without excluding them.

Additionally, we report results for the effect of criminalization by several ways of clustering. In our main analysis, we cluster the standard errors at the country-cohort level to consider possible shocks that are the same across all observations in a given country and cohort. Figure 4.8 presents results by clustering at the region-ethnicity level, village/cluster, and country level. However, the baseline result remains robust for both cases, suggesting even more precise estimates for clustering at the village level. Clustering at region-ethnicity and country levels separately lead to almost similar results.

Moreover, we estimate the effect of criminalization using DHS and MICS data sets separately. Figure 4.9 presents the reform exposure effects on the probability of ever-experiencing FGC, showing significant and strong effects, using both DHS and MICS. Further, the impact is stronger among females in the top tercile of the pre-reform FGC prevalence than the middle tercile for both data sets.

Revising the event-study estimated effect of criminalization by including all the available old cohorts in the data would not alter our results in the main analyses. Appendix Figure A4.3 presents the criminalization effect for each cohort and terciles of the pre-reform FGC prevalence on ever-experiencing FGC, including all cohorts. In this case, we consider those aged 47 at reform, the oldest cohort at reform available, as a control group. Similar to the main analysis, the reform's effect increases with the pre-reform FGC prevalence, moving from bottom to top terciles. Also, the effect of criminalization is insignificant among females in the bottom tercile. Moreover, the impacts become smaller for older cohorts at reform.

Further restricting the control group to those females in the bottom tercile, as shown in Appendix Figure A4.4, we compare each similar cohort in the top and middle tercile of the pre-reform FGC prevalence region-ethnicity groups with their counterparts in the bottom tercile. As a result, we find that the stronger impacts of the reform on FGC practice appear among the youngest cohorts in both top and middle terciles, same to the results of the main analyses. Also, the policy impact on each cohort is stronger among the females in the top tercile than the middle tercile.

Finally, although we do not focus on examining the effectiveness of non-criminal laws, we report an event-study estimated effect of non-criminal laws in Appendix Figure A4.5. Notably, non-criminal laws seem to be ineffective,

except for few young cohorts at reform. However, we also observe insignificant results for the very youngest cohorts, possibly due to data problems.

4.4 Potential mechanisms for criminalization to affect female genital cutting

Estimating the effect of criminalizing FGC on the practice, we find that the reform substantially reduces the likelihood of ever-experiencing FGC. Given the deep-rooted nature of the FGC custom, this finding is surprising and contrary to several studies suggesting that modern laws are ineffective in changing social norms (e.g., André and Platteau 1998; Platteau 2009). However, there could be two mechanisms through which FGC criminalization can be effective. The first one is through the fear of arrests and prosecutions expecting these criminal laws to be enforced although actual enforcement cases are generally limited.²⁰ Because people perceive the laws as potentially enforceable due to the rumors and gossips in the community and fear of possible punishment if they practice FGC. As a result, the risk of being caught cutting raises the cost of FGC (Camilotti 2016; Shell-Duncan et al. 2013).

The second mechanism is in line with the findings of Aldashev et al. (2012a, 2012b), suggesting that criminalizing FGC can be effective even with low enforcement. The reforms serve those who are willing to abandon FGC practice as a helping hand from outside. In this case, the policies can enhance the bargaining power of the victims by creating an exit option out of the custom trap (Aldashev et al. 2012a, 2012b). Thus, these criminal laws against FGC help those who do not support FGC resist social pressure while abandoning the practice (Camilotti 2016; Shell-Duncan et al. 2013).

Moreover, the criminalization policy is more effective in the region-ethnic groups with a higher pre-reform FGC prevalence rate. This finding seems to support the argument that FGC is a social coordination norm. As a result, each community (or region-ethnic group in our case) has a threshold of pre-reform FGC prevalence rate to switch from practicing to not practicing FGC. Considering FGC as a social coordination norm, the substantial reduction in the likelihood of ever-experiencing FGC in these region-ethnic groups comes probably due to a sufficiently large number

²⁰ For the reports about each country's enforcement experience, see <https://www.28toomany.org/research-resources/>.

of families who agree to abandon FGC induced by the reform (Mackie 1996). The result suggests that the region-ethnic factors are still crucial for the households to decide whether to practice FGC, although several studies (e.g., Bellemare et al. 2015; Efferson et al. 2015; Novak 2020) increasingly show the dominant role of individual and household level factors.

Additionally, our findings show that criminalizing FGC is more effective for younger cohorts at reform. If the reform is at a younger age of their daughters, sufficiently younger than the prime marriageability age of their daughters, then parents might more likely to abide by the criminalization and consider alternative pre-marital investment instead of FGC to increase their daughters' value at the marriage market, such as education and early marriage, to be examined in Chapter 5. However, if the reform is at an older age of their daughters, sufficiently close to or older than the prime marriageability age of their daughters, then it would be difficult for the parents to consider other options as they hurry to find husbands for their daughters. As a result, they might take the risk of potential punishment to maintain the value of their daughters, leading the reform to be less effective.

CHAPTER 5

THE IMPACT OF CRIMINALIZING FEMALE GENITAL CUTTING ON WELL-BEING

5.1 Descriptive analysis for the female well-being outcome variables

We present summary statistics for female well-being outcome variables in Table 5.1. We categorize the variables into respondents' reproductive health, marriage and fertility, education and work status, general health, and spousal education and work status, and family wealth. About 11 percent of the women ever experience a genital problem, either a genital discharge or ulcer, in the last 12 months before the interview. Moreover, about 24 percent of women lost at least one child, respectively.

Most of the women in our sample ever-married and began childrearing before the interview, about 78 and 72 percent of females, respectively. Moreover, child marriage and fertility are common among females in our sample. About 34 and 14 percent of the respondents married and gave birth, respectively, before they turned 18.²¹

To observe the level of education after experiencing FGC, we focus on secondary education or more level for the analysis of female education. About 27 percent of females have secondary education or more. Moreover, the literacy rate among females in our sample is 44 percent. In addition, almost 60 percent of the women in the sample engaged in working.

Looking into the body-mass index (BMI) of females as a general health indicator, 26 and 11 percent of females in our sample are over-weighted and underweight, respectively.²² Moreover, a large share of females, about 43 percent, are anemic in the sample. The data on spousal attributes come from DHS only, showing that 44 percent of females

²¹ Also, the data show the availability of marriage practice and childrearing even before age 15, although not reported for the sake of brevity, about 14 and 5 percent of those females reporting the age at marriage. However, the age at marriage is missing for almost 30 percent of the ever-married sample in criminalizing countries.

²² We use the body mass index (BMI) from DHS to compute the proportion of underweight and overweight females from the sample. BMI is a measure of acute nutritional status calculated as weight in kilograms divided by the square of height in meters, and the normal values range from 18.5 to 24.9 (Rutstein and Rojas 2006). The underweight proportion includes those females with less than 18.5 BMI, whereas the overweight takes those with a BMI of 25 or more. Anemic can be either severe, moderate, or mild anemic based on the hemoglobin level (mean = 12 dc/l) for the female.

married to husbands with primary education or more. Further, about 22 percent of the females with observable spousal data matched to husbands with secondary education or more. Moreover, about 92 percent of the respondents with spousal data have working husbands, and 17 percent of women have husbands with skilled type work. Also, one-third of the respondents live in a poor family, classified as poorest or poor based on the quintile distribution of household asset holdings.²³

5.2 Empirical specification for the reduced-form analysis of criminalization effect on female well-being

5.2.1 Difference-in-difference event-study approach

In this sub-section, we examine the effect of criminalizing FGC on female well-being, applying an event-study analysis. Specifically, we investigate the criminalization impact per cohort. In this analysis, we allow for the effect of criminalization to vary with each individual's age at reform, in addition to the variation in the pre-reform FGC prevalence rate region-ethnicity groups and across countries. The strategy is to contrast females across terciles of pre-reform FGC prevalence; top versus bottom and middle versus bottom. We compare those females aged a at reform in the top tercile, for example, with those females age a in the bottom tercile. Similarly, we compare those females aged a at reform in the middle tercile with those females aged a in the bottom tercile. For this purpose, we omit the cohorts in the bottom tercile as reference groups. As a result, our policy variables include the interactions of top and middle terciles of the pre-reform FGC rate with a series of dummies indicating the age at reforms. The reduced-form equation to be estimated is, therefore,

$$Y_{acgit} = \sum_{a=1}^{22} (d_a \times T_g) \alpha_{1a} + \sum_{a=1}^{22} (d_a \times M_g) \alpha_{2a} + \alpha_4 P_g + \gamma_{ac} + X'_{acgit} \pi + \varepsilon_{acgit} \quad (5.1)$$

²³ Based on the DHS wealth index constructed using family asset holdings, households are classified into five categories, including the poorest, poorer, middle, richer, and richest. The type of asset holding for the construction of the index includes house floor type, refrigerator ownership, electricity and water supply and sanitation, vehicle type, agricultural land ownership, persons per sleeping room ownership of radio, television, and telephone, having domestic servant, and other country-specific items (Rutstein and Rojas 2006). We construct a lower wealth dummy for those poorest and poorer and a higher wealth dummy for those richer and richest quantiles.

where Y indicates female well-being outcomes, including health, education, and family formation. The variables T and M are *dummy* indicators if females are in the top and middle terciles of the pre-reform FGC prevalence regions. d is a dummy indicating whether the female is aged a at reform. To focus on young cohorts, we include 22 age-at-reform dummies interacting with top and middle tercile dummies of the pre-reform FGC prevalence region-ethnicity groups. We omit those cohorts in the bottom tercile as reference groups. We consider each coefficient α_{1a} as an estimate of the criminalization effect on a given cohort's outcome in the top tercile and α_{2a} in the middle tercile. P represents an indicator for pre-reform FGC prevalence rate across regions-by-ethnicities. The coefficient α_3 indicates the effect of pre-reform differential prevalence across regions-by-ethnicities on the current prevalence rate.

5.2.2 Difference-in-difference approach comparing young and old cohorts at reform

In this sub-section, we categorize the cohorts into young and old based on the degree that the reform affected them. As shown from the event-study analysis in Chapter 4, young and old cohorts include those aged 1 to 22 and 23 to 36 at reform, respectively. Instead of examining the criminalization effect on outcomes per cohort, we estimate the effect by contrasting young and old cohorts. Also, we consider the variation in the effectiveness of the reform across region-ethnicity groups within countries. As a result, our difference-in-difference reduced-form specification will be:

$$Y_{acgit} = \alpha_1 Young_a \times T_g + \alpha_2 Young_a \times M_g + \alpha_3 P_g + \gamma_{ac} + X'_{acgit} \pi + \varepsilon_{acgit} \quad (5.2)$$

Here, *Young* is a dummy indicating whether the individual belongs to the young cohort, aged 1 to 22 at reform. The parameters of interest include the coefficients α_1 and α_2 , measuring the causal effects of the policy reform on female well-being, including several human capital and socio-economic outcomes. To interpret the coefficients as causal impacts of FGC policies, we assume no differential trends in FGC prevalence and well-being outcomes between partially-affected and least-affected females before the reform.

5.3 Results for the reduced-form estimated effects of criminalization on female well-being

5.3.1 Impact of criminalization on health outcomes

In this sub-section, we report the estimated effects of FGC criminalization on females' reproductive health for each cohort. The reproductive health indicators include genital problems, discharge or ulcer, and child death. Figure 5.1 presents the effect of FGC criminalization for each young cohort on the probability of experiencing a genital problem within the year before the interview and ever-experiencing child death. The results in both top and middle terciles show that criminalization does not significantly affect the chance of experiencing genital problems. Moreover, the reform does not seem to affect the likelihood of experiencing child death, although we observe some significant cohorts in the middle tercile. Also, we obtain similar overall insignificant results by focusing on the comparison between young and old cohorts, as shown in panel B of Table 5.2. However, Wagner (2015) finds a significant effect of FGC on genital problems, which may be due to omitted variable bias.

Notably, the stronger impacts of the reform on child death appear among the youngest cohorts in the top tercile, particularly for those aged 1 to 5 at reform. Moreover, for these youngest cohorts, the policy impact on each cohort is stronger among the females in the top tercile than the middle tercile.

For general health outcomes, we pick the body-mass index (BMI) and anemia status indicators. Figure 5.5 and columns (3) to (5) of panel B in Table 5.2 present the estimated effects of criminalization on being underweight, overweight, and anemic. The criminalization effect is insignificant on the likelihood of being underweight and anemic, although it slightly increases the chance of being overweight. Overall, our findings do not show a convincing significant effect of the reform on the general health outcomes, supporting Wagner (2015)'s findings. Further, we present the falsification tests for the results on own health outcomes in Table 5.4, suggesting that the effects on all the outcomes pass the tests.

5.3.2 Impact of criminalization on marriage and fertility outcomes

In this sub-section, we present the estimated effects of criminalization on marriage and fertility. Figures 5.2 reports the per-cohort criminalization effect on overall marriage and fertility. Notably, criminalization leads to an increase in the probability of marrying and childrearing. The effects are more pronounced among the youngest cohorts, particularly for those aged below 16 at reform. This shows that the reform leads to a more increase in the probability of early marriage than overall marriage and fertility. To further examine the effect on early marriage and fertility, we report a result showing the effect of the reform on the probability of marrying and childrearing before age 18. Figure 5.3 provides evidence that criminalization leads to early marriage and fertility before age 18.

Columns (3) to (5) of panel A in Table 5.2 provide further evidence for the criminalization effect on marriage and fertility outcomes using young-old cohort contrasts. Young cohorts aged 1 to 22 at reform in the top tercile, for example, are 6.8 and 4.9 percentage points (about 7 and 5 percent of the control means) more likely to marry and start childrearing than old cohorts aged 23 to 36 at reform. Also, for cohorts in the middle tercile, criminalization increases the chance of marriage and fertility by 3.6 percentage points, about 4 to 5 percent of the control means. Further, criminalization substantially increases the likelihood of marrying before age 18 for cohorts in both top and middle terciles. Specifically, the reform increases the risk of early marriage by approximately 8 and 5 percentage points for cohorts in the top and bottom terciles, about 20 and 13 percent of the control means. However, these results are contrary to the findings by Wagner (2015), suggesting that cut women marry before uncut women and are more fertile. Moreover, we present the falsification tests for the results on marriage and fertility outcomes in Table 5.4, suggesting that the effects on all the outcomes pass the tests.

5.3.3 Impact of criminalization on own educational outcomes

For own-educational outcomes, we focus on secondary education or more to compare their education level after the prime period for FGC practice is almost over. To address sample selection concerns, we construct a dummy if the

respondent has a secondary education level or more, taking zero otherwise, including for those with no education. Also, we include a literacy indicator representing a more than secondary education level, or at least able to read and write for lesser education level. Figure 5.4 shows the results for the per-cohort criminalization effect on secondary or more education and literacy. Overall, the reform seems to decrease the chance of having secondary education or more and being literate.

Focusing on comparing the young and old cohorts, columns (1) and (2) of panel A in Table 5.2 present additional evidence for the estimated effects of criminalization on own educational outcomes. For the cohorts in the top tercile, for example, the reform leads to a decrease in the probability of achieving at least secondary education and being literate by approximately 4 and 8 percentage points, about 20 to 21 percent of control means. Additionally, we present the falsification tests for the results on their own educational outcomes in Table 5.4, suggesting that the effects on all the outcomes pass the tests.

5.3.4 Impact of criminalization on spousal attributes and family wealth

In this sub-section, we examine the effect of criminalizing FGC on several spousal attributes and household wealth. In these analyses, sample selection bias is a challenge because we necessarily focus on ever-married women. To minimize the selection bias, we restrict the sample to those women aged 30 and above at survey, whose rates of ever-marriage are very high (more than 96 percent). Figure 5.6 shows the results for the per-cohort criminalization effect on primary and secondary education or more. Restricting the sample to females aged 30 to 49 at the survey, we find that criminalization decreases the chance of matching to better-educated husbands, particularly for cohorts in the top tercile. In Figure 5.7, we provide the results from similar analyses using the whole sample aged 15 to 49 at the survey, suggesting similar findings.

Table 5.3 provides further evidence for the effects of FGC criminalization on spousal attributes and family wealth, focusing on young and old cohorts. Panels A and B report the effect of the reform using the whole ever-married sample aged 15 to 49 and the ever-married sample aged 30 and above. Column (1) in panel A, for example, suggest that

criminalization for cohorts in the top tercile decreases the likelihood of marrying husbands having primary education or more by approximately 5 percentage points, about 11 percent of the control mean. Further, criminalization decreases the probability of finding husbands with secondary education or more by approximately 3 percentage points for cohorts in the top tercile, about 12 percent of the control mean.

Moreover, columns (3) and (4) of Table 5.3 present the estimated effects of the reform on the partners' working status and engagement in a skilled type of work. The results suggest that criminalization does not have a significant effect on the probability of having working husbands. However, the reform decreases the chance of matching to a partner with a skilled type of work. Columns (5) and (6) present the reform effect on family wealth, suggesting that criminalization leads to end up with a poor family. Also, Figures 5.6 to 5.11 provide additional evidence by investigating the reform effect for each cohort, suggesting similar findings. Overall, criminalization decreases spousal quality for ever-married women. We present the falsification tests for the results on spousal attributes in Table 5.5, suggesting that the effects on all the outcomes pass the tests.

5.4 Potential mechanisms for criminalization to affect female well-being

Given the rat-race nature of the marriage market in Africa, many parents engage in alternative pre-marital investments for their daughters, including FGC, education, and early marriage. Education can be one of the substitutes for FGC in the marriage market (García-Hombrados and Salgado 2019; Kudo 2018). Because parents perceive both FGC and education increase their daughters' quality or bride price in the context of marriage transfers (García-Hombrados and Salgado 2019; Lafortune 2013; Mackie 1996). Another alternative strategy to maintain the quality of their daughters before marriage is to accept marriage proposals when their daughters are young, leading to child marriage. Men prefer to marry young brides as the perceived quality of women decreases by age (Wahhaj 2018).

In the context of FGC criminalization increasing the cost of FGC, two options are left for parents; to educate their daughters or rush to get a match to their daughters at a young age. However, our findings related to the effect of criminalization on education suggest that the reform decreases the likelihood of having secondary education or above

and being literate. This result, in contrast to García-Hombrados and Salgado (2019), shows that parents do not substitute education for FGC, possibly due to the higher relative cost of education. Instead, they may opt for arranging a marriage for their daughters at a young age. In line with this argument, the reform increases the probability of early marriage and fertility. Thus, education is not the mechanism for the effect of the reform (or FGC) on early marriage; however, early marriage can be an alternative substitute for FGC in the marriage market.

For the negative effect of criminalization on female educational achievement, we can think of two mechanisms. First, abandoning FGC may expose to stigmatization, discrimination, and peer isolation at school for the uncut girls, leading to a loss of interest in schooling and poor educational achievement (Pesambili 2013; Pesambili and Mkumbo 2018). Second, since the reform exposes females to early marriage and fertility, they have to drop out of school when they are ready to marry, leading to low educational achievement (Field and Ambrus 2008; Jensen and Thornton 2003).

Related to the reform effect of increasing the likelihood of being overweight, the stress and tension associated with the feeling of stigma in abandoning FGC (Pesambili and Mkumbo 2018) and their low educational achievement may be a possible mechanism. Brunello, Fabbri, and Fort (2013) suggest that education affects the BMI of females by reducing the probability of being obese or overweight. Also, education is associated with health-related behaviors such as smoking, alcohol consumption, and physical exercise, which matter for general well-being (e.g., Amin, Behrman, and Spector 2013; Brunello et al. 2013; and Conti, Heckman, and Urzua 2010).

Finally, we find that criminalizing FGC reduces the chance of meeting a quality spouse with more education and higher skill occupations. This could be related to shorter searching time and hurrying to find any partner at the expense of spousal quality in early marriage options due to the reform (Vogl 2013). Further, given the women's low educational and labor market achievement and low spousal quality, it is not surprising they end up with low-wealth families.

CHAPTER 6

CONCLUSION

6.1 Summary

This dissertation shows that laws against widely accepted social customs matter at least in part in the eradicating efforts of such practices. The analyses emphasize estimating the effect of criminalizing FGC in African countries using nationally representative cross-sectional data sets from pooling African Demographic and Health Surveys and Multiple Indicator Cluster Surveys. The method in this study employs a difference-in-difference approach in which the reform effect varies by birth cohort and within-country region-ethnicity groups. The identification strategy includes both cohort-by-cohort contrasts across pre-reform FGC prevalence rate region-ethnicity groups and comparing young with old cohorts.

The results in chapter 3 suggest that criminalization substantially decreases the likelihood of ever-experiencing female genital cutting, particularly for youngest cohorts at reform and cohorts in the highest pre-reform FGC prevalence rate region-ethnicity groups. Chapter 4 of this dissertation empirically estimates the criminalization effect on several female well-being outcomes. The findings suggest that the reform increases the risk of early marriage and fertility at the expense of quality marriage and lowers educational attainment. However, the reform does not have significant effects on reproductive health outcomes. The results in both chapters pass several falsification tests, supporting the validity of this dissertation's research design.

6.2 Policy implications

The results in chapter 4 of this dissertation suggest the effectiveness of FGC criminalization, having important policy implications aiming to eradicate FGC, such as promoting national bans on the practice. Even with limited actual

enforcement, the law can be effective by changing people's perception of the reform as potentially enforceable, leading to the fear of punishment. Also, the reform might work through enhancing the bargaining and decision-making power of those women against FGC. As a result, policymakers may have to consider such possible mechanisms while trying to improve the effectiveness of the interventions. Moreover, the findings in this chapter show that the effect of criminalizing FGC is more pronounced for young females at reform and those females in the highest pre-reform FGC prevalence region-ethnicity groups. Thus, policymakers may use this result as evidence whether to consider the cohort-specific and region-ethnicity variations in the reactions to the FGC bans.

Moreover, policymakers may have to consider the possible interactions of FGC with other traditions such as marriage customs, particularly early marriage. Because one of the findings in chapter 5 of this dissertation is that criminalizing FGC leads girls to the risk of early marriage and fertility as parents tend to substitute early marriage for FGC to take advantage of at least the young age of their daughters. Child marriage and adolescent pregnancy, by themselves, affect the girls' human capital development adversely and worsens gender inequality. Further, child marriage is usually at the expense of spousal quality due to a shorter searching period, negatively affecting the girls' marriage returns.

An additional finding in this chapter of the dissertation is that criminalizing FGC lowers educational attainment, suggesting that parents do not substitute education for FGC in the presence of reform. Because the possible mechanism for lower education could be stigmatization and peer isolation associated with FGC abandoning, educational policymakers should think over such concerns while promoting female education. Overall, policymakers suggesting bans on FGC should take into account the overall well-being of the women abandoning the practice instead of merely focusing on the eradication of the custom.

6.3 Limitations and future research

Except for Shell-Duncan et al. (2013)'s tabulation of criminal and other types of laws against FGC, almost no evidence is available showing clear demarcation of criminal and non-criminal laws and socio-political factors determining reform types. Although this dissertation attempts to differentiate and compare between criminal and non-criminal laws in

chapter 3, it focuses on examining the effect of only criminal laws because of the overall unclear association between the non-criminal laws and age at reform.

The data for non-criminalizing countries taken as a whole do not show a lower FGC prevalence for those younger cohorts at reform. However, it is possible to investigate the trend for FGC prevalence for each country separately. For Chad, Kenya, and Nigeria, out of the 6 non-criminalizing countries, the data show a decreasing trend in FGC prevalence, moving from older to younger cohorts at reform. As a result, investigating how the non-criminal laws have affected the recent prevalence of FGC in these countries could be an additional possible analysis. Particularly for Nigeria, because several states have adopted FGC bans since 1999, it would be interesting to examine the reform effect exploiting the rollout of the reform across the states.

Moreover, questions related to FGC might be sensitive, leading the respondents to misreport their actual status. If the criminalization leads the respondents to under-report their true status of cutting (or deny their cutting), then the estimated effect of the reform will be negatively biased, suggesting the estimate is an upper bound. However, over-reporting might also be possible due to the feeling of the stigma associated with FGC abandoning. As a result, the direction of bias for the estimated effect of criminalization may be confusing. It could be interesting to examine the causal impact of FGC on female well-being using the reform as an instrument to address the measurement error issue.

Further, in this dissertation, the sample for the main analyses includes those women who are partially affected and least affected by the reform while excluding the fully affected cohorts with zero or negative age at reform. Because the fully affected cohorts are very small, coming from only 5 countries with criminal bans. As more and more DHS and MICS latest surveys are coming, however, it may be possible to obtain more samples for those born in the year of reform or after. As a result, it would also be interesting to compare the fully affected females with those partially affected.

Because age at cutting information is missing for lots of surveys, this dissertation does not examine the effect of criminalization on cutting before several age thresholds, rather focuses on ever-experiencing FGC only. However, limiting the sample to only those who report the age at cutting, it may be possible to extend the analysis for the cutting by several age thresholds.

Additionally, this dissertation uses the women's DHS and MICS data to examine how the reform affects family formation rather than the daughter-mother data lacking such outcome variables. However, using the daughter-mother data may be better to construct the first stage because women generally do not make FGC decisions by themselves. Whether to conform to the community's traditional practice could be a critical determinant for parents to decide whether to cut their daughters. Using the women's data may not hold constant maternal conformity. As a result, the observed relationship between criminalization and women's genital cutting status might be spurious even given the same region, ethnicity, and birth cohort.

In other words, the estimated effect of the reform might be biased without controlling for mothers' desire towards the continuation of the FGC practice. On one hand, mothers' attitude towards the continuation of the custom might matter on the prevalence of FGC among their daughters. If they are conservative and desire the continuation of the practice, then they are more likely to engage in cutting their daughters before the reform. On the other hand, their attitude might determine the way they respond to criminalization. That is, if they are conservative they might more likely be risk-takers to cut their daughters. They might less likely to consider people's rumors about the enforceability of the reform and fear the punishment of being caught practicing. Moreover, they are less likely to take advantage of the law to increase their bargaining and decision-making power to abandon the practice. As a result, the estimated effect of the reform might be biased downwards because conformed (conservative) mothers might more likely to engage in their daughters' cutting and less likely to abide by the laws against FGC.

If we assume that the FGC criminalization policy had no effect on female well-being outcomes other than by reducing the likelihood of ever-experiencing FGC, one can use this policy to construct instrumental variables estimates of the impact of FGC on well-being outcomes. This method could solve both limitations related to measurement error due to the self-reporting bias and omitted maternal conformity information. However, the researcher should examine the possibility of cases whereby the reform affects the outcomes indirectly through other mechanisms.

REFERENCES

- Acemoglu, Daron and Matthew O. Jackson. 2017. "Social Norms and the Enforcement of Laws." *Journal of the European Economic Association* 15(2):245–95.
- Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development : An Empirical Investigation." *American Economic Review* 91(5):1369–1401.
- Aldashev, Gani, Imane Chaara, Jean-Philippe Platteau, and Zaki Wahhaj. 2012a. "Formal Law as a Magnet to Reform Custom." *Economic Development and Cultural Change*.
- Aldashev, Gani, Imane Chaara, Jean-Philippe Platteau, and Zaki Wahhaj. 2012b. "Using the Law to Change the Custom." *Journal of Development Economics* 97:182–200. Retrieved (<http://dx.doi.org/10.1016/j.jdeveco.2011.03.001>).
- Amin, Vikesh, Jere R. Behrman, and Tim D. Spector. 2013. "Does More Schooling Improve Health Outcomes and Health Related Behaviors ? Evidence from U . K . Twins." *Economics of Education Review* 35:134–48. Retrieved (<http://dx.doi.org/10.1016/j.econedurev.2013.04.004>).
- André, Catherine and Jean Philippe Platteau. 1998. "Land Relations under Unbearable Stress: Rwanda Caught in the Malthusian Trap." *Journal of Economic Behavior and Organization* 34(1):1–47.
- Antonazzo, Monica. 2003. "Problems with Criminalizing Female Genital Cutting." *Peace Review* 15(4):471–77. Retrieved (<https://doi.org/10.1080/1040265032000156663>).
- Ashraf, Nava, Natalie Bau, Nathan Nunn, and Alessandra Voena. 2020. "Bride Price and Female Education Nava Ashraf Natalie Bau Alessandra Voena." *Journal of Political Economy* 128(2):591–641.
- Balachandran, Aswini A., Swapna Duvalla, Abdul H. Sultan, and Ranee Thakar. 2018. "Are Obstetric Outcomes Affected by Female Genital Mutilation ?" *Int Urogynecol J* 29:339–44.
- Banerjee, Abhijit and Lakshmi Iyer. 2000. "History , Institutions , and Economic Performance : The Legacy of

- Colonial Land Tenure Systems in India.” *American Economic Review* 95(4):1190–1213.
- Bazarkulova, Dana and Janice Compton. 2021. “Marriage Traditions and Investment in Education: The Case of Bride Kidnapping.” *Journal of Comparative Economics* 49(1):147–63. Retrieved (<https://doi.org/10.1016/j.jce.2020.07.005>).
- Becker, Gary S. 1968. “Crime and Punishment : An Economic Approach.” *Journal of Political Economy* 76(2):169–217. Retrieved (<https://doi.org/10.1086/259394>).
- Bellemare, Marc F., Lindsey Novak, and Tara L. Steinmetz. 2015. “All in the Family: Explaining the Persistence of Female Genital Cutting in West Africa.” *Journal of Development Economics* 116:252–65.
- Belloc, Marianna and Samuel Bowles. 2013. “The Persistence of Inferior Cultural-Institutional Conventions.” *The American Economic Review* 103(3):93–98. Retrieved (<http://www.jstor.org/stable/23469709>).
- Berg, Rigmor C. and Eva Denison. 2013. “A Tradition in Transition: Factors Perpetuating and Hindering the Continuance of Female Genital Mutilation/Cutting (FGM/C) Summarized in a Systematic Review.” *Health Care for Women International* 34(10):837–59.
- Berg, Rigmor C. and Eva Denison. 2012. “Effectiveness of Interventions Designed to Prevent Female Genital Mutilation / Cutting : A Systematic Review.” *Studies in Family Planning* 43(2):135–46.
- Berg, Rigmor C. and Vigdis Underland. 2013. “The Obstetric Consequences of Female Genital Mutilation / Cutting : A Systematic Review and Meta-Analysis.” *Obstetrics and Gynecology International* 1–15. Retrieved (<http://dx.doi.org/10.1155/2013/496564>).
- Berg, Rigmor C., Vigdis Underland, Jan Odgaard-Jensen, Atle Fretheim, and Gunn E. Vist. 2014. “Effects of Female Genital Cutting on Physical Health Outcomes: A Systematic Review and Meta-Analysis.” *BMJ Open* 4(11).
- Bloom, David E., Michael Kuhn, and Klaus Prettnner. 2020. “The Contribution of Female Health to Economic Development.” *Economic Journal* 130(630):1650–77.
- Boyle, Elizabeth Heger and Sharon E. Preves. 2000. “National Politics as International Process : The Case of Anti-

- Female-Genital-Cutting Laws.” *Law & Society Review* 34(3):703–37. Retrieved (<http://www.jstor.org/stable/3115141>).
- Brunello, Giorgio, Daniele Fabbri, and Margherita Fort. 2013. “The Causal Effect of Education on Body Mass : Evidence from Europe.” *Journal of Labor Economics* 31(1):195–223. Retrieved (<https://www.jstor.org/stable/10.1086/667236>).
- Burdett, Ken and Melvyn G. Coles. 1997. “Marriage and Class.” *The Quarterly Journal of Economics* 112(1):141–68. Retrieved (<https://doi.org/10.1162/003355397555154>).
- Cameron, A. Colin and Douglas L. Miller. 2015. “A Practitioner’s Guide to Cluster- Robust Inference.” *Journal of Human Resources* 50(2):317–72.
- Camilotti, Giulia. 2016. “Interventions to Stop Female Genital Cutting and the Evolution of the Custom : Evidence on Age at Cutting in Senegal.” *Journal of African Economies* 25(1):133–58.
- De Cao, Elisabetta and Clemens Lutz. 2018. “Sensitive Survey Questions : Measuring Attitudes Regarding Female Genital Cutting Through a List Experiment *.” *Oxford Bulletin of Economics and Statistics* 80(5):305–9049.
- Chen, Stacey H., Yen Chien Chen, and Jin Tan Liu. 2019. “The Impact of Family Composition on Educational Achievement.” *Journal of Human Resources* 54(1):122–70.
- Chow, Vinci and Eva Vivalt. 2021. “Challenges in Changing Social Norms: Evidence from Interventions Targeting Child Marriage in Ethiopia.” *Journal of African Economies* 0(0):1–28.
- Chuang, Erica, Pascaline Dupas, Elise Huillery, and Juliette Seban. 2021. “Sex, Lies, and Measurement: Consistency Tests for Indirect Response Survey Methods.” *Journal of Development Economics* 148.
- Conti, Gabriella, James Heckman, and Sergio Urzua. 2010. “The Education-Health Gradient.” *American Economic Review: Papers & Proceedings* 100(2):234–38.
- Cook, R. J., B. M. Dickens, and M. F. Fathalla. 2002. “Female Genital Cutting (Mutilation/ Circumcision) : Ethical and Legal Dimensions.” *International Journal of Gynecology and Obstetrics* 79:281–87.

- Deaton, Angus. 2018. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Reissue Ed. Washington, DC: World Bank.
- Diop, Nafissatou J. and Ian Askew. 2009. “The Effectiveness of a Community-Based Education Program on Abandoning Female Genital Mutilation / Cutting in Senegal.” *Studies in Family Planning* 40(4):307–18.
- Duflo, Esther. 2001. “Schooling and Labor Market Consequences of School Construction in Indonesia : Evidence from an Unusual Policy Experiment.” *American Economic Review* 91(4):795–813.
- Efferson, Charles, Sonja Vogt, Amy Elhadi, Hilal El Fadil Ahmed, and Ernst Fehr. 2015. “Female Genital Cutting Is Not a Social Coordination Norm.” *Science* 349(6255).
- Elmusharaf, Susan, Nagla Elhadi, and Lars Almroth. 2006. “Reliability of Self Reported Form of Female Genital Mutilation and WHO Classification: Cross Sectional Study.” *British Medical Journal* 333(7559):124. Retrieved (<https://www.bmj.com/content/333/7559/124>).
- Engelsma, Brian, Gerry Mackie, and Brandon Merrell. 2020. “Unprogrammed Abandonment of Female Genital Mutilation/Cutting.” *World Development* 129(104845). Retrieved (<https://doi.org/10.1016/j.worlddev.2019.104845>).
- Field, Erica and Attila Ambrus. 2008. “Early Marriage , Age of Menarche , and Female Schooling Attainment in Bangladesh.” *Journal of Political Economy* 116(5):881–930.
- Field, Erica, Seema Jayachandran, and Rohini Pande. 2010. “Do Traditional Institutions Constrain Female Entrepreneurship ? A Field Experiment on Business Training in India.” *American Economic Review: Papers & Proceedings* 100(2):125–29.
- García-Hombrados, Jorge and Edgar Salgado. 2019. *Female Genital Cutting and Education: Theory and Causal Evidence from Senegal*. London.
- Gibson, Mhairi A., Eshetu Gurm, Beatriz Cobo, Mari´a M. Rueda, and Isabel M. Scott. 2018. “Indirect Questioning Method Reveals Hidden Support for Female Genital Cutting in South Central Ethiopia.” *PLoS ONE* 13(5):1–14.

Retrieved (<https://doi.org/10.1371/journal.pone.0193985>).

Giuliano, Paola. 2020. “Gender and Culture.” *Oxford Review of Economic Policy* 36(4):944–61.

Grose, Grace Rose, Katherine A. Roof, Daniel C. Semenza, Ximena Leroux, and Kathryn M. Yount. 2019. “Mental Health , Empowerment , and Violence Against Young Women in Lower- Income Countries : A Review of Reviews.” *Aggression and Violent Behavior* 46:25–36. Retrieved (<https://doi.org/10.1016/j.avb.2019.01.007>).

Hancioglu, Attila and Fred Arnold. 2013. “Measuring Coverage in MNCH: Tracking Progress in Health for Women and Children Using DHS and MICS Household Surveys.” *PLoS Medicine* 10(5).

Howard, Janet A. and Mhairi A. Gibson. 2017. “Frequency-Dependent Female Genital Cutting Behaviour Confers Evolutionary Fitness Benefits.” *Nature Ecology and Evolution* 1(49):1–6. Retrieved (<http://dx.doi.org/10.1038/s41559-016-0049>).

Jacoby, Hanan G. and Ghazala Mansuri. 2010. “Watta Satta: Bride Exchange and Women’s Welfare in Rural Pakistan.” *American Economic Review* 100(4):1804–25.

Jayachandran, Seema. 2015. “The Roots of Gender Inequality in Developing Countries.” *Annu. Rev. Econ.* 7:63–88.

Jensen, Robert and Rebecca Thornton. 2003. “Early Female Marriage in the Developing World.” *Gender and Development* 11(2):9–19.

Jordá, Vanesa and José María Sarabia. 2015. “International Convergence in Well-Being Indicators.” *Social Indicators Research* 120(1):1–27.

Kaplan, Adriana et al. 2013. “Female Genital Mutilation / Cutting in The Gambia : Long-Term Health Consequences and Complications during Delivery and for the Newborn.” *International Journal of Women’s Health* 5:323–31.

Klein, Elliot, Elizabeth Helzner, Michelle Shayowitz, Stephan Kohlhoff, and Tamar A. Smith-norowitz. 2018. “Female Genital Mutilation : Health Consequences and Complications — A Short Literature Review.” *Obstetrics and Gynecology International* 1–7. Retrieved (<https://doi.org/10.1155/2018/7365715>).

- Konu, Anne and Matti Rimpela. 2002. "Well-Being in Schools: A Conceptual Model." *Health Promotion International* 17(1):79–87.
- Kouba, Leonard J. and Judith Muasher. 1985. "Female Circumcision in Africa : An Overview." *African Studies Review* 28(1):95–110. Retrieved (<http://www.jstor.org/stable/524569>).
- Krivenko, Ekaterina Yahyaoui. 2015. "Rethinking Human Rights and Culture Through Female Genital Surgeries." *Human Rights Quarterly* 37(1):107–36.
- Kudo, Yuya. 2021. "Does Criminalizing Discriminatory Cultural Practices Improve Women's Welfare? A Simple Model of Levirate Marriage in Africa." *Economics Letters* 199:109728. Retrieved (<https://doi.org/10.1016/j.econlet.2021.109728>).
- Kudo, Yuya. 2018. *Female Genital Cutting and Long-Term Adjustment of Marriage Markets : Evidence from West Africa*. Retrieved (<http://hdl.handle.net/2344/00050654>).
- Lafortune, Jeanne. 2013. "Making Yourself Attractive : Pre-Marital Investments and the Returns to Education in the Marriage Market." *American Economic Journal: Applied Economics* 5(2):151–78.
- Lowes, Sara and Nathan Nunn. 2018. "Bride Price and the Wellbeing of Women." Pp. 117–38 in *Towards Gender Equity in Development*, edited by S. Anderson, L. Beaman, and J.-P. Platteau. Oxford: Oxford University Press.
- Mackie, Gerry. 1996. "Ending Footbinding and Infibulation : A Convention Account." *American Sociological Review* 61(6):999–1017. Retrieved (<https://www.jstor.org/stable/2096305>).
- McGavock, Tamara. 2021. "Here Waits the Bride? The Effect of Ethiopia's Child Marriage Law." *Journal of Development Economics* 149:102580.
- Merry, Sally Engle. 2003. "Human Rights Law and the Demonization of Culture (And Anthropology Along the Way)." *Political and Legal Anthropology Review* 26(1):55–76.
- Morison, Linda et al. 2001. "The Long-Term Reproductive Health Consequences of Female Genital Cutting in Rural Gambia : A Community-Based Survey." *Tropical Medicine and International Health* 6(8):643–53.

- Mulongo, Peggy, Caroline Hollins Martin, and Sue McAndrew. 2014. "The Psychological Impact of Female Genital Mutilation/Cutting (FGM/C) on Girls/Women's Mental Health: A Narrative Literature Review." *Journal of Reproductive and Infant Psychology* 32(5):469–85. Retrieved (<http://dx.doi.org/10.1080/02646838.2014.949641>).
- Nollenberger, Natalia, Núria Rodríguez-Planas, and Almudena Sevilla. 2016. "The Math Gender Gap : The Role of Culture." *American Economic Review* 106(5):257–61. Retrieved (<http://dx.doi.org/10.1257/aer.p20161121>).
- Novak, Lindsey. 2020. "Persistent Norms and Tipping Points: The Case of Female Genital Cutting." *Journal of Economic Behavior and Organization* 177:433–74.
- OHCHR. 1995. *Harmful Traditional Practices Affecting the Health of Women and Children*.
- Pesambili, Joseph C. and Kitila A. K. Mkumbo. 2018. "Implications of Female Genital Mutilation on Girls ' Education and Psychological Wellbeing in Tarime , Tanzania." *Journal of Youth Studies* 21(8):1111–26.
- Pesambili, Joseph Christopher. 2013. "Consequences of Female Genital Mutilation on Girls ' Schooling in Tarime , Tanzania : Voices of the Uncircumcised Girls on the Experiences , Problems and Coping Strategies." *Journal of Education and Practice* 4(16):109–20.
- Platteau, Jean-philippe. 2009. "Institutional Obstacles to African Economic Development : State , Ethnicity , and Custom." *Journal of Economic Behavior & Organization* 71:669–89.
- Platteau, Jean-Philippe, Giulia Camilotti, and Emmanuelle Auriol. 2018. "Eradicating Women-Hurting Customs: What Role for Social Engineering?" Pp. 319–56 in *Towards Gender Equity in Development*, edited by S. Anderson, L. Beaman, and J.-P. Platteau. Oxford: Oxford University Press.
- La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny. 1998. "Law and Finance." *Journal of Political Economy* 106(6):1113–55.
- La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny. 1999. "The Quality of Government." *The Journal of Law, Economics, & Organization* 15(1):222–79.

- Ruderman, Rachel. 2013. "Female Circumcision : The Ethics of Harm Reduction Policies." *The Michigan Journal of Public Affairs* 10:95–107.
- Rutstein, Shea Oscar and Guillermo Rojas. 2006. *Guide to DHS Statistics: Demographic and Health Surveys Methodology*. Calverton, Maryland.
- Saith, Ruhi and Barbara Harriss-White. 1999. "The Gender Sensitivity of Well-Being Indicators." *Development and Change* 30(3):465–97.
- Schoellman, Todd and Michele Tertilt. 2006. "Marriage Laws and Growth in Sub-Saharan Africa." *American Economic Review* 96(2):295–98.
- Shell-Duncan, Bettina. 2008. "From Health to Human Rights: Female Genital Cutting and the Politics of Intervention." *American Anthropologist* 110(2):225–36. Retrieved (<http://www.jstor.org/stable/27563985>).
- Shell-Duncan, Bettina, Katherine Wander, Ylva Hernlund, and Amadou Moreau. 2013. "Legislating Change ? Responses to Criminalizing Female Genital Cutting in Senegal." *Law & Society Review* 47(4):803–35. Retrieved (<http://www.jstor.org/stable/43670360>).
- Ssenyonjo, Manisuli. 2007. "Culture and the Human Rights of Women in Africa: Between Light and Shadow." *Journal of African Law* 51(1):39–67.
- Taba, A. H. 1980. "Female Circumcision." *Tropical Doctor* 10(January):21–23.
- Tertilt, Michele. 2005. "Polygyny, Fertility, and Savings." *Journal of Political Economy* 113(6):1341–71.
- Tertilt, Michèle. 2006. "Polygyny , Women's Rights , and Development." *Journal of the European Economic Association* 4(2/3):523–30.
- UNICEF. 2013. *Female Genital Mutilation / Cutting : A Statistical Overview and Exploration of the Dynamics of Change*. New York.
- Vogl, Tom S. 2013. "Marriage Institutions and Sibling Competition : Evidence from South Asia." *The Quarterly*

Journal of Economics 128(3):1017–72. Retrieved (<https://doi.org/10.1093/qje/qjt011>).

Wagner, Natascha. 2015. “Female Genital Cutting and Long-Term Health Consequences – Nationally Representative Estimates across 13 Countries.” *The Journal of Development Studies* 51(3):226–46. Retrieved (<http://dx.doi.org/10.1080/00220388.2014.976620>).

Wahhaj, Zaki. 2018. “An Economic Model of Early Marriage.” *Journal of Economic Behavior and Organization* 152:147–76.

WHO. 2016. *WHO Guidelines on the Management of Health Complications from Female Genital Mutilation*.

Yoder, P. Stanley and Shanxiao Wang. 2013. *Female Genital Cutting : The Interpretation of Recent DHS Data. DHS Comparative Reports No. 33*. Calverton, Maryland, USA.

Young, H. Peyton. 2015. “The Evolution of Social Norms.” *Annual Review of Economics* 7(1):359–87. Retrieved (<http://www.annualreviews.org/doi/10.1146/annurev-economics-080614-115322>).

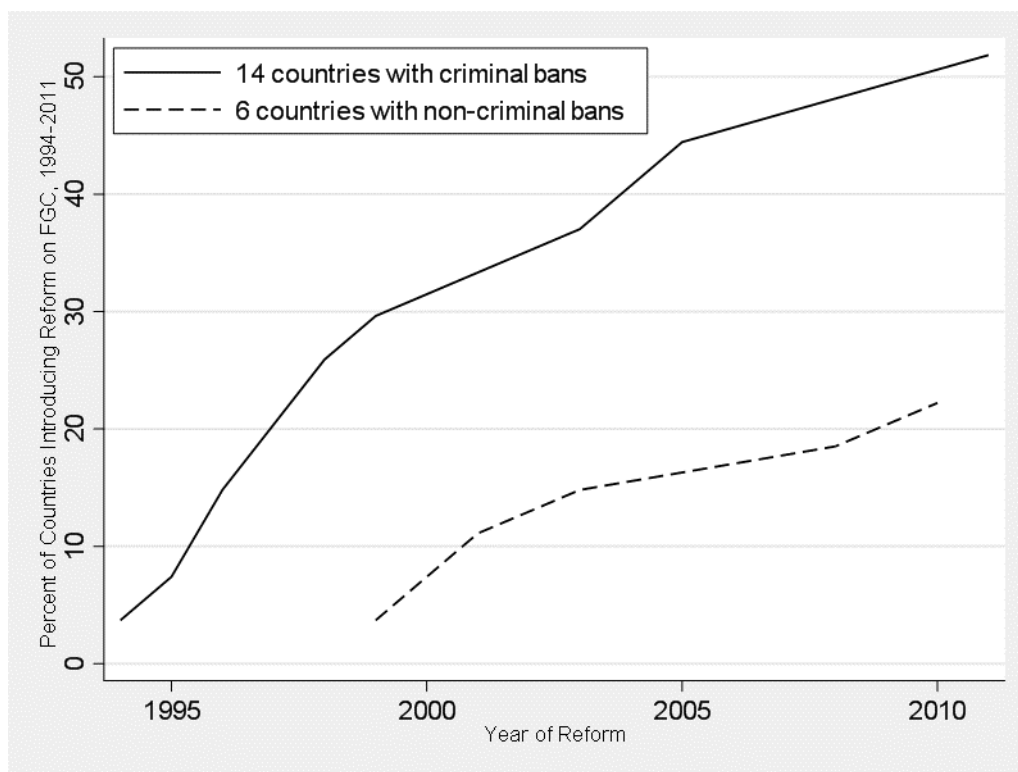


Figure 3.1: Rollout of criminal and non-criminal ban on female genital cutting, 1994-2011.

Notes: The solid and dashed line in this figure indicate the cumulative percentage of countries with criminal and non-criminal ban on FGC respectively, from 1994 to 2011. The numerator includes the African countries with criminal or non-criminal laws against FGC and the denominator includes all of the countries in Africa where FGC is prevalent. The figures exclude Eritrea and Somalia due to lack of post-reform survey although having ban on 2007 and 2012, respectively.

Source: Authors' computations from Shell-Duncan et al. (2013) and UNICEF (2013).

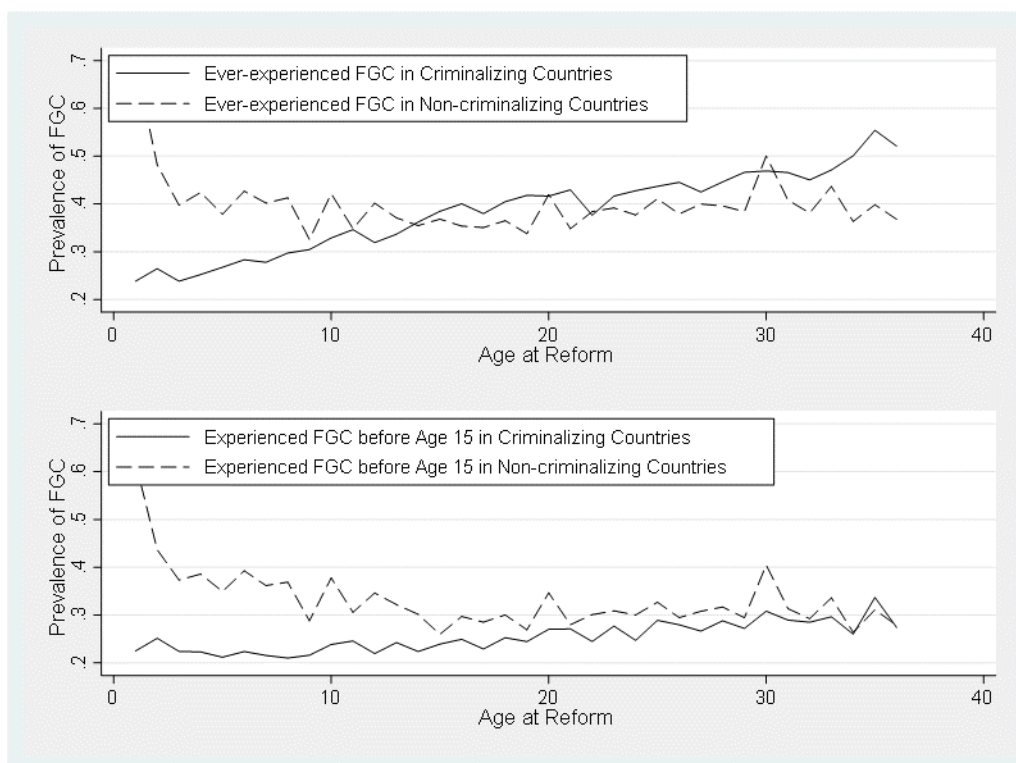


Figure 3.2: Prevalence of FGC and FGC before age 15 by cohort and reform type.

Notes: The solid and dashed lines in the top figure indicate the mean prevalence of ever-experiencing FGC in criminalizing and non-criminalizing countries/states, respectively. The solid and dashed lines in the bottom figure indicate the mean prevalence of FGC below age 15 for criminalizing and non-criminalizing countries/states, respectively. The sample size for criminalizing and non-criminalizing countries is 392,951 and 148,885 for both graphs at the top and bottom.

Source: Authors' graphical analysis using DHS and MICS.

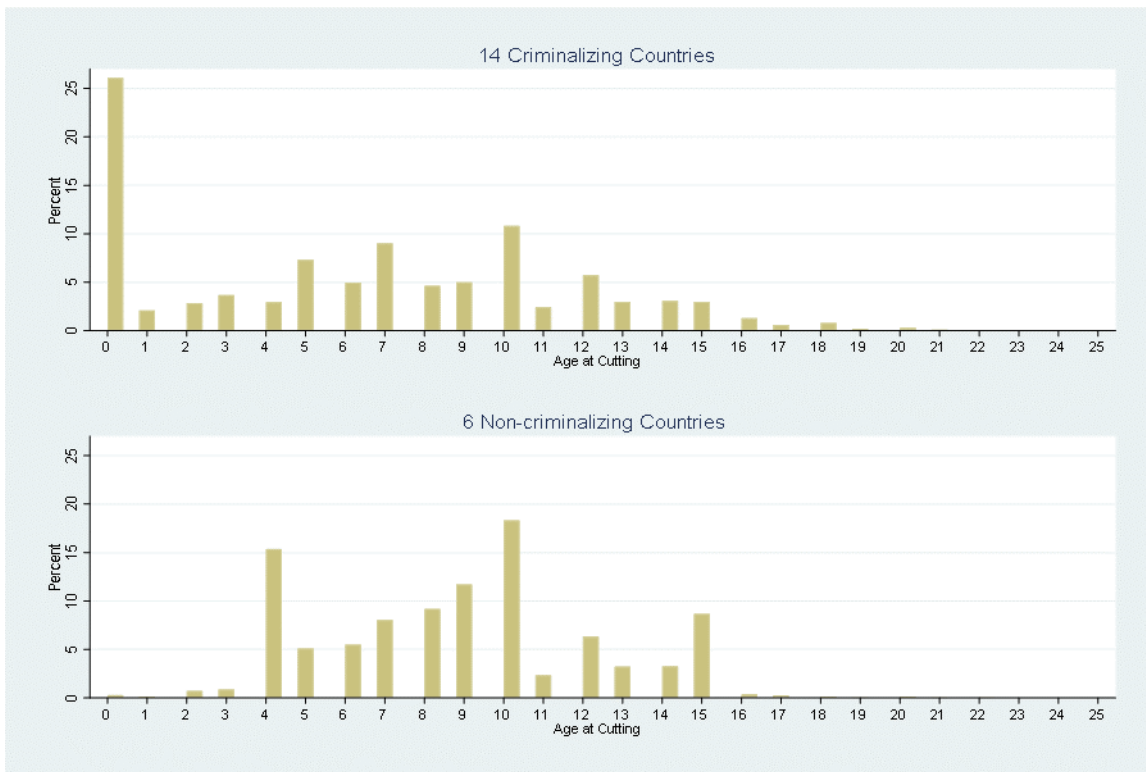


Figure 3.3: Age at cutting distributions by reform type.

Notes: The histogram shows the percent distributions for the female age at genital cutting separately for criminalizing and non-criminalizing countries, except the countries without age at cutting information. This figure uses frequency weights approximated by rounding the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018).

Source: Authors' computation using DHS and MICS.

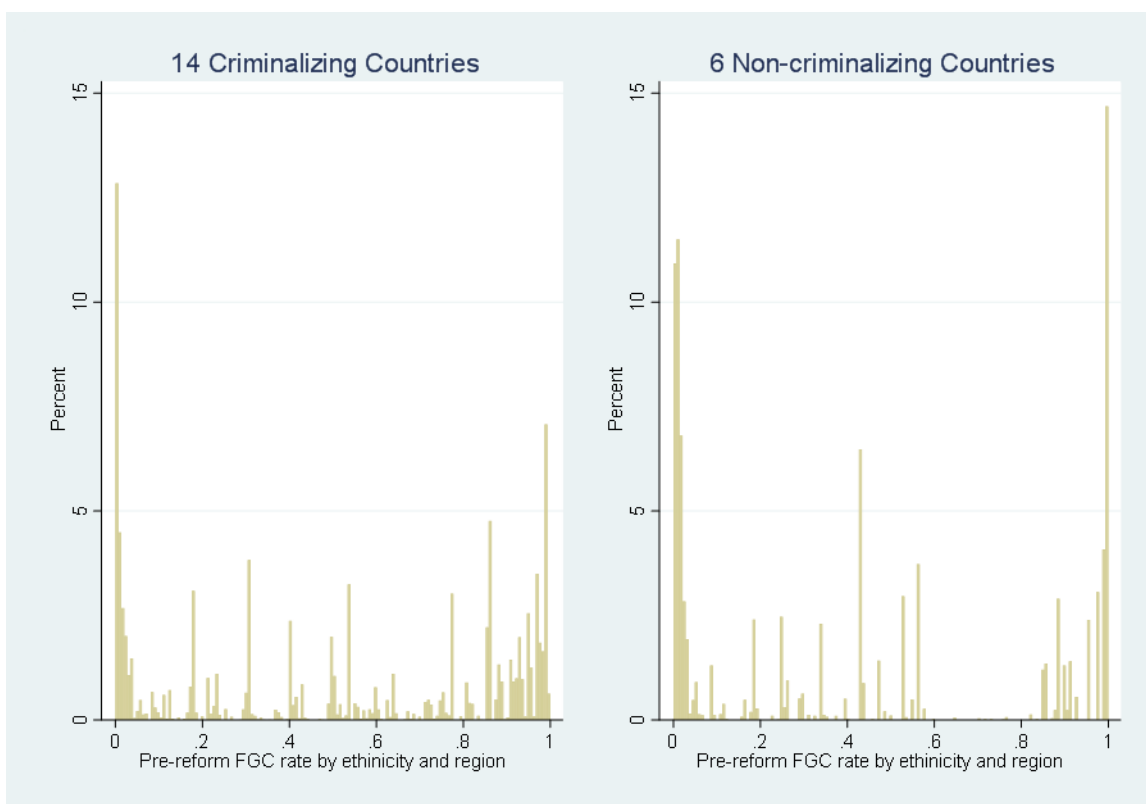


Figure 3.4: Distributions for pre-reform FGC prevalence rate by region-ethnicity groups.
Notes: The histogram shows the percent distributions for the pre-reform FGC prevalence rate in the ethnicity-by-region group. This figure uses frequency weights approximated by rounding the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018).

Source: Authors' graphical analysis using DHS and MICS.

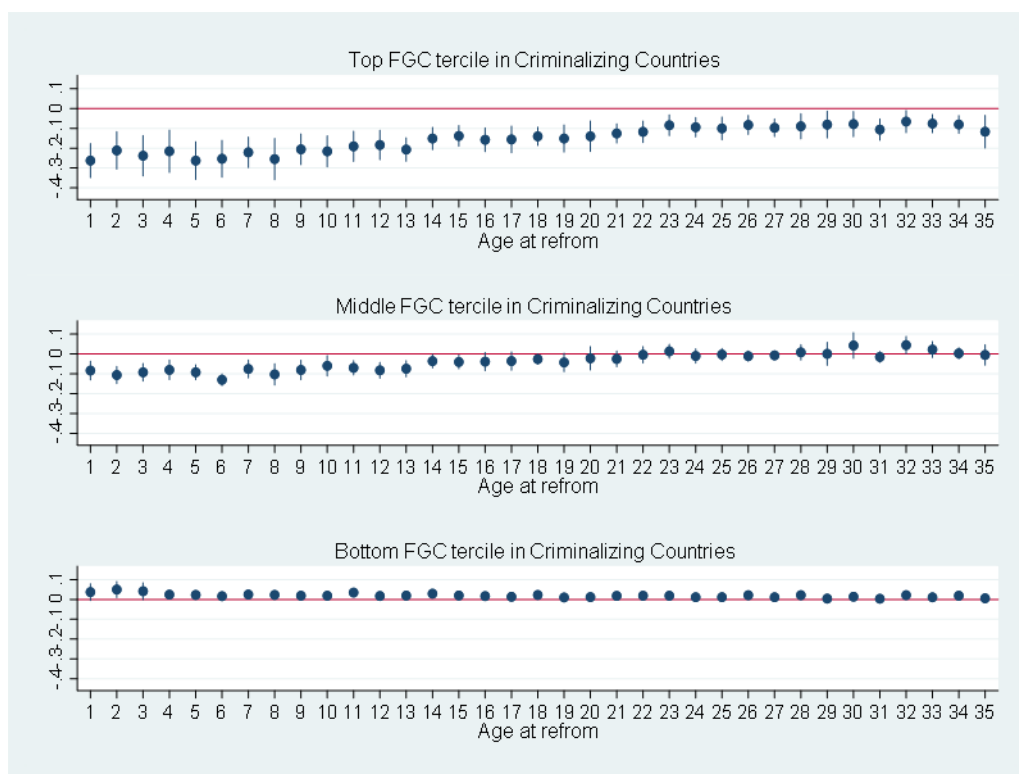


Figure 4.1: Event-study estimated effect of criminalization on ever-experiencing FGC by cohort and pre-reform FGC prevalence region-ethnicity groups.

Notes: The figure presents the coefficients of interactions of age at reform dummies with dummies for the top, middle, and bottom terciles of the pre-reform FGC prevalence rate in the region-ethnicity groups. In this regression, we omit those females aged 36 at reform as a reference group. The sample includes only the partially- and least-affected cohorts in criminalizing countries, excluding the fully-affected cohorts. The specification includes basic controls, demographics, and region fixed effect. "Basic controls" include fixed effects for the year of interview and country by age at reform. "Demographics" include the full set of dummies for rural and religion. We cluster standard errors at the country-cohort level. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

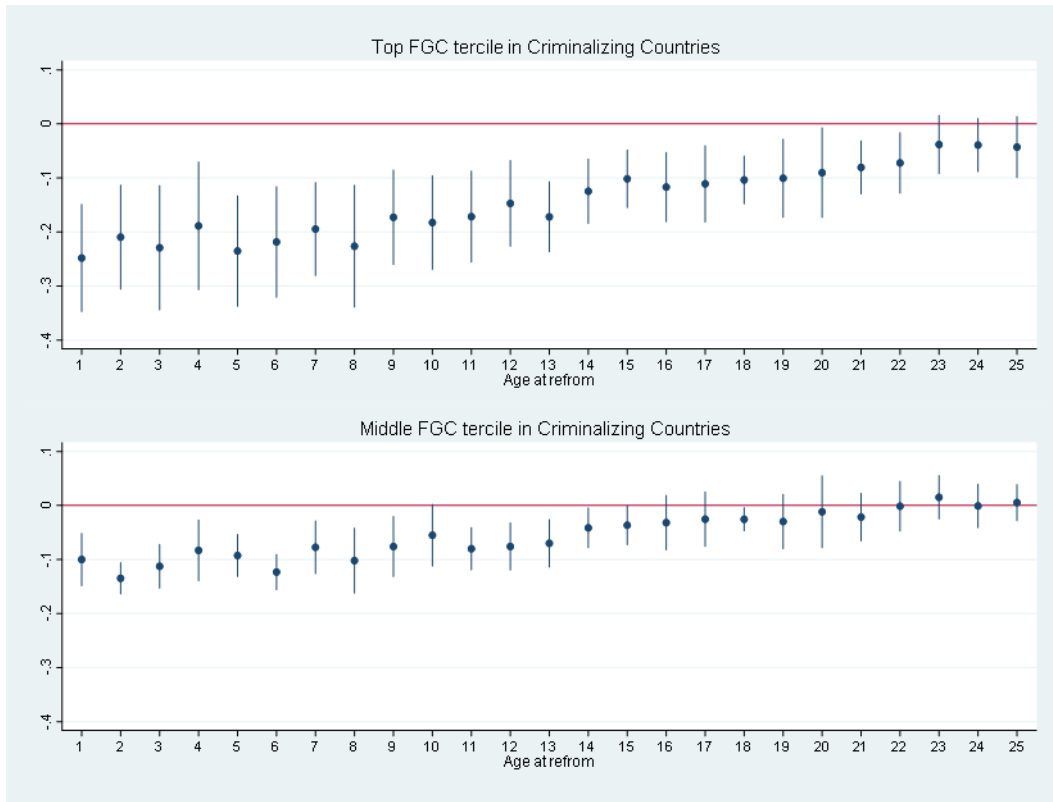


Figure 4.2: Event-study estimated effect of criminalization on ever-experiencing FGC for each cohort.

Notes: The top and bottom panels of the figure present the coefficients of interactions of age at reform dummies with dummies for the top and middle terciles of the pre-reform FGC prevalence rate in the region-ethnicity groups. In this analysis, we omit those females in the bottom tercile as the control group. The sample includes those females aged 15 to 49 at the survey and partially- and least-affected cohorts in criminalizing countries, excluding the fully-affected cohorts. The specification includes basic controls, demographics, and region fixed effect. "Basic controls" include fixed effects for the year of interview and country by age at reform. "Demographics" include the full set of dummies for rural and religion. We cluster standard errors at the country-cohort level. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

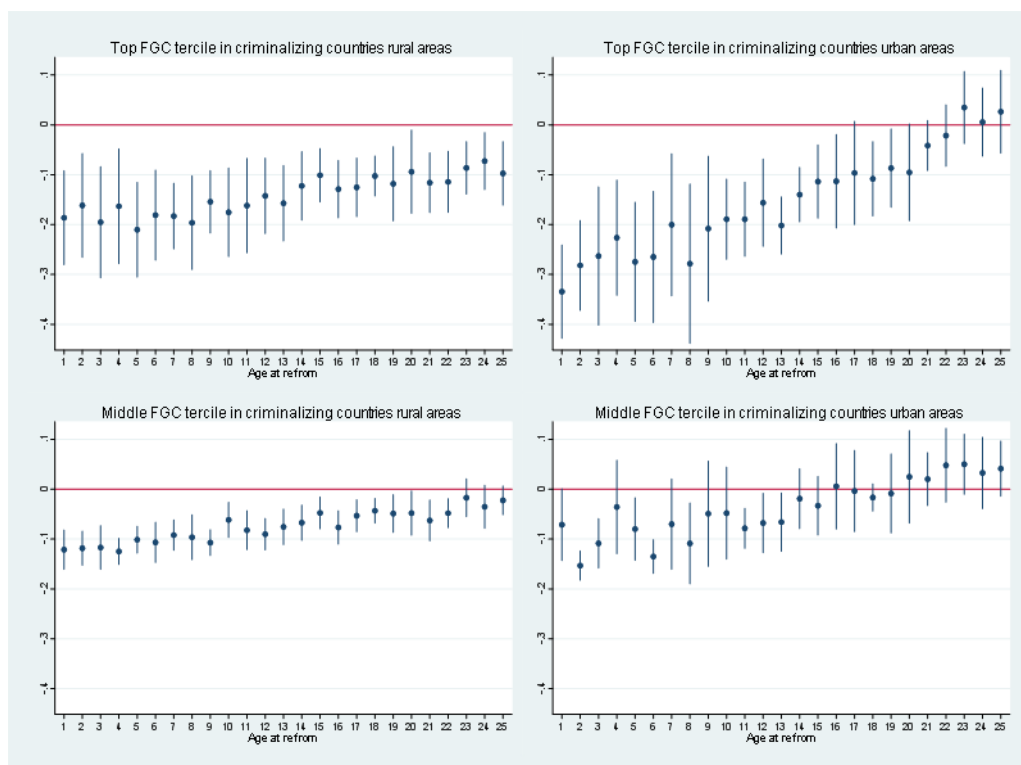


Figure 4.3: Event-study estimated effect of criminalization on ever-experiencing FGC by residence (rural/urban).

Notes: The top and bottom panels of the figure present the coefficients of interactions of age at reform dummies with dummies for the top and middle terciles of the pre-reform FGC prevalence rate in the region-ethnicity groups. The left and right panels report the results for the rural and urban areas. In this analysis, we omit those females in the bottom tercile as the control group. The sample includes those females aged 15 to 49 at the survey and partially- and least-affected cohorts in criminalizing countries, excluding the fully-affected cohorts. The specification includes basic controls, demographics, and region fixed effect. "Basic controls" include fixed effects for the year of interview and country by age at reform. "Demographics" include the full set of dummies for rural and religion. We cluster standard errors at the country-cohort level. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

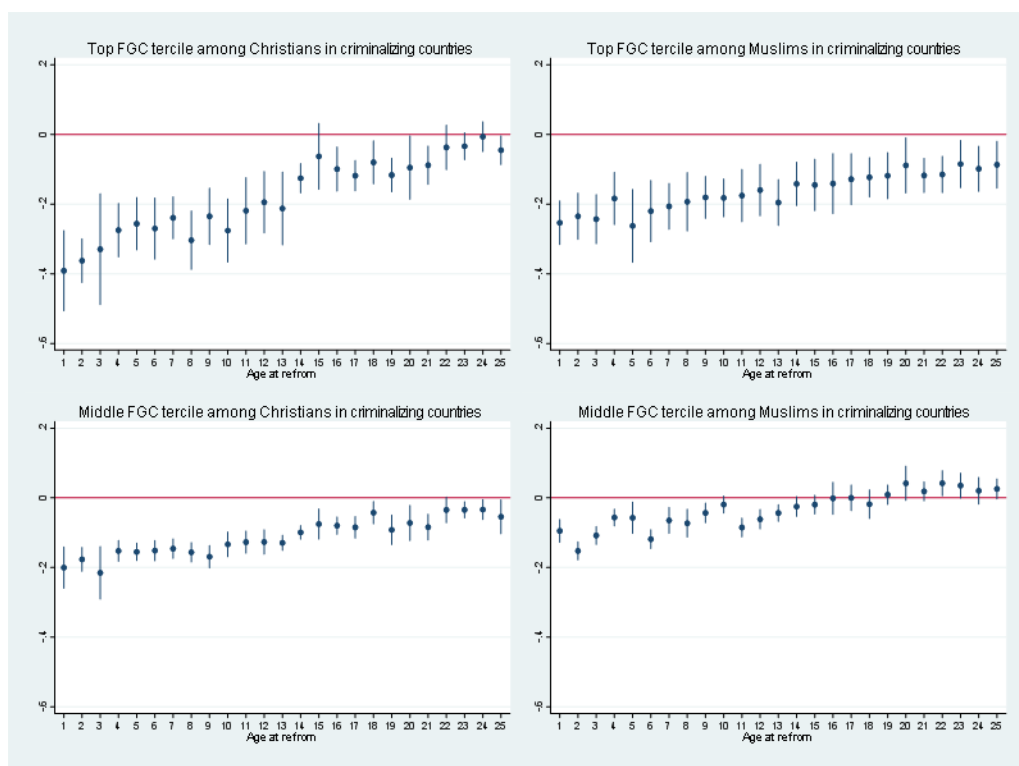


Figure 4.4: Event-study estimated effect of criminalization on ever-experiencing FGC by religion.

Notes: The top and bottom panels of the figure present the coefficients of interactions of age at reform dummies with dummies for the top and middle terciles of the pre-reform FGC prevalence rate in the region-ethnicity groups. The left and right panels report the results for Christians and Muslims. In this analysis, we omit those females in the bottom tercile as the control group. The sample includes those females aged 15 to 49 at the survey and partially- and least-affected cohorts in criminalizing countries, excluding the fully-affected cohorts. The specification includes basic controls, demographics, and region fixed effect. "Basic controls" include fixed effects for the year of interview and country by age at reform. "Demographics" include the full set of dummies for rural and religion. We cluster standard errors at the country-cohort level. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

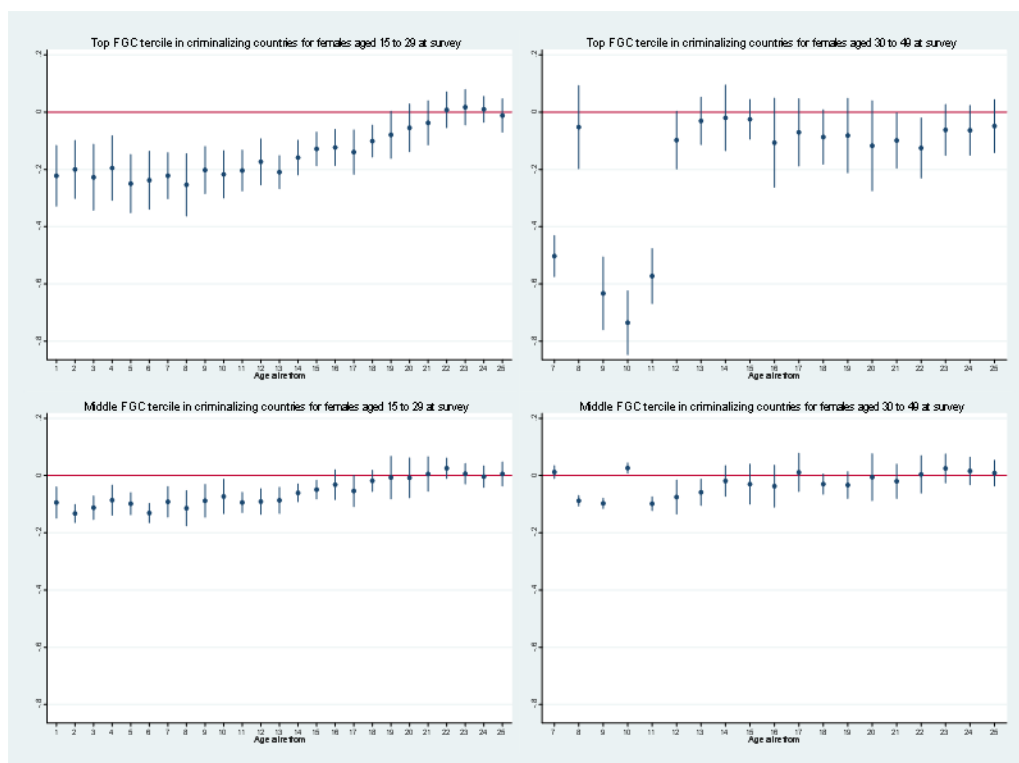


Figure 4.5: Event-study estimated effect of criminalization on ever-experiencing FGC by age at survey.

Notes: The top and bottom panels of the figure present the coefficients of interactions of age at reform dummies with dummies for the top and middle terciles of the pre-reform FGC prevalence rate in the region-ethnicity groups. The left and right panels report the results for those females aged 15 to 29 and aged 30 to 49 at interview, respectively. In this analysis, we omit those females in the bottom tercile as the control group. The sample includes those partially- and least-affected cohorts in criminalizing countries, excluding the fully-affected cohorts. The specification includes basic controls, demographics, and region fixed effect. "Basic controls" include fixed effects for the year of interview and country by age at reform. "Demographics" include the full set of dummies for rural and religion. We cluster standard errors at the country-cohort level. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

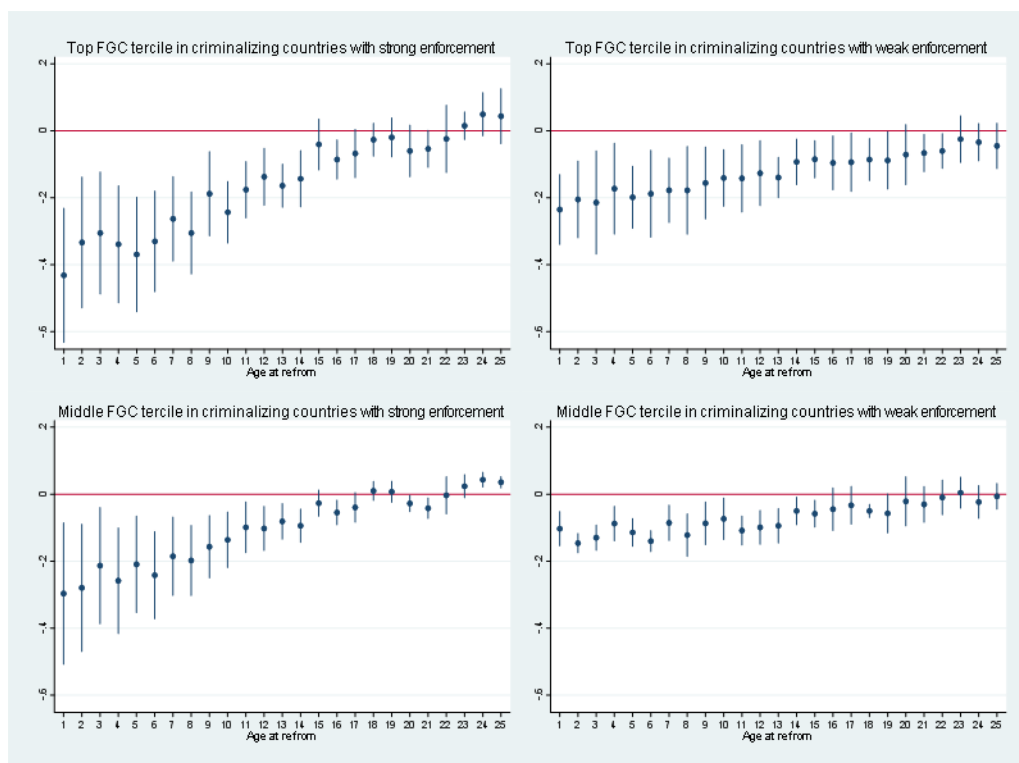


Figure 4.6: Event-study estimated effect of criminalization on ever-experiencing FGC by degree of enforcement.

Notes: The top and bottom panels of the figure present the coefficients of interactions of age at reform dummies with dummies for the top and middle terciles of the pre-reform FGC prevalence rate in the region-ethnicity groups. The left and right panels report the results for strong and weak enforcement of criminalization. In this analysis, we omit those females in the bottom tercile as the control group. The sample includes those females aged 15 to 49 at the survey and partially- and least-affected cohorts in criminalizing countries, excluding the fully-affected cohorts. The specification includes basic controls, demographics, and region fixed effect. "Basic controls" include fixed effects for the year of interview and country by age at reform. "Demographics" include the full set of dummies for rural and religion. We cluster standard errors at the country-cohort level. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

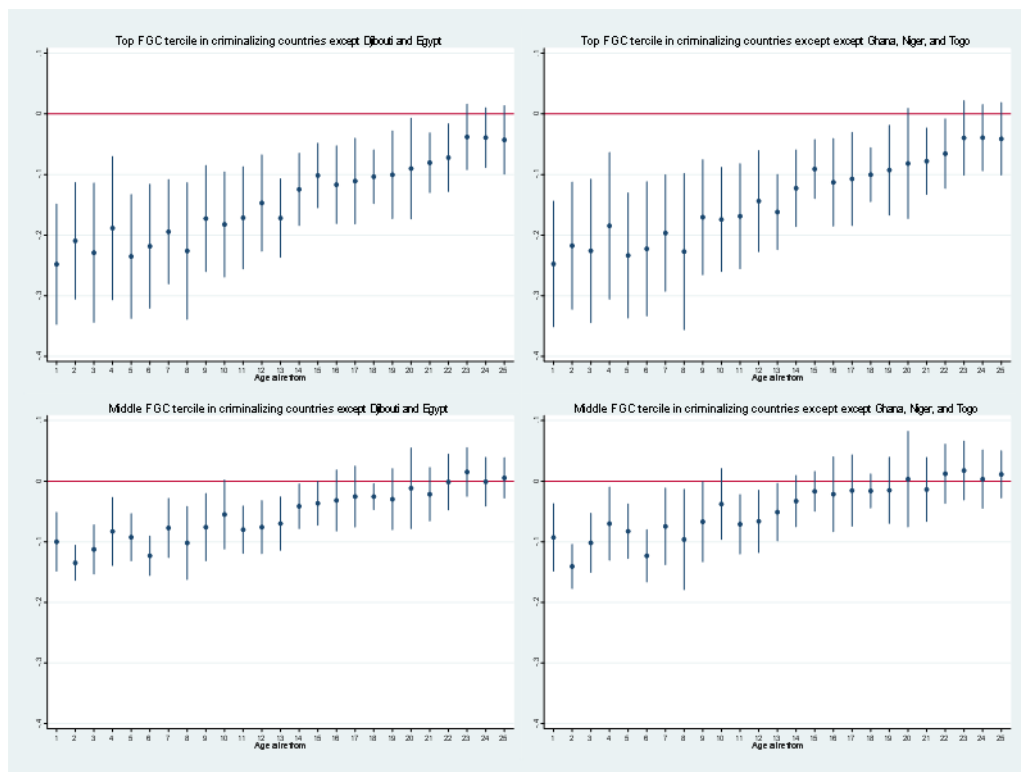


Figure 4.7: Event-study estimated effect of criminalization on ever-experiencing FGC by excluding countries with less within-country variation in the pre-reform FGC rate.

Notes: The top and bottom panels of the figure present the coefficients of interactions of age at reform dummies with dummies for the top and middle terciles of the pre-reform FGC prevalence rate in the region-ethnicity groups. The left and right panels report the results for excluding Djibouti and Egypt, and Ghana, Niger, and Togo, respectively. In this analysis, we omit those females in the bottom tercile as the control group. The sample includes those females aged 15 to 49 at the survey and partially- and least-affected cohorts in criminalizing countries, excluding the fully-affected cohorts. The specification includes basic controls, demographics, and region fixed effect. "Basic controls" include fixed effects for the year of interview and country by age at reform. "Demographics" include the full set of dummies for rural and religion. We cluster standard errors at the country-cohort level. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

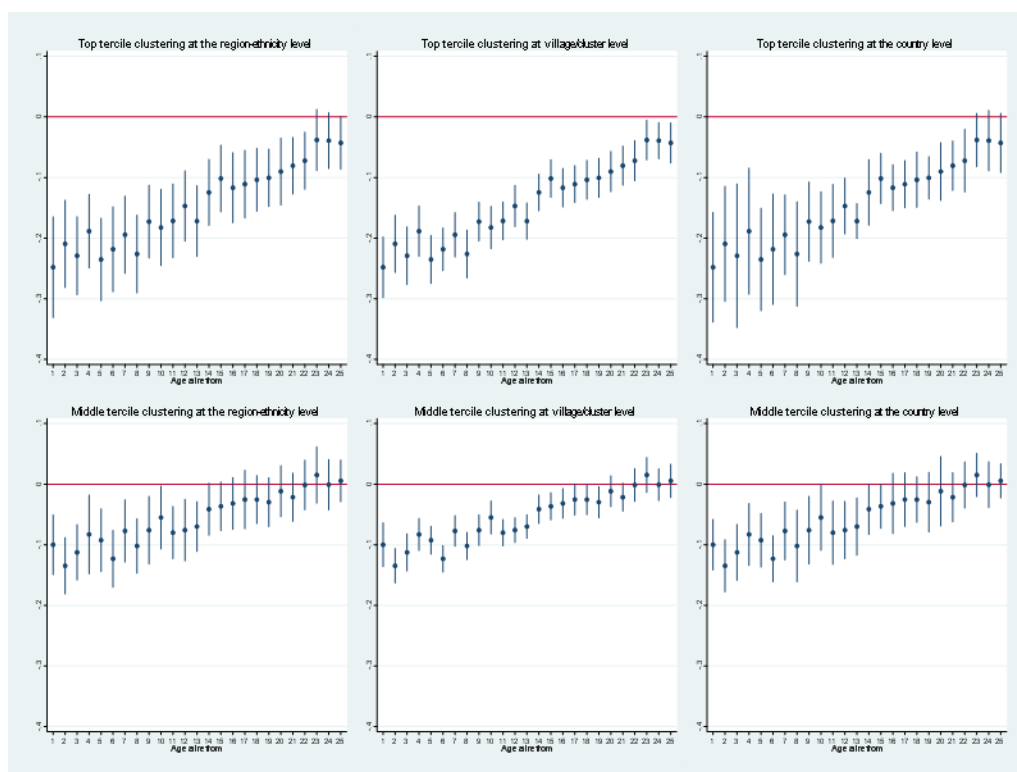


Figure 4.8: Event-study estimated effect of criminalization on ever-experiencing FGC by different clustering.

Notes: The top and bottom panels of the figure present the coefficients of interactions of age at reform dummies with dummies for the top and middle terciles of the pre-reform FGC prevalence rate in the region-ethnicity groups. The left, middle, and right panels report the results for clustering at the region-ethnicity, village/cluster, and country levels. In this analysis, we omit those females in the bottom tercile as the control group. The sample includes those females aged 15 to 49 at the survey and partially- and least-affected cohorts in criminalizing countries, excluding the fully-affected cohorts. The specification includes basic controls, demographics, and region fixed effect. "Basic controls" include fixed effects for the year of interview and country by age at reform. "Demographics" include the full set of dummies for rural and religion. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

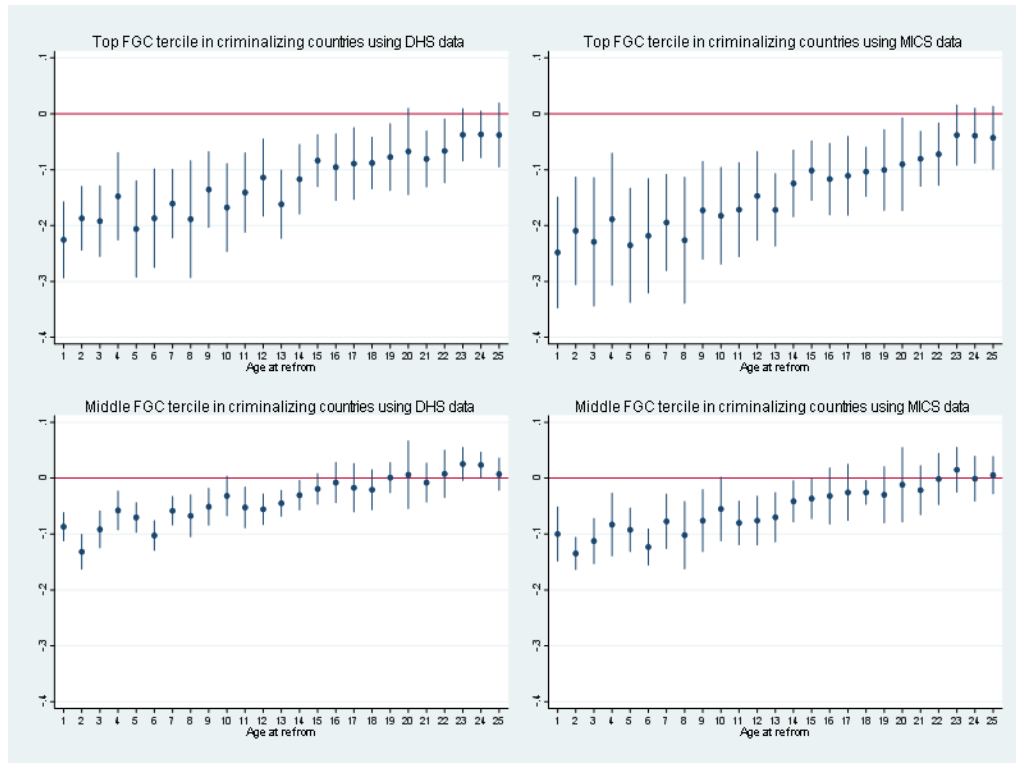


Figure 4.9: Event-study estimated effect of criminalization on ever-experiencing FGC by data type.

Notes: The top and bottom panels of the figure present the coefficients of interactions of age at reform dummies with dummies for the top and middle terciles of the pre-reform FGC prevalence rate in the region-ethnicity groups. The left and right panels report the results for the analyses using DHS and MICS respectively. In this analysis, we omit females in the bottom tercile as reference groups. The sample includes those females aged 15 to 49 at the survey and partially- and least-affected cohorts in criminalizing countries, excluding the fully-affected cohorts. The specification includes basic controls, demographics, and region fixed effect. "Basic controls" include fixed effects for the year of interview and country by age at reform. "Demographics" include the full set of dummies for rural and religion. We cluster standard errors at the country-cohort level. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' tabulation using DHS and MICS.

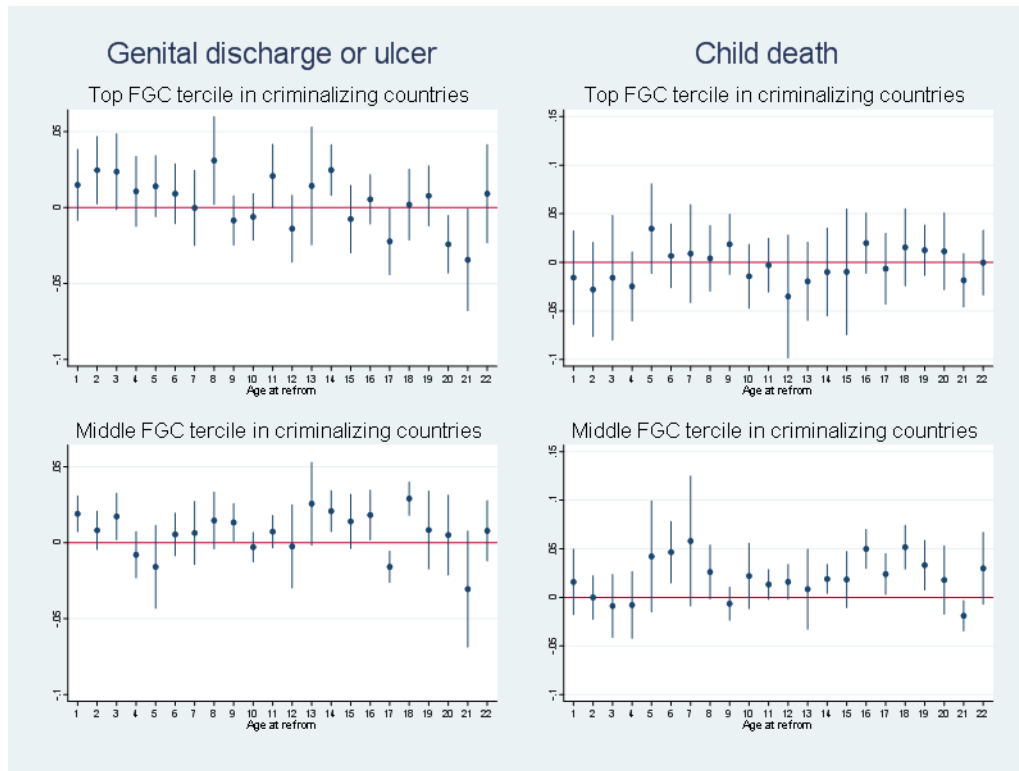


Figure 5.1: Reduced-form estimated effect of FGC criminalization on genital problems and child death among females aged 15 to 49 at the interview.

Notes: The left and right panels of the figure present the reduced-form event-study estimated effects of FGC criminalization on genital discharge or ulcer, and child death. The top and bottom panels of the figure report the results for the effect of criminalization among females in the top and middle terciles of the pre-reform FGC prevalence rate region-ethnicity groups. In addition to the interactions of pre-reform FGC prevalence terciles with age at reform dummies, the regression includes fixed effects for the year of interview, country by age at reform, rural, region, and religion. We omit those females in the bottom tercile as control group. The analysis in this figure uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.



Figure 5.2: Reduced-form estimated effect of FGC criminalization on marriage and fertility among females aged 15 to 49 at the interview.

Notes: The left and right panels of the figure present the reduced-form event-study estimated effects of FGC criminalization on marriage and fertility before interview. The top and bottom panels of the figure report the results for the effect of criminalization among females in the top and middle terciles of the pre-reform FGC prevalence rate region-ethnicity groups. In addition to the interactions of pre-reform FGC prevalence terciles with age at reform dummies, the regression includes fixed effects for the year of interview, country by age at reform, rural, region, and religion. We omit those females in the bottom tercile as control group. The analysis in this figure uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

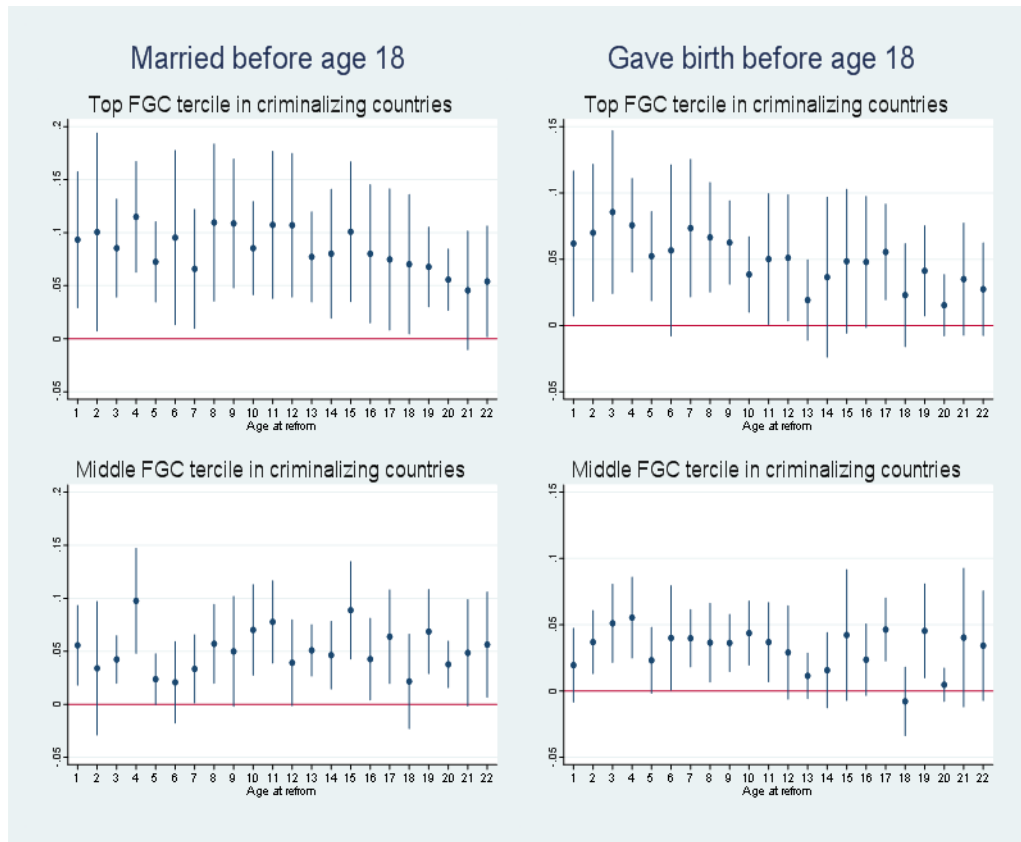


Figure 5.3: Reduced-form estimated effect of FGC criminalization on early marriage and fertility among females aged 15 to 49 at the interview.

Notes: The left, middle, and right panels of the figure present the reduced-form event-study estimated effects of FGC criminalization on marriage and fertility before age 18. The top and bottom panels of the figure report the results for the effect of criminalization among females in the top and middle terciles of the pre-reform FGC prevalence rate region-ethnicity groups. In addition to the interactions of pre-reform FGC prevalence terciles with age at reform dummies, the regression includes fixed effects for the year of interview, country by age at reform, rural, region, and religion. We omit those females in the bottom tercile as control group. The analysis in this figure uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

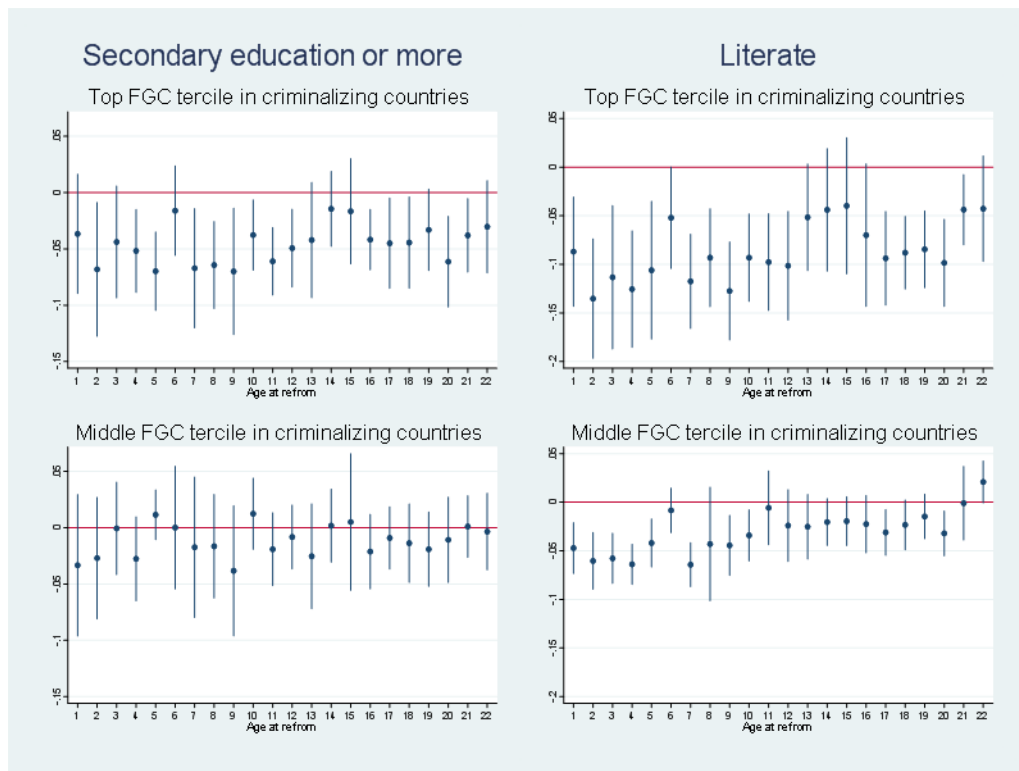


Figure 5.4: Reduced-form estimated effect of FGC criminalization on own education among females aged 15 to 49 at the interview.

Notes: The left and right panels of the figure present the reduced-form event-study estimated effects of FGC criminalization on secondary education and literacy. The top and bottom panels of the figure report the results for the effect of criminalization among females in the top and middle terciles of the pre-reform FGC prevalence rate region-ethnicity groups. In addition to the interactions of pre-reform FGC prevalence terciles with age at reform dummies, the regression includes fixed effects for the year of interview, country by age at reform, rural, region, and religion. We omit those females in the bottom tercile as control group. The analysis in this figure uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

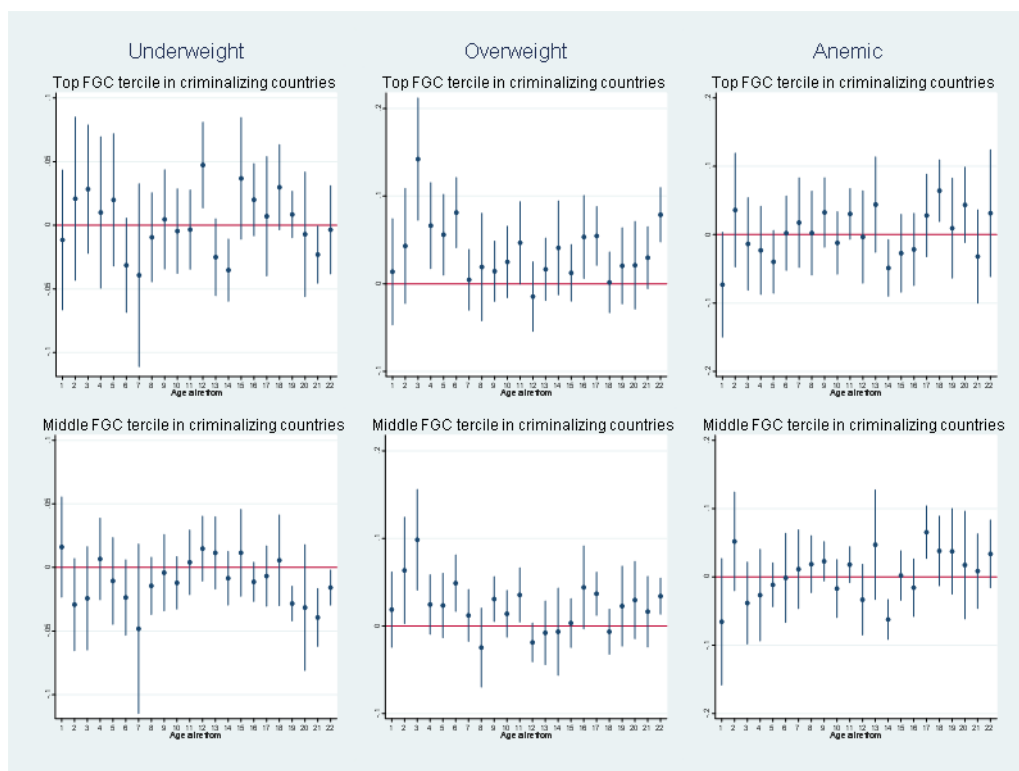


Figure 5.5: Reduced-form estimated effect of FGC criminalization on own general health among females aged 15 to 49 at the interview.

Notes: The left, middle, and right panels of the figure present the reduced-form event-study estimated effects of FGC criminalization on being underweight, over-weight, and anemic. The top and bottom panels of the figure report the results for the effect of criminalization among females in the top and middle tertiles of the pre-reform FGC prevalence rate region-ethnicity groups. In addition to the interactions of pre-reform FGC prevalence tertiles with age at reform dummies, the regression includes fixed effects for the year of interview, country by age at reform, rural, region, and religion. We omit those females in the bottom tertile as control group. The analysis in this figure uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

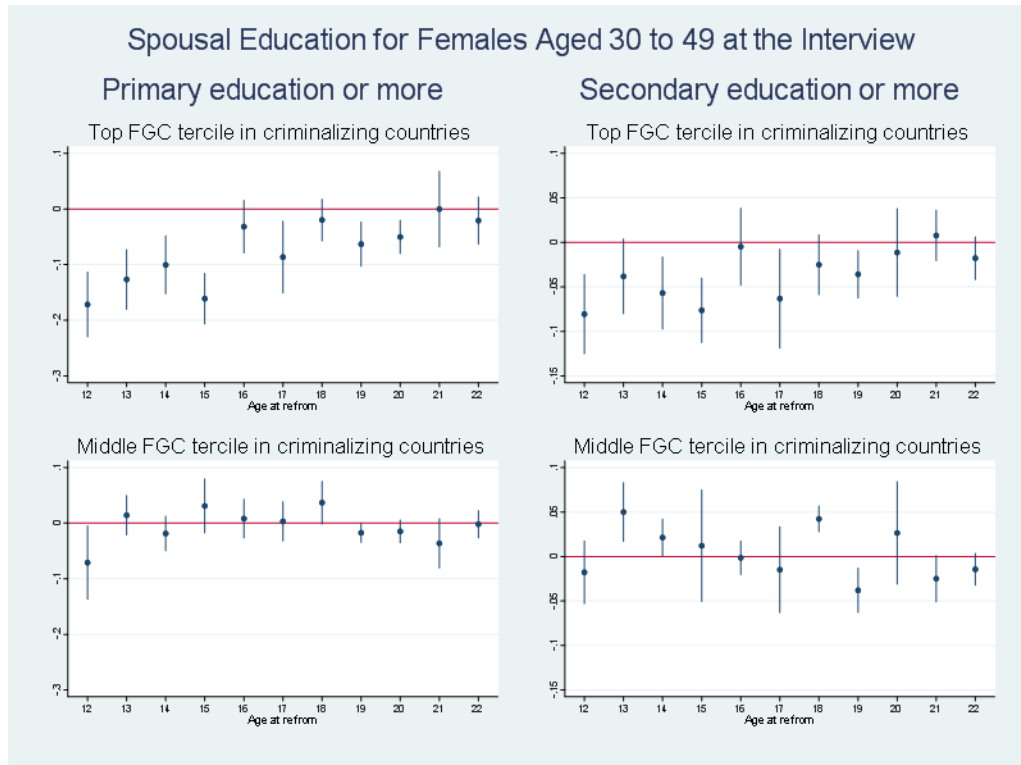


Figure 5.6: Reduced-form estimated effect of FGC criminalization on spousal education among females aged 30 to 49 at the interview.

Notes: The left and right panels of the figure present the reduced-form event-study estimated effects of FGC criminalization on spousal primary or more education, and secondary or more education, respectively. The top and bottom panels of the figure report the results for the effect of criminalization among females in the top and middle terciles of the pre-reform FGC prevalence rate region-ethnicity groups. In addition to the interactions of pre-reform FGC prevalence terciles with age at reform dummies, the regression includes fixed effects for the year of interview, country by age at reform, rural, region, and religion. We omit those females in the bottom tercile as control group. The analysis in this figure uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

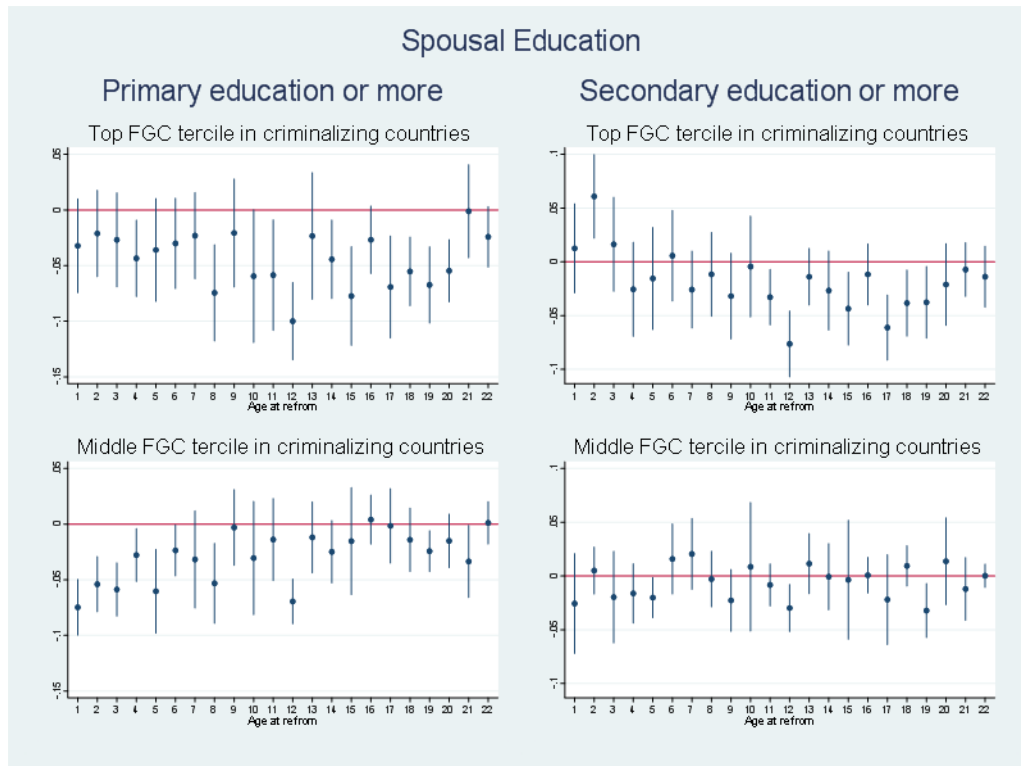


Figure 5.7: Reduced-form estimated effect of FGC criminalization on spousal education among females aged 15 to 49 at the interview.

Notes: The left and right panels of the figure present the reduced-form event-study estimated effects of FGC criminalization on spousal primary or more education, and secondary or more education, respectively. The top and bottom panels of the figure report the results for the effect of criminalization among females in the top and middle terciles of the pre-reform FGC prevalence rate region-ethnicity groups. In addition to the interactions of pre-reform FGC prevalence terciles with age at reform dummies, the regression includes fixed effects for the year of interview, country by age at reform, rural, region, and religion. We omit those females in the bottom tercile as control group. The analysis in this figure uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

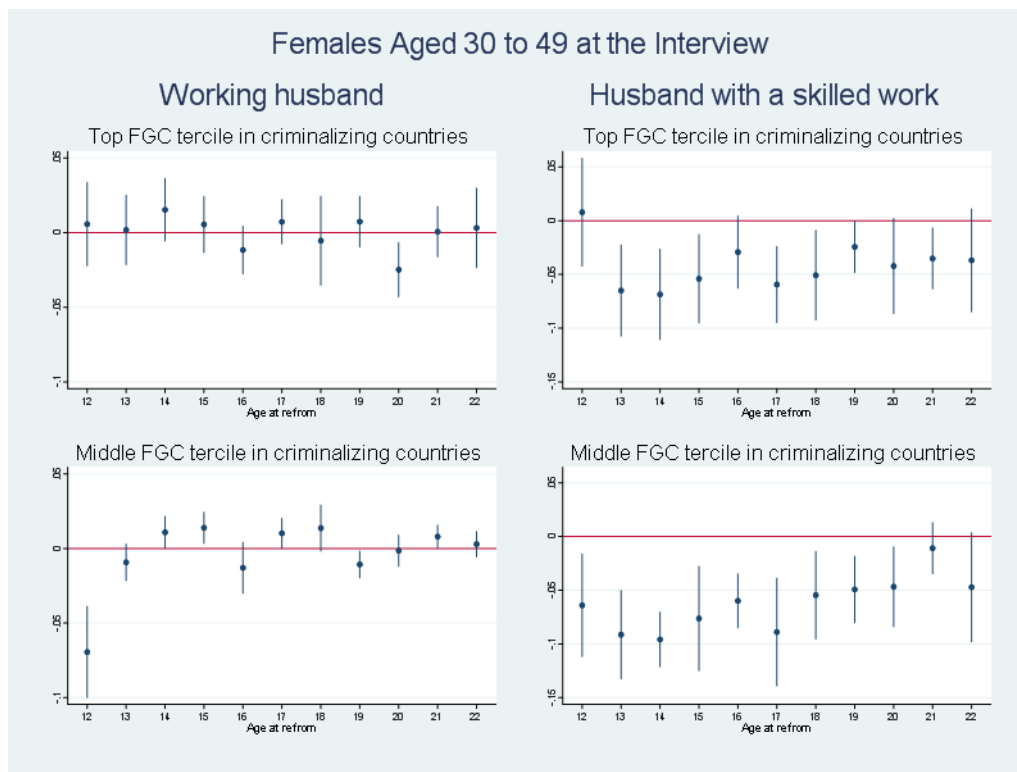


Figure 5.8: Reduced-form estimated effect of FGC criminalization on spousal occupation among females aged 30 to 49 at the interview.

Notes: The left and right panels of the figure present the reduced-form event-study estimated effects of FGC criminalization on spousal work-status and skilled work, respectively. The top and bottom panels of the figure report the results for the effect of criminalization among females in the top and middle terciles of the pre-reform FGC prevalence rate region-ethnicity groups. In addition to the interactions of pre-reform FGC prevalence terciles with age at reform dummies, the regression includes fixed effects for the year of interview, country by age at reform, rural, region, and religion. We omit those females in the bottom tercile as control group. The analysis in this figure uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.



Figure 5.9: Reduced-form estimated effect of FGC criminalization on spousal occupation among females aged 15 to 49 at the interview.

Notes: The left and right panels of the figure present the reduced-form event-study estimated effects of FGC criminalization on spousal work-status and skilled work, respectively. The top and bottom panels of the figure report the results for the effect of criminalization among females in the top and middle terciles of the pre-reform FGC prevalence rate region-ethnicity groups. In addition to the interactions of pre-reform FGC prevalence terciles with age at reform dummies, the regression includes fixed effects for the year of interview, country by age at reform, rural, region, and religion. We omit those females in the bottom tercile as control group. The analysis in this figure uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

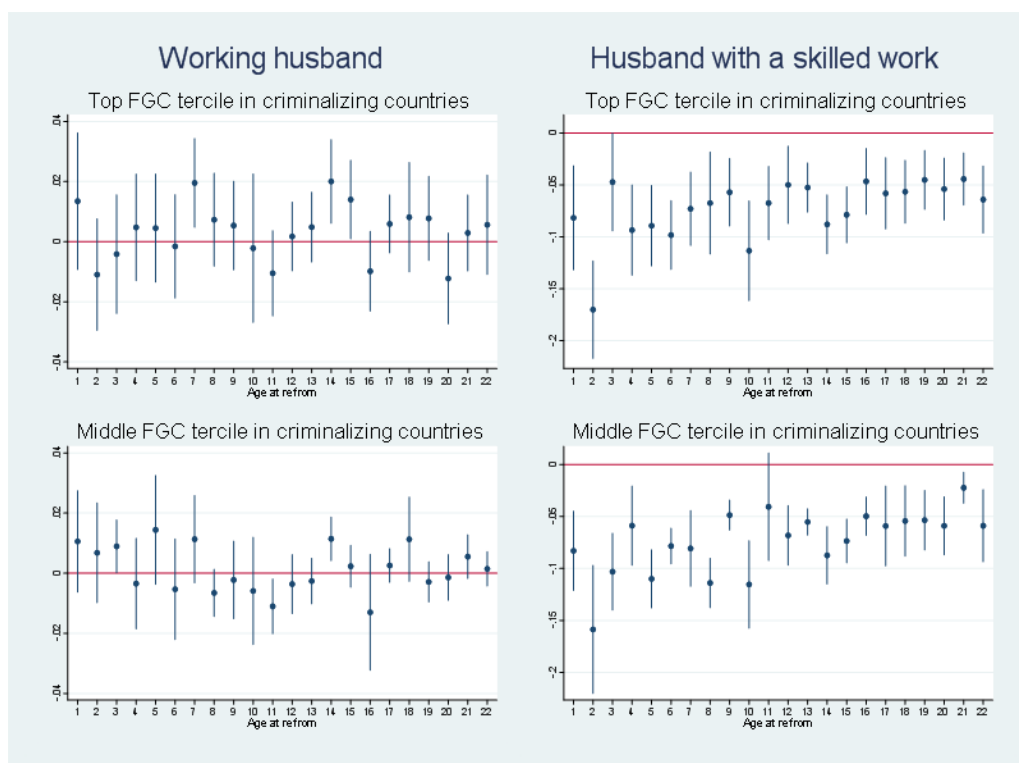


Figure 5.10: Reduced-form estimated effect of FGC criminalization on family wealth among females aged 30 to 49 at the interview.

Notes: The left and right panels of the figure present the reduced-form event-study estimated effects of FGC criminalization on having low and high family wealth, respectively. The top and bottom panels of the figure report the results for the effect of criminalization among females in the top and middle terciles of the pre-reform FGC prevalence rate region-ethnicity groups. In addition to the interactions of pre-reform FGC prevalence terciles with age at reform dummies, the regression includes fixed effects for the year of interview, country by age at reform, rural, region, and religion. We omit those females in the bottom tercile as control group. The analysis in this figure uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.



Figure 5.11: Reduced-form estimated effect of FGC criminalization on family wealth among females aged 15 to 49 at the interview.

Notes: The left and right panels of the figure present the reduced-form event-study estimated effects of FGC criminalization on having low and high family wealth, respectively. The top and bottom panels of the figure report the results for the effect of criminalization among females in the top and middle terciles of the pre-reform FGC prevalence rate region-ethnicity groups. In addition to the interactions of pre-reform FGC prevalence terciles with age at reform dummies, the regression includes fixed effects for the year of interview, country by age at reform, rural, region, and religion. We omit those females in the bottom tercile as control group. The analysis in this figure uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

Table 3.1: Data availability of African countries with an FGC ban as of 2013, year and type of reform, and post-reform surveys.

| Reform surveys. | | | | | | |
|--|--------------------------------|------|-------------------------|----------------------|------------------------------------|---|
| Country (state) | Pre-reform FGC prevalence rate | | Year of reform | Type of reform | Degree of enforcement ^f | Post-reform Data |
| | Mean | SD | | | | |
| <i>Panel A: Implementing criminal bans</i> | | | | | | |
| Benin | 0.15 | 0.36 | 2003 | Criminal legislation | Weak | 2006 DHS, 2011-12 DHS, 2014 MICS |
| Burkina Faso | 0.73 | 0.45 | 1996 | Criminal legislation | Strong | 1998-99 DHS, 2003 DHS, 2006 MICS, 2010 DHS |
| Cote d'Ivoire | 0.69 | 0.46 | 1998 | Criminal legislation | Strong | 1998-99 DHS, 2005 AIS, 2006 MICS, 2011-12 DHS, 2016 MICS |
| Egypt | 0.98 | 0.16 | 2008 | Criminal law | Strong | 2014 DHS |
| Ethiopia | 0.77 | 0.42 | 2004 | Criminal law | Strong | 2005 DHS, 2016 DHS |
| Ghana | 0.06 | 0.24 | 1994 | Criminal law | Weak | 2003 DHS, 2006 MICS, 2011 MICS, 2017-18 MICS |
| Niger | 0.05 | 0.22 | 2003 | Criminal legislation | Strong | 2006 DHS, 2012 DHS |
| Senegal | 0.29 | 0.46 | 1999 | Criminal law | Weak | 2005 DHS, 2010-11 DHS, 2014 Cont. DHS, 2015 Cont. DHS, 2016 Cont. and MIS, 2017 Cont. DHS |
| Tanzania | 0.20 | 0.40 | 1998 | Criminal law | Strong | 2004-05 DHS, 2010 DHS, 2015-16 DHS |
| Togo | 0.06 | 0.23 | 1998 | Criminal law | Weak | 2006 MICS, 2010 MICS, 2013-14 DHS, 2017 DHS |
| <i><u>Only MICS data available after reform</u></i> | | | | | | |
| Central African Republic | 0.41 | 0.49 | <u>1996^a</u> | Criminal law | Weak | 2000 MICS, 2006 MICS, 2010 MICS |
| Djibouti | 0.93 | 0.25 | 1995 | Criminal law | Weak | 2006 MICS |
| Guinea Bissau | 0.43 | 0.50 | 2011 | Criminal law | Strong | 2014 MICS |
| Mauritania | 0.71 | 0.45 | 2005 | Criminal law | Weak | 2007 MICS, 2011 MICS, 2015 MICS |

(continues to the next page)

Table 3.1: Continued.

| Country (state) | Pre-reform FGC prevalence rate | | Year of reform | Content of the reform | Degree of enforcement ^f | Post-reform Data |
|--|--------------------------------|------|------------------------------|-----------------------------------|------------------------------------|--|
| | Mean | SD | | | | |
| <i>Panel B: Implementing non-criminal bans</i> | | | | | | |
| Chad | 0.45 | 0.50 | 2003 | Reproductive health law | Weak | 2004 DHS, 2010 MICS, 2014-15 DHS |
| Guinea | 0.98 | 0.13 | <u>2000^b</u> | Reproductive health law | Strong | 2005 DHS, 2012 DHS, 2016 MICS, 2018 DHS |
| Kenya | 0.33 | 0.47 | 2001 | Children's Act | Strong | 2003 DHS, 2008-09 DHS, 2014 DHS |
| Nigeria | 0.44 | 0.50 | <u>1999-2009^c</u> | Laws prohibiting FGC | Weak | 2007 MICS, 2008 DHS, 2011 MICS, 2013 DHS, 2016-17 MICS, 2018 DHS |
| Uganda | 0.01 | 0.12 | 2010 | Law banning FGC | Strong | 2011 DHS, 2016 DHS |
| Sudan | 0.92 | 0.28 | <u>2008-2009^d</u> | Prevention of FGC Act | Weak | 2014 MICS |
| <i>Countries excluded due to data limitations:</i> | | | | | | |
| Eritrea | | | <u>2007^e</u> | Criminal law | Strong | na |
| Somalia | | | 2012 | Provisional Constitutional decree | Weak | na |

Notes: Of the 22 African countries having an FGC ban, only 20 have available DHS/MICS data for analysis. FGC is also prevalent in Cameroon, Gambia, Liberia, Mali, and Sierra Leone, not on this list because those countries have not imposed any ban as of 2013. (a)(b) Both the Central African Republic and Guinea have introduced FGC bans earlier in 1966 and 1965, respectively (Shell-Duncan et al. 2013, UNICEF 2013). This study focused on the later reforms to observe the younger cohorts comparable with those in other countries. (c) Since 1999, several states have rolled out an FGC ban in Nigeria (see <https://www.28toomany.org/nigeria/>). These states include Edo, Cross River, Ebonyi, Enugu and Bayelsa, and Rivers, having FGC bans since 1999, 2000, 2001, 2004, and 2009. We exclude the 1990 and 2003 surveys due to a lack of information regarding the state of residence. (d) Only two states in Sudan (South Kordofan and Gadaref) introduced an FGC ban in 2008 and 2009 (Shell-Duncan et al. 2013). (e) Although already criminalized FGC since 2007, Eritrea has no public-use DHS available for analysis. (f) "Strong" and "Weak" enforcement represent those countries with above and below-median number of arrests and prosecutions in the concerned countries as of 2017 since the years of reform, respectively.

Source: Authors' compilations from DHS, MICS, UNICEF (2013), Shell-Duncan et al. (2013), and <https://www.28toomany.org/nigeria/>.

Table 3.2: Can the pre-reform country and region-ethnicity characteristics explain the timing of reform adoption and the type of reform?

| | Dependent variables | | | | | | |
|--------------------------------------|----------------------------------|-------------------|-------------------|-------------------------|----------------------------|--|----------------------------------|
| | Enactment year of FGC ban | | | | | | A dummy for criminal law |
| | 20 countries with any ban on FGC | | | | 14 criminalizing countries | 6 non-criminalizing countries ^a | 20 countries with any ban on FGC |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Pre-reform FGC prevalence rate | -0.615 (2.590) | -7.363 (7.266) | -6.573 (6.570) | -1.510 (2.461) | -4.912 (2.308) | -0.088 (0.112) | -0.265 (0.210) |
| Pre-reform FGC prevalence, squared | | 7.159 (5.469) | 6.316 (4.588) | 1.725 (2.307) | 2.243 (1.537) | 0.084 (0.111) | -0.126 (0.238) |
| 1989-93 log of population in million | | | 0.986 (0.873) | 0.156 (3.057) | 8.079 (2.319) | | 0.336 (0.161) |
| 1989-93 ratio of foreign aid to GNI | | | | 0.262 (0.263) | 0.813 (0.180) | 1.195 (0.002) | 0.015 (0.016) |
| 1989-93 fertility rate | | | | 5.013 (2.238) | -1.523 (1.549) | 11.670 (0.002) | -0.303 (0.167) |
| 1989-93 mortality rate | | | | -0.012 (0.023) | 0.038 (0.038) | -0.089 (0.000) | 0.000 (0.004) |
| 1989-93 percent of urban population | | | | 0.021 (0.090) | 0.167 (0.103) | 1.159 (0.001) | 0.019 (0.010) |
| Former French colony | | | | -1.658 (3.735) | -0.593 (2.720) | | -0.431 (0.331) |
| Former British colony | | | | 2.092 (3.576) | -1.651 (4.064) | | -1.658 (0.402) |
| Former French colony × foreign aid | | | | -0.091 (0.232) | 0.334 (0.181) | -1.127 (0.001) | 0.014 (0.014) |
| Former British colony × foreign aid | | | | -0.078 (0.267) | -0.648 (0.219) | | 0.031 (0.019) |
| Adjusted R-squared | 0.002 | 0.024 | 0.064 | 0.570 | 0.765 | 0.991 | 0.548 |
| Observations | 781 | 781 | 781 | 781 | 540 | 241 | 781 |

Notes: The data are at the region-ethnicity level and the dependent variable is the enactment year of the FGC reform per country/state in this study. The "ratio of foreign aid to GNI" is the so-called "official development assistance grants and loans" divided by the gross national income (GNI) of a country. Standard errors clustered at the country level are in parentheses. (a) The 6 non-criminalizing countries include 4 countries and 8 states/regions in 2 countries. Regressing the timing of the reform on the covariates for non-criminalizing countries, Stata omits the log of population size and former French colony. However, we omit the British colony dummy deliberately because France and Britain were the only colonizers for the non-criminalizing countries. Among the criminalizing countries, however, we take the countries colonized by Portugal or not colonized as a reference group, and include both French and British colony dummies in the regression.

Source: The data for the 1989-93 country characteristics come from the world development indicator (WDI). Bertocchi and Canova (2002) provide information on "Colonial ties." The sources of data for the pre-reform FGC prevalence rate are DHS and MICS.

Table 3.3: Summary statistics for country attributes by reform type

| | Criminalizing countries | | Non-criminalizing countries | | Criminalizing minus non- criminalizing countries | Standard error |
|---|----------------------------|-----------------------|--------------------------------|-----------------------|---|-------------------|
| | Sample mean | Standard deviation | Sample mean | Standard deviation | | |
| <i>1989-93 country characteristics:</i> | | | | | | |
| Total fertility rate | 6.41 | 0.66 | 6.67 | 0.58 | -0.26 | 0.38 |
| Infant mortality rate | 97.85 | 18.89 | 98.06 | 25.05 | -0.22 | 13.95 |
| Foreign aid (percent of GNI) | 13.73 | 6.03 | 14.49 | 3.72 | -0.76 | 1.75 |
| Percent of urban population | 31.60 | 12.00 | 18.95 | 5.99 | 12.65 | 4.11 |
| Population (in million) | 12.94 | 13.71 | 19.70 | 20.20 | -6.76 | 7.26 |
| <i>Colonial ties:</i> | | | | | | |
| French colony=1 | 0.72 | 0.45 | 0.39 | 0.49 | 0.33 | 0.28 |
| Britain colony=1 | 0.18 | 0.38 | 0.61 | 0.49 | -0.44 | 0.27 |
| Pre-reform FGC prevalence | 0.40 | 0.37 | 0.36 | 0.40 | 0.04 | 0.15 |
| Number of observations | 540 | | 241 | | 781 | |

Notes: The data are at the region-ethnicity level. The analysis sample includes females born between 1948 and 2003, excluding the pre-reform surveys and the fully affected cohorts. This table uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft, et al. (2018). We cluster standard errors at the country level. "1989-93 country characteristics" are the average of the country characteristics from 1989-1993, before Ghana's FGC ban in 1994, which is the earliest reform considered in this study. GNI stands for gross national income.

Source: Author's calculations using DHS from 1998 to 2018, MICS from 2000 to 2018, and WDI from 1989 to 1993. Bertocchi and Canova (2002) provide information on "Colonial ties."

Table 3.4: Correlation between FGC experience and non-reporting of the age at cutting

| | Dependent variable | | | | |
|-----------------------------|---|--------------------------|--------------------------|---------------------------------|---------------------------------|
| | =Age at cutting not reported [Mean = 0.740] | | | | |
| | All 20 countries with FGC ban | | | 14 Criminalizing countries | 6 Non-criminalizing countries |
| | (1) | (2) | (3) | (4) | (5) |
| Pre-reform FGC rate | -0.525 (0.034) | -0.541 (0.035) | -0.464 (0.039) | -0.521 (0.044) | -0.334 (0.046) |
| Pre-reform FGC rate squared | -0.047 (0.023) | -0.006 (0.024) | 0.042 (0.026) | 0.113 (0.028) | -0.108 (0.046) |
| Adjusted R-squared | 0.507 | 0.511 | 0.529 | 0.491 | 0.616 |
| Observation | 541,706 | 541,706 | 541,706 | 392,821 | 148,885 |
| Specifications: | | | | | |
| Basic controls | Yes | Yes | Yes | Yes | Yes |
| Demographics | No | Yes | Yes | Yes | Yes |
| Region fixed effect | No | No | Yes | Yes | Yes |

Notes: This table reports the association between the pre-reform FGC prevalence rate across regions by ethnicities and non-reporting of age at cutting. The sample includes women who are partially-affected or least-affected by the reform while excluding the fully-affected cohorts existing in only seven countries. This table uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). "Basic controls" include fixed effects for the year of interview and country by age at reform. "Demographics" include the full set of dummies for rural and religion. Standard errors at the country-cohort levels are in parenthesis.

Source: Authors' tabulation using DHS and MICS.

Table 4.1: African countries with FGC criminalization as of 2013 and had DHS/MICS respondents affected fully or partially by the reform.

| Country (state) | Year of reform | Youngest available cohort | Oldest available cohort | Birth year of DHS/MICS respondents | | | | | | | |
|--------------------------|----------------|---------------------------|-------------------------|------------------------------------|--------|----------------------------|--------|------------------------|-----|------|--|
| | | | | Least-affected cohorts | | Partially-affected cohorts | | Fully-affected cohorts | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | | |
| Benin | 2003 | 1999 | 1956 | 1967 | - 1980 | 1981 | - 1999 | na | - | na | |
| Burkina Faso | 1996 | 1995 | 1948 | 1960 | - 1973 | 1974 | - 1995 | na | - | na | |
| Central African Republic | 1996 | 1995 | 1950 | 1960 | - 1973 | 1974 | - 1995 | na | - | na | |
| Cote d'Ivoire | 1998 | 2001 | 1948 | 1962 | - 1975 | 1976 | - 1997 | 1998 | - | 2001 | |
| Djibouti | 1995 | 1991 | 1957 | 1959 | - 1972 | 1973 | - 1991 | na | - | na | |
| Egypt | 2008 | 1999 | 1964 | 1972 | - 1985 | 1986 | - 1999 | na | - | na | |
| Ethiopia | 2004 | 2001 | 1949 | 1968 | - 1981 | 1982 | - 2001 | na | - | na | |
| Ghana | 1994 | 2003 | 1953 | 1958 | - 1971 | 1972 | - 1993 | 1994 | - | 2003 | |
| Guinea Bissau | 2011 | 1999 | 1964 | 1975 | - 1988 | 1989 | - 1999 | na | - | na | |
| Mauritania | 2005 | 2000 | 1957 | 1969 | - 1982 | 1983 | - 2000 | na | - | na | |
| Niger | 2003 | 1997 | 1956 | 1967 | - 1980 | 1981 | - 1997 | na | - | na | |
| Senegal | 1999 | 2002 | 1955 | 1963 | - 1976 | 1977 | - 1998 | 1999 | - | 2002 | |
| Tanzania | 1998 | 2001 | 1954 | 1962 | - 1975 | 1976 | - 1997 | 1998 | - | 2001 | |
| Togo | 1998 | 2002 | 1956 | 1962 | - 1975 | 1976 | - 1997 | 1998 | - | 2002 | |

Notes: "Least-affected cohorts" are females whose age at the reform is 23 or more, while "Partially-affected cohorts" include those below age 23 at the reform. "Fully-affected cohorts" cover those born after the reform. This table uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). Column (1) = the year of reform. Column (2) = the youngest available cohort in the surveys. Column (3)=the oldest available cohort in the surveys. Column (4) = Column (1)-36. Column (5) = Column (1)-23. Column (6)=Column (5)+1. Column (7) = Column (6)+21 if Column (2) > Column (6)+21, and Column (7) = Column (2) if Column (2) <= Column (6)+21.

Source: Authors' tabulation using DHS and MICS.

Table 4.2: The effect of FGC criminalization on the probability of ever-experiencing FGC for each cohort

| Age at reform | Dependent variable = An indicator for ever-cut (Mean=0.347) | | | | | | | |
|---------------|---|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| | (1) | | (2) | | (3) | | (4) | |
| | Top tercile | Middle tercile | Top tercile | Middle tercile | Top tercile | Middle tercile | Top tercile | Middle tercile |
| 1 | -0.177 (0.062) | -0.097 (0.029) | -0.172 (0.066) | -0.092 (0.034) | -0.211 (0.041) | -0.108 (0.027) | -0.248 (0.050) | -0.100 (0.025) |
| 2 | -0.139 (0.059) | -0.131 (0.018) | -0.133 (0.064) | -0.123 (0.021) | -0.171 (0.040) | -0.140 (0.011) | -0.209 (0.049) | -0.135 (0.015) |
| 3 | -0.187 (0.084) | -0.100 (0.028) | -0.186 (0.093) | -0.096 (0.034) | -0.209 (0.063) | -0.114 (0.022) | -0.229 (0.059) | -0.113 (0.021) |
| 4 | -0.155 (0.096) | -0.078 (0.041) | -0.154 (0.104) | -0.073 (0.046) | -0.174 (0.070) | -0.090 (0.037) | -0.189 (0.060) | -0.083 (0.029) |
| 5 | -0.192 (0.080) | -0.107 (0.035) | -0.193 (0.089) | -0.100 (0.042) | -0.212 (0.059) | -0.104 (0.029) | -0.235 (0.052) | -0.093 (0.020) |
| 6 | -0.190 (0.088) | -0.141 (0.037) | -0.194 (0.095) | -0.137 (0.042) | -0.199 (0.064) | -0.135 (0.026) | -0.218 (0.052) | -0.123 (0.017) |
| 7 | -0.145 (0.078) | -0.090 (0.047) | -0.150 (0.087) | -0.091 (0.054) | -0.171 (0.055) | -0.091 (0.038) | -0.194 (0.044) | -0.077 (0.025) |
| 8 | -0.197 (0.088) | -0.122 (0.050) | -0.204 (0.096) | -0.121 (0.056) | -0.207 (0.067) | -0.117 (0.040) | -0.226 (0.057) | -0.102 (0.031) |
| 9 | -0.131 (0.077) | -0.092 (0.050) | -0.134 (0.084) | -0.090 (0.056) | -0.149 (0.052) | -0.090 (0.038) | -0.173 (0.045) | -0.076 (0.028) |
| 10 | -0.130 (0.070) | -0.064 (0.042) | -0.132 (0.076) | -0.061 (0.048) | -0.153 (0.050) | -0.063 (0.036) | -0.182 (0.044) | -0.055 (0.029) |
| 11 | -0.123 (0.070) | -0.091 (0.037) | -0.127 (0.077) | -0.087 (0.042) | -0.145 (0.048) | -0.086 (0.023) | -0.171 (0.043) | -0.080 (0.020) |
| 12 | -0.108 (0.068) | -0.091 (0.041) | -0.112 (0.075) | -0.087 (0.046) | -0.123 (0.045) | -0.083 (0.028) | -0.147 (0.040) | -0.076 (0.022) |
| 13 | -0.125 (0.057) | -0.079 (0.036) | -0.132 (0.065) | -0.076 (0.042) | -0.145 (0.037) | -0.075 (0.026) | -0.172 (0.033) | -0.070 (0.022) |
| 14 | -0.082 (0.052) | -0.051 (0.033) | -0.088 (0.058) | -0.048 (0.039) | -0.098 (0.030) | -0.046 (0.021) | -0.125 (0.030) | -0.041 (0.019) |
| 15 | -0.055 (0.047) | -0.048 (0.031) | -0.062 (0.053) | -0.045 (0.037) | -0.074 (0.027) | -0.039 (0.020) | -0.102 (0.027) | -0.036 (0.018) |
| 16 | -0.077 (0.055) | -0.044 (0.041) | -0.082 (0.058) | -0.040 (0.044) | -0.090 (0.032) | -0.033 (0.026) | -0.117 (0.033) | -0.032 (0.026) |
| 17 | -0.070 (0.057) | -0.036 (0.039) | -0.079 (0.064) | -0.035 (0.046) | -0.085 (0.034) | -0.025 (0.026) | -0.111 (0.036) | -0.025 (0.026) |
| 18 | -0.068 (0.040) | -0.030 (0.023) | -0.076 (0.044) | -0.027 (0.026) | -0.079 (0.020) | -0.023 (0.010) | -0.104 (0.023) | -0.025 (0.011) |

(continues to the next page)

Table 4.2: Continued.

| Dependent variable = An indicator for ever-cut (Mean=0.347) | | | | | | | | |
|---|-------------------|-------------------|-------------------|-------------------|--------------------------|-------------------|--------------------------|-------------------|
| | (1) | | (2) | | (3) | | (4) | |
| Age at reform | Top tercile | Middle tercile | Top tercile | Middle tercile | Top tercile | Middle tercile | Top tercile | Middle tercile |
| 19 | -0.057 (0.055) | -0.030 (0.037) | -0.063 (0.064) | -0.026 (0.044) | -0.075 (0.031) | -0.026 (0.023) | -0.100 (0.037) | -0.030 (0.026) |
| 20 | -0.049 (0.058) | -0.016 (0.041) | -0.057 (0.064) | -0.015 (0.047) | -0.067 (0.037) | -0.006 (0.031) | -0.090 (0.042) | -0.012 (0.034) |
| 21 | -0.048 (0.042) | -0.029 (0.033) | -0.051 (0.044) | -0.024 (0.036) | -0.057 (0.020) | -0.014 (0.020) | -0.080 (0.025) | -0.022 (0.022) |
| 22 | -0.042 (0.042) | -0.003 (0.028) | -0.050 (0.048) | -0.002 (0.034) | -0.051 (0.023) | 0.008 (0.019) | -0.072 (0.028) | -0.001 (0.023) |
| 23 | -0.004 (0.043) | 0.016 (0.027) | -0.016 (0.048) | 0.014 (0.032) | -0.017 (0.021) | 0.027 (0.017) | -0.038 (0.027) | 0.015 (0.020) |
| 24 | -0.011 (0.041) | 0.005 (0.032) | -0.018 (0.046) | 0.007 (0.036) | -0.018 (0.021) | 0.011 (0.016) | -0.039 (0.025) | -0.001 (0.021) |
| 25 | -0.007 (0.044) | 0.008 (0.026) | -0.015 (0.047) | 0.010 (0.029) | -0.030 (0.021) | 0.019 (0.011) | -0.043 (0.029) | 0.005 (0.017) |
| Adjusted R-squared | 0.529 | | 0.535 | | 0.544 | | 0.546 | |
| Sample size | 392,933 | | 392,933 | | 392,933 | | 392,933 | |
| Specifications: | | | | | | | | |
| Year of survey FE | Yes | | Yes | | Yes | | Yes | |
| Country by cohort FE | Yes | | Yes | | Yes | | Yes | |
| Pre-reform FGC prevalence rate | Yes | | Yes | | Yes | | Yes | |
| Religion and rural dummies | No | | Yes | | Yes | | Yes | |
| Region FE | No | | No | | Yes | | Yes | |
| Region-specific cohort trend | No | | No | | No | | Yes | |

Notes: Columns (1), (2), (3) and (4) present separate regressions of the effect of criminalization on FGC, using different specifications. All the four specifications include the interactions of each cohort with "top" and "middle" terciles of pre-reform FGC prevalence rate, in addition to the "basic controls" and pre-reform FGC prevalence rate at region-ethnicity group. Column (1) only includes fixed effects for the year of interview and country by age at reform. In column (2), we add demographics, including the full set of dummies for rural and religion. Also, we capture the region fixed effect in column (3). Finally, in column (4), we add the region-specific cohort trend. The sample includes women who are partially-affected or least-affected by the reform while excluding the fully-affected cohorts (with zero or negative age at reform), making below 2 percent of our sample. Standard errors at country-cohort levels are in parentheses. Source: Authors' computation using DHS and MICS.

Table 4.3: Unconditional reform effect in a difference-in-difference design: Correlation between FGC criminalization and the probability of ever-experiencing FGC.

| | Dependent variable = Ever experiencing FGC | | | | |
|--|--|--------------------------|--------------------------|--------------------------|--------------------------|
| | Pre-reform FGC prevalence rate | | | Top minus Bottom | Middle minus Bottom |
| | Top tercile | Middle tercile | Bottom tercile | | |
| | (1) | (2) | (3) | (4) | (5) |
| Panel A: Experiment of Interest | | | | | |
| Young (Aged 1 to 22 at reform) | 0.728 (0.015) | 0.317 (0.016) | 0.014 (0.002) | 0.714 (0.015) | 0.303 (0.016) |
| Old (Aged 23 to 36 at reform) | 0.849 (0.011) | 0.424 (0.018) | 0.023 (0.002) | 0.826 (0.012) | 0.400 (0.019) |
| Young minus Old | -0.121 (0.019) | -0.107 (0.024) | -0.009 (0.003) | -0.112 (0.019) | -0.097 (0.024) |
| Number of observations | 129,420 | 132,472 | 131,041 | 260,460 | 263,401 |
| Panel B: Falsification Experiment | | | | | |
| Young (Aged 23 to 29 at reform) | 0.865 (0.012) | 0.446 (0.023) | 0.027 (0.004) | 0.838 (0.013) | 0.419 (0.023) |
| Old (Aged 30 to 36 at reform) | 0.893 (0.010) | 0.493 (0.025) | 0.036 (0.005) | 0.857 (0.012) | 0.457 (0.026) |
| Young minus Old | -0.028 (0.016) | -0.047 (0.034) | -0.009 (0.007) | -0.019 (0.018) | -0.037 (0.035) |
| Number of observations | 42,688 | 42,823 | 42,920 | 85,625 | 85,599 |

Notes: The "young" cohorts include those females aged 1 to 22 at reform, and as a result, more likely to be partially affected by the reform. The "old" cohorts include those females whose age at reform was 23 or more, and as a result, they are least affected by the reform. "Top tercile", "middle tercile", and "bottom tercile" represent the females' region-ethnicity groups by pre-reform FGC prevalence rate. The sample includes both the partially-affected and least-affected cohorts while excluding the fully-affected cohorts. We use the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). Clustered standard errors at country-cohort levels are in parentheses.

Source: Authors' tabulation using DHS and MICS.

Table 4.4: The effect of FGC criminalization on the probability of ever-experiencing FGC.

| | Dependent variable = An indicator for ever-cut | | | |
|--|--|--------------------------|--------------------------|--------------------------|
| | (1) | (2) | (3) | (4) |
| Panel A: Experiment of interest by comparing females aged 1 to 22 with those aged 23 to 36 at reform | | | | |
| Age 1 to 22 at reform * Top tercile of pre-reform FGC rate | -0.109 (0.013) | -0.111 (0.014) | -0.122 (0.010) | -0.127 (0.013) |
| Age 1 to 22 at reform * Middle tercile of pre-reform FGC rate | -0.070 (0.008) | -0.066 (0.009) | -0.065 (0.007) | -0.052 (0.008) |
| Adjusted R-squared | 0.529 | 0.534 | 0.543 | 0.546 |
| Sample size | 392,933 | | | |
| Mean FGC among females aged 23 to 36 at reform | 0.429 | | | |
| Panel B: Falsification experiment by comparing females aged 23 to 29 with those aged 30 to 36 at reform | | | | |
| Age 23 to 29 at reform * Top tercile of pre-reform FGC rate | 0.012 (0.019) | 0.008 (0.020) | -0.010 (0.015) | -0.023 (0.022) |
| Age 23 to 29 at reform * Middle tercile of pre-reform FGC rate | 0.007 (0.013) | 0.013 (0.014) | 0.004 (0.012) | 0.009 (0.016) |
| Adjusted R-squared | 0.563 | 0.568 | 0.575 | 0.576 |
| Sample size | 128,431 | | | |
| Mean FGC among females aged 30 to 36 at reform | 0.465 | | | |
| Specifications: | | | | |
| Year of interview/survey fixed effect | Yes | Yes | Yes | Yes |
| Country by age at reform fixed effect | Yes | Yes | Yes | Yes |
| Pre-reform FGC prevalence rate at region-ethnicity level | Yes | Yes | Yes | Yes |
| Religion and rural dummies | No | Yes | Yes | Yes |
| Region fixed effect | No | No | Yes | Yes |
| Region-specific cohort trend | No | No | No | Yes |

Notes: Panel A presents the estimated results for the effect of criminalization, comparing "young" cohorts (aged 1 to 22 at reform) with "old" cohorts (aged 23 to 36 at reform). Panel B presents falsification results using "old" cohorts. Regressions in each panel for all columns include the interaction of "young" cohort with "top" and "middle" terciles of pre-reform FGC prevalence rate, in addition to the covariates. Column (1) for each panel includes fixed effects for the year of interview and country by age at reform, and pre-reform FGC prevalence rate. In column (2), we include the full set of dummies for rural and religion. Also, we capture the region fixed effect in column (3). Finally, we control for the region-specific cohort trend in column (4). The sample includes women who are partially-affected and least-affected by the reform while excluding the fully-affected cohorts (with zero or negative age at reform and very small sample size). Standard errors at country-cohort levels are in parentheses.

Source: Authors' tabulation using DHS and MICS.

Table 4.5: The heterogeneous effects of FGC criminalization on the probability of ever-experiencing FGC by residence and religion.

| | Dependent variable = An indicator for ever-cut | | | |
|--|--|--------------------------|---------------------------------|--------------------------------|
| | Residence | | Religion | |
| | Rural | Urban | Christian | Muslim |
| | (1) | (2) | (3) | (4) |
| Panel A: Experiment of interest by comparing females aged 1 to 22 with those aged 23 to 36 at reform | | | | |
| Age 1 to 22 at reform * Top tercile of pre-reform FGC rate | -0.110 (0.014) | -0.156 (0.018) | -0.159 (0.018) | -0.132 (0.018) |
| Age 1 to 22 at reform * Middle tercile of pre-reform FGC rate | -0.064 (0.007) | -0.044 (0.012) | -0.106 (0.009) | -0.026 (0.012) |
| Adjusted R-squared | 0.578 | 0.493 | 0.427 | 0.593 |
| Sample size | 232,798 | 160,135 | 115,174 | 168,330 |
| Mean FGC among females aged 23 to 36 at reform | 0.470 | 0.370 | 0.293 | 0.549 |
| Panel B: Falsification experiment by comparing females aged 23 to 29 with those aged 30 to 36 at reform | | | | |
| Age 23 to 29 at reform * Top tercile of pre-reform FGC rate | -0.019 (0.022) | -0.027 (0.028) | -0.048 (0.025) | -0.035 (0.029) |
| Age 23 to 29 at reform * Middle tercile of pre-reform FGC rate | 0.002 (0.014) | 0.003 (0.022) | -0.048 (0.016) | 0.051 (0.017) |
| Adjusted R-squared | 0.614 | 0.519 | 0.489 | 0.636 |
| Sample size | 77,756 | 50,675 | 39,364 | 53,232 |
| Mean FGC among females aged 30 to 36 at reform | 0.506 | 0.406 | 0.313 | 0.592 |
| Specifications: | | | | |
| Year of interview/survey fixed effect | Yes | Yes | Yes | Yes |
| Country by age at reform fixed effect | Yes | Yes | Yes | Yes |
| Pre-reform FGC prevalence rate at region-ethnicity level | Yes | Yes | Yes | Yes |
| Rural indicator | No | No | Yes | Yes |
| Religion dummies | Yes | Yes | No | No |
| Region fixed effect | Yes | Yes | Yes | Yes |
| Region-specific cohort trend | Yes | Yes | Yes | Yes |

Notes: Panel A presents the estimated results for the effect of criminalization by residence and religion, comparing "young" cohorts (aged 1 to 22 at reform) with "old" cohorts (aged 23 to 36 at reform). Panel B presents falsification results using "old" cohorts. Regressions in each panel for all columns include the interaction of "young" cohort with "top" and "middle" terciles of pre-reform FGC prevalence rate, in addition to the covariates. All columns for each panel include fixed effects for the year of interview and country by age at reform, pre-reform FGC prevalence rate, region fixed effect, and region-specific cohort trend. We exclude residence in columns (1) and (2), and religion in columns (3) and (4). The sample includes women who are partially-affected and least-affected by the reform while excluding the fully-affected cohorts (with zero or negative age at reform and very small sample size). Standard errors at country-cohort levels are in parentheses.

Source: Authors' tabulation using DHS and MICS.

Table 5.1: Summary statistics for outcome variables

| | Mean | Std. dev. | Obs. |
|---|------|--------------|---------|
| <i>Own reproductive health:</i> | | | |
| Genital ulcer or discharge | 0.11 | 0.31 | 225,512 |
| Any dead child = 1 | 0.24 | 0.43 | 392,933 |
| <i>Marriage and birth outcomes:</i> | | | |
| Married before interview = 1 | 0.78 | 0.42 | 392,933 |
| Married before age 18 = 1 | 0.34 | 0.47 | 392,933 |
| Have child before interview = 1 | 0.72 | 0.45 | 392,933 |
| Have child before age 18 = 1 | 0.14 | 0.35 | 392,933 |
| <i>Own education and work:</i> | | | |
| Secondary education or more = 1 | 0.27 | 0.44 | 392,933 |
| Literate ^a = 1 | 0.44 | 0.50 | 236,113 |
| Working = 1 | 0.61 | 0.49 | 245,566 |
| <i>Own General Health:</i> | | | |
| Underweight = 1 | 0.11 | 0.32 | 140,089 |
| Overweight = 1 | 0.26 | 0.44 | 140,089 |
| Anemia = 1 | 0.43 | 0.49 | 104,854 |
| <i>Partner's education and occupation:</i> | | | |
| Primary education or more = 1 | 0.44 | 0.50 | 184,607 |
| Secondary education or more = 1 | 0.22 | 0.41 | 184,607 |
| Working = 1 | 0.92 | 0.26 | 191,040 |
| Skilled work=1 | 0.17 | 0.38 | 191,040 |
| <i>Family Wealth:</i> | | | |
| Low wealth = 1 | 0.33 | 0.47 | 392,933 |
| High wealth = 1 | 0.43 | 0.50 | 392,933 |

Notes: The sample includes those females born between 1958 and 2001, and as a result, partially-and least-affected by FGC criminalization. (a) DHS defines literate women as "those who attended schooling higher than the secondary level or who can read a whole sentence or part of a sentence" (Croft et al. 2018, p. 3.8). We consider sampling weights based on the guide to DHS statistics to adjust for variations in probability of selection and interview due to design. "Underweight" takes 1 for those females with BMI less than 18.5, whereas the overweight include those with BMI of 25 or more.

Source: Authors' computation using DHS and MICS.

Table 5.2: The reduced-form estimated effect of criminalizing FGC on education, marriage, and health outcomes.

| | (1) | (2) | (3) | (4) | (5) |
|--|-----------------------------------|--------------------------|-------------------------|-----------------------------|---------------------------------|
| | Outcome variables | | | | |
| | Own education | | Marriage and fertility | | |
| | Secondary education or more | Literate | Ever- married | Married before age 18 | Having at least one child |
| Panel A: Education and marriage outcomes | | | | | |
| Age 1 to 22 at reform * Top tercile of pre-reform FGC rate | -0.042 (0.011) | -0.080 (0.021) | 0.068 (0.012) | 0.081 (0.017) | 0.049 (0.010) |
| Age 1 to 22 at reform * Middle tercile of pre-reform FGC rate | -0.011 (0.007) | -0.027 (0.009) | 0.036 (0.007) | 0.052 (0.009) | 0.036 (0.006) |
| Adjusted R-squared | 0.295 | 0.3 | 0.407 | 0.166 | 0.408 |
| Sample size | 392,933 | 236,113 | 392,933 | 356,967 | 392,933 |
| Mean among those aged 23 to 36 at reform | 0.203 | 0.385 | 0.959 | 0.385 | 0.940 |
| | Reproductive health | | General health | | |
| | Genital discharge or ulcer | Child death | Underweight | Over- weight | Anemia |
| Panel B: Health outcomes | | | | | |
| Age 1 to 22 at reform * Top tercile of pre-reform FGC rate | 0.000 (0.007) | -0.001 (0.010) | 0.001 (0.010) | 0.035 (0.012) | 0.009 (0.016) |
| Age 1 to 22 at reform * Middle tercile of pre-reform FGC rate | 0.005 (0.005) | 0.022 (0.006) | -0.012 (0.008) | 0.018 (0.009) | 0.010 (0.012) |
| Adjusted R-squared | 0.077 | 0.186 | 0.082 | 0.347 | 0.074 |
| Sample size | 225,512 | 283,386 | 140,089 | 140,089 | 104,854 |
| Mean among those aged 23 to 36 at reform | 0.111 | 0.417 | 0.092 | 0.363 | 0.416 |

Notes: Panels A and B present the reduced-form estimated results for the effect of criminalization on education and marriage, and health outcomes. The child health analysis in panel B limits the sample to those females who gave birth to at least one child. Regressions in each panel for all columns include the interaction of "young" cohort (aged 1 to 22 at reform) with "top" and "middle" terciles of pre-reform FGC prevalence rate, in addition to the controls. The controls include fixed effects for the year of interview and country by age at reform, pre-reform FGC prevalence rate, and full set of dummies for rural, religion, and region. Standard errors at country-cohort levels are in parentheses.

Source: Authors' tabulation using DHS and MICS.

Table 5.3: The reduced-form estimated effect of criminalizing FGC on spousal attributes and family wealth

| | Outcome variables | | | | | |
|---|---------------------------|-----------------------------|--------------------|--------------------------|-------------------------|--------------------------|
| | Spousal education | | Spousal occupation | | Family asset | |
| | Primary education or more | Secondary education or more | Working | Skilled work | Low wealth | High wealth |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Panel A: Ever-married women aged 15 to 49 at the interview | | | | | | |
| Age 1 to 22 at reform * Top tercile of pre-reform FGC rate | -0.047 (0.015) | -0.027 (0.012) | 0.004 (0.005) | -0.064 (0.012) | 0.063 (0.013) | -0.034 (0.010) |
| Age 1 to 22 at reform * Middle tercile of pre-reform FGC rate | -0.022 (0.007) | -0.006 (0.007) | 0.001 (0.003) | -0.066 (0.008) | 0.046 (0.007) | -0.031 (0.007) |
| Adjusted R-squared | 0.38 | 0.312 | 0.054 | 0.11 | 0.308 | 0.408 |
| Sample size | 184,607 | 184,607 | 182,030 | 182,030 | 307,475 | 307,475 |
| Mean among those aged 23 to 36 at reform | 0.444 | 0.234 | 0.972 | 0.159 | 0.348 | 0.403 |
| Panel B: Ever-married women aged 30 to 49 at the interview | | | | | | |
| Age 1 to 22 at reform * Top tercile of pre-reform FGC rate | -0.040 (0.019) | -0.021 (0.014) | -0.002 (0.008) | -0.042 (0.014) | 0.063 (0.018) | -0.027 (0.016) |
| Age 1 to 22 at reform * Middle tercile of pre-reform FGC rate | 0.000 (0.009) | 0.001 (0.010) | 0.001 (0.006) | -0.056 (0.012) | 0.047 (0.010) | -0.028 (0.011) |
| Adjusted R-squared | 0.382 | 0.304 | 0.066 | 0.094 | 0.318 | 0.423 |
| Sample size | 096,381 | 096,381 | 095,228 | 095,228 | 160,644 | 160,644 |
| Mean among those aged 23 to 36 at reform | 0.444 | 0.234 | 0.972 | 0.159 | 0.348 | 0.403 |

Notes: Panels A and B present the reduced-form estimated results for the effect of criminalization on spousal attributes and household wealth for ever-married women, separately for those aged 15 to 49, and 30 to 49 at interview. The analysis in panel B limits the sample to those aged 30 and above at interview, about 96 percent of them ever-married, to minimize the sample selection bias related to focusing only on ever-married sample. Regressions in each panel for all columns include the interaction of "young" cohort (aged 1 to 22 at reform) with "top" and "middle" terciles of pre-reform FGC prevalence rate, in addition to the covariates. The covariates include fixed effects for the year of interview and country by age at reform, pre-reform FGC prevalence rate, full set of dummies for rural, religion, and region.. The sample includes women who are partially-affected and least-affected by the reform while excluding the fully-affected cohorts (with zero or negative age at reform and very small sample size). Standard errors at country-cohort levels are in parentheses.

Source: Authors' computation using DHS and MICS.

Table 5.4: Falsification test for the reduced-form estimated effect of criminalizing FGC on own education, marriage, and health

| | Outcome variables | | | | |
|--|-----------------------------|-------------------|------------------------|-----------------------|---------------------------|
| | Education | | Marriage and fertility | | |
| | Secondary education or more | Literate | Ever-married | Married before age 18 | Having at least one child |
| | (1) | (2) | (3) | (4) | (5) |
| Panel A: Education and marriage outcomes | | | | | |
| Age 23 to 29 at reform * Top tercile of pre-reform FGC rate | -0.004 (0.014) | -0.040 (0.027) | -0.001 (0.007) | 0.037 (0.023) | 0.015 (0.008) |
| Age 23 to 29 at reform * Middle tercile of pre-reform FGC rate | -0.003 (0.013) | -0.005 (0.021) | -0.002 (0.004) | 0.028 (0.015) | 0.010 (0.006) |
| Adjusted R-squared | 0.313 | 0.328 | 0.118 | 0.163 | 0.071 |
| Sample size | 128,431 | 075,658 | 128,431 | 128,431 | 128,431 |
| Mean among those aged 30 to 36 at reform | 0.196 | 0.352 | 0.972 | 0.392 | 0.956 |
| | | | | | |
| | Reproductive health | | General health | | |
| | Genital discharge or ulcer | Child death | Underweight | Over-weight | Anemia |
| Panel B: Health outcomes | | | | | |
| Age 23 to 29 at reform * Top tercile of pre-reform FGC rate | -0.020 (0.014) | 0.010 (0.017) | -0.007 (0.017) | -0.013 (0.026) | 0.026 (0.032) |
| Age 23 to 29 at reform * Middle tercile of pre-reform FGC rate | -0.005 (0.013) | 0.000 (0.011) | -0.007 (0.010) | -0.018 (0.025) | 0.023 (0.022) |
| Adjusted R-squared | 0.088 | 0.163 | 0.071 | 0.408 | 0.07 |
| Sample size | 72,617 | 120,751 | 53,178 | 53,178 | 33,998 |
| Mean among those aged 30 to 36 at reform | 0.104 | 0.470 | 0.095 | 0.393 | 0.421 |

Notes: Panels A and B present the reduced-form estimated results for the effect of criminalization on education and marriage, and health outcomes. The child health analysis in panel B limits the sample to those females who gave birth to at least one child. Regressions in each panel for all columns include the interaction of "young" cohort (aged 1 to 22 at reform) with "top" and "middle" terciles of pre-reform FGC prevalence rate, in addition to the covariates. The controls include fixed effects for the year of interview and country by age at reform, pre-reform FGC prevalence rate, and full set of dummies for rural, religion, and region. The sample includes women who are partially-affected and least-affected by the reform while excluding the fully-affected cohorts (with zero or negative age at reform and very small sample size). Standard errors at country-cohort levels are in parentheses. Source: Authors' tabulation using DHS and MICS.

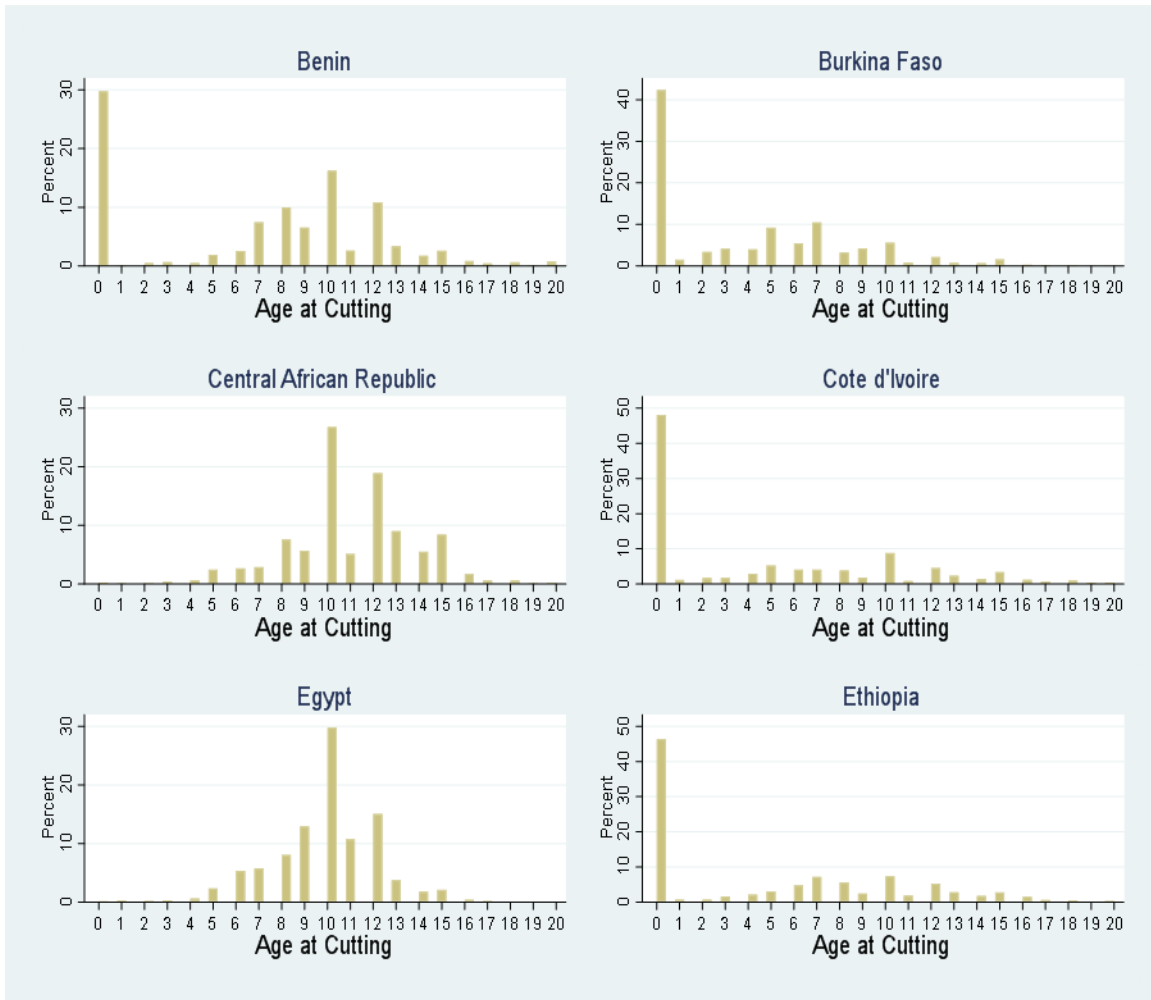
Table 5.5: Falsification test for the reduced-form estimated effect of criminalizing FGC on spousal attributes and family wealth

| | Outcome variables | | | | | |
|---|---------------------------|-----------------------------|--------------------|-------------------|------------------|-------------------|
| | Spousal education | | Spousal occupation | | Family asset | |
| | Primary education or more | Secondary education or more | Working | Skilled work | Low wealth | High wealth |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Panel A: Ever-married women aged 15 to 49 at the interview | | | | | | |
| Age 23 to 29 at reform * Top tercile of pre-reform FGC rate | -0.024 (0.020) | -0.012 (0.016) | 0.005 (0.013) | -0.008 (0.018) | 0.025 (0.016) | -0.007 (0.017) |
| Age 23 to 29 at reform * Middle tercile of pre-reform FGC rate | 0.008 (0.015) | 0.012 (0.015) | 0.003 (0.011) | -0.013 (0.016) | 0.005 (0.010) | 0.002 (0.011) |
| Adjusted R-squared | 0.388 | 0.34 | 0.07 | 0.101 | 0.299 | 0.408 |
| Sample size | 074,735 | 074,735 | 073,714 | 073,714 | 123,464 | 123,464 |
| Mean among those aged 30 to 36 at reform | 0.430 | 0.239 | 0.969 | 0.146 | 0.350 | 0.394 |
| Panel B: Ever-married women aged 30 to 49 at the interview | | | | | | |
| Age 23 to 29 at reform * Top tercile of pre-reform FGC rate | -0.023 (0.020) | -0.009 (0.017) | 0.005 (0.014) | -0.002 (0.019) | 0.025 (0.017) | -0.001 (0.017) |
| Age 23 to 29 at reform * Middle tercile of pre-reform FGC rate | 0.014 (0.015) | 0.015 (0.016) | 0.003 (0.012) | -0.011 (0.018) | 0.010 (0.010) | 0.001 (0.010) |
| Adjusted R-squared | 0.395 | 0.341 | 0.072 | 0.098 | 0.305 | 0.415 |
| Sample size | 066,579 | 066,579 | 065,659 | 065,659 | 110,541 | 110,541 |
| Mean among those aged 30 to 36 at reform | 0.430 | 0.239 | 0.969 | 0.146 | 0.350 | 0.394 |

Notes: Panels A and B present the reduced-form estimated results for the effect of criminalization on spousal attributes and household wealth for ever-married women, separately for those aged 15 to 49, and 30 to 49 at interview. The analysis in panel B limits the sample to those aged 30 and above at interview, about 96 percent of them ever-married, to minimize the sample selection bias related to focusing only on ever-married sample. Regressions in each panel for all columns include the interaction of "young" cohort (aged 1 to 22 at reform) with "top" and "middle" terciles of pre-reform FGC prevalence rate, in addition to the full-set of controls. The controls include fixed effects for the year of interview and country by age at reform, pre-reform FGC prevalence rate, and full set of dummies for rural, religion, and region. The sample includes women who are partially-affected and least-affected by the reform while excluding the fully-affected cohorts (with zero or negative age at reform and very small sample size). Standard errors at country-cohort levels are in parentheses.

Source: Authors' tabulation using DHS and MICS.

APPENDIX A



(continues to the next page)

Figure A3.1: Pre-reform distribution of age at female genital cutting for criminalizing countries.

Notes: The histogram shows the percent distributions for the female age at genital cutting for criminalizing countries, except the countries without age at cutting information. This figure uses frequency weights approximated by rounding the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018).

Source: Authors' graphical analysis using DHS and MICS.

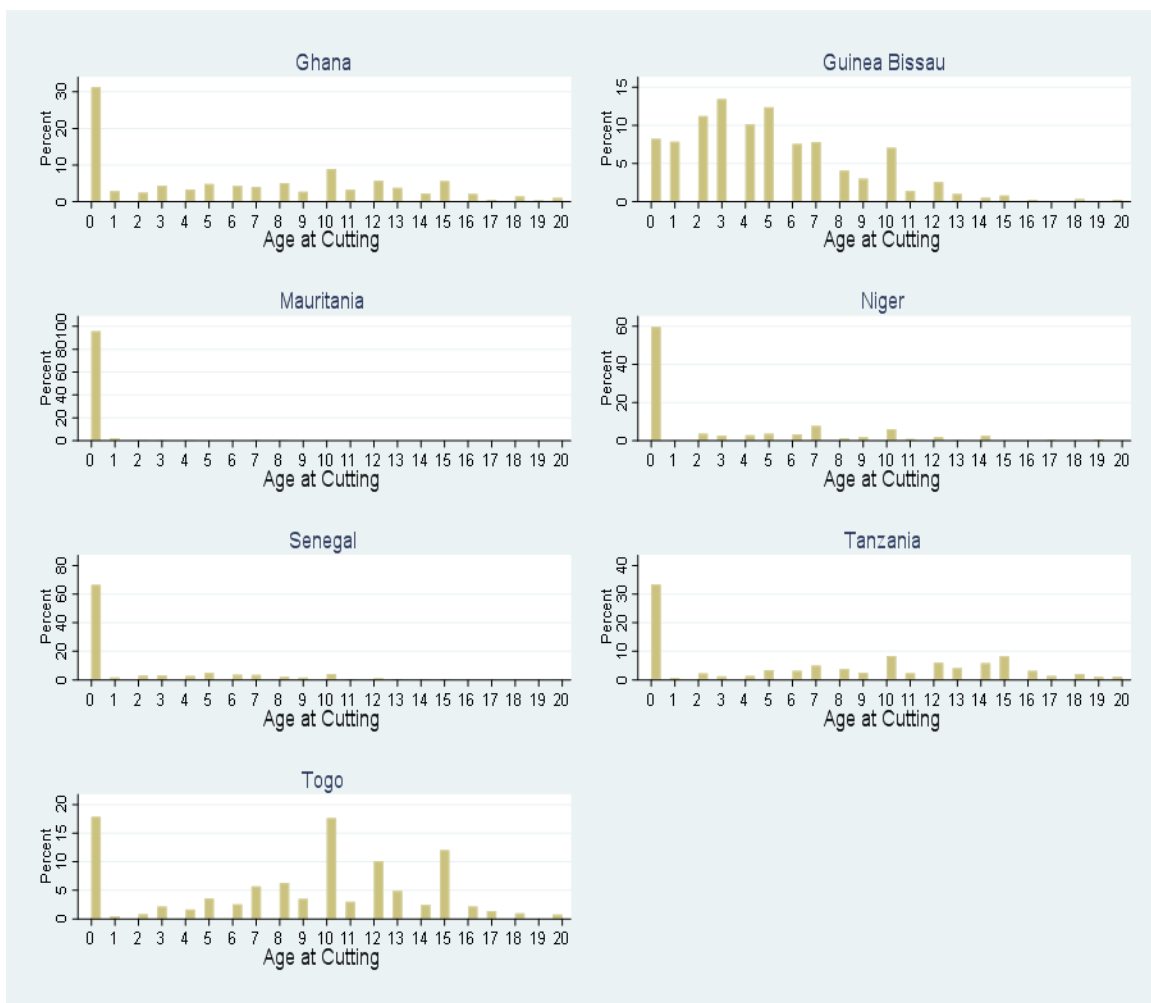


Figure A3.1: Pre-reform distribution of age at female genital cutting for criminalizing countries (continued).

Notes: The histogram shows the percent distributions for the female age at genital cutting for criminalizing countries, except the countries without age at cutting information. This figure uses frequency weights approximated by rounding the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018).

Source: Authors' graphical analysis using DHS and MICS.

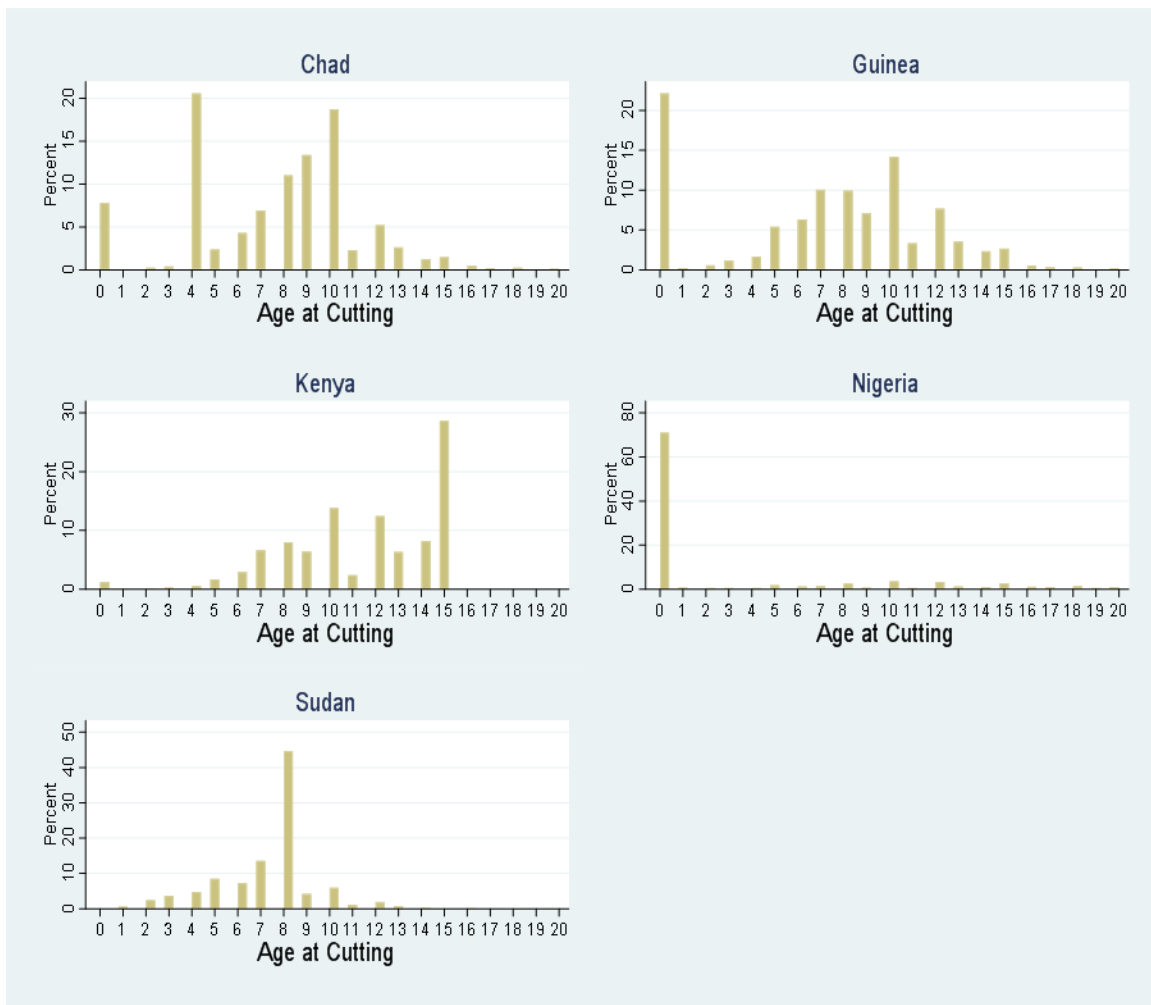
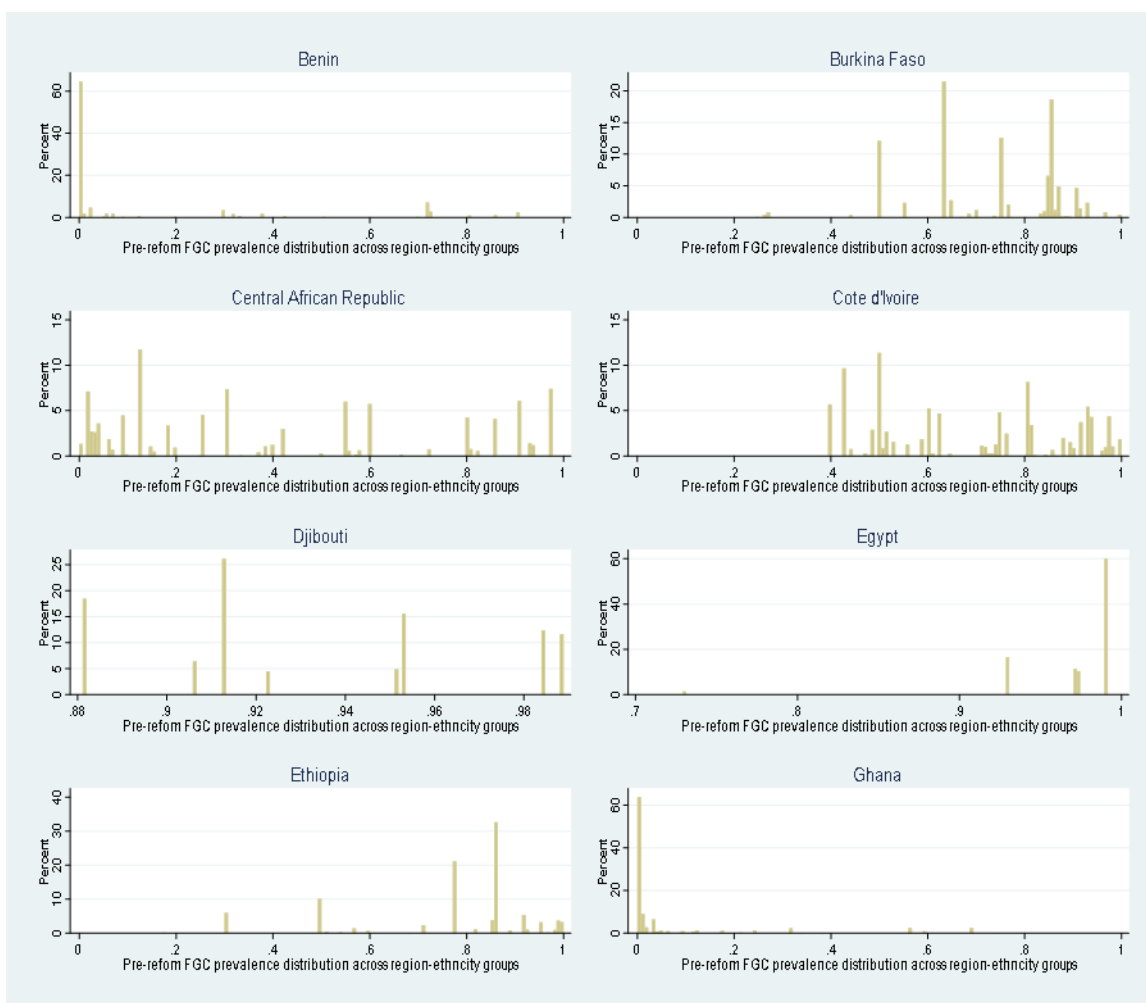


Figure A3.2: Pre-reform distribution of age at female genital cutting for non-criminalizing countries.

Notes: The histogram shows the percent distributions for the female age at genital cutting for non-criminalizing countries, except the countries without age at cutting information. This figure uses frequency weights approximated by rounding the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018).

Source: Authors' graphical analysis using DHS and MICS.



(continues to the next page)

Figure A3.3: Pre-reform FGC prevalence rate distribution across region-ethnicity groups for criminalizing countries.

Notes: The histogram shows the percent distributions for the pre-reform FGC prevalence rate distributions across regions by ethnicities for criminalizing countries. This figure uses frequency weights approximated by rounding the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018).

Source: Authors' graphical analysis using DHS and MICS.

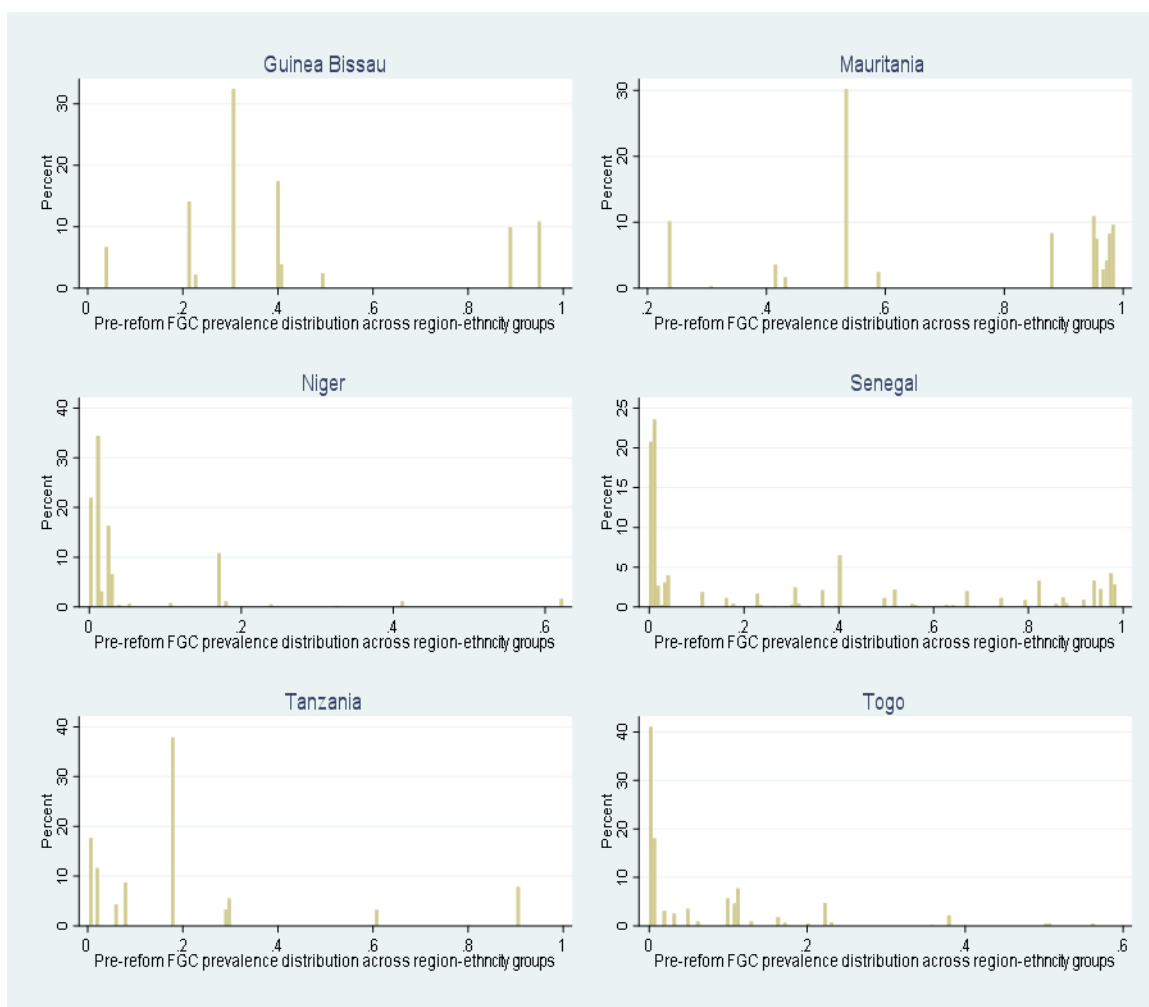


Figure A3.3: Pre-reform FGC prevalence rate distribution across region-ethnicity groups for criminalizing countries (continued).

Notes: The histogram shows the percent distributions for the pre-reform FGC prevalence rate distributions across regions by ethnicities for criminalizing countries. This figure uses frequency weights approximated by rounding the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018).

Source: Authors' graphical analysis using DHS and MICS.

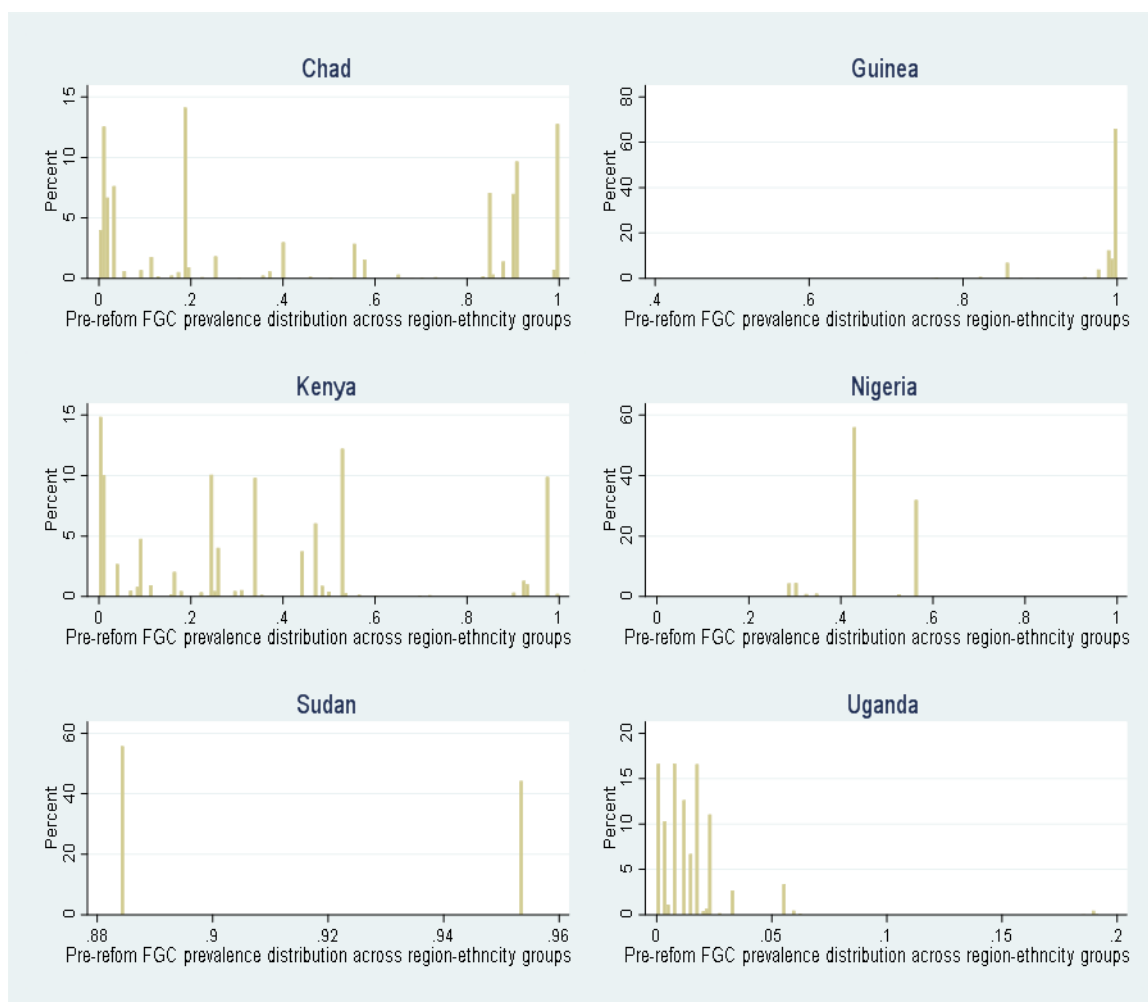


Figure A3.4: Pre-reform FGC prevalence rate distribution across region-ethnicity groups for non-criminalizing countries.

Notes: The histogram shows the percent distributions for the pre-reform FGC prevalence rate distributions across regions by ethnicities for non-criminalizing countries. This figure uses frequency weights approximated by rounding the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018).

Source: Authors' graphical analysis using DHS and MICS.

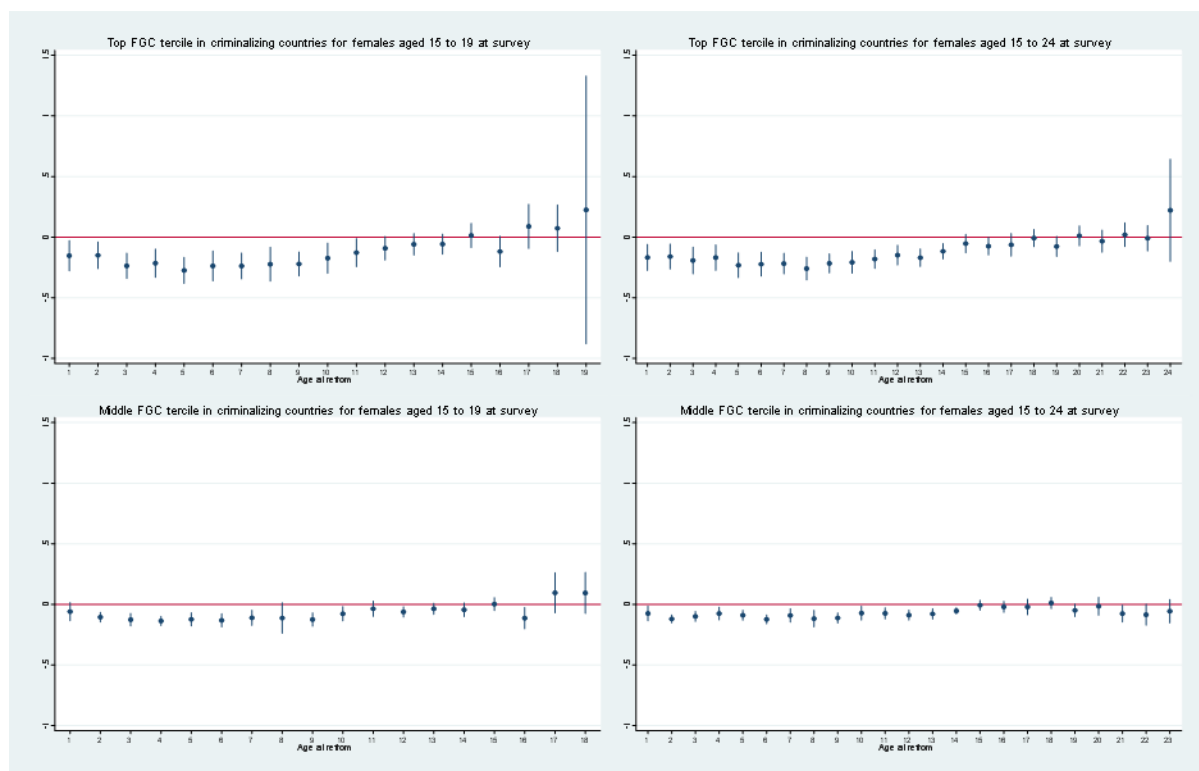


Figure A4.1: Event-study estimated effect of criminalization on ever-experiencing FGC for young age females at the survey.

Notes: The top and bottom panels of the figure present the coefficients of interactions of age at reform dummies with dummies for the top and middle terciles of the pre-reform FGC prevalence rate in the region-ethnicity groups. The left and right panels report the results for those females aged 15 to 19 and aged 15 to 24 at interview, respectively. In this analysis, we omit females in the bottom tercile as reference groups. The sample includes those partially- and least-affected cohorts in criminalizing countries, excluding the fully-affected cohorts. The specification includes basic controls, demographics, and region fixed effect. "Basic controls" include fixed effects for the year of interview and country by age at reform. "Demographics" include the full set of dummies for rural and religion. We cluster standard errors at the country-cohort level. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' tabulation using DHS and MICS.

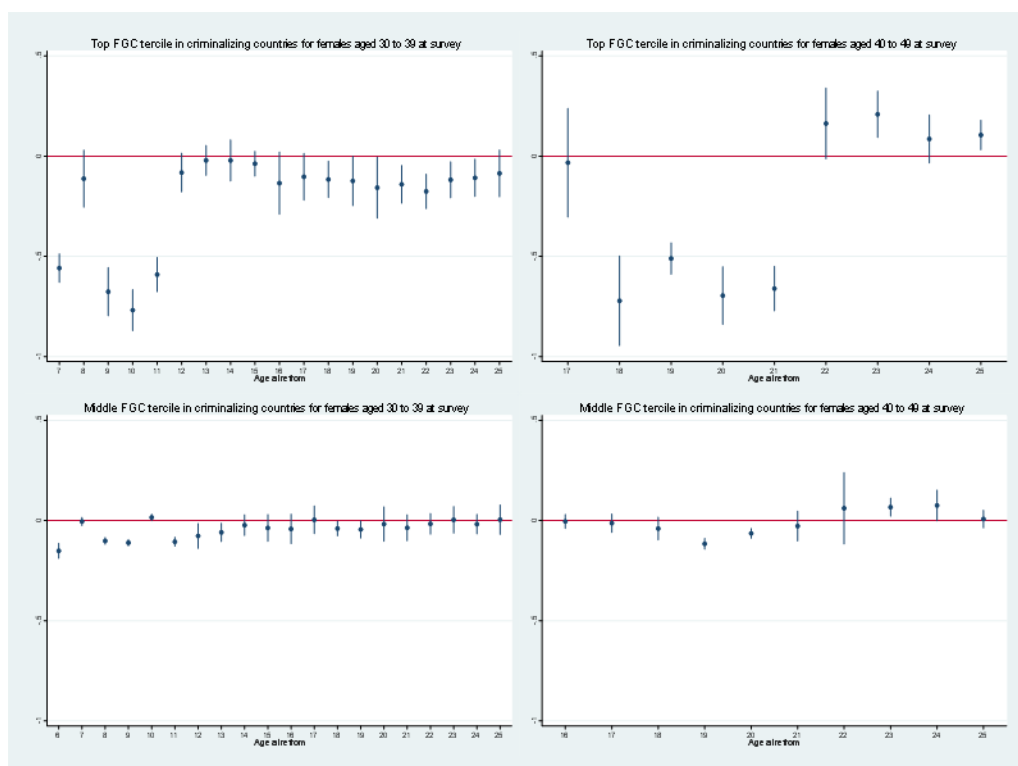


Figure A4.2: Event-study estimated effect of criminalization on ever-experiencing FGC for old age females at the survey.

Notes: The top and bottom panels of the figure present the coefficients of interactions of age at reform dummies with dummies for the top and middle terciles of the pre-reform FGC prevalence rate in the region-ethnicity groups. The left and right panels report the results for those females aged 30 to 39 and aged 40 to 49 at interview, respectively. In this analysis, the control group includes those aged 23 to 36 at the reform in the top and middle terciles, and those aged 1 to 36 in the bottom tercile. The sample includes those partially- and least-affected cohorts in criminalizing countries, excluding the fully-affected cohorts. The specification includes basic controls, demographics, and region fixed effect. "Basic controls" include fixed effects for the year of interview and country by age at reform. "Demographics" include the full set of dummies for rural and religion. We cluster standard errors at the country-cohort level. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' tabulation using DHS and MICS.

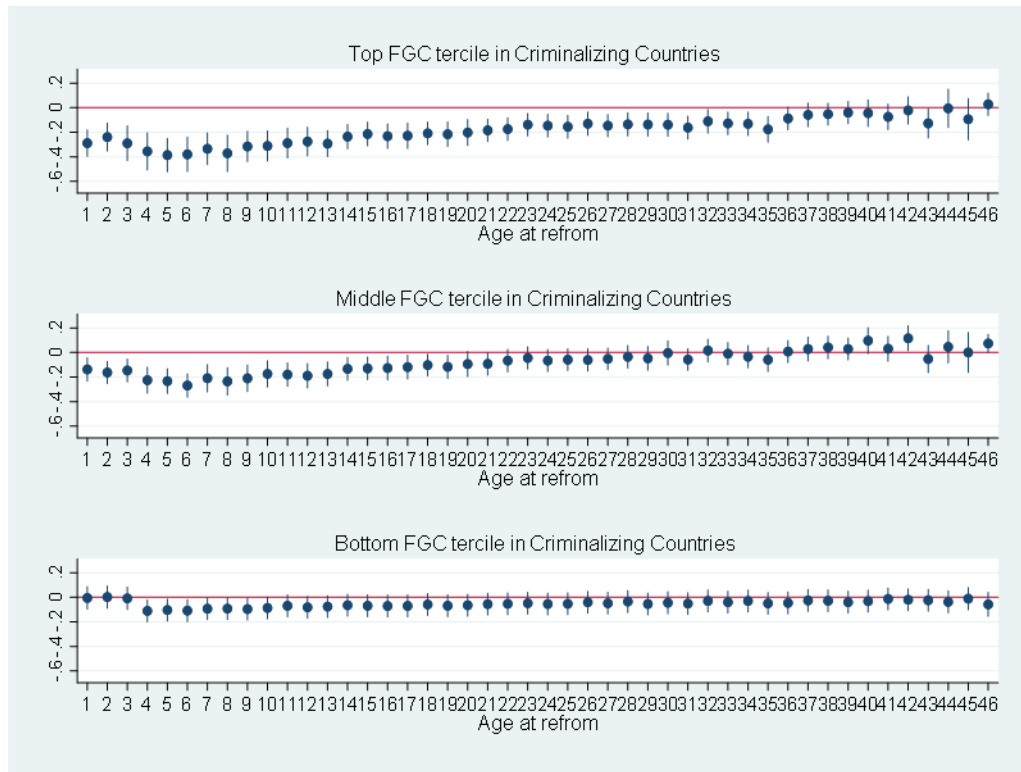


Figure A4.3: Event-study estimated effect of criminalization on ever-experiencing FGC including very old cohorts at reform.

Notes: The figure presents the coefficients of interactions of age at reform dummies with dummies for the top, middle, and bottom terciles of the pre-reform FGC prevalence rate in the region-ethnicity groups. In this regression, we omit those females aged 47 at reform as a reference group. The sample includes only the partially- and least-affected cohorts in criminalizing countries, excluding the fully-affected cohorts. The specification includes basic controls, demographics, and region fixed effect. "Basic controls" include fixed effects for the year of interview and country by age at reform. "Demographics" include the full set of dummies for rural and religion. We cluster standard errors at the country-cohort level. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

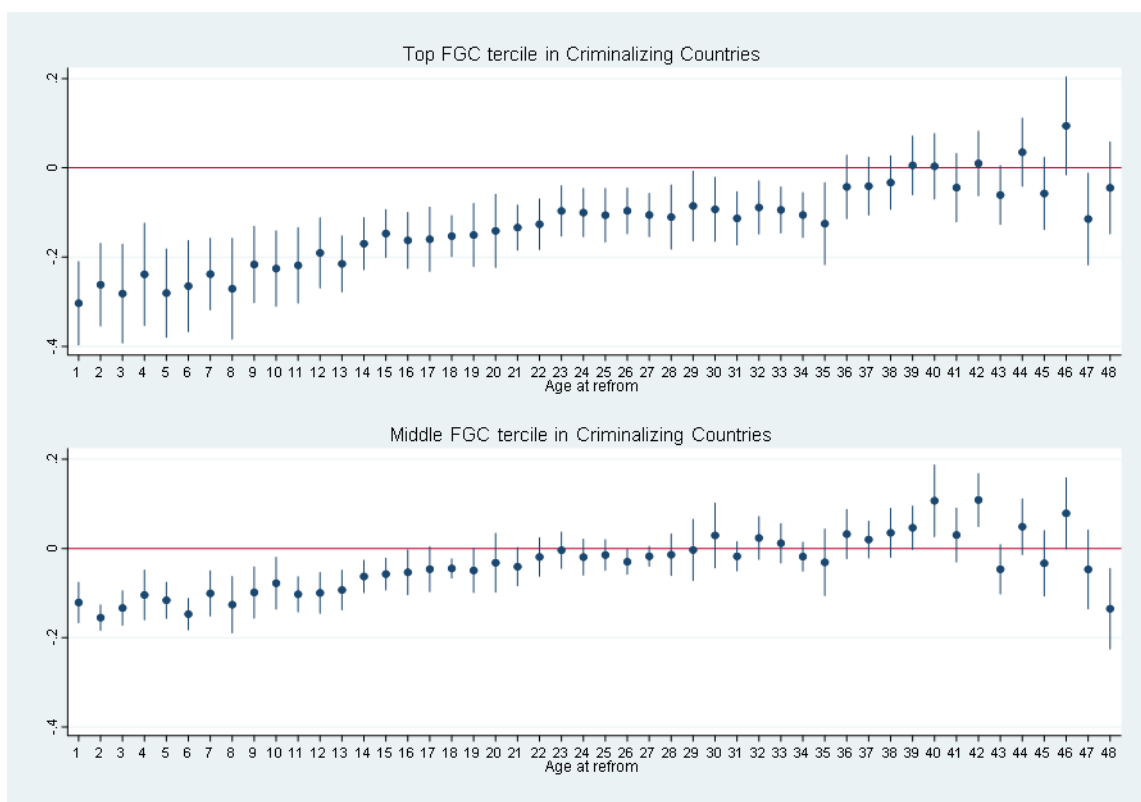


Figure A4.4: Event-study estimated effect of criminalization on ever-experiencing FGC using all the available cohorts at reform.

Notes: The top and bottom panels of the figure present the coefficients of interactions of age at reform dummies with dummies for the top and middle terciles of the pre-reform FGC prevalence rate in the region-ethnicity groups. In this analysis, we omit those females in the bottom tercile as the control group. The sample includes those females aged 15 to 49 at the survey and partially- and least-affected cohorts in criminalizing countries, excluding the fully-affected cohorts. The specification includes basic controls, demographics, and the region fixed effect. "Basic controls" include fixed effects for the year of interview and country by age at reform. "Demographics" include the full set of dummies for rural and religion. We cluster standard errors at the country-cohort level. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

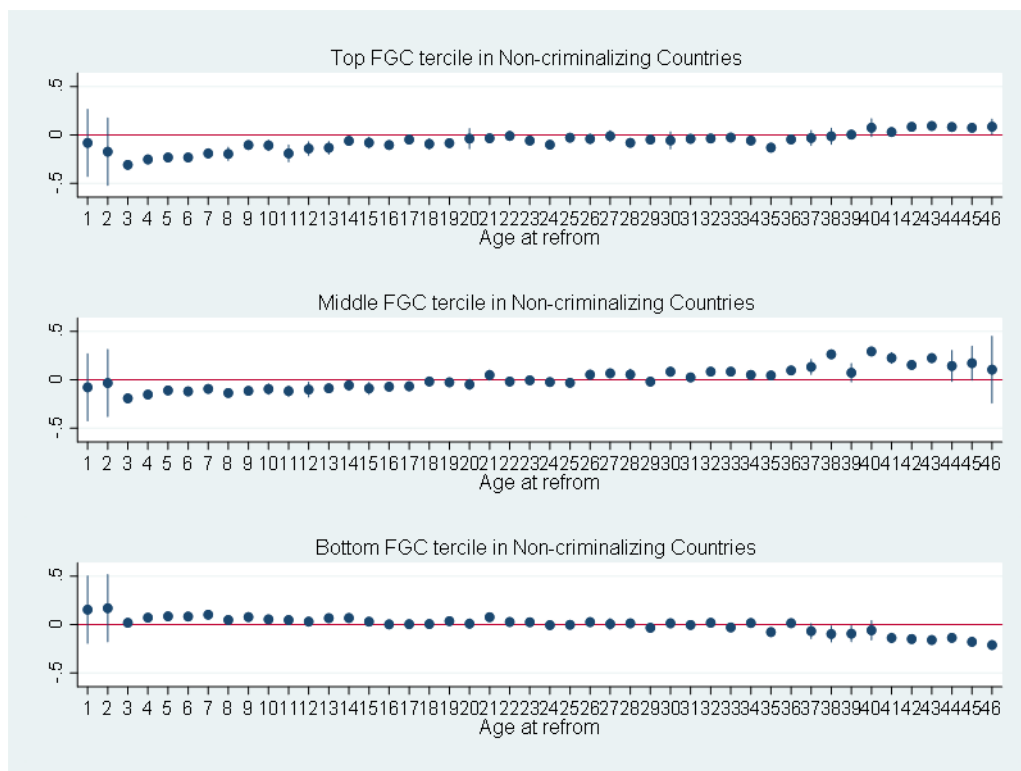


Figure A4.5: Event-study estimated effect of non-criminal laws on ever-experiencing FGC using all the available cohorts at reform.

Notes: The figure presents the coefficients of interactions of age at reform dummies with dummies for the top, middle, and bottom tertiles of the pre-reform FGC prevalence rate in the region-ethnicity groups. In this regression, we omit those females aged 47 at reform as a reference group. The sample includes only the partially- and least-affected cohorts in non-criminalizing countries, excluding the fully-affected cohorts. The specification includes basic controls, demographics, and region fixed effect. "Basic controls" include fixed effects for the year of interview and country by age at reform. "Demographics" include the full set of dummies for rural and religion. We cluster standard errors at the country-cohort level. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

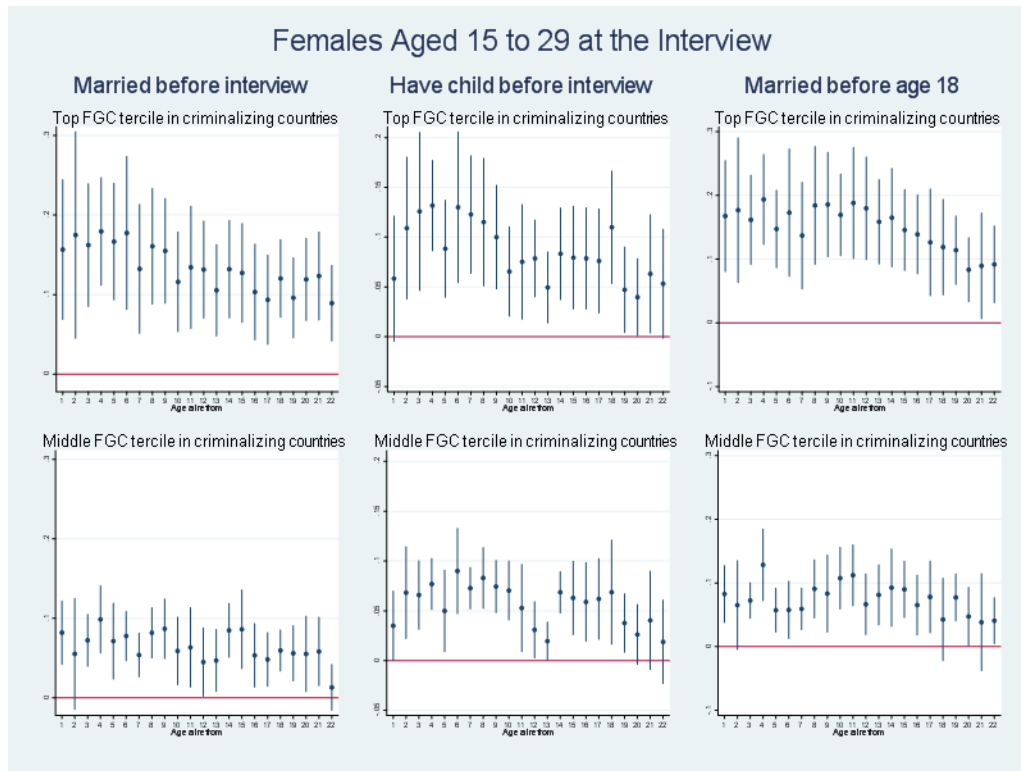


Figure A5.1: Reduced-form estimated effect of FGC criminalization on marriage and fertility among females aged 15 to 29 at the interview.

Notes: The left, middle, and right panels of the figure present the reduced-form event-study estimated effects of FGC criminalization on marriage and fertility before interview, and marriage before age 18. The top and bottom panels of the figure report the results for the effect of criminalization among females in the top and middle terciles of the pre-reform FGC prevalence rate region-ethnicity groups. In addition to the interactions of pre-reform FGC prevalence terciles with age at reform dummies, the regression includes fixed effects for the year of interview, country by age at reform, rural, region, and religion. We omit those females in the bottom tercile as control group. The analysis in this figure uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

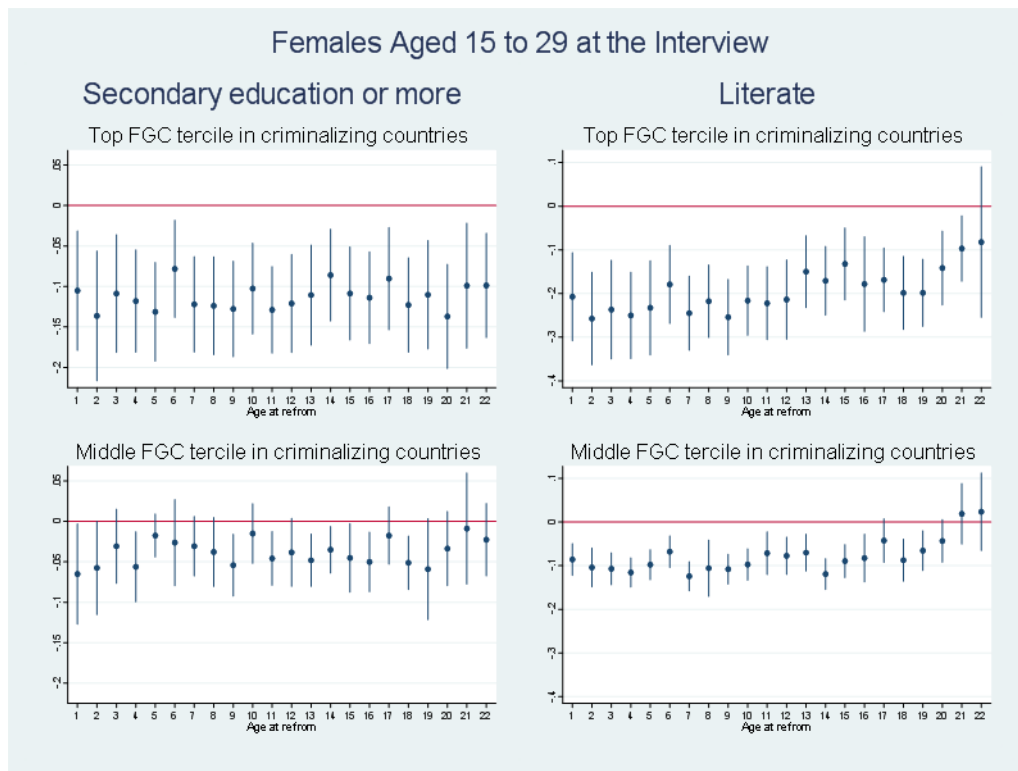


Figure A5.2: Reduced-form estimated effect of FGC criminalization on own education among females aged 15 to 29 at the interview.

Notes: The left and right panels of the figure present the reduced-form event-study estimated effects of FGC criminalization on secondary education and literacy. The top and bottom panels of the figure report the results for the effect of criminalization among females in the top and middle terciles of the pre-reform FGC prevalence rate region-ethnicity groups. In addition to the interactions of pre-reform FGC prevalence terciles with age at reform dummies, the regression includes fixed effects for the year of interview, country by age at reform, rural, region, and religion. We omit those females in the bottom tercile as control group. The analysis in this figure uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

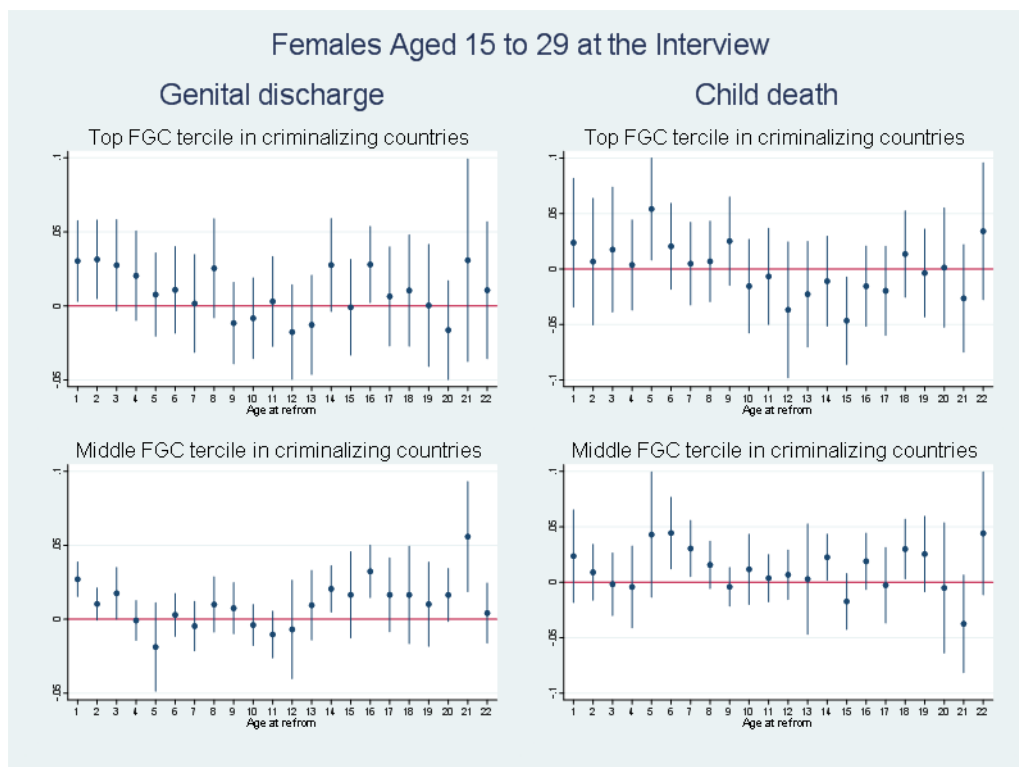


Figure A5.3: Reduced-form estimated effect of FGC criminalization on genital discharge and child health among females aged 15 to 29 at the interview.

Notes: The left and right panels of the figure present the reduced-form event-study estimated effects of FGC criminalization on genital discharge and child health. The top and bottom panels of the figure report the results for the effect of criminalization among females in the top and middle terciles of the pre-reform FGC prevalence rate region-ethnicity groups. In addition to the interactions of pre-reform FGC prevalence terciles with age at reform dummies, the regression includes fixed effects for the year of interview, country by age at reform, rural, region, and religion. We omit those females in the bottom tercile as control group. The analysis in this figure uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

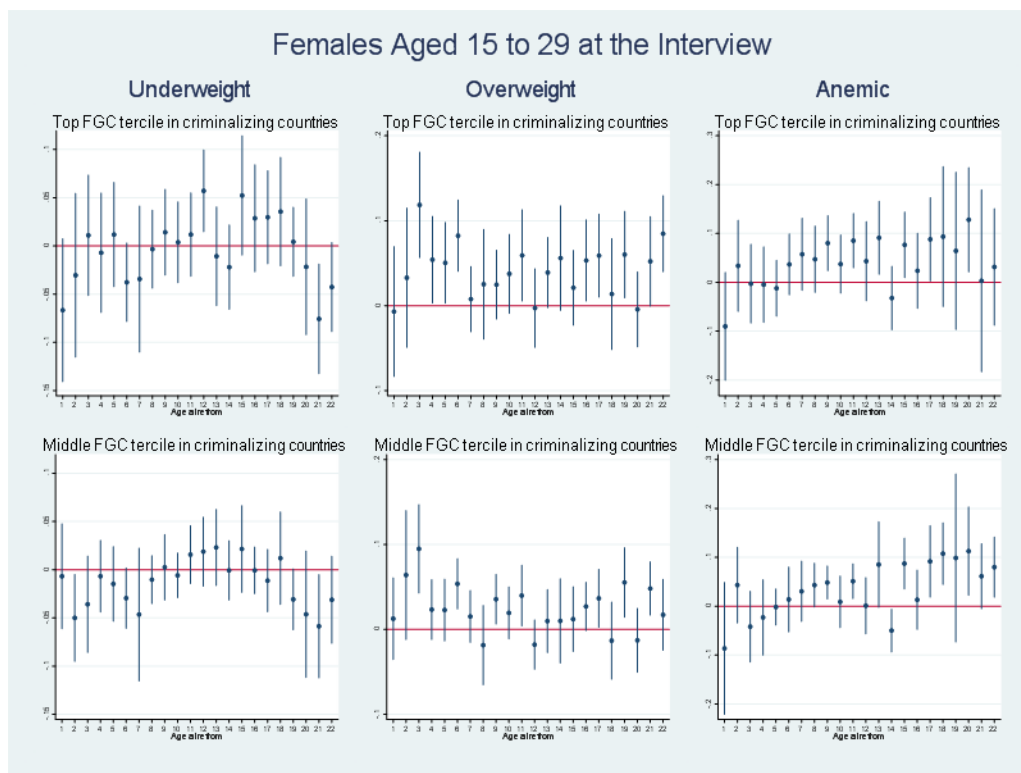


Figure A5.4: Reduced-form estimated effect of FGC criminalization on own general health among females aged 15 to 29 at the interview.

Notes: The left, middle, and right panels of the figure present the reduced-form event-study estimated effects of FGC criminalization on being underweight, overweight, and anemic. The top and bottom panels of the figure report the results for the effect of criminalization among females in the top and middle terciles of the pre-reform FGC prevalence rate region-ethnicity groups. In addition to the interactions of pre-reform FGC prevalence terciles with age at reform dummies, the regression includes fixed effects for the year of interview, country by age at reform, rural, region, and religion. We omit those females in the bottom tercile as control group. The analysis in this figure uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval. The markers are for the coefficients, and horizontal spikes are for a 95-percent confidence interval.

Source: Authors' graphical analysis using DHS and MICS.

APPENDIX B

Table B3.1: African countries with an FGC ban as of 2013, pre-reform surveys and FGC prevalence rate at country level, and construction of pre-reform FGC prevalence rate by region-ethnicity groups.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|--|----------------------|---|------|-----------------------------------|--|---------------------|---------------------|
| | | Pre-reform FGC prevalence rate | | Pre-reform DHS/MICS surveys | Construction of pre-reform prevalence rate by region- ethnicity (policy intensity) | | |
| Country (state) | Year of reform | Mean | SD | | Data source | Minimum age used | Maximum age used |
| <i>Implementing criminal laws:</i> | | | | | | | |
| Benin | 2003 | 0.15 | 0.36 | 2001 DHS | 2001 DHS | 15 | 37 |
| Burkina Faso | 1996 | 0.73 | 0.45 | NA | 1998- 99 DHS | 18 | 40 |
| Central African Republic | 1996 | 0.69 | 0.46 | 1994-95 DHS | 1994- 95 DHS | 15 | 37 |
| Cote d'Ivoire | 1998 | 0.98 | 0.16 | <u>1994 DHS^a</u> | 2005 DHS | 23 | 45 |
| Djibouti | 1995 | 0.77 | 0.42 | NA | 2006 MICS | 27 | 49 |
| Egypt | 2008 | 0.06 | 0.24 | 1995 DHS, 2000 DHS | 2000 DHS | 15 | 37 |
| Ethiopia | 2004 | 0.05 | 0.22 | 2000 DHS | 2000 DHS | 15 | 37 |
| Ghana | 1994 | 0.29 | 0.46 | NA | 2003 DHS | 25 | 47 |
| Guinea Bissau | 2011 | 0.20 | 0.40 | 2006 MICS | 2006 MICS | 15 | 37 |
| Mauritania | 2005 | 0.06 | 0.23 | NA | 2007 MICS | 18 | 40 |
| Niger | 2003 | 0.41 | 0.49 | 1998 DHS | 1998 DHS | 15 | 37 |
| Senegal | 1999 | 0.93 | 0.25 | NA | 2005 DHS | 22 | 44 |
| Tanzania | 1998 | 0.43 | 0.50 | 1996 DHS | 1996 DHS | 15 | 37 |
| Togo | 1998 | 0.71 | 0.45 | NA | 2006 MICS | 24 | 46 |

(continues to the next page)

Table B3.1: Continued.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|----------------------|---|------|-----------------------------------|---|---------------------|---------------------|
| | | Pre-reform FGC prevalence rate | | | Construction of pre-reform prevalence rate by region-ethnicity (policy intensity) | | |
| Country (state) | Year of reform | Mean | SD | Pre-reform DHS/MICS surveys | Data source | Minimum age used | Maximum age used |
| <i>Implementing non-criminal laws:</i> | | | | | | | |
| Chad | 2003 | 0.45 | 0.50 | NA | 2004 DHS | 17 | 39 |
| Guinea | 2000 | 0.98 | 0.13 | 1999 DHS | 1999 DHS | 15 | 37 |
| Kenya | 2001 | 0.33 | 0.47 | 1998 DHS | 1998 DHS | 15 | 37 |
| Nigeria | 1999- 2009 | 0.44 | 0.50 | NA | 2007 MICS | 24 | 46 |
| Bayelsa | 2004 | 0.41 | 0.49 | | 2007 MICS | 19 | 41 |
| Cross River | 2000 | 0.42 | 0.49 | | 2007 MICS | 23 | 45 |
| Ebonyi | 2001 | 0.61 | 0.49 | | 2007 MICS | 22 | 44 |
| Edo | 1999 | 0.46 | 0.50 | | 2007 MICS | 24 | 46 |
| Enugu | 2004 | 0.47 | 0.50 | | 2007 MICS | 19 | 41 |
| Rivers | 2009 | 0.29 | 0.46 | | 2007 MICS | 15 | 37 |
| Sudan | 2008- 2009 | 0.92 | 0.28 | 2000 MICS | 2000 MICS | 15 | 37 |
| South Gordofan | 2008 | 0.95 | 0.22 | | 2000 MICS | 15 | 37 |
| Gadarif | 2009 | 0.89 | 0.31 | | 2000 MICS | 15 | 37 |
| Uganda | 2010 | 0.01 | 0.12 | NA | 2011 DHS | 17 | 39 |

Notes: "Least-affected cohorts" are females whose age at the reform is 23 or more, while "partially-affected cohorts" include those below age 23 at the reform. "Fully-affected cohorts" cover those born after the reform. This table uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). Column (1) = the year of reform. Column (2) and (3) = the mean and standard deviations for the pre-reform FGC prevalence rate at country level. Column (4) = the available pre-reform DHS/MICS surveys. Column (5) = data source used to construct pre-reform FGC prevalence at the region-by-ethnicity level. For those countries with pre-reform surveys, we consider the recent pre-reform survey to construct the FGC prevalence, except for Cote d'Ivoire. (a) To construct the policy intensity for Cote d'Ivoire, this study uses the 2005 post-reform survey, instead of a pre-reform survey in 1994, which has inconsistent regional coding with later surveys. Column (6)=15 for those countries with a pre-reform survey, and column (6)=15+[column(5)-(column(1)-1)] for those without a pre-reform survey. Age 15 is the youngest age cohort available in every survey for DHS/MICS. In column (6), we subtract 1 from column(1) to consider the year just before the reform and subtract the result from column(5). Then, we add the difference to 15 to construct comparable cohorts to those who are more likely to be affected by the reform later post-reform (aged 1 to 22 at reform). Column (7) = Column (6)+22.

Source: Authors' computation using DHS and MICS.

Table B4.1: African countries with non-criminal FGC bans as of 2013 and had DHS/MICS respondents affected fully or partially by the reform.

| | | | | Birth year of DHS/MICS respondents | | | | | | | | |
|--------------------|----------------------|---------------------------------|-------------------------------|------------------------------------|-----|------|-----------------------------------|-----|---------------------------|------|---|------|
| Country (state) | Year of reform | Youngest available cohort | Oldest available cohort | Least-affected cohorts | | | Partially- affected cohorts | | Fully-affected cohorts | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | | | |
| Chad | 2003 | 2000 | 1951 | 1967 | - | 1980 | 1981 | - | 2000 | na | - | na |
| Guinea | 2000 | 2003 | 1955 | 1964 | - | 1977 | 1978 | - | 1999 | 2000 | - | 2003 |
| Kenya | 2001 | 1999 | 1953 | 1965 | - | 1978 | 1979 | - | 1999 | na | - | na |
| Nigeria | 1999- 2009 | 2003 | 1957 | | | | | | | | | |
| Bayelsa | 2004 | 2003 | 1957 | 1968 | - | 1981 | 1982 | - | 2003 | na | - | na |
| Cross River | 2000 | 2003 | 1957 | 1964 | - | 1977 | 1978 | - | 1999 | 2000 | - | 2003 |
| Ebonyi | 2001 | 2003 | 1958 | 1965 | - | 1978 | 1979 | - | 2000 | 2001 | - | 2003 |
| Edo | 1999 | 2003 | 1958 | 1963 | - | 1976 | 1977 | - | 1998 | 1999 | - | 2003 |
| Enugu | 2004 | 2003 | 1957 | 1968 | - | 1981 | 1982 | - | 2003 | na | - | na |
| Rivers | 2009 | 2003 | 1957 | 1973 | - | 1986 | 1987 | - | 2003 | na | - | na |
| Sudan | 2008- 2009 | | | | | | | | | | | |
| South Gordofan | 2008 | 1999 | 1965 | 1972 | - | 1985 | 1986 | - | 1999 | na | - | na |
| Gadarif | 2009 | 1999 | 1965 | 1973 | - | 1986 | 1987 | - | 1999 | na | - | na |
| Uganda | 2010 | 2001 | 1961 | 1974 | - | 1987 | 1988 | - | 2001 | na | - | na |

Notes: "Least-affected cohorts" are females whose age at the reform is 23 or more, while "Partially-affected cohorts" include those below age 23 at the reform. "Fully-affected cohorts" cover those born after the reform. This table uses the DHS and MICS sampling weights to adjust for differences in the probability of selection and interview due to sampling design, as suggested by Croft et al. (2018). Column (1) = the year of reform. Column (2) = the youngest available cohort in the surveys. Column (3)=the oldest available cohort in the surveys. Column (4) = Column (1)-36. Column (5) = Column (1)-23. Column (6)=Column (5)+1. Column (7) = Column (6)+21 if Column (2) > Column (6)+21, and Column (7) = Column (2) if Column (2) <= Column (6)+21.

Source: Authors' calculation using DHS and MICS.