Persistent Norms and Tipping Points: Female Genital Cutting in Burkina Faso*

Lindsey Novak†

March 23, 2017

Abstract

Female genital cutting (FGC) is prevalent across many parts of Africa, the Middle East, and Asia. This practice can have profound negative effects on women's physical and psychological well-being. Even with these negative effects, FGC is often perpetuated for generations. The prevailing theory in the study of FGC is that it is a social coordination norm—that is, households will abandon FGC if and only if a sufficient proportion of households within the community agree to abandon the practice. Under this theory, if a sufficient number of community members agree to abandon FGC, a tipping point is reached and the rate of FGC should fall to zero. Recent empirical evidence rejects that theory. I draw on Schelling's (1978) model of critical mass to contribute to this important debate and generate a new data-supported theory that has important implications for the types of policies that should be introduced. Using a dataset of more than 6,000 women born between 1949 and 1995 in Burkina Faso, I show that households within a community have heterogeneous preferences for FGC such that each household therein may require a different proportion of community members to abandon FGC before they decide to also reject the practice. I show that the presence of this heterogeneity makes the existence of a tipping point far from guaranteed and that stable interior equilibria in FGC rates are possible. My theoretical and empirical findings suggest that individuals and households are in fact able to deviate from an entrenched, gender-biased social norm and that policies aimed at reducing the prevalence of FGC should target changing individual and household, rather than village-level, preferences.

JEL Classification Codes: D19, I15, O10, Z1

Keywords: Social Norms, Development, Female Genital Cutting, Sexual and Reproductive Health

*I would like to thank Marc Bellemare, Paul Glewwe, Marie Baguet, Jeta Rudi, Joel Cuffey, and Adam Morris for extensive comments on earlier drafts of this paper. I am also grateful for useful feedback from Deepankar Basu, Sonja Vogt, Charles Efferson, and Sara Lowes, as well as seminar participants at the Western Economic Association International conference, Agricultural & Applied Economics Association meeting, North Eastern Universities Development Consortium conference, and the University of Minnesota Trade & Development seminar. I would like to thank Joshua Duchene for his excellent research assistance. I am grateful to the Department of Economics and the Doctoral Dissertation Fellowship from the University of Minnesota for the funding to conduct this research.

†PhD Candidate, Department of Applied Economics, University of Minnesota, 332 Ruttan Hall, 1994 Buford Ave, Saint Paul, MN 55108, novak352@umn.edu.

1 Introduction

Each year more than three million girls undergo female genital cutting (WHO, 2012). Female genital cutting (FGC)[1]—a practice wherein a woman's genitalia are partially or totally removed for nonmedical reasons—has no documented health benefits, and this procedure can have profound negative health effects on the women subjected to the practice. Women who have undergone FGC are more than twice as likely to experience birthing complications (Jones et al., 1999; Banks et al., 2006) and 25 percent more likely to contract sexually transmitted diseases (Wagner 2014). They are also more likely to experience anxiety, depression, and marital conflict (Dorkenoo, 1999). These complications can create barriers to women working effectively both inside and outside of the household, which can contribute to economic underdevelopment in communities that practice FGC. Yet the practice persists because of beliefs that FGC will discourage infidelity among women subjected to the practice (Shell-Duncan and Hernlund, 2001), as well as beliefs that women who have undergone FGC attain higher standards of beauty, cleanliness, and femininity (Toubia and Sharief, 2003).

Why does FGC persists as a norm when there are no medical benefits and there are many potential physical and psychological costs? I propose a new theoretical explanation for this phenomenon, and I test this explanation using observational data from Burkina Faso's Demographic and Health Survey. Additionally, I investigate whether there is a tipping point in the proportion of community members abandoning FGC.

Gender-biased harmful practices, including FGC, child marriage, and sex-selective abortion, have become a significant topic of public and political discourse. Many gender-biased harmful practices are a result of entrenched social norms which, once in place, are very dif-

[1]Female genital cutting (FGC) is also referred to as "female circumcision" and "female genital mutilation." I use the terminology FGC because it makes larger a distinction between FGC and male circumcision than does the term female circumcision, and it is more value-neutral than the term "mutilation." Throughout this manuscript I use the terminology FGC, "cut," and "cutting" interchangeably.

ficult to alter without external stimuli such as public policy interventions. Whether those interventions can change individual behavior or whether they must be targeted toward changing community-wide beliefs and behavior is the subject of an important debate (Powell, 2017). If FGC is a social coordination norm—that is, communities will abandon FGC if and only if a sufficient proportion of households within the community agree to abandon the practice—as posited by Mackie (1996), deviating from the social norm is so socially costly that individuals or households will not do so on their own. Many development organizations and governments working to reduce the prevalence of FGC design interventions with the belief that FGC is a social coordination norm. Agencies and governments operating under this assumption advocate for public declarations by community members that they will abandon FGC if a sufficient number of community members also agree to abandon the practice.

Recent empirical findings show that FGC is not a social coordination problem in Sudan (Efferson et al., 2015) and that individual- and household-level factors contribute to a larger share of the practice's persistence than do community-level factors in West Africa (Bellemare, Novak, & Steinmetz, 2015). This suggest that development agencies could broaden the set of policy interventions in order to address the support for and the perpetuation of FGC. Specifically, agencies could design policies aimed to change individual- or household-level preferences.

This paper's contribution is threefold. First, recent empirical findings that FGC is not a social coordination norm suggest the need for a new theoretical explanation for why FGC persists as a practice in many communities. I posit a new theory that highlights preference heterogeneity. I hypothesize that while households may wait to abandon FGC until a certain proportion of community members have already abandoned the practice, each household may have a different threshold, i.e. it may require a different proportion of community members to abandon FGC before it also abandons the practice.

Second, I test this theory with observational data from Burkina Faso, and I show that households do have heterogeneous thresholds. This heterogeneity provides important insights into why FGC persists; in a community with heterogeneous preferences there may not be a tipping point in the rate of FGC, and there may in fact be stable internal equilibrium. The distribution of household preferences has a large influence on the effectiveness of policies aimed at reducing the incidence of FGC (Platteau, Camilotti, and Auriol, 2017).

Third, I explore whether there is a tipping point in the proportion of community members practicing FGC beyond which the prevalence of the practice should shrink to zero or if there are stable internal equilibria in the rates of FGC. I show that some communities in my data have a stable equilibrium at low levels of FGC, suggesting that achieving at a low rate of FGC may be feasible; completely eliminating FGC, however, may be more difficult.

The strength of my approach comes from the long time period I analyze using three cross-sectional datasets from 1998, 2003, and 2010 from Burkina Faso that include women born between 1949 and 1995. The use of this six-decade timeframe is important when considering an intergenerational problem and long-term dynamics of FGC as well as the external validity of my estimates.

The remainder of this paper is organized as follows. In section 2 I provide background information on FGC. Section 3 provides a conceptual framework for analysis, and section 4 introduces the data and descriptive statistics. Section 5 discusses the empirical framework and estimation strategy, and section 6 present the results of the analysis. In section 7 I interpret the results and conclude.

2 Background

The practice of FGC is concentrated in 29 countries across parts of Africa, Asia, and the Middle East (UNICEF 2013), as well as among immigrant communities from those parts of the world now living in other countries. The World Health Organization (2012) classifies FGC into four types. Clitoridectomy (Type I) includes any partial or total removal of the clitoris, excision (Type II) includes partial or total removal of the clitoris and the labia minora, and infibulation (Type III) consists of narrowing the vaginal opening by sewing or stitching the labia together. Type IV includes all other procedures including pricking, piercing, incising, scraping or cauterizing the female genitalia for non-medical reasons.

In Burkina Faso, the vast majority of girls who undergo FGC, experience either clitoridectomy or excision. In the data, less than four percent of girls who undergo FGC experience infibulation. This figure is comparable to Jones, et.al. (1999), who show, using gynecological exams, that five percent of women in their Burkina Faso sample had undergone infibulation. The majority of procedures occur at very young ages in Burkina Faso. In the data, more than 94 percent of girls who undergo the procedure are cut before the age of 11, and approximately 42 percent of procedures occur during infancy.

The cultural norms surrounding FGC vary widely. Most scholars speculate that the practice originated as a way to reduce premarital and extramarital sex and thus ensure the paternity of children (Mackie, 1996; Boyle, 2005; Dorkenoo, 1999). Though FGC arose and may persist in part in response to male preferences, today women are primarily responsible for the decision to cut their daughters (Mackie, 1996; Toubia & Sharief, 2003). Undergoing FGC is seen as a rite of passage, or a way to join the society of women in the community (Toubia & Sharief, 2003). Some societies view the clitoris as a masculine part of a woman's body that must be removed in order for the girl to be fully female (Gruenbaum, 2000; Shell-

Duncan & Hernlund, 2001). Other societies believe that girls who have not undergone FGC are "unclean," and these girls are not allowed to wash dishes or touch certain items.

Burkina Faso is well suited to studying how and why the norm of FGC persists or wanes. The rate of FGC in Burkina Faso, while still high, fell substantially during the 60-year period for which I have data. This heterogeneity in rates of FGC by year of birth allows me to explore how household decisions change when faced with different rates of FGC. Further, it is possible to speculate about the future trends in countries where the rate of FGC remains high. Guinea, Mali, and Sierre Leone are three such countries, where rates of FGC are respectively 96 percent, 89 percent, and 88 percent of adult women (UNICEF, 2013).

3 Theoretical Model

In order to analyze a household's decision of whether its daughter should undergo FGC, I present an adapted version of the social coordination norm model proposed by Mackie (1996) and add to it Schelling's (1978) model of critical mass.

3.1 Social Coordination Norm

Consider a household's decision of whether its daughter will undergo FGC. I refer to the household as a single unit and remain agnostic about the decision-making process within the household because I am able observe only the final decision made by the household (i.e., the girl's FGC status) and the constraints face by the household, but not the intrahousehold bargaining process. Assume that the household knows the extent to which community members practice FGC. This is not an unreasonable assumption, because the FGC procedure is often performed in conjunction with a ceremony and community members often know who participated in such a ceremony (Cloward, 2016).

Let $r \in [0, 1]$ represent the rate of FGC at the community level, and let $g \in \{0, 1\}$ represent the daughter's FGC status where $g = 1$ denotes a girl who has undergone FGC. A household takes r as given and chooses between two strategies; (i) to cut its daughter ($g = 1$) and (ii) to not cut its daughter ($g = 0$). Let $s_{gi}(r)$ represent the social cost associated with strategy g for household i when faced with FGC rate r. This social cost is a function of r because social sanctions delivered by the community are related to the proportion of community members practicing FGC.

I first normalize to zero the payoff accruing to a household that does not cut its daughter in a community in which the cutting rate equals zero, that is $s_{0i}(0) = 0$. Then, let $s_{0i}(r) \geq 0$ and $s_{1i}(r) \geq 0$. These social costs include reduced marriage prospects of the girl (Wagner, 2014) and reduced acceptance into the community (Toubia & Sharief, 2003). Additionally, $s_{0i}(r)$ may include perceived disregard for religious edicts (Cloward, 2016) and inferior perceived beauty and femininity (Shell-Duncan & Hernlund, 2000). Finally, let $c_i \geq 0$ be the perceived costs—monetary, psychological, and physical—associated with cutting the girl.

Assume that the payoff accruing to a household that abstains from FGC is monotonically decreasing in the proportion of community members that practice FGC, and that the payoff accruing to a household that practices FGC is monotonically increasing in the proportion of community members that practice FGC.

Assumption 1. $\frac{\partial s_{0i}(r)}{\partial r} < 0$, *i.e. the payoff accruing to a household that abstains from practicing FGC is monotonically decreasing in the proportion of community members that practices FGC.*

Assumption 2. $\frac{\partial s_{1i}(r)}{\partial r} > 0$, *i.e. the payoff accruing to a household that practices FGC is monotonically increasing in the proportion of community members that practices FGC.*

It is not necessary to assume a functional form for the payoff functions. It is, however, important to consider the number of times the payoff functions for the two strategies could cross. Assumptions 1 and 2 allow us to conclude that the payoff functions can cross at most once in the space of FGC rates.

Lemma 1. *Under assumptions 1 and 2, the potential payoffs accruing to households under each strategy cross at most once in the space of community FGC rates.*

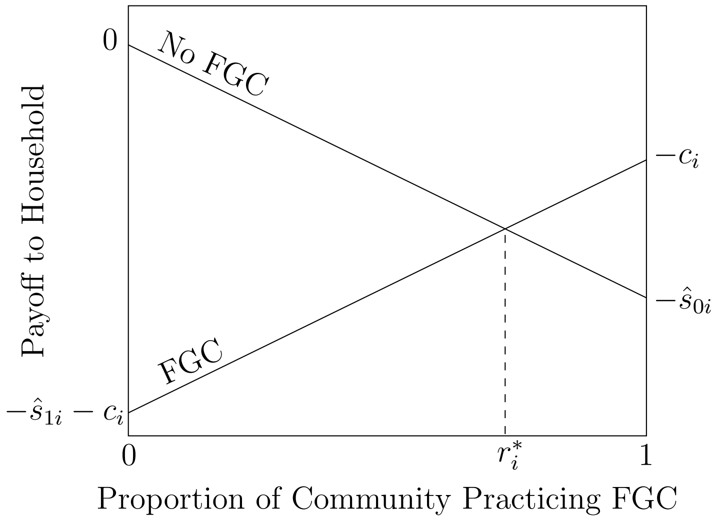
Figure 1 shows this graphically by displaying the potential payoff accruing to a household if it practices FGC or abstains from FGC, dependent on the proportion of community members practicing FGC. Let $s_{0i}(1) = \hat{s}_{0i}$ and $s_{1i}(0) = \hat{s}_{1i}$. Assume that $s_{1i}(1) = 0$ because there is no social cost of practicing FGC in a community in which all members practice FGC. Most households receive a higher payoff if they adhere to local norms. That is, in most cases, a household in a community in which all members practice FGC will achieve a higher payoff if the daughter undergoes FGC (i.e., $-c_i > -\hat{s}_{0i}$). Conversely, a household in a community in which no members practice FGC will achieve a higher payoff if the daughter does not undergo FGC (i.e., $0 > -\hat{s}_{1i} - c_i$).

For ease of exposition, and without loss of generality, I use a payoff structure that is a linear function of the rate of FGC in the community. Figure 1 shows that when the proportion of community members practicing FGC is above r_i^*, the household derives a higher payoff from practicing FGC than from abstaining. This situation flips if the proportion of practicing households is below r_i^*.

Proposition 1. *If $r < r_i^*$, household i will abstain from FGC for its daughter, and if $r > r_i^*$, household i will choose to have its daughter undergo FGC.*

Proof. Normalize $s_{0i}(0)$ to zero, and assume that $s_{1i}(1) = 0$. Let $s_{0i}(1) = \hat{s}_{0i}$ and $s_{1i}(0) = \hat{s}_{1i}$. If Lemma 1 is satisfied, then the household has a single indifference point at r_i^*.

Figure 1: Household $i's$ Payoff Structure



□

3.2 The Threshold Model of Collective Behavior

Mackie (1996) assumes homogeneous households within a community such that each household is facing the same costs and benefits of cutting. This implies that $r_i^* = r^* \ \forall \, i \in \mathcal{H}$, where $\mathcal{H}$ is the set of households in the community. According to this model, if a community can gather at least $1 - r^*$ community members to declare that they will abandon FGC, all community members will abandon FGC. That is, r^* is the *tipping point* for this community. This is the crux of the policies that arrange public declarations for the abandonment of FGC. Bellemare, Novak, & Steinmetz (2015) and Efferson et al. (2015) show that highly heterogeneous cutting rates—between zero and one—exist across communities. These findings are inconsistent with the hypotheses generated by Mackie's theory because if social sanctions prevent households from deviating from the norm and if a tipping point is guaranteed in every community, one would expect to see rates of FGC that are either very close to zero or very close to one.

Hypothesis 1. *Households within a community have heterogeneous thresholds* r_i^*.

If instead each individual values the benefits and costs to cutting differently, there are as many variations to figure 1 as there are community members. This implies that each individual will require a different proportion of community members to abandon FGC before deciding to switch from practicing to abandoning FGC. This indifference point—referred to in this paper as a threshold—is the proportion of community members that must abandon FGC in order for the household to abandon the practice. In figure 1, r_i^* is this household's threshold. In this paper I test whether households in a given community have heterogenous thresholds.

3.3 Tipping Point or Stable Internal Equilibria

A key insight from Efferson et al. (2015) is that if thresholds are heterogeneous within a community, a tipping point as proposed in Mackie (1996) may not exist. Schelling (1978, p. 105) makes the point that in the face of heterogeneous thresholds (what he refers to as "cross-over points"), multiple stable equilibria may exist, some of which are interior solutions. Drawing on Schelling's model of critical mass, I investigate whether there is a tipping point for the practice of FGC in communities in Burkina Faso. To do this, I analyze the cumulative distribution function (CDF) of community member thresholds.

Continue to assume that a household observes the rate of FGC in the community before making the decision for their daughter. Let $f(r_i^*)$ be the probability density function (PDF) and $F(r)$ be the CDF of the thresholds of the community members. Then, $F(r) = \int_0^r f(r_i^*)dr_i^*$ represents the proportion of community members with a threshold that is lower than or equal to r.

If all households have the same threshold, the PDF of community thresholds would be a vertical line at the level r^*, and the CDF of the proportion of households favoring FGC would resemble curve 1 in figure 2. In this curve, no household favors FGC if the rate of FGC is below r^*, and all households favor FGC for their daughter if the rate is above r^*. Alternatively, if thresholds are heterogeneous, the CDF of the proportion of households favoring FGC could resemble curve 2 (or a myriad of other curves). In this particular rendering of a CDF in a community with heterogeneous household thresholds, a small proportion of household will practice FGC even if they believe that no one else will. The proportion of community members favoring FGC is higher if the proportion of community members practicing FGC is larger. In this hypothetical community there are some household that will abstain from practicing FGC even if they expect every other household in the community will practice FGC.

Figure 2: Hypothetical Community CDFs

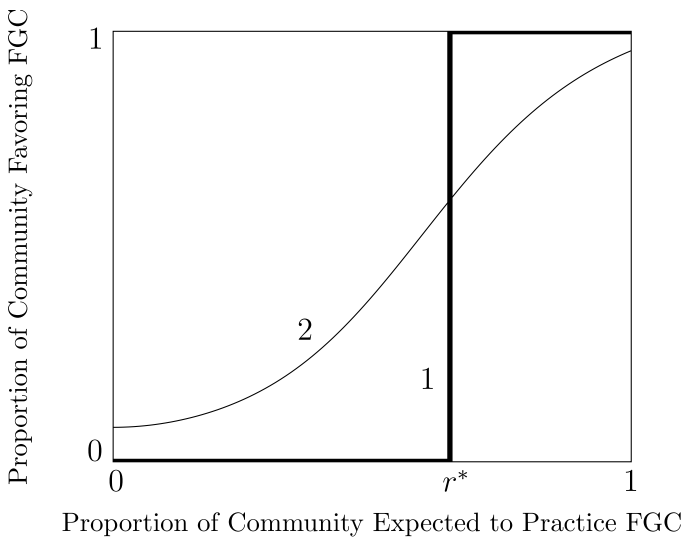
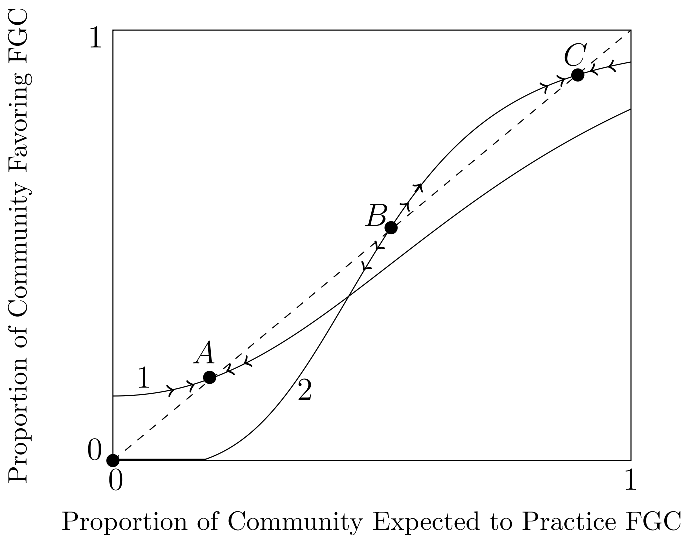


Figure 3 shows hypothetical CDFs of the proportion of community household favoring FGC as a function of the proportion of community household that are expected to practice FGC. The dashed line is the 45-degree line.

Figure 3: Hypothetical Community CDFs



Any point at which the CDF crosses the 45-degree line, or $F(r) = r$, is an equilibrium—either stable or unstable. Curve 1 of figure 3 shows a community in which a small portion of households value FGC sufficiently that they will practice FGC even if no other household practices. Because these households practice FGC, they will draw a few more community households with a low threshold into practicing FGC until the community reaches point A. If instead the rate of FGC is higher than the rate at point A, there are fewer households favoring FGC (as a function of practicing households) than there are practicing households. This would lead these households to abandon FGC, pushing the equilibrium to point A. Thus, point A is a stable equilibrium, and it is the only equilibrium on curve 1.

If instead the community's distribution of thresholds resembles curve 2, there are three equilibria, two of which are stable. If the proportion of households practicing FGC is below point B, fewer households favor FGC than there are households expected to practice, so the proportion of households practicing FGC will fall to zero. If instead, the proportion of household practicing FGC is above point B, more households will be drawn to practicing

FGC until point C is reached. If the proportion of households practicing FGC is higher than at point C, some will decide to stop practicing, and the proportion of practicing households will return to point C. A community with thresholds, r_i^* of community members distributed as shown in curve 2 has two stable equilibria—one at a high rate of FGC and one at zero, and one unstable equilibrium at point B.

Thus, with heterogeneous rates of cutting it is possible that there is a tipping point, as is illustrated with curve 2 at point B. It is also possible that a tipping point does not exist, as is illustrated with curve 1. Thus, if community members have heterogeneous thresholds, the existence of a tipping point is not guaranteed and it is possible that there is a stable equilibrium at an interior rate of cutting (i.e., $r \in (0, 1)$). Below, I use data from Burkina Faso to determine which phenomenon is most prevalent in communities across the country.

3.4 Estimating the Cumulative Distribution Function

A household's threshold is unobservable. It is a function of individual preferences and intra-household bargaining power. Additionally, each individual may have difficulty articulating their threshold if asked via survey. Fortunately, for the purpose of estimating a community's CDF, it is not necessary to identify a household's threshold. Instead, by revealed preference, it is possible to recover the community CDF from the behavior of the community members. In fact, the CDF at rate r is exactly equal to the proportion of households that opt to practice FGC when faced with rate of FGC r.

Let $\pi(r)$ be the proportion of households that opt to practice FGC when faced with rate of FGC r, and let $g_i = 1$ if the girl undergoes FGC. Then, $\pi(r) = F(r)$.

Proposition 2. *$F(r) = \pi(r)$.*

Proof. The probability that girl i undergoes FGC when her parents are faced with rate r

is equal to the proportion of households that opt to practice FGC when faced with rate r, or $\pi(r) = \mathbb{P}(g_i = 1|r)$. Furthermore, the probability that a girl undergoes FGC when faced with rate r is equal to the probability that her household's threshold is less than r, or $\mathbb{P}(y_i = 1|r) = \mathbb{P}(r_i^* < r)$. And, $\mathbb{P}(r_i^* < r) = \int_0^r f(r_i^*)dr_i^* = F(r)$.

$$\pi(r) = \mathbb{P}(y_i = 1|r) = \mathbb{P}(r_i^* < r) = \int_0^r f(r_i^*)dr_i^* = F(r)$$

□

Thus, it is possible to recover the CDF of household thresholds by estimating the rate of FGC in a cohort given the rate of FGC of the previous cohort, because $\pi(r) = F(r)$.

In the next section I discuss the data used in this analysis.

4 Data and Descriptive Statistics

I use publicly available data from the Demographic and Health Survey (DHS) for Burkina Faso. I use three cross-sectional datasets collected in 1998, 2003, and 2010. The DHS includes female respondents aged 15 to 49 at the time of survey, thus I have data for women born between 1949 and 1995. A respondent provides information on her health and her children's health, along with many characteristics of her partner, if relevant, and the household in which she lives. I identify in which department and province a household resides by using the geographic coordinate information provided by the DHS. There are 45 provinces in Burkina Faso[2] and 351 departments. Provinces are the second administrative level in Burkina Faso, one level below region, one level above departments, and two levels above villages. Combining

[2]In 1998 there were 30 provinces. The provinces were redrawn between 1998 and 2003. I use the geographic coordinates of the villages sampled in 1998 to determine in which modern province the village lies.

three cross-sectional datasets facilitates the investigation of a much longer timeframe than any one dataset provides, which is important when considering long-term dynamics in FGC.

Women report their own FGC status and the FGC status of their daughters. These reports are susceptible to reporting bias. In order to determine the likelihood and extent of misreporting, I compare my data to data from Burkina Faso collected in 1998 in which gynecological exams are used to estimate the prevalence of FGC (Jones et al., 1999). This study includes women attending rural clinics in two provinces of Burkina Faso (Bazèga and Zoundwéog) who are between 15 and 55 at the time of data collection. They find that 93 percent of women in their sample have undergone FGC. Restricting my sample to the rural areas in the two specified provinces and weighting rates of FGC by cohort according to the proportion of their sample in each cohort, I find a rate of FGC of 89 percent. This four percentage point difference in the rate of FGC may be even smaller given that the sample used by Jones et al. (1999) includes women up to the age of 55 in 1998 while my data include women up to the age of 49, and this cohort has the highest observed rate of FGC. Further, their data were collected in clinics, and it is possible that the women attending the clinics have more health complications than the average woman in these areas, and some of these health complications may have been the result of FGC. While this does not rule out the possibility of reporting bias, I argue that it is not a large concern in this setting.

Table 1 shows descriptive statistics for the three samples used in my analyses. These three cross-sections provide information from 24,474 women born between 1949 and 1995 who report their FGC status (shown in column 1). In many specifications, I use data only from the 6,873 women who have a daughter who is old enough to have undergone FGC and who report their daughter's FGC status (column 2). This sample is restricted to women with a daughter who is 11 years or older because 94 percent of women who underwent FGC were cut before the age of 11. Including daughters younger than 11 would risk counting girls as uncut when in fact they will undergo FGC at a later date. In a small number of analyses I include only the 5,934 women who report their daughter's FGC status and who themselves underwent FGC (column 3).

Approximately 75 percent of women surveyed have undergone FGC. Only 28 percent of surveyed women are educated—bluntly defined here as having attended at least some formal primary school. Almost 76 percent have been or are currently married, the average year of birth of respondents is 1981, 25 percent of respondents are Catholic and 59 percent are Muslim, 19 percent of respondents own a television and 73 percent own a radio. Fifty-four percent of respondents are from the Mossi ethnic group, the remaining 46 percent belong to 10 other ethnic groups. Thirty-one percent of respondents live in an urban area.

Thirty-eight percent of women who have a daughter 11 years of age or older state that their daughter has undergone FGC, while 44 percent of women who have undergone FGC themselves have a daughter who has undergone FGC. Due to data limitations, I use the information on the woman's oldest daughter's FGC status only. The data collected in 1998 and 2003 include information on the oldest daughter only, data collected in 2010 include data on all daughters and show that there is little variation in FGC status of a woman's daughters. Thus, using the woman's oldest daughter only should yield very similar results to using information on multiple daughters. The women included in the daughter analyses (columns 2 and 3) are, on average, older, less educated, and have a higher rate of FGC than the full sample of women. This is unsurprising since these are women with relatively old daughters, and the rate of FGC has been decreasing over time in Burkina Faso while the rate of education has been rising.

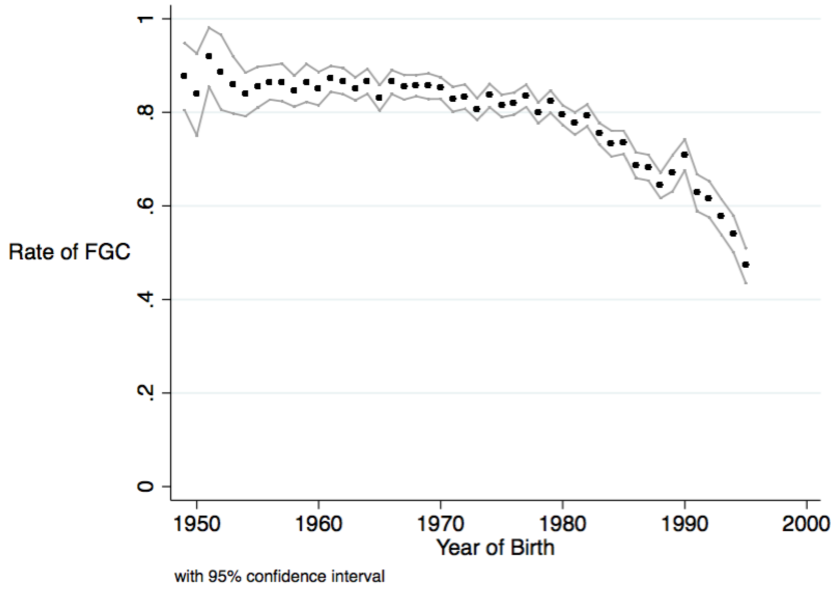
Figure 4 displays the rate of FGC by year of birth. The substantial amount of heterogeneity in rates of FGC over time in Burkina Faso which makes Burkina Faso the ideal place to analyze changes in the practice of FGC.

Table 1: Descriptive Statistics

	(1) Full Sample	(2) Respondents with Daughters	(3) Cut Respondents with Daughters
Respondent Underwent FGC	0.749	0.863	1.000
	(0.003)	(0.004)	(0.000)
Daughter Underwent FGC		0.383	0.437
		(0.006)	(0.006)
Respondent Educated	0.280	0.104	0.103
	(0.003)	(0.004)	(0.004)
Respondent Ever Married	0.755	1.000	1.000
	(0.003)	(0.000)	(0.000)
Respondent's Partner Educated		0.125	0.126
		(0.004)	(0.004)
Average Rate of FGC in Province Cohort	0.745	0.633	0.654
	(0.001)	(0.002)	(0.002)
Average Rate of FGC in Province-Ethnic Cohort	0.744	0.637	0.663
	(0.001)	(0.003)	(0.003)
Respondent in Polygamous Marriage		0.568	0.567
		(0.006)	(0.006)
Respondent Year of Birth	1981	1967	1967
	(0.045)	(0.086)	(0.092)
Respondent's Partner's Year of Birth		1955	1956
		(0.137)	(0.145)
Daughter's Year of Birth		1988	1989
		(0.080)	(0.087)
Religion: Catholic	0.246	0.240	0.226
	(0.003)	(0.005)	(0.005)
Protestant	0.062	0.054	0.046
	(0.002)	(0.003)	(0.003)
Muslim	0.592	0.580	0.610
	(0.003)	(0.006)	(0.006)
Other	0.100	0.126	0.118
	(0.002)	(0.004)	(0.004)
Household Owns TV	0.189	0.117	0.119
	(0.003)	(0.004)	(0.004)
Household Owns Radio	0.733	0.702	0.708
	(0.003)	(0.006)	(0.006)
Ethnic Group: Bobo	0.037	0.032	0.032
	(0.001)	(0.002)	(0.002)
Dioula	0.026	0.028	0.030
	(0.001)	(0.002)	(0.002)
FulFulde/Peul	0.067	0.057	0.061
	(0.002)	(0.003)	(0.003)
Gourmatche	0.062	0.062	0.055
	(0.002)	(0.003)	(0.003)
Gourounsi	0.047	0.042	0.031
	(0.001)	(0.002)	(0.002)
Lobi	0.045	0.029	0.031
	(0.001)	(0.002)	(0.002)
Mossi	0.542	0.588	0.598
	(0.003)	(0.006)	(0.006)
Senoufo	0.049	0.049	0.053
	(0.001)	(0.003)	(0.003)
Touareg/Bella	0.012	0.011	0.005
	(0.001)	(0.001)	(0.001)
Bissa	0.040	0.041	0.044
	(0.001)	(0.002)	(0.003)
Other	0.073	0.059	0.061
	(0.002)	(0.003)	(0.003)
Household is in Urban Area	0.307	0.206	0.209
	(0.003)	(0.005)	(0.005)
Observations	24,474	6,873	5,934

Standard errors in parentheses

Figure 4: Rate of FGC in Burkina Faso by Year of Birth



4.1 Identifying the Community & Cohort

In order to estimate the rate of FGC at the community level, I first identify the year of birth of the respondent and the year of birth of the respondent's oldest daughter. Year of birth of respondents and daughters (of different women) do overlap, given that my data provide such a long timeframe. I combine information on self-reported FGC status of the respondents (using the full sample of 24,474 women) with the reported FGC status of daughters (6,873 daughters). I then estimate the rate of FGC in the daughter's cohort within her community.

I define cohort as the girls born five years prior to the birth of the daughter in question. I use the information on girls born five years prior because I assume that the household is able to observe the rate of FGC for girls in preceding years of birth before making the decision for its own daughter. This definition of cohort has the added benefit of minimizing the reflection problem (Manski, 1993).[3]

[3]The reflection problem arises when attempting to identify the effect of a group's behavior on an individual when that individual's behavior simultaneously affects the group's behavior.

Correctly identifying the community is especially important in this type of analysis. Ideally, I would have information on the social network of observed households. In the absence of this data, I use multiple definitions of the community and analyze the consistency, or lack thereof, of my results between definitions. I use the marriage market as a proxy for the community because people are more likely to interact with individuals in their marriage pool as well as see them as the relevant reference group for norms. In Burkina Faso, inter-ethnic marriage are uncommon (Breusers, Nederlof, Van Rheenen, 1998), thus I use ethnic group in every definition of community. It is common for Burkinabe women to move out of their natal village for marriage (Henry, Schoumaker, Beauchemin, 2004), thus I do not use village as the definition of community.

For robustness I define the community in four ways. First, all girls in the same province within the same ethnic group. In order to determine if province is too large to accurately capture the community, I narrow the definition of community by looking at only those girls in the same department. Recall that there are 45 provinces and 351 departments in Burkina Faso, so each province includes 7.8 departments, on average. Thus, the second definition of community is all girls in the same department within the same ethnic group. While households in rural communities are aware of the FGC status of other girls in the community, this assumption may be less reasonable in urban areas. Thus, the third definition of community is all girls in the same province within the same ethnic group who live in the rural area. Lastly, religion plays a key role in the marriage market. I focus on Catholics and Muslims because they are the most common religions in Burkina Faso. Thus, the fourth definition of community is all girls in the same province within the same ethnic group and same religious group.

5 Empirical Framework

In this section I explain my estimation strategy, relate it to the ideal experiment for comparison, and I discuss threats to validity of my estimates. I begin by testing Hypothesis 1, that is, whether thresholds are heterogeneous among households in Burkina Faso. Second, I test whether there is a tipping point or stable internal equilibrium in the rate of FGC in communities in Burkina Faso.

5.1 Tipping Point or Stable Internal Equilibrium

I begin by analyzing whether thresholds are heterogenous among households in Burkina Faso by plotting the CDF of the proportion of households that decide to cut their daughter as a function of the proportion of community members practicing FGC (where community is defined in the four ways specified above). As stated in Proposition 2 it is not possible to observe each household's threshold r_i^*. Instead, I am able to directly observe the community CDF by estimating the proportion of households opting to practice FGC as a function of the previous cohort's rate of FGC, r.

I perform a kernel-weighted local polynomial smoothing regression of the rate of FGC in community k for girls born in year t on the rate of FGC in community k among girls in the cohort c born between $t-1$ and $t-6$. I use an an Epanechnikov kernel with a polynomial smooth of degree two because higher order polynomials perform better at the boundary points than lower order polynomials (Fan and Gijbels, 1996).

$$r_{tk} = f(r_{c\in[t-6,t-1],k}) + \eta_{tk} \tag{1}$$

If the CDFs resembles curve 1 in figure 2, households have homogenous thresholds. Otherwise, we can conclude that thresholds are heterogeneous among households.

I then turn to analyzing whether there is a tipping point in communities or if there are stable internal equilibria in the rate of FGC. I do this by determining how the CDFs estimated using equation 1 interact with the 45-degree line, and I compare this to figure 3.

5.2 Estimation Strategy

I estimate the likelihood that a girl will undergo FGC based on the rate of FGC in her community within her cohort and her household characteristics. I estimate equation 2 for girls who have a mother who has undergone FGC. I restrict the sample to girls with mothers who have undergone FGC because it is rare for a woman who has not undergone FGC to have a daughter who has undergone FGC.

$$y_{itk} = \beta_0 + \beta_1 r_{ck} + \beta_2 m_{itk} + \beta_3 f_{itk} + \beta_4 \mathbf{X}_{itk} + \beta_5 \mathbf{d}_t + \beta_6 \mathbf{d}_m + \beta_7 \mathbf{d}_f + \beta_8 \mathbf{s} + \epsilon_{itk} \tag{2}$$

The subscripts denote girl i born in year t, whose relevant cohort is cohort $c \in [t-6, t-1]$ whose mother is in cohort m, whose father is in cohort f, and who lives in community k. Let y_{itk} equal one if the respondent's daughter has undergone FGC and zero if her daughter has not undergone FGC. Let r_{ck} denote the rate of FGC in the girl's community-cohort, let m_{itk} denote the level of education of the girl's mother, and f_{itk} denote the level of education of the girl's father. Let $\mathbf{X}_{itk}$[4] be a vector of other control variables, $\mathbf{d}_t$ is a vector of the girl's year of birth fixed effects, $\mathbf{d}_m$ is a vector of mother's year of birth fixed effects, $\mathbf{d}_f$ is a vector of father's year of birth fixed effects, and $\mathbf{s}$ is a vector of survey wave fixed effects. Finally, ϵ_{itk} is an error term with mean zero.

I estimate equation 2 using ordinary least squares (OLS) regression. Because y_{itk} is binary,

[4] Boldface is used to denote vectors.

my use of OLS implies that each equation I estimate is a linear probability model (LPM). In estimating an LPM rather than a logit or a probit model, I follow the recommendations of Angrist and Pischke (2009). The primary benefits of using LPMs are (i) LPMs do not rely on distributional assumptions for the error term that are required by logit and probit estimators, and (ii) LPMs do a much better job than probit models of handling a large number of fixed effects (Angrist & Pischke, 2009 page 98). The primary drawback to using LPMs is that LPMs produce errors that are heteroskedastic. I use robust standard errors in all estimations in order to address this concern. An additional drawback of LPMs is that LPMs can predict a likelihood of FGC that is outside of the $[0, 1]$ interval. This is not a large concern in this case, as I am not attempting to forecast the likelihood that a girl will undergo FGC. In a series of robustness checks, I estimate equation 2 using logit regression (see table A2).

5.3 Ideal Dataset and Threats to Validity

In order to appreciate the threats of validity of my approach, it is useful to imagine the ideal dataset for addressing this question. Assume first that in each community there is a set time when all girls in a given community undergo FGC, if their household decides to practice FGC.

Now, imagine that a researcher can go to each household within a community to ascertain the household's exact threshold. The researcher should then determine which households have girls who are near the age of undergoing FGC, let's call them group 1. Assume the researcher knows the rate of FGC among the previous cohort, then given the distribution of household thresholds among group 1, the researcher could determine what the rate of FGC will be in group 1. Based on the distribution of thresholds among the households of girls born in the subsequent year, call them group 2, the researcher could determine what the rate of FGC will be in the subsequent period. The household threshold is a complicated function of individual preferences for FGC and intrahousehold bargaining power. Moreover, it is likely difficult for individuals to articulate their preference for FGC.

A much more reliable approach involves observing a households revealed preference. Consider instead a dataset in which there are 100 communities. Each community has an identical distribution of thresholds. Each community, however, has a different rate of FGC (imagine that this is, for example, a result of interventions in some communities). One could then observe the rate of FGC among group 1 in all 100 communities. From this information, a research could infer the distribution of thresholds within the 100 communities. Unfortunately, it is not possible for the observer to guarantee that the distribution of thresholds is identical in each of these communities.

Instead, imagine a world in which each household in a community has 10 daughters, each born one year apart. In year one, a research could give false information to each household regarding the rate of FGC in the community. The research could then observe the choice of the household made to cut or refrain from cutting their daughter in group 1. In the second year, the researcher could return to the household and give new false information regarding the rate of FGC and observe whether the household cuts its daughter in group 2. Assuming that all households believe the researchers each year, and if the falsely reported rate of FGC is in increments of 10 percentage points, the researcher could create bounds on each household's threshold within 10 percentage points. Finding such a community is unlikely, moreover this strategy relies on the assumption that the household's threshold does not change within that 10 year timeframe.

These research designs are not feasible. The data that I use are as close to ideal as is possible because I use data on revealed preference, and I observe heterogeneity in the rate of FGC within a given community. It is, however, important to highlight key limitations of the data. First, all of the research designs above assume that the researcher has correctly identified the relevant community. In the ideal dataset, one would ask respondents who they view as the relevant community or who they interact with on a regular basis in order to identify the social network. In lieu of that information, I have identified the most likely definition of community according the literature, and in order to alleviate some of this concern, I vary the definition of community to determine if my results are consistent across definition.

Second, I assume that the distribution of thresholds within a community do not change over the period in which I observe them. This is a strong assumption since households observed at a later date may, for example, have more information about the health complications associated with FGC or may have experienced a religious leader publicly denounce FGC. More fundamentally, each year the households that make the decision to cut their daughters are different households than the year before or the year after. If thresholds are randomly allocated among households with daughters of different ages, the fact that I observe different households does not bias my measurement of the threshold distribution. If instead thresholds are correlated with age of the daughter, this could have some effect on the observed distribution of thresholds. If for example, people with high thresholds had daughters born in 1990, while people with low thresholds had daughters born in 1995, this could make the CDF appear volatile. The use of the local polynomial smoothing regression addresses this issue.

Third, I assume that households have a clear picture of the proportion of households that practice FGC. This is reasonable for rural communities as the FGC procedure is typically celebrated and households, and particularly women in those households, are aware of who has undergone FGC. This, however, may be a strong assumption for urban communities. Thus, I conduct robustness checks in which I analyze only rural communities.

6 Results

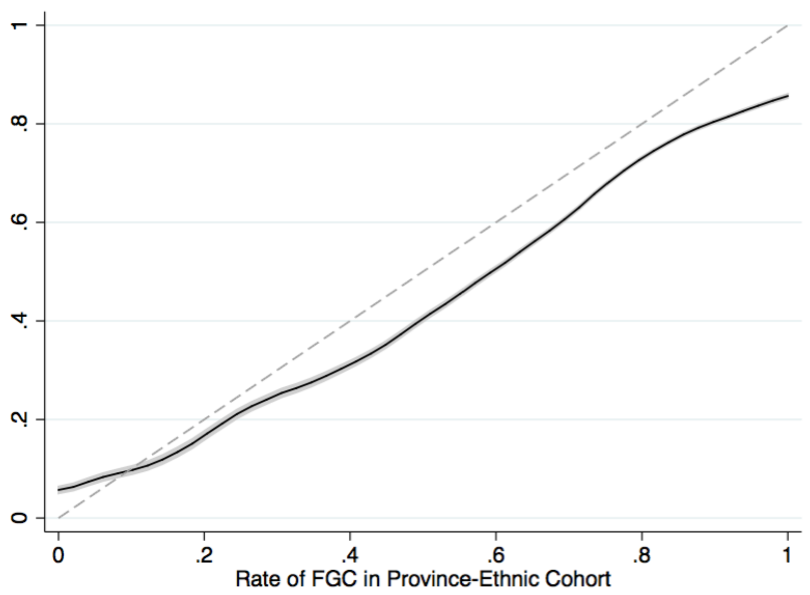
I begin by examining whether households have heterogeneous thresholds. I then analyze whether there is a tipping point or stable internal equilibria in the rate of FGC in communities in Burkina Faso.

6.1 Thresholds and Tipping Point

Figure 5 displays the CDF of household thresholds for all of Burkina Faso using the ethnic group within the province as the definition of community. One can see that within Burkina Faso thresholds are heterogeneous because the CDF does not resemble curve 1 in figure 2, and in fact looks very similar to curve 2 in figure 2.

Figure 5: CDF for Burkina Faso

Community defined as Province-Ethnic Cohort, with 95% confidence interval



Because the CDF crosses the 45-degree line from above, figure 5 indicates that there is stable interior equilibrium in the rate of FGC in Burkina Faso. This suggests that the rate of FGC may remain close to 10% in Burkina Faso for some time. It is, however, a bit misleading to analyze the CDF for Burkina Faso as a whole. Instead, we should draw conclusions from community-level figures.

I now turn to analyzing whether there is a tipping point or stable internal equilibrium in the proportion of community members practicing FGC in communities in Burkina Faso. Figures 6, 7, and 8 show the CDFs for a select sample of three province-ethnic groups. Panel a in each figure is the CDF for the entire province-ethnic group, panel b is the CDF for the province-ethnic group in the rural area only, panels c and, if applicable,[5] d include data from the department-ethnic groups from departments within the given province. Panel e shows the CDF for Catholics in the province-ethnic group, and panel f shows the CDF for Muslims in the province-ethnic group.

The CDFs for the Mossi in Mouhoun (figure 6) suggest that a stable low-rate of FGC is possible among this community. The rural communities appear to behave similarly to the province as a whole. The observed department includes only low rates of FGC, suggesting that this department may be different from the 12 other departments in Mouhoun. Catholics have a lower rate of FGC than Muslims in Mohoun.

The CDFs for the Senoufo in Ganzourgo (figure 7) show that the observed rates of FGC are higher among this province-ethnic group than among the Mossi in Mouhoun. Panel a suggest that this group may have had a tipping point at a rate of FGC of approximately 95 percent. The rural community behaves similarly to the province as a whole, though the very low rates of FGC observed in panel a appear to occur the urban area. The observed department (of which is 78 percent of surveyed households lie in an urban area) behaves

[5] I include only communities with 50 or more girls for whom the decision to practice FGC is observed. This excludes a large number of department-ethnic groups as well as some province-ethnic groups.

similarly to the province as a whole, but more similarly to the rural area. Lastly, Catholics and Muslims behave similarly to one another among the Senoufou in Ganzourgo.

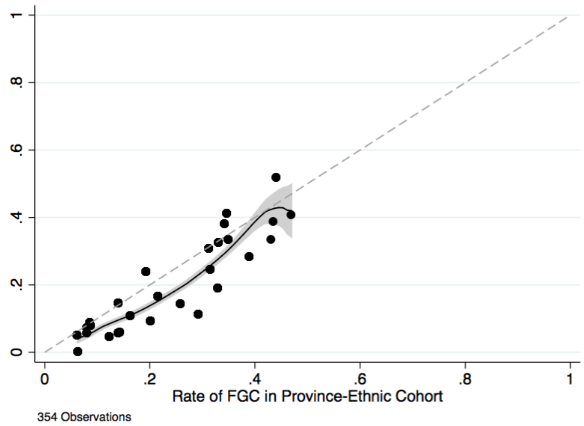
The CDFs for the Mossi in Ioba (figure 8) suggest that there is a tipping point in the rate of FGC among this province-ethnic group. The rural area behaves similarly to the full province-ethnic group. The two departments have little data for drawing conclusions, but the first department (panel c) appears to behave very similarly to the rural community (though 79 percent of surveyed households in this department lie in an urban area). The Catholic and Muslim households behave similarly to each other and to the broader province-ethnic group.

6.2 Likelihood That Girl Undergoes FGC

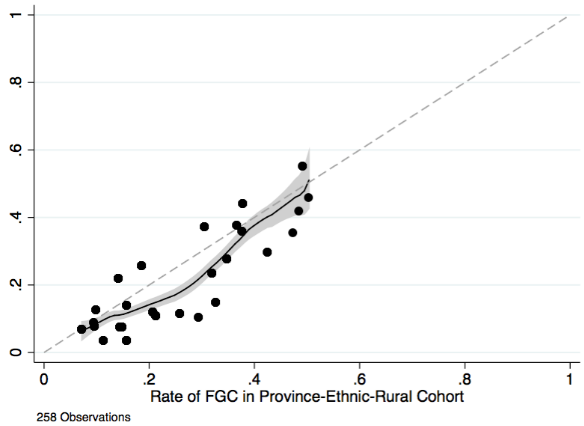
Table 2 presents the results of estimating equation 2 using five definitions of the girl's community-cohort. Cohort is defined as all girls born five years prior to the girl in question. The definition of community is always narrowed to the girl's ethnic group. Column 1 defines the girl's community-cohort as all girls who live in the same province, column 2 defines the girl's community-cohort as all girls who live in the same department. Column 3 estimates equation 2 defines the girl's community-cohort as all girls who live in the province and is restricted to those living in a rural area. Columns 4 and 5 defines the girl's community-cohort as all girls who live in the same province, are part of the same religious group.

Figure 6: CDFs from the Mossi in Mouhoun
with 95% confidence interval

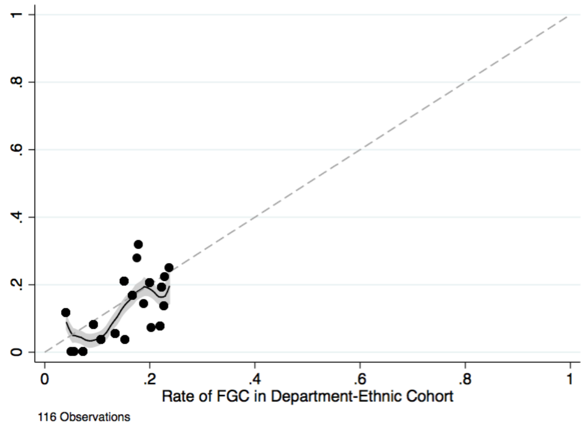
(a) All Province-Ethnic Group



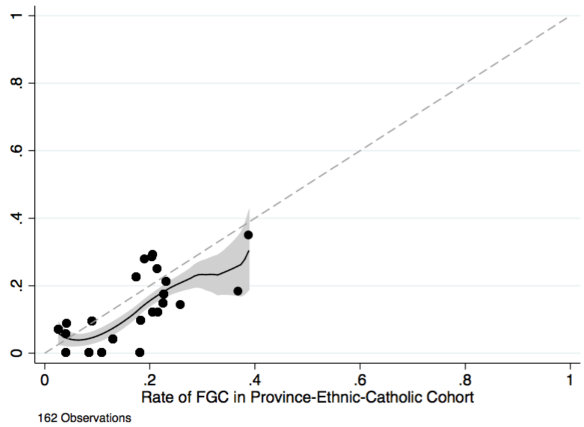
(b) Rural Province-Ethnic Group



(c) Department 1-Ethnic Group



(e) Catholic-Province-Ethnic Group



(f) Muslim-Province-Ethnic Group

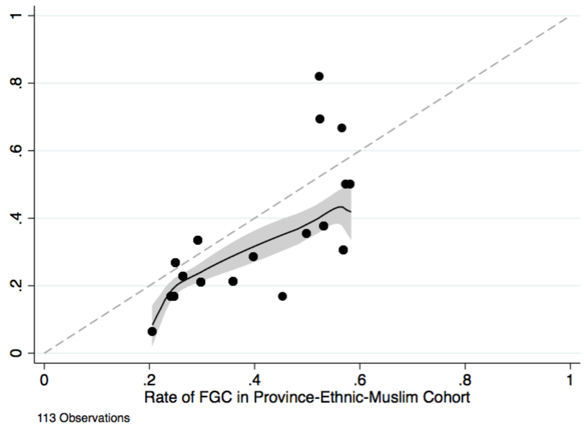
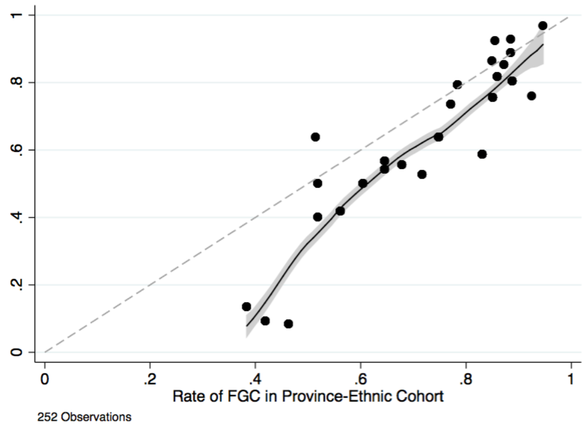
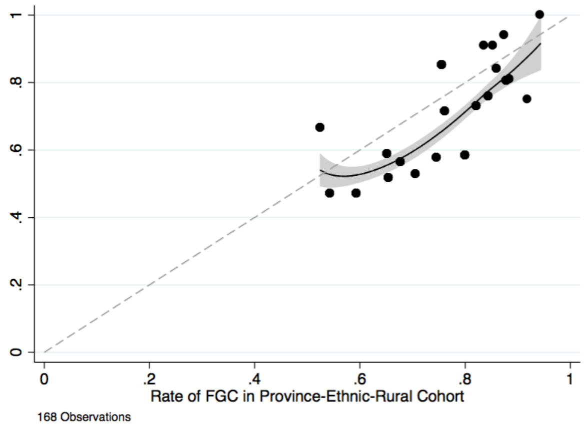


Figure 7: CDFs from the Senoufo in Ganzourgo
with 95% confidence interval

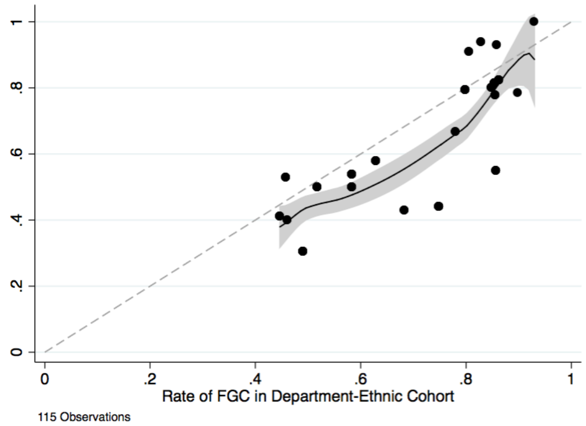
(a) All Province-Ethnic Group



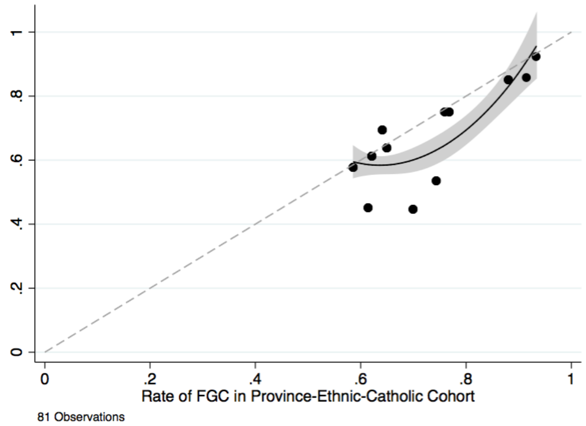
(b) Rural Province-Ethnic Group



(c) Department 1-Ethnic Group



(e) Catholic-Province-Ethnic Group



(f) Muslim-Province-Ethnic Group

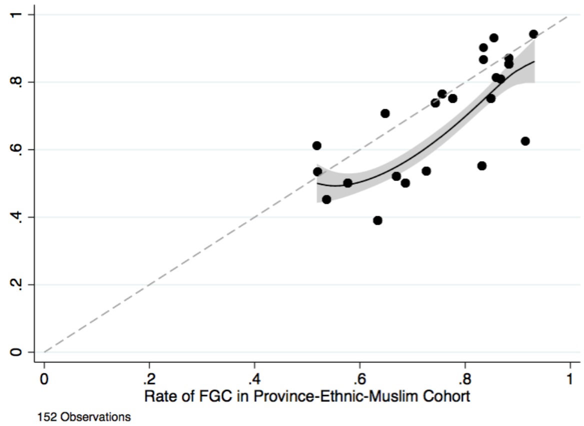
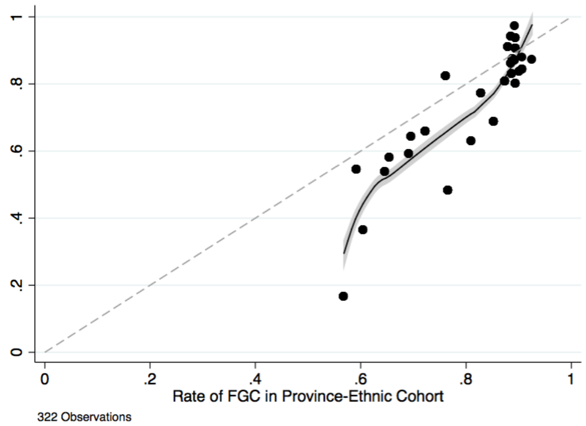
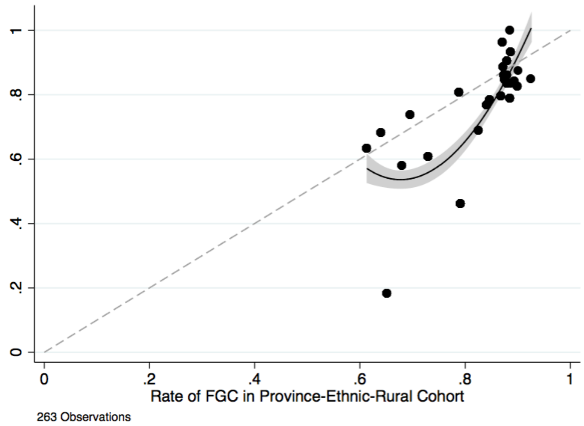


Figure 8: CDFs from the Mossi in Ioba
with 95% confidence interval

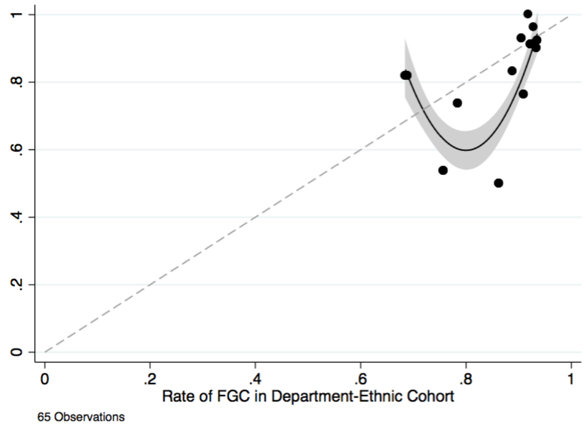
(a) All Province-Ethnic Group



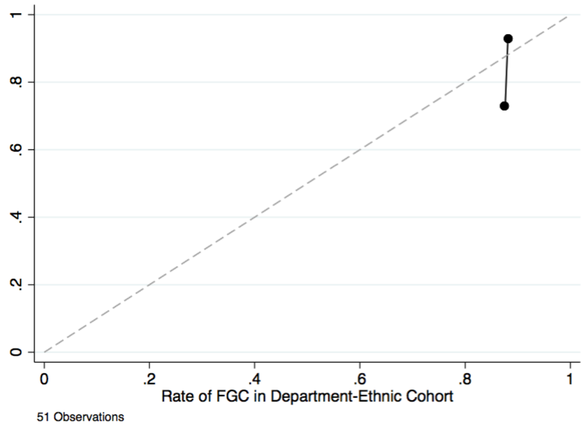
(b) Rural Province-Ethnic Group



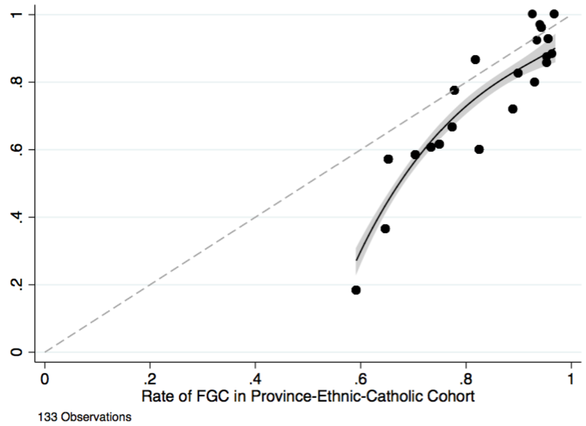
(c) Department 1-Ethnic Group



Department 2-Ethnic Group



(e) Catholic-Province-Ethnic Group



(f) Muslim-Province-Ethnic Group

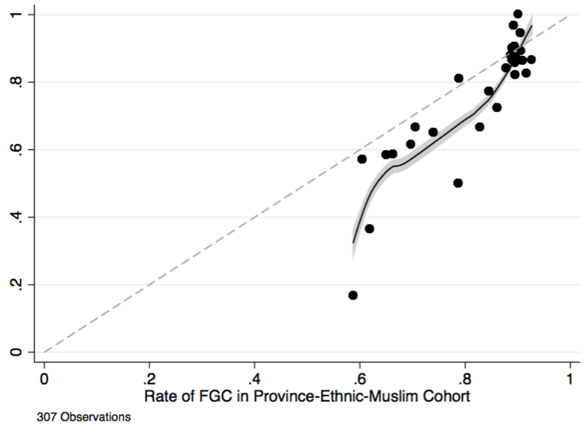


Table 2: LPM Estimation Results for Whether Girl Undergoes FGC

	(1) Province-Ethnic Cohort	(2) Department-Ethnic Cohort	(3) Province-Ethnic Rural Cohort	(4) Province-Ethnic Catholic Cohort	(5) Province-Ethnic Muslim Cohort
Rate of FGC in Community	0.580***	0.568***	0.588***	0.436***	0.548***
	(0.048)	(0.034)	(0.044)	(0.070)	(0.066)
Urban Household	-0.037*	-0.037*		0.030	-0.056**
	(0.020)	(0.021)		(0.039)	(0.026)
Religion [Omitted: Catholic]					
Protestant	-0.034	-0.049	-0.020		
	(0.030)	(0.041)	(0.045)		
Muslim	0.076***	0.076***	0.096***		
	(0.015)	(0.017)	(0.021)		
Other	0.064**	0.048	0.061		
	(0.029)	(0.035)	(0.038)		
Constant	0.511***	0.583***	0.615***	0.692*	1.347***
	(0.161)	(0.187)	(0.193)	(0.356)	(0.191)
Observations	5,934	4,523	3,493	1,079	2,818
R-squared	0.215	0.237	0.228	0.297	0.213

Controls included: Ethnicity, religion, education level of the mother, education level of the father, proxies for wealth, rural indicator, polygamy indicator, mother's year of birth FE, father's year of birth FE daughter's year of birth FE
Standard Errors Clustered at the Province Level
***p<0.01 **p<0.05 *p<0.1

Table 2 shows that the rate of FGC in the girl's community cohort plays an important role in determining the girl's likelihood of undergoing FGC, however the likelihood that a girl undergoes FGC is not fully determined by the rate of FGC in her community-cohort. Coefficients range between 0.436 and 0.588 suggesting that a girl living in a community in which $r = 1$ is between 43.6 and 58.8 percentage points more likely to undergo FGC than a girl living in a community in which no one practices FGC. These two extremes ($r = 0$ and $r = 1$) are uncommon in my data. A more useful conceptualization of the results is that for each 10 percentage point increase in the rate of FGC, a girl is between 4.4 and 5.9 percentage points more likely to undergo FGC.

The coefficients on the rate of FGC within the community (variably defined) are similar but not identical across definitions of the community. Following Pei, Pischke, and Schwandt (2016) I perform a Hausman (1978) specification test to determine whether these coefficients are statistically similar. Table 3 reports the Hausman test statistics. Recall that the null hypothesis of the Hausman test is that the two coefficients are similar. In six of the reported cases, the null is rejected at the 5 percent level, and the null is rejected in seven of cases at the 10 percent level. The coefficients on the rate of FGC in the province-ethnic cohort, the rate of FGC on the province-ethnic-rural cohort, and the rate of FGC in the province-ethnic cohort among Muslims are statistically similar to each other.

Table 3: Hausman Specification Test Results Between Coefficient on Rate of FGC in the Girl's Community

	(1) Province-Ethnic Cohort	(2) Department-Ethnic Cohort	(3) Province-Ethnic Rural Cohort	(4) Province-Ethnic Catholic Cohort
Department-Ethnic Cohort	5.46			
	(0.020)			
Province-Ethnic Rural Cohort	0.07	4.16		
	(0.786)	(0.041)		
Province-Ethnic Catholic Cohort	17.54	4.08	13.18	
	(0.000)	(0.043)	(0.000)	
Province-Ethnic Muslim Cohort	0.02	3.19	0.00	8.54
	(0.883)	(0.074)	(0.973)	(0.004)

χ^2 value reported with p-value in parentheses

6.3 Discussion

According to the regression results and Hausman tests, religion appears to be an important component of the definition of community. This is shown to be true in the CDFs for the Mossi in Mouhoun, though it is less apparent among the other two featured province-ethnic groups. The rural communities appear to behave similarly to the province as a whole. Further delineating the definition of community into the department rather than the province seems to be important according to the regression results, though the data limitations make drawing conclusions from the CDFs difficult.

Overall, it is clear that communities do behave differently than one another. In particular, some communities appear to have a tipping point in the rate of FGC (this is most clearly shown among the Mossi in Ioba), while other communities have a stable interior equilibrium in the rate of FGC (as is suggested by the CDFs for the Mossi in Mouhoun).

6.4 Limitations

I discussed the limitations of the research design in subsection 5.3. In addition to those limitations, there are some limitations of the data. First, I am able to observe only the girls who are 10 years of age or older who have a mother who is 49 year of age or younger at the time of survey. Thus, the external validity of my results should be tested by including younger women in the analysis.

Additionally, my analyses rely on direct reporting of the respondent's FGC status and daughter's FGC status. Obtaining physician records would lead to more precise estimates. I discussed the possible extent of measurement error in section 4. Finally, it is important to be aware of the differences between the context of Burkina Faso between 1949 and 2010 and other settings. The rate of FGC has declined substantially in Burkina Faso during the observation period, while other countries have not seen a similar decline.

7 Discussion and Conclusions

I contribute to the current debate in the literature regarding whether FGC is a social coordination norm (Powell, 2017). If FGC is a social coordination norm, as was posited by Mackie (1996), then the presence of a tipping point in the rate of FGC in a community is guaranteed. Recent empirical evidence question this assertion by showing that individual and household factors explain a substantially larger share of the persistence of FGC than do community-level factors (Bellemare, Novak, & Steinmetz, 2015) and that rates of FGC within a community are often between zero and one (Efferson et al., 2015). I propose a new theoretical explanation for why FGC persists. I show that heterogenous thresholds—where threshold is defined as the proportion of community members practicing FGC that makes a household indifferent between cutting and not cutting its daughter—among households make the presence of a tipping point far from guaranteed.

Using data from Burkina Faso that include women born between 1949 and 1995, I show that households have heterogeneous thresholds. Some communities in Burkina Faso do have a tipping point, and that tipping point has likely been reached in those communities. There are communities, however, that have a stable equilibrium in the rate of FGC in their community. This suggests that eliminating FGC from these communities may be extremely difficult.

The strength of my research design comes from the six-decade long timeframe provided by the three cross-sections of the Demographic and Health Survey data. These data allow me to observe the same community over an extended period of time, observe a household's revealed preference for FGC (the decision to cut its daughter), and observe variation in the constraints faced by households over time (changes in the rate of FGC). Additionally, I am able to define the community in multiple ways in order to determine if my results are sensitive to the definition of community and further narrow my definition of community based on those results.

That said, community is imperfectly defined in my analysis. Ideally, I would have complete social network data from these communities. Additionally, I rely on reported FGC status of the mother and daughter. Data from gynecological exams or other more objective measures of FGC would be preferable. Finally, Burkina Faso is only 1 of 29 countries in which FGC is practiced. The social norms surrounding FGC are heterogeneous between communities and certainly between countries. More research is needed to determine how these findings hold in other settings.

My findings show that households are able to deviate from the social norm and are willing to do so at different rates of FGC. This suggest that interventions that target village-level behavior may be inefficient compared to interventions that target households and individuals. Platteau, Camilotti, & Auriol (2017) demonstrate that the distribution of household preferences (what they refer to as an aversion coefficient) greatly influences the way in which a community will respond to laws or interventions aimed at decreasing the prevalence of FGC or other harmful norms. If a community has a tipping point at a high rate of FGC, the most efficient strategy is to target those most willing to deviate from the norm. Alternatively, if a community has a stable equilibrium, targeting those who are the least willing to abandon FGC could be the most efficient strategy because once those with a low threshold abandon FGC others are likely to follow.

More research is needed to identify the distribution of thresholds in communities in which FGC is practiced. Additionally, little research has been done that rigorously identifies the impact of policies aimed at reducing FGC. Camilotti (2015a,b) shows that one NGO's programs have had a small effect in the reduction of FGC and that laws that ban FGC have had an adverse effect—namely households are cutting their daughters at an earlier age. More studies are needed to show the impacts of interventions aimed at curbing FGC and how these interventions interact with the distribution of household thresholds.

8 Bibliography

Angrist, J.D., Pischke, J.S., 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

Baguet, M., Novak, L.K., 2016. "Change of Heart: Explaining Shifting Opinions of Female Genital Cutting in West Africa." *Working Paper.* University of Minnesota.

Banks, E., Meirik, O., Farley, T., Akande, O. 2006. "Female Genital Mutilation and Obstetric Outcome: WHO Collaborative Prospective Study in Six African Countries." *The Lancet* 367: 1835-1841.

Bellemare, M., Novak, L., Steinmetz, T., 2015. "All in the Family: Explaining the Persistence of Female Genital Cutting in West Africa." *Journal of Development Economics.* 116 (2015) 252-265.

Boyle, E.H., 2005. *Female Genital Cutting: Cultural Conflict in the Global Community.* Johns Hopkins University Press.

Boyle, E. H., Carbone-López, K., 2016. "Movement Frames and African Women's Explanations for Opposing Female Genital Cutting." *International Journal of Comparative Sociology.* 47 (6) 435-465.

Breusers, M., Nederlof, S., Van Rheenen, T. (1998). "Conflict or symbiosis? Disentangling farmer-herdsman relations: The Mossi and Fulbe of the Central Plateau, Burkina Faso." *The Journal of Modern African Studies*,36 (3), 357-380

Camilotti, G., 2015a. "Changing Female Genital Cutting: Evidence from Senegal." Working Paper, University of Namur.

Camilotti, G., 2015b. "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) 133-158.

Cloward, K., 2016. *When Norms Collide: Local Responses to Activism against Female Genital Mutilation and Early Marriage.* Oxford University Press.

Dorkenoo, E., 1999. "Combating Female Genital Mutilation: An Agenda for the Next Decade." *Women's Stud Q.* 27 (1/2), 87-97

Efferson, C., Vogt, S., Elhadi, A., Ahmed, H.E.F., Fehr, E., 2015. "Female Genital Cutting is Not a Social Coordination Norm." *Science.* 349 (6255), 1446-1447

Fan, J. and Gijbels, I., 1996. *Local Polynomial Modeling and its Applications: Monographs on Statistics and Applied Probability.* 66 (66). CRC Press.

Granovetter, M., 1978. "Threshold Models of Collective Behavior." *The American Journal of Sociology.* 83 (6), 142-1443

Gruenbaum, E., 2000. *The Female Circumcision Controversy: An Anthropological Perspective.* University of Pennsylvania Press; 1st edition.

Henry, S., Schoumaker, B., Beauchemin, C., 2004. "The Impact of Rainfall on the First Out-Migration: A Multi-level Event-History Analysis in Burkina Faso." *Population and the Environment* 25 (5), 423Ð460

Jones, H., Diop, N., Askew, I., Kabore, I., 1999. "Female Genital Cutting Practices in Burkina Faso and Mali and their Negative Health Outcomes." *Studies in Family Planning.* 30 (3), 219-230.

Hausman, J., 1978. "Specification Tests in Econometrics." *Econometrica.* 46 (6), 1251-1271.

Mackie, G., 1996. "Ending Footbinding and Infibulation: A Convention Account." *American Sociological Review.* 61 (6), 999-1017

Mackie, G., 2003. "Female Genital Cutting: A Harmless Practice?" *International Journal of the Analysis of Health.* 17(2), 135-158.

Manski, C., 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies.* 60 (3), 531-542.

Modrek, S., Liu, J.X., 2013. "Exploration of Pathways Related to the Decline in Female Circumcision in Egypt." *BMC Public Health.* 13:921.

Obermeyer, C.M., 1999. "Female Genital Surgeries: The Known, the Unknown, and the Unknowable." *International Journal for the Analysis of Health.* 13 (1). 79-106.

Pei, Z., Pischke, J. S., Schwandt, H., 2017. "Poorly Measured Confounders are More Useful on the Left Than on the Right." *National Bureau of Economic Research Working Paper Series.*

Platteau, J. P., Camilotti, G., Auriol, E. 2017. "Eradicating Women-Hurting Customs: What Role for Social Engineering?" *Gender and Development.* Oxford University Press.

Powell, K., 2017. "A Nagging Persistence." *Nature Human Behaviour.* 1 (26)

Schelling, T., 1978. *Micromotives and Macrobehavior.* W. W. Norton & Company; 1st Edition.

Shell-Duncan, B., Hernlund, Y., 2001. *Female "Circumcision" in Africa: Culture, Controversy, and Change.* Lynne Rienner Publishers; 1 edition.

Toubia, N.F., Sharief, E.H., 2003. "Female Genital Mutilation: Have We Made Progress?" *International Journal of Gynecology & Obstetrics.* 82 (3), 251-261.

UNICEF, 2013. Female Genital Mutilation/Cutting: A Statistical Overview and Exploration of the Dynamics of Change. http://data.unicef.org/corecode/uploads/document6/uploaded_pdfs/corecode/FGMC_Lo_res_Final_26.pdf (Last accessed June 15, 2016)

US Department of Health and Human Services, 2009. Female Genital Cutting. http://womenshealth.gov/publications/our-publications/fact-sheet/female-genital-cutting.html. (Last accessed May 27, 2016).

Wagner, N., 2014. "Why Female Genital Cutting Persists." *Journal of Development Studies*. 51 (3), 226-246

Williams, L., Sobieszczy, T., 1997. "Attitudes Surrounding the Continuation of Female Circumcision in the Sudan: Passing the Tradition to the Next Generation." *J. Marriage Fam.* 59 (4), 966-981.

Wooldridge, J.M., 2002. *Econometrics of Cross-Section and Panel Data.* MIT Press, Cambridge, MA.

World Health Organization, 2012. Female Genital Mutilation, Fact Sheet #241. http://www.who.int/mediacentre/factsheets/fs241/en/ (Last accessed November 4, 2015).

Table A1: LPM Estimation Results for Whether Girl Undergoes FGC—Squared Term for Rate of FGC in Community Included

	(1) Province-Ethnic Cohort	(2) Department-Ethnic Cohort	(3) Province-Ethnic Rural Cohort	(4) Province-Ethnic Catholic Cohort	(5) Province-Ethnic Muslim Cohort
Rate of FGC in Community	0.286*	0.311**	0.281	0.395*	0.083
	(0.147)	(0.140)	(0.184)	(0.200)	(0.250)
Rate of FGC in Community^2	0.264*	0.229*	0.274	0.039	0.388*
	(0.145)	(0.124)	(0.182)	(0.200)	(0.210)
Urban Household	-0.036*	-0.034		0.030	-0.055**
	(0.020)	(0.021)		(0.039)	(0.026)
Religion [Omitted: Catholic]					
Protestant	-0.035	-0.048	-0.019		
	(0.030)	(0.041)	(0.045)		
Muslim	0.075***	0.076***	0.096***		
	(0.015)	(0.017)	(0.021)		
Other	0.062**	0.045	0.060		
	(0.029)	(0.035)	(0.038)		
Constant	0.567***	0.640***	0.681***	0.699*	1.478***
	(0.164)	(0.183)	(0.203)	(0.362)	(0.211)
Observations	5,934	4,523	3,493	1,079	2,818
R-squared	0.215	0.238	0.229	0.297	0.214

Controls included: Ethnicity, religion, proxies for wealth, rural indicator, polygamy indicator, mother's year of birth FE, father's year of birth FE daughter's year of birth FE

Standard Errors Clustered at the Province Level

***p<0.01 **p<0.05 *p<0.1

Table A2: Logit Estimation Results for Whether Girl Undergoes FGC—Marginal Effects Evaluated at the Mean

	(1) Province-Ethnic Cohort	(2) Department-Ethnic Cohort	(3) Province-Ethnic Rural Cohort	(4) Province-Ethnic Catholic Cohort	(5) Province-Ethnic Muslim Cohort
Rate of FGC in Community	0.767***	0.758***	0.782***	0.550***	0.699***
	(0.062)	(0.042)	(0.066)	(0.099)	(0.133)
Observations	5908	4501	3459	1035	2790
Pseudo R-squared	0.170	0.192	0.177	0.135	0.191

Controls included: Ethnicity, religion, proxies for wealth, rural indicator, polygamy indicator, mother's year of birth FE, father's year of birth FE daughter's year of birth FE

Standard Errors Clustered at the Province Level

***p<0.01 **p<0.05 *p<0.1