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

Lindsey Novak*

Abstract

This paper develops a new data-supported theory for why the practice of female genital cutting (FGC) persists and tests this theory with a dataset of 24,000 women born across six decades in Burkina Faso. I find that households within a community have heterogeneous preferences for FGC and that this heterogeneity makes the existence of a tipping point in the rate of FGC far from guaranteed. These findings are contrary to the prevailing theory in the FGC literature and provide guidance for which policies can change the practice. Additionally, this theory has implications for the evolution of other social phenomena.

JEL Classification Codes: D19, I15, J16, O15, Z13

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

*Assistant Professor, Department of Economics, Colby College, 5237 Mayflower Hill, Waterville, Maine 04901, `lindsey.novak@colby.edu`.

1 Introduction

Female genital cutting is practiced in parts of Africa, the Middle East, and Asia, and each year more than three million girls undergo the practice (World Health Organization, 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, 2015). 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.

Why does FGC persists when there are no medical benefits and there are many potential physical and psychological costs? I propose a new theoretical explanation for this phenomenon that allows for agents to have heterogenous preferences for practicing FGC and that allows for preference interaction among community members (Manski, 2000; Bikhchandani et al., 1998; Young, 2015). I draw on Schelling (1978)'s theoretical social interaction models to investigate why FGC persists in some communities but has subsided in others. Using observational data from Burkina Faso's Demographic and Health Survey I investigate whether there is a tipping point or a stable internal equilibrium in the proportion of community members abandoning FGC in communities in Burkina Faso. The strength of my approach comes from the long time period I analyze using three cross-sectional datasets from 1998, 2003, and 2010 that include women and their daughters born who were between 1949 and 2010. The use of this six-decade timeframe is important when studying an intergenerational problem and the

[1] Female genital cutting (FGC) is also referred to as "female circumcision" and "female genital mutilation." I use the terminology FGC because it makes a larger 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.

long-term dynamics of FGC as well as increasing the external validity of my estimates. 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).

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 difficult 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), then a tipping point in the rate of FGC is guaranteed to exist in a community, and deviating from the social norm is so socially costly that individuals or households will not do so on their own. 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 et al., 2015).

The distribution of household preferences has a large influence on the effectiveness of policies aimed at reducing the incidence of FGC (Platteau et al., 2017). 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. The findings in this paper suggests that development agencies should broaden the set of policy interventions in order to address the support for and the perpetuation of FGC. Specifically, agencies should design policies aimed to change individual- or household-level preferences.

The idea that heterogenous preferences among community members may lead to multiple equilibria is certainly not new in economics, though it is a relatively new idea that is gaining traction in the study of FGC. Efferson et al. (2015) show theoretically that one possible explanation for the lack of a tipping point in the rate of FGC in communities in Sudan is that households have heterogeneous preferences. Platteau et al. (2017) use a theoretical model to show how a law banning FGC may affect communities differently based on the distribution of preferences for FGC among community members. To date, no study has empirically examined how community behavior toward FGC is based on the distribution of household preferences.

Outside of the study of FGC there has been a recent resurgence of interest in the theoretical work in Schelling (1978) (Ioannides, 2017; Bayer and Timmins, 2007). This resurgence has particularly focused on the study of racial segregation in neighborhoods both using simulations (Zhang, 2011) and using data from the United States (Card et al., 2008, 2011; Ananat, 2011). Card et al. (2008) show that there is a tipping point the racial makeup of neighborhoods in the United States. No study has empirically found multiple equilibria in a Schelling (1978) social interaction model. This paper fills this gap in the literature.

The contribution of this paper 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 a stable internal equilibrium.

Third, this paper is the first to empirically show community level outcomes as predicted by Schelling (1978) in a setting in which a tipping point does not always exist. I show that some communities in my data have a tipping point, while others have a stable equilibrium at low levels of FGC. This suggests that achieving a low rate of FGC may be feasible; completely eliminating FGC, however, may be more difficult.

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 presents and interprets the results of the analysis. In section 7 I discuss 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 the United States, Britian, and other developed countries (Black and Debelle, 1995; Jones et al., 1997; Newsweek, 2017; National Public Radio, 2004). 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. Less than four percent of girls who undergo FGC experience infibulation.[2]

[2] This estimate from the Demographic and Health Survey data used in this paper 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. 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). Even today, some women are cut because of beliefs that FGC will discourage infidelity (Shell-Duncan and Hernlund, 2001). Though FGC arose and may persist in part in response to male preferences, today women are typically the primary decision-makers when it comes to cutting their daughters (Mackie, 1996; Toubia and Sharief, 2003). Undergoing FGC is seen as a rite of passage, or a way to join the society of women in the community (Toubia and Sharief, 2003). Others see FGC as a sacred norm that is important for social cohesion (Tetlock et al., 2017). 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 and 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 (Molloy, 2013).

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 within their community. 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

To analyze a household's decision of whether its daughter should undergo FGC, I present a formalized preference interaction model that shows how an agent's preference ordering of the two alternatives—cut or not cut a girl—depends on the behaviors of others in the community. This is an adapted version of the social coordination norm model proposed by Mackie (1996). I then add to this model Schelling (1978)'s model of social interactions which shows how macro-level outcomes in a community depend on the distribution of individual preferences.

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 to observe the final decision made by the household (i.e., the girl's FGC status) and the constraints faced by the household, but I am not able to observe 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 know who participated in such a ceremony (Cloward, 2016). Additionally, assume that households in the sample are norm-takers such that no one household can change the norm.

Assumption 1. *Households are able to observe the rate of FGC in their community.*

Assumption 2. *Households in the sample are norm-takers.*

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) not to 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 social cost accruing to a household that does not cut its daughter in a community in which the cutting rate equals zero, that is $s_{0i}(0) = 0$. Assume also that $s_{1i}(1) = 0$ because there is no social cost of practicing FGC in a community in which all members practice FGC. Then, let $s_{0i}(r) \geq 0$ and $s_{1i}(r) \geq 0$ be differentiable functions. These social costs include reduced marriage prospects of the girl (Wagner, 2015) and reduced acceptance into the community (Toubia and Sharief, 2003). Additionally, $s_{0i}(r)$ may include perceived disregard for religious edicts (Cloward, 2016) and inferior perceived beauty and femininity (Shell-Duncan and Hernlund, 2001). Finally, let $c_i \geq 0$ be the perceived non-social costs—monetary, psychological, and physical—associated with cutting the girl. Then the payoff accruing to a household that abstains from cutting its daughter is $-s_{0i}(r)$, and the payoff accruing to a household that cuts is daughter is $-s_{1i}(r) - c_i$.

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

Assumption 3. $\frac{\partial s_{0i}(r)}{\partial r} > 0$, *i.e. the social cost accruing to a household that abstains from practicing FGC is monotonically increasing in the proportion of community members that practices FGC.*

Assumption 4. $\frac{\partial s_{1i}(r)}{\partial r} < 0$, *i.e. the social cost accruing to a household that practices FGC is monotonically decreasing in the proportion of community members that practices FGC.*

It is not necessary to assume a functional form for the social cost functions. It is, however, important to consider the number of times the payoff functions for the two strategies could cross. Assumptions 3 and 4 allow us to conclude that the payoff functions can cross at most once in the space of FGC rates. This single-crossing guarantees that there is exactly one r at which a household is indifferent between cutting and not cutting its daughter.

[Figure I about here]

Lemma 1. *Under assumptions 3 and 4, the potential payoffs accruing to households under each strategy, $g \in \{0, 1\}$, cross at most once in the space of community FGC rates.*

Proof. Let $h(r) = -s_{0i}(r) + [s_{1i}(r) + c_i]$ for $0 \leq r \leq 1$.

First, notice that $h(r)$ is equal to the difference between the payoff associated with not cutting the girl and the payoff associated with cutting the girl. That is, $-s_{0i}(r) - [-s_{1i}(r) - c_i] = -s_{0i}(r) + [s_{1i}(r) + c_i] = h(r)$. Then, observe that to prove Lemma 1 it is sufficient to show that there is at most one point $a \in [0, 1]$ such that $h(a) = 0$.

Since $s_{0i}(r)$ is monotonically increasing in the space $r \in [0, 1]$, $-s_{0i}(r)$ is monotonically decreasing. Further, since $s_{1i}(r)$ is monotonically decreasing in the space $r \in [0, 1]$, then $s_{1i}(r)+c_i$ is also monotonically decreasing in $r \in [0, 1]$. Therefore $h(r) = -s_{0i}(r)+[s_{1i}(r)+c_i]$ is monotonically decreasing as it is the sum of two monotonically decreasing functions. Since monotonically decreasing functions are injective (that is, if $h(x) = h(y)$, then $x = y$), there can be at most one $a \in [0, 1]$ for which $h(a) = 0$. To see this, imagine that for some other a_0, $h(a_0) = 0$, then because h is injective, $a_0 = a$. □

Figure I shows this graphically by displaying the potential costs accruing to a household if it practices or abstains from FGC, depending on the proportion of community members practicing FGC. Let $s_{0i}(1) = \hat{s}_{0i}$ and $s_{1i}(0) = \hat{s}_{1i}$. 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 cost structure that is a linear function of the rate of FGC in the community.

Figure I 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^*.

3.2 The Threshold Model of Collective Behavior

Mackie (1996) implicitly 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 et al. (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 single 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.

If instead each household values the benefits and costs to cutting differently, there are as many variations to figure I as there are households. This implies that each household could 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 practice FGC in order to induce the household to decide to cut its daughter. In figure I, 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 social interaction phenomena, 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 its 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 II. 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 households 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 households that will abstain from practicing FGC even if they expect every other household in the community will practice FGC.

[Figure II about here]

Figure III shows hypothetical CDFs of the proportion of community households 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. Any point at which the CDF crosses the 45-degree line, or $F(r) = r$, is an equilibrium—either stable or unstable. The equilibrium is stable if the CDF approaches the 45-degree line from above, while the equilibrium is unstable if the CDF approaches the 45-degree line from below.

[Figure III about here]

Curve 1 of figure III 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 distributed as shown in curve 2 has two stable equilibria—one at a high rate of FGC (point C) and one at zero, and one unstable equilibrium at point B.

Thus, if households have heterogeneous thresholds 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. 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 intrahousehold bargaining power. Additionally, each individual may have difficulty articulating their individual 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. One additional assumption is needed. That is, a household uses the rate of FGC among girls in the previous cohort (born in the year(s) prior to their daughter) as a best approximation for the rate r upon which they base their decision to cut the girl or not.

Assumption 5. *Household i uses the rate of FGC among girls born in the community in the year(s) prior to girl i as a basis for predicting the rate of FGC in girl i's cohort.*

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

Proposition 1. $F(r) = \pi(r)$.

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 $\mathbb{P}(g_i = 1|r) = \mathbb{P}(r_i^* < r)$. And, $\mathbb{P}(r_i^* < r) = \int_0^r f(r_i^*)dr_i^* = F(r)$. Thus,

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

□

Therefore, 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[3] 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 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

[3] 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.

Zoundwéog) who were between 15 and 55 years of age 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. The difference between the reported rate of FGC and the rate of FGC estimated in Jones et al. (1999) may be even smaller than this four percentage point difference would suggest. The data used in Jones et al. (1999) include women who were as old as 55 years in 1998 while the DHS data include women up to the age of 49 in 1998. The older cohorts in the DHS data have the highest reported rates of FGC, and the rate of FGC among women born earlier are likely similar or higher. Further, the data used by Jones et al. (1999) 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 (Wagner, 2015; Jones et al., 1999; Banks et al., 2006). While I cannot rule out the possibility of reporting bias, I argue that misreporting is not a large concern in this setting.

Table I shows descriptive statistics for the two 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.

[Table I about here]

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. 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 respondent's daughters.[4] The women who have daughters (column 2) 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 mature daughters, and the rate of FGC has been decreasing over time in Burkina Faso while the rate of education has been rising. The rate of FGC among the respondents' daughters is just over 38 percent. This sharp difference in the rate of FGC between mothers and daughters is consistent with the overall trend in the rate of FGC in Burkina Faso.

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

[Figure IV about here]

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

[4] Including only the woman's oldest daughter in the analysis has the benefit of mitigating the risk of measuring a household's reluctance to make different decisions for different daughters.

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 or fewer prior to the birth of the daughter in question. I use the information on girls born within the previous five years 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 (see assumptions 1 and 5). This definition of cohort has the added benefit of minimizing the reflection problem (Manski, 1993).[5]

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 is uncommon (Breusers et al., 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, though women most often marry within their natal province (Henry et al., 2004). Thus I do not use village as the definition of community but instead use the respondent's province.

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 very likely to be aware of the FGC status of other girls in the community, this

[5]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.

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 a 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, explain how it compares to the ideal experiment, and I discuss threats to the validity of my estimates. I begin by testing the hypothesis that 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 Estimation Strategy

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 that practiced FGC in the previous cohort (where community is defined in the four ways specified above). As stated in proposition 1 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 an Epanechnikov 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 cohort c, where the cohort is all girls born between years $t-6$ and $t-1$. I use an an Epanechnikov kernel with a polynomial smooth of degree three 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 II, 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 III. If the CDF crosses the 45-degree line from above, this suggests that there is a stable equilibrium in the rate of FGC in the community. If the CDF crosses the 45-degree line from below, the community has an unstable equilibrium, or a tipping point.

5.2 Coefficient Comparisions

I estimate the likelihood that a girl will undergo FGC based on the rate of FGC in her cohort within her community—where community is defined in four different ways. I then determine if different definitions of community are statistically different from one another. I estimate equation 2 and then run a series of Hausmann tests to compare the resulting β_1 coefficients.

$$y_{itk} = \beta_0 + \beta_1 r_{ck} + \epsilon_{itk} \tag{2}$$

Let y_{itk} equal one if girl i born in year t and community k has undergone FGC and zero if she has not undergone FGC. I estimate equation 2 using ordinary least squares (OLS) regression. Because y_{itk} is binary, 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 as well as adding squared and cubed terms for r_{ck} to the LPMs (see tables A1, A2, and A3 in the appendix).

Following Pei et al. (2016), I then perform a Hausman (1978) specification test to determine whether the coefficients produced using the various definitions of community are statistically similar.

5.3 Ideal Dataset and Threats to Validity

In order to appreciate the threats to validity of my approach, it is useful to imagine the ideal dataset for addressing this question and understand the obstacles to obtaining such a dataset. Assume first that in each community there is a set age 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 ask the household to state its 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 individual preference for FGC. Thus, eliciting a household's threshold is unlikely to reveal reliable results.

A more reliable approach to eliciting a household's threshold involves devising a method to observe a household's revealed preference for FGC. 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 that took place in some communities and not in others). One could then observe the rate of FGC among group 1 in all 100 communities. From this information, a researcher 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 researcher could give false information to each household regarding the rate of FGC in the community. The researcher could then observe the choice 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 researcher each year, the researcher could create bounds on each household's threshold. Finding a community where all households have a sufficient number of girl children is unlikely. Further, it is unlikely that individuals would believe the researcher who is providing false information, and such a research project is unlikely to be approved by an Institutional Review Board.

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 because I observe that rate of FGC over a long time period. 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, may have experienced a legal ban on 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.

Fourth, I use the rate of FGC in the previous cohort as the household's expected rate of FGC among its daughter's cohort. In fact, a household is forming a belief regarding how many community members will cut their daughters. It is difficult to precisely predict how households will respond to such uncertainty (Li et al., 2017). The rate of FGC in the previous cohort is an important piece of information that a household considers, but it is not the sole piece of relevant information that the household takes into account.

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. Finally, I examine the various definitions of community to determine which are statistically equivalent and which are different.

6.1 Thresholds and Tipping Point

Figure V 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 II, and in fact looks very similar to curve 2 in figure II.

[Figure V about here]

Because the CDF crosses the 45-degree line from above, figure V 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 without a stimulus to change household thresholds. 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 VI, VII, and VIII 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,[6] 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.

[Figure VI about here]

The CDFs for the Mossi in Mouhoun (figure VI) suggest that a stable equilibrium at a low-rate of FGC exists in 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.

[Figure VII about here.]

The CDFs for the Senoufo in Ganzourgo (figure VII) 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 in the urban areas. The observed department (of which 78 percent of surveyed households lie in an urban area) behaves 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 VIII) 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

[6] 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.

79 percent of surveyed households in this department lie in an urban area). The Muslim households behave similarly to the broader province-ethnic group.

[Figure VIII about here.]

6.2 Likelihood That a Girl Undergoes FGC

Table II presents the results of estimating equation 2 using five definitions of the girl's community-cohort. Cohort is defined as all girls born in the five years prior to the year of birth of the girl in question. Column 1 defines the girl's community-cohort as all girls who live in the same province and are part of the same ethnic group, column 2 defines the girl's community-cohort as all girls who live in the same department and are part of the same ethnic group. Column 3 estimates equation 2 defines the girl's community-cohort as all girls who live in the province, are part of the same ethnic group, 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 ethnic group, are part of the same religious group.

[Table II about here.]

Table II 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 0.687 and 0.899 suggesting that a girl living in a community in which $r = 1$ is between 68.7 and 89.9 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 6.9 and 9.0 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 et al. (2016) I perform a Hausman (1978) specification test to determine whether these coefficients are statistically similar. Table III reports the Hausman test statistics. The null hypothesis of the Hausman test is that the difference between two coefficients is statistically zero. 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. This suggests that using province-ethnic group as the definition of community is a good approximation for those living in rural areas and Muslims, though it may misrepresent other community members.

[Table III about here.]

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. Though, defining the community as the department-ethnic group seems to be inferior to defining the community as the province-ethnic group.

Overall, it is clear that communities do behave differently from each other. 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, my analyses rely on direct reporting of the respondent's FGC status and daughter's FGC status. Obtaining physician records would lead to more accurate estimates. I discussed the possible extent of measurement error in section 4, and measurement error does not appear to be a large concern in this sample. Second, 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. Thus, the behavior in other countries in which FGC is practiced could be different from what is found in this study.

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 questions 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 et al., 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. 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 whether 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 households. 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 suggests that interventions that target village-level behavior may be inefficient compared to interventions that target households and individuals. Platteau et al. (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.

References

Ananat, E. O. (2011). The wrong side(s) of the tracks: The causal effects of racial segregation on urban poverty and inequality. *American Economic Journal: Applied Economics 3*(2), 34–66.

Banks, E., O. Meirik, T. Farley, O. Akande, et al. (2006). Female genital mutilation and obstetric outcome: Who collaborative prospective study in six African countries. *The Lancet 367*(9525), 1835.

Bayer, P. and C. Timmins (2007). Estimating equilibrium models of sorting across locations. *The Economic Journal 117*(518), 353–374.

Bellemare, M., L. Novak, and T. Steinmetz (2015). All in the family: Explaining the persistence of female genital cutting in West Africa. *Journal of Development Economics 116*, 252–265.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *The Journal of Economic Perspectives 12*(3), 151–170.

Black, J. and G. Debelle (1995). Female genital mutilation in Britain. *BMJ: British Medical Journal 310*(6994), 1590.

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

Breusers, M., S. Nederlof, and T. Van Rheenen (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*(03), 357–380.

Camilotti, G. (2015a). Changing female genital cutting: Evidence from Senegal. Unpublished Paper.

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.

Card, D., A. Mas, and J. Rothstein (2008). Tipping and the dynamics of segregation. *The Quarterly Journal of Economics 123*(1), 177–218.

Card, D., A. Mas, and J. Rothstein (2011). Are mixed neighborhoods always unstable? Two-sided and one-sided tipping. In *Neighborhood and Life Chances: How Place Matters in Modern America*, pp. 237–256. University of Pennsylvania Press.

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. *Womens Studies Quarterly 27*(12), 87–97.

Efferson, C., S. Vogt, A. Elhadi, H. Ahmed, and E. Fehr (2015). Female genital cutting is not a social coordination norm. *Science 6255*(349), 1446–1447.

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

Gruenbaum, E. (2000). *The Female Circumcision Controversy: An Anthropological Perspective.* University of Pennsylvania Press.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica 46*(6), 1251–1271.

Henry, S., B. Schoumaker, and C. Beauchemin (2004). The impact of rainfall on the first out-migration: A multi-level event-history analysis in Burkina Faso. *Population and Environment 25*(5), 423–460.

Ioannides, Y. M. (2017). Review of climbing Mount Laurel: The struggle for affordable housing and social mobility in an American suburb by Douglas S. Massey et al. *Journal of Economic Literature 55*(2), 609–620.

Jones, H., N. Diop, I. Askew, and I. Kabore (1999). Female genital cutting practices in Burkina Faso and Mali and their negative health outcomes. *Studies in Family Planning 3*(30), 219–230.

Jones, W. K., J. Smith, B. Kieke Jr, and L. Wilcox (1997). Female genital mutilation. Female circumcision. Who is at risk in the US? *Public Health Reports 112*(5), 368.

Li, Z., G. Loomes, and G. Pogrebna (2017). Attitudes to uncertainty in a strategic setting. *The Economic Journal 127*(601), 809–826.

Mackie, G. (1996). Ending footbinding and infibulation: A convention account. *American Sociological Review 61*(6), 999–1017.

Manski, C. F. (2000, September). Economic analysis of social interactions. *Journal of Economic Perspectives 14*(3), 115–136.

Molloy, A. (2013). *However long the night: Molly Melching's Journey to help millions of African women and girls triumph.* HarperOne.

National Public Radio (2004). Atlanta female circumcision case stirs concerns. `http://www.npr.org/templates/story/story.php?storyId=1763682`. (last accessed October 22, 2017).

Newsweek (2017). Nearly 100 girls may have had genitals cut by doctors in Michigan, prosecutor says. `http://www.newsweek.com/genital-mutilation-girls-michigan-doctor-623295`. (last accessed October 22, 2017).

Pei, Z., J.-S. Pischke, and H. Schwandt (2016). Poorly measured confounders are more useful on the left than on the right. Technical report, National Bureau of Economic Research.

Platteau, J., E. Auriol, and G. Camilotti (2017). Eradicating women-hurting customs: What role for social engineering. In *Gender and Development*. Oxford University Press.

Powell, K. (2017). A nagging persistence. *Nature Human Behaviour 1*, 0026.

Schelling, T. C. (1978). *Micromotives and Macrobehavior*. New York: WW Norton & Co.

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

Tetlock, P. E., B. A. Mellers, and J. P. Scoblic (2017). Sacred versus pseudo-sacred values: How people cope with taboo trade-offs. *American Economic Review 107*(5), 96–99.

Toubia, N. and E. Sharief (2003). Female genital mutilation: Have we made progress? *International Journal of Gynecology and Obstetrics 3*(82), 251–261.

UNICEF (2013). Female genital mutilation/cutting: A statistical overview and exploration of the dynamics of change. United Nations Children's Fund, New York.

Wagner, N. (2015). Female genital cutting and long-term health consequences–Nationally representative estimates across 13 countries. *The Journal of Development Studies 51*(3), 226–246.

World Health Organization (2012). Female genital mutilation, Fact sheet #241. `http://www.who.int/mediacentre/factsheets/fs241/en/`. (last accessed October 22, 2017).

Young, H. P. (2015). The evolution of social norms. *The Annual Review of Economics 7*, 359–387.

Zhang, J. (2011). Tipping and residential segregation: a unified Schelling model. *Journal of Regional Science 51*(1), 167–193.

Table I: Descriptive Statistics

	(1) Full Sample	(2) Respondents with Daughters
Respondent Underwent FGC	0.749	0.863
	(0.003)	(0.004)
Daughter Underwent FGC		0.383
		(0.006)
Respondent Ever Attended School	0.280	0.104
	(0.003)	(0.004)
Respondent Ever Married	0.755	1.000
	(0.003)	(0.000)
Respondent's Partner Educated		0.125
		(0.004)
Average Rate of FGC in Province Cohort	0.745	0.633
	(0.001)	(0.002)
Average Rate of FGC in Province-Ethnic Cohort	0.744	0.637
	(0.001)	(0.003)
Respondent in Polygamous Marriage		0.568
		(0.006)
Respondent Year of Birth	1981	1967
	(0.045)	(0.086)
Respondent's Partner's Year of Birth		1955
		(0.137)
Daughter's Year of Birth		1988
		(0.080)
Religion: Catholic	0.246	0.240
	(0.003)	(0.005)
Protestant	0.062	0.054
	(0.002)	(0.003)
Muslim	0.592	0.580
	(0.003)	(0.006)
Other	0.100	0.126
	(0.002)	(0.004)
Household Owns TV	0.189	0.117
	(0.003)	(0.004)
Household Owns Radio	0.733	0.702
	(0.003)	(0.006)
Ethnic Group: Bobo	0.037	0.032

Continued on next page

Table I (continued)

	(1) Full Sample	(2) Respondents with Daughters
	(0.001)	(0.002)
Dioula	0.026	0.028
	(0.001)	(0.002)
FulFulde/Peul	0.067	0.057
	(0.002)	(0.003)
Gourmatche	0.062	0.062
	(0.002)	(0.003)
Gourounsi	0.047	0.042
	(0.001)	(0.002)
Lobi	0.045	0.029
	(0.001)	(0.002)
Mossi	0.542	0.588
	(0.003)	(0.006)
Senoufo	0.049	0.049
	(0.001)	(0.003)
Touareg/Bella	0.012	0.011
	(0.001)	(0.001)
Bissa	0.040	0.041
	(0.001)	(0.002)
Other	0.073	0.059
	(0.002)	(0.003)
Household is in Urban Area	0.307	0.206
	(0.003)	(0.005)
Observations	24,474	6,873

Standard deviations in parentheses

Table II: 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	0.851***	0.814***	0.851***	0.687***	0.899***
	(0.050)	(0.043)	(0.045)	(0.058)	(0.061)
Constant	-0.142***	-0.127***	-0.139***	-0.061	-0.187***
	(0.037)	(0.031)	(0.032)	(0.037)	(0.049)
Observations	8,566	6,643	5,146	1,679	3,776
R-squared	0.148	0.167	0.146	0.152	0.115

Unconditional linear probability model. Standard Errors Clustered at the Province Level.
***p<0.01 **p<0.05 *p<0.1

Table III: 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	12.41			
	(0.000)			
Province-Ethnic Rural Cohort	0.32	22.43		
	(0.5724)	(0.000)		
Province-Ethnic Catholic Cohort	24.50	3.27	18.36	
	(0.000)	(0.071)	(0.000)	
Province-Ethnic Muslim Cohort	1.79	17.07	2.22	20.65
	(0.181)	(0.000)	(0.136)	(0.000)

χ^2 value reported with p-value in parentheses

Figure I: Household $i's$ Payoff Structure

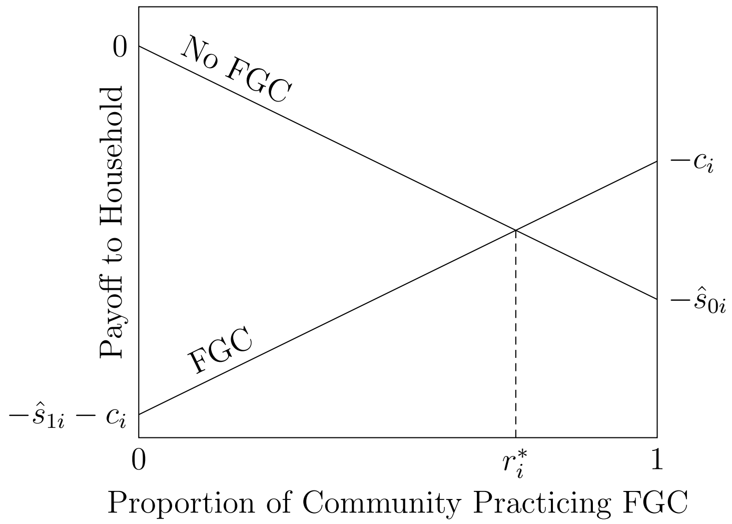


Figure II: Hypothetical Community CDFs

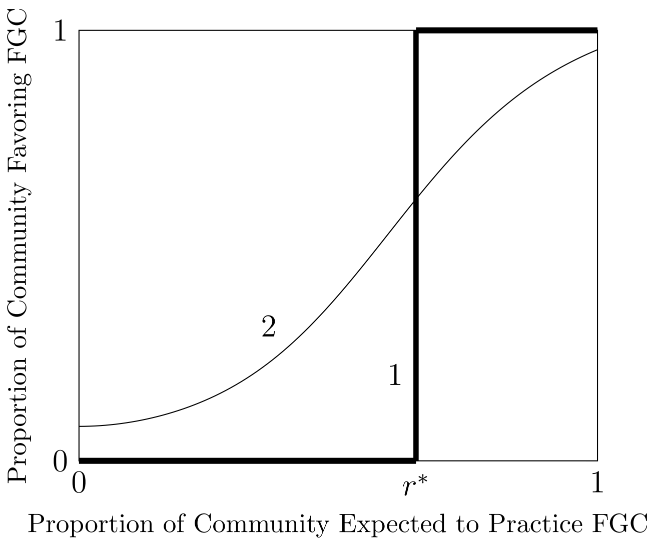


Figure III: Hypothetical Community CDFs

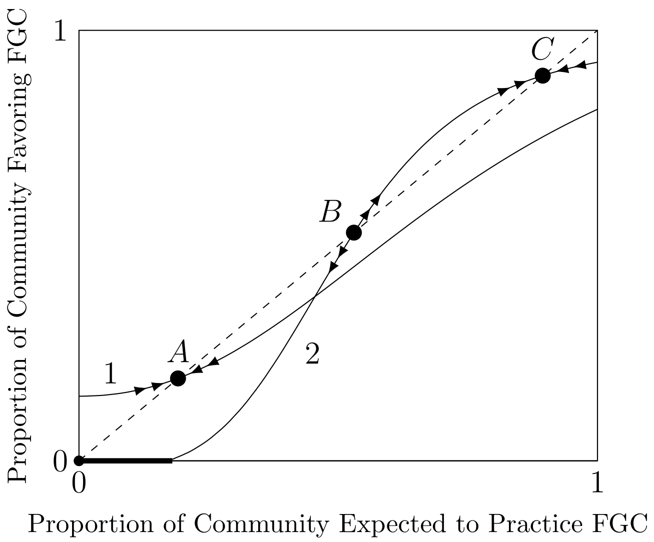


Figure IV: Rate of FGC in Burkina Faso by Year of Birth

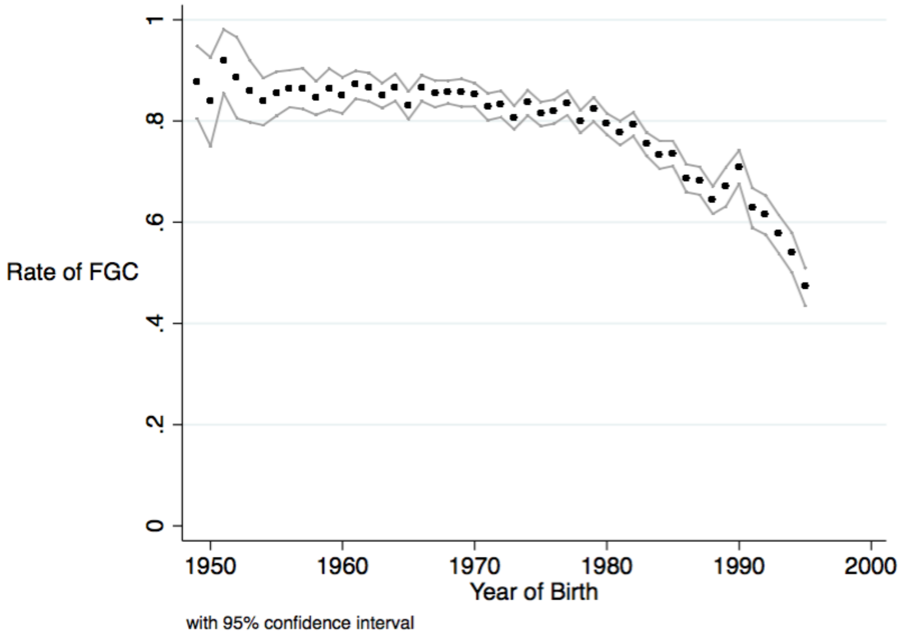


Figure V: CDF for Burkina Faso

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

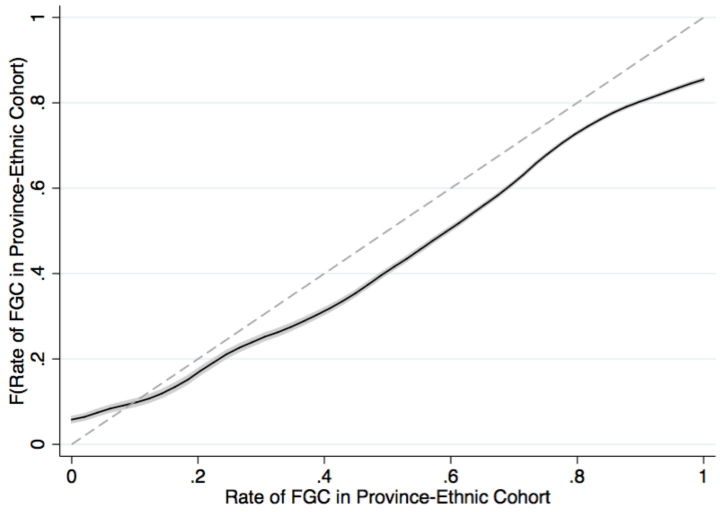
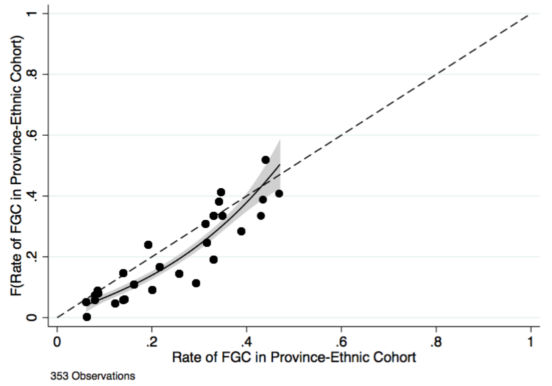
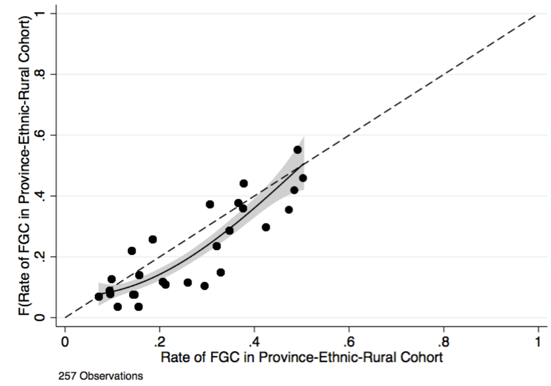


Figure VI: CDFs from the Mossi in Mouhoun

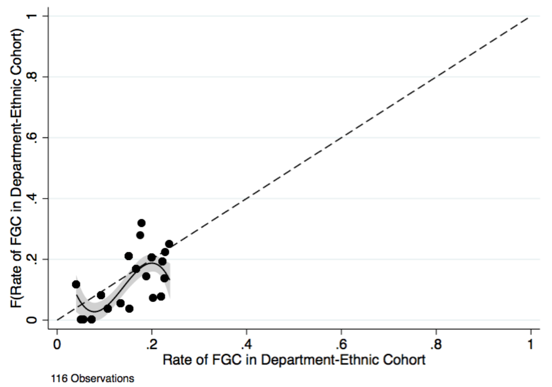
Local polynomial regressions using Epanechnikov kernel
with bandwidth 0.4, with 99% confidence interval



(a) All Province-Ethnic Group

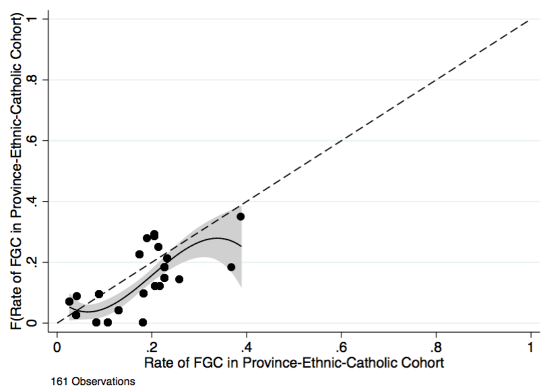


(b) Rural Province-Ethnic Group

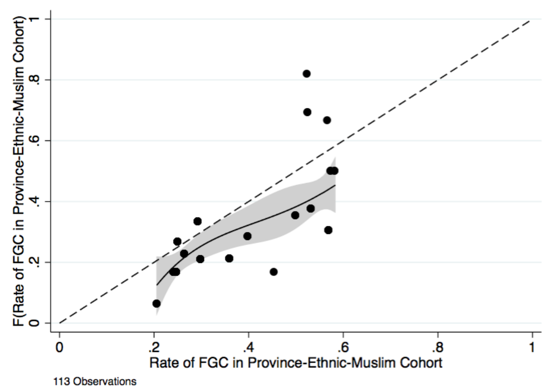


(c) Department 1-Ethnic Group

(d)



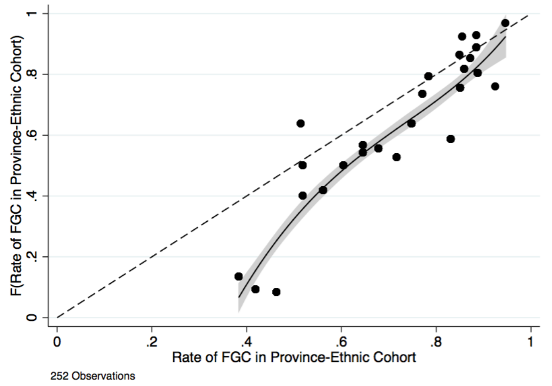
(e) Catholic-Province-Ethnic Group



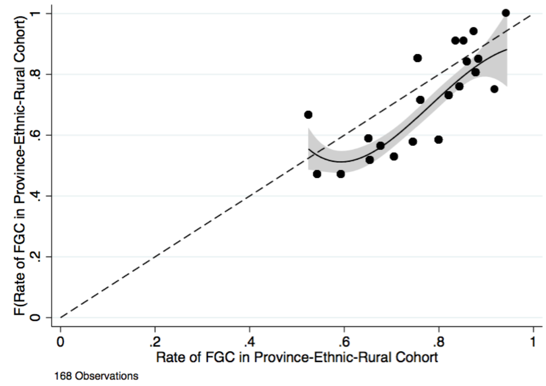
(f) Muslim-Province-Ethnic Group

Figure VII: CDFs from the Senoufo in Ganzourgo

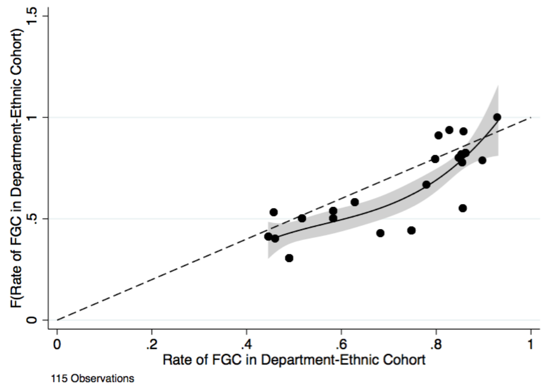
Local polynomial regressions using Epanechnikov kernel
with bandwidth 0.4, with 99% confidence interval



(a) All Province-Ethnic Group



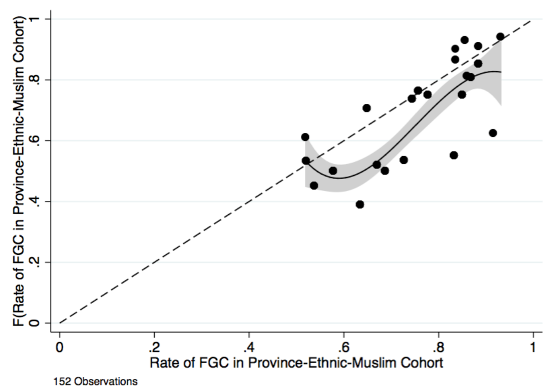
(b) Rural Province-Ethnic Group



(c) Department 1-Ethnic Group

(d)

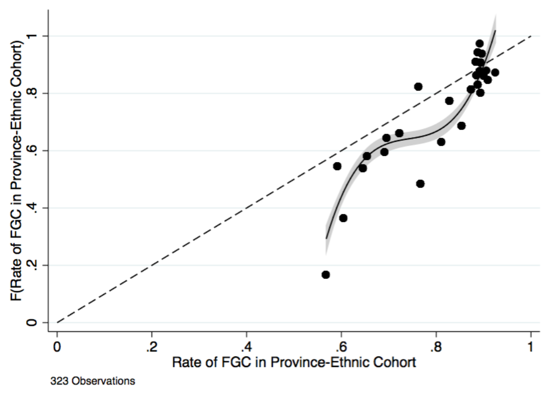
(e)



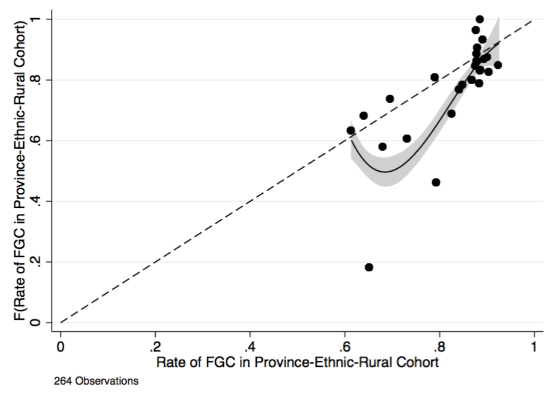
(f) Muslim-Province-Ethnic Group

Figure VIII: CDFs from the Mossi in Ioba

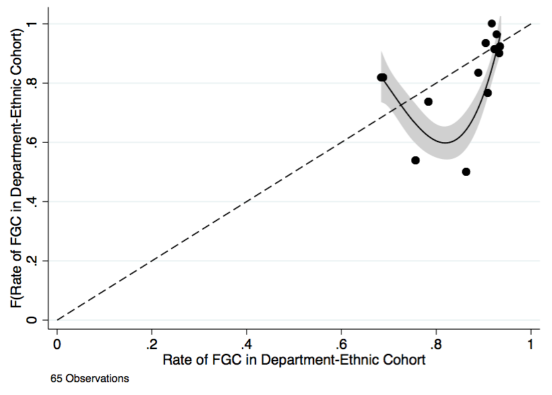
Local polynomial regressions using Epanechnikov kernel
with bandwidth 0.4, with 99% confidence interval



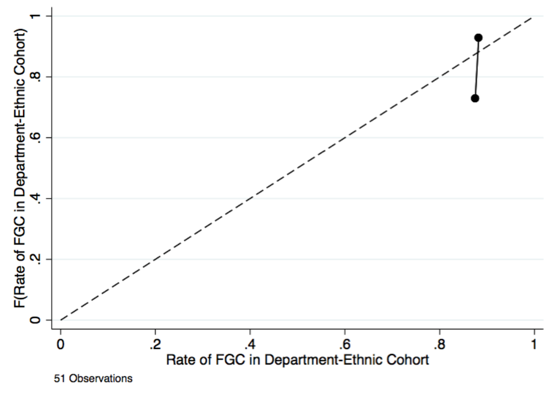
(a) All Province-Ethnic Group



(b) Rural Province-Ethnic Group

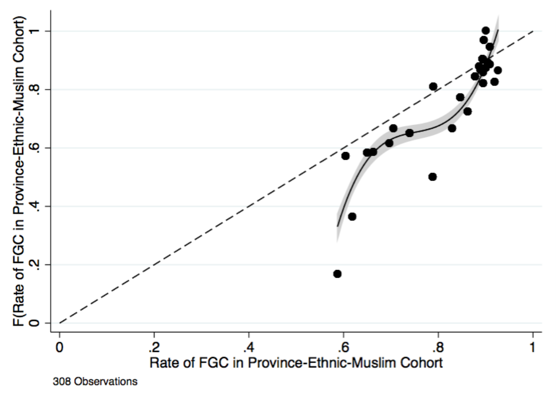


(c) Department 1-Ethnic Group



(d) Department 2-Ethnic Group

(e)



(f) Muslim-Province-Ethnic Group

Appendix

Table A1: LPM Estimation Results for Whether Girl Underwent FGC—Cube 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	0.263***	0.299***	0.227**	0.291***	0.079
	(0.071)	(0.050)	(0.103)	(0.103)	(0.136)
Rate of FGC in Community^3	0.534***	0.469***	0.552***	0.415***	0.645***
	(0.067)	(0.047)	(0.092)	(0.101)	(0.116)
Constant	0.046*	0.027	0.064*	0.041	0.115**
	(0.024)	(0.019)	(0.033)	(0.034)	(0.050)
Observations	8,566	6,643	5,146	1,679	3,776
R-squared	0.158	0.178	0.157	0.161	0.128

Unconditional linear probability model. Standard Errors Clustered at the Province Level.
***p<0.01 **p<0.05 *p<0.1

Table A2: LPM Estimation Results for Whether Girl Underwent FGC—Square 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	-0.097	-0.043	-0.171	-0.029	-0.471*
	(0.129)	(0.085)	(0.180)	(0.153)	(0.245)
Rate of FGC in Community^2	0.828***	0.759***	0.879***	0.685***	1.107***
	(0.117)	(0.081)	(0.156)	(0.149)	(0.206)
Constant	0.083***	0.062***	0.108**	0.078**	0.183***
	(0.030)	(0.021)	(0.043)	(0.034)	(0.066)
Observations	8,566	6,643	5,146	1,679	3,776
R-squared	0.157	0.177	0.156	0.162	0.128

Unconditional linear probability model. Standard Errors Clustered at the Province Level.
***p<0.01 **p<0.05 *p<0.1

Table A3: 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	0.896***	0.875***	0.889***	0.686***	0.978***
	(0.035)	(0.032)	(0.031)	(0.043)	(0.055)
Observations	8,566	6,643	5,146	1,679	3,776
Pseudo R-squared	0.122	0.141	0.120	0.133	0.093

Unconditional logit regression results. Standard Errors Clustered at the Province Level.
***p<0.01 **p<0.05 *p<0.1