A Stepping Stone Approach to Norm Transitions*

Selim Gulesci[1] Sam Jindani[2] Eliana La Ferrara[3]
David Smerdon[4] Munshi Sulaiman[5] H. Peyton Young[6]

8 April 2023

Abstract

We propose a model to study when an intermediate action can serve as a *stepping stone* that enables the elimination of a harmful norm. While the intermediate action may facilitate the first "step", it may also become a new norm. We derive intuitive conditions for stepping stones, which depend on the relative size of social penalties and intrinsic utility benefits. We propose an econometric approach to testing whether an intermediate action is a stepping stone, and apply it to original data on female genital cutting in Somalia. The analysis shows that the intermediate action may become the new norm.

* We thank Dan Andenberg, Sam Bowles, Rossella Calvi, Jean-Paul Carvalho, Jessica Leight, Debraj Ray and seminar participants at Berkeley, Bocconi, Cambridge, CEU, Essex, Harvard, Luiss, NHH, NUS, NYU, Oxford, Rice, the 2019 IFO Workshop on Gender-Based Violence, the 2020 BREAD Conference on Behavioral Economics & Development, the 2020 ThReD Conference, the 2021 CESifo conference on Behavioral Economics, the 2021 CEPR-ENS workshop on Political Economy, Cultural Economics and Gender and the 2023 IOG conference in Chicago for useful comments. Julien Manili and Pedro de Souza Ferreira provided outstanding research assistance. We also thank Beatrice Bonini, Mattia Chiapello, Matteo Courthoud, Enrico Guizzardi, Vrinda Kapoor, Jack Melbourne, Giovanni Pisauro and Armin Satzger for excellent work on the data. An earlier version of this paper was circulated under the title "A Stepping Stone Approach to Understanding Harmful Norms". La Ferrara acknowledges financial support from MIUR FARE grant J42F17000280001.

1 Trinity College Dublin; gulescis@tcd.ie.
2 National University of Singapore; sam.jindani@nus.edu.sg.
3 Harvard University, CEPR, NBER and LEAP; elaferrara@hks.harvard.edu.
4 University of Queensland; d.smerdon@uq.edu.au.
5 BRAC; munshi.sulaiman@brac.net.
6 University of Oxford; peyton.young@economics.ox.ac.uk.

1 Introduction

Harmful norms often persist despite high costs for individuals and society, and despite the presence of legislation against them. The conventional approach taken by governments and NGOs is to push for the outright abandonment of these norms. In practice, such an approach is often ineffective. Dowry and early marriage persist in South Asia despite being outlawed (Anderson, 2007; Ambrus and Field, 2008; Corno et al., 2020). Female genital cutting is widespread in Africa, and yet many governments have passed laws against it (28 Too Many, 2018). Historically, footbinding in China and dueling in Europe persisted for centuries despite repeated attempts by governments to extirpate the practices (Mackie, 1996; Nye, 1993).

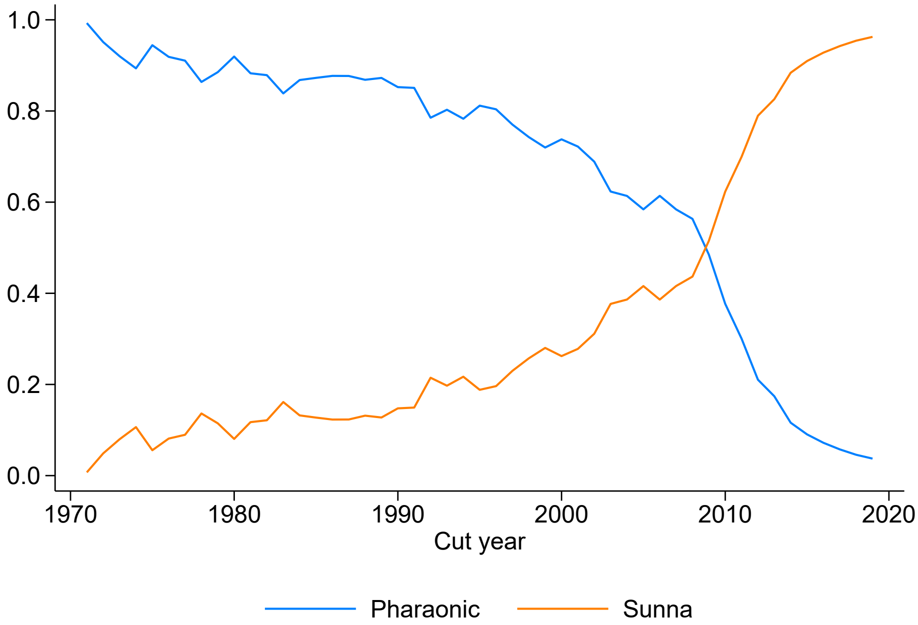
For the purpose of eliminating a harmful norm in the long run, does the existence of a "mildly harmful" alternative facilitate the transition, compared to a binary choice scenario? In the examples above, the intermediate alternative could be a cap on dowry payments, a lower minimum age for marriage, or a less invasive form of female genital cutting. Leaving aside moral considerations (on which our analysis is silent), the answer to this question is not trivial. On the one hand, people who may be reluctant to completely abandon a harmful practice may be persuaded instead to adopt a less harmful variant; once this first step is taken, it may be easier to take further steps, and eventually eliminate the practice altogether. On the other hand, precisely because the variant is less costly, incentives to abandon it are lower, and it may become a new absorbing norm.

In this paper we propose a general model of norm dynamics that sheds light on this trade-off. We provide intuitive conditions under which an intermediate action acts as a *stepping stone* that facilitates the elimination of a harmful norm. We develop an econometric approach to testing whether the conditions for a stepping stone are satisfied in practice, and apply this to the case of female genital cutting (FGC) in Somalia, using original survey data.

FGC in Somalia is particularly germane for two reasons. First, as we discuss below, the prevalence of FGC is almost universal. Second, FGC in Somalia takes two forms: a highly invasive one referred to as "Pharaonic" and a milder one referred to as "Sunna". Figure 1 shows the type of FGC performed in Somalia by year, based on survey data we collected. Until the early 1990s, the vast majority of girls and women who were cut received Pharaonic circumcision. Since then, the proportion of girls and women cut with Sunna has rapidly increased, so that by 2010 the majority of cuts performed were Sunna. Today, Sunna has almost completely displaced Pharaonic. An important question is whether Sunna will persist or whether there will be a further transition from Sunna to a norm of no cutting. This paper provides both a theoretical and an econometric framework in which this question and similar ones may be approached.

We propose a discrete choice model in which agents choose between a set of actions to

Figure 1: Type of FGC by year of cutting



Source: Authors' calculations on originally collected data.

maximize a utility that has two components: an *intrinsic utility* of the action and a *social utility* that is linearly decreasing in the shares of players who choose actions different from the individual's. The latter term captures social pressure for conformity and is commonly used in models with social interactions (Akerlof, 1997; Brock and Durlauf, 2001). Players receive opportunities to update their actions at random intervals and their utilities are subject to shocks. The proportion of players at each action, or *state* of the process, determines the choice probabilities of an updating agent.

In our model, we assume three actions (L, M, and H, for low, medium and high intrinsic utility, respectively) and logit perturbations (McFadden, 1974; McKelvey and Palfrey, 1995; Blume, 1993). The model can easily be generalised to n actions and other forms of perturbations.

We say that the intermediate action M is a *stepping stone* if, starting from the harmful norm L, updating agents initially tend to choose M but then tend to choose H once sufficiently many agents have switched to M. We derive necessary and sufficient conditions for the intermediate action to be a stepping stone. Intuitively, these require that M be a good "social substitute" for both L and H, in the sense that the social penalties for going from L to M and then from M to H are small relative to the corresponding gains in intrinsic utility. Failure to satisfy these conditions implies either that the intermediate norm is not attractive enough and will fail to destabilize the costly norm L, or that it is too attractive – in which case it will become absorbing. Furthermore, we show that a necessary condition for the existence of stepping stones is the *reverse triangle inequality*:

the sum of the sanction parameters for going from L to M and from M to H must not exceed the sanction for going directly from L to H. In other words, stepping stones require a form of convexity of sanctions.

We also characterize the dynamics of stepping-stone transitions. We highlight two features that emerge from the analysis. First, an increase in the proportion of agents playing H does not necessarily mean that the process will converge to the high-utility norm; that is, it is possible for the proportion of agents playing H to increase even if M remains dominant. Second, if M is a stepping stone, the process may come quite close to the intermediate norm before moving towards the high-utility norm; that is, the fact that most agents are playing M does not imply that M is absorbing.

In the last part of the paper we develop a general econometric approach to testing whether an intermediate action is a stepping stone and apply it to FGC in Somalia, using originally collected data. Our data cover 4,130 households from 141 communities and include information on whether female respondents and their daughters are cut, the type of cut, the year in which they were cut, their beliefs about the expected health costs of FGC, and their beliefs about community sanctions against those who deviate from the practice. In this application, the agents are the parents, who are making decisions on behalf of their daughters. In the context of our model, we interpret Pharaonic circumcision as action L, Sunna as M, and Uncut as H (where the ranking reflects the health consequences of the three actions). Because we know the year and type of each cut, our data allow us to derive empirical estimates of the state (the stock of agents who have each type of cut in a given year) and choice probabilities (the flow of agents who choose each type of cut in a given year).

We present three tests for whether an intermediate action is a stepping stone. The first examines whether the above-mentioned *reverse triangle inequality* – a necessary condition for a stepping stone – holds. The second consists in estimating the choice probability functions and simulating the resulting process. The third involves assessing the size of the flow into Uncut relative to the size of the flow into Sunna – another condition implied by our model.[7] In the case of FGC in Somalia, all three tests suggest that Sunna may not be a stepping stone, and instead seems to be absorbing. Of course, this assumes the parameters governing the process do not change. Future interventions or exogenous shocks that increase the perceived attractiveness of Uncut relative to Sunna may alter the outcome.

One of the implications of our analysis is that policies promoting an intermediate action may have adverse consequences. In particular, when evaluating a policy that would disincentivize a harmful social norm, one should consider whether this might lead to the

7 Neither the second nor the third test is reliant on the model specification (specifically, the assumptions of linear sanctions and logit perturbation).

permanent adoption of an intermediate norm rather than the intended abandonment of the harmful norm. In the context of FGC, if Sunna proves to be absorbing, then it may be more difficult to dislodge than Pharaonic. However, our results also indicate that a relatively small shift in the proportion of people choosing not to cut at all could tip the process and prevent Sunna from becoming absorbing, thus highlighting a possible role for policy interventions. Even though our empirical focus is on FGC, our framework is general and can be used to analyze historical norms like dueling, as well as contemporary norms like child marriage or smoking.

Our paper relates to multiple literatures. First, we contribute to the theoretical literature on the evolution of social norms.[8] Our paper is particularly close to the approaches of Akerlof (1997) and Brock and Durlauf (2001), who explicitly model conformity motives and their effect on norms.[9] Akerlof considers a continuum of actions and Brock and Durlauf consider two actions, whereas we allow for an intermediate action and focus on how this affects the dynamics. Another difference is that we study out-of-equilibrium dynamics and convergence to equilibrium, whereas Akerlof and Brock and Durlauf focus on equilibrium outcomes.

Our paper also relates to work in evolutionary game theory that analyzes how systems of interacting agents converge to normative behaviors from out-of-equilibrium conditions (Young, 1993, 1998; Kandori et al., 1993; Blume, 1993, 1995; Bowles, 2006; Newton, 2020).[10] This framework has been used to study, for example, the evolution of property rights (Bowles and Choi, 2013, 2019), as well as Pareto-inferior institutions and cultures (Belloc and Bowles, 2013). A difference with this literature is that we focus on intermediate-run (rather than long-run) dynamics, which are more relevant for the applications we have in mind.[11]

More broadly, we speak to a growing literature in economics that has studied the persistence and welfare effects of gender norms (e.g., Alesina et al., 2013; Ashraf et al., 2020), including norms around fertility and female labor force participation (Fernández and Fogli, 2009; Fernández, 2013), and signalling norms such as veiling among Muslim women (Carvalho, 2013). Only recently has female genital cutting come to the forefront

8 Lewis (1969) and Schelling (1978) were among the first to think of norms in game-theoretic terms. For overviews of the subject, see Young (1998, 2015); Bowles (2004); and Bicchieri (2005).

9 Related work on social interactions includes the literatures on networks (Jackson, 2008; Goyal, 2012) and identity (Akerlof and Kranton, 2000).

10 For book-length treatments of evolutionary game theory and its applications see Weibull (1995); Samuelson (1997); Young (1998); Vega-Redondo (1996); Bowles (2004); and Sandholm (2010).

11 It is worth noting that the concept of stepping stones has been studied in evolutionary game theory, but with a different interpretation. Most prominently, Ellison (2000) and Norman (2009) showed that intermediate actions can speed up long-run transitions when shocks become vanishingly small. In contrast, we focus on how intermediate actions affect the intermediate-run dynamics with non-vanishing noise. Stepping stones are also of interest in other economic contexts, such as dynamic mechanism design (Ely and Szydlowski, 2020), and in biology (Kimura and Weiss, 1964; Kimura, 1983).

of the analysis. Among others, Bellemare et al. (2015) study the individual and household level correlates of attitudes towards FGC in West Africa, while Becker (2022) and Corno et al. (2020) trace the origins of FGC to historical characteristics of societies. Mackie (1996) was the first to interpret FGC as an equilibrium of a game with social interactions. Subsequent work has extended and tested this theory (Shell-Duncan et al., 2011; Efferson et al., 2015; Bicchieri and Marini, 2016; Kudo, 2019; Novak, 2020; Efferson et al., 2020). However, the above literature does not formalize the dynamics of the adjustment process and it is silent on the role that intermediate actions may play.

Finally, our empirical results speak to an emerging literature on the evaluation of policies aimed at eradicating FGC (e.g., Diop et al., 2004; UNICEF, 2008; Camilotti, 2016; Vogt et al., 2016; Platteau et al., 2018; García-Hombrados and Salgado, 2023). We contribute to this literature by calling attention to the role of intermediate actions.

The paper is organized as follows. Section 2 presents our theoretical model, section 3 contains our empirical application, and section 4 concludes. Proofs are contained in the Appendix.

2 A model of harmful social norms

Following Akerlof (1997) and Brock and Durlauf (2001), we consider a model in which agents choose an action to maximize a utility that has two components: (i) the *intrinsic utility* of the action for the individual, and (ii) a *social utility* that is decreasing in the shares of agents who choose actions different from the individual's. Agents' utilities are subject to shocks, leading to heterogeneity in individual choice.

2.1 Model setup

Consider a population of players who can choose from actions L, M, and H (for low, medium, and high intrinsic utility, respectively). In the case of FGC in Somalia, L represents Pharaonic circumcision, M represents Sunna, and H represents not cutting (the ranking reflects the health costs of each action). Let $A = \{L, M, H\}$. There is a continuum of agents. Let $p_i \in [0, 1]$ denote the proportion of agents playing action $i \in A$ and $p = (p_L, p_M, p_H)$ denote the *state* of the process. Let $\Delta = \{p \in \mathbb{R}^3_+ : \sum_{i \in A} p_i = 1\}$ denote the two-dimensional simplex. Let $p^i \in \Delta$ denote the state in which all players choose $i \in A$. We will refer to p^i as a *norm*.

Time is continuous and agents update their actions via independent Poisson arrival processes with unit expectation. Thus, each agent updates once per unit of time in expectation. The utility of an agent playing action i consists of an intrinsic, a social, and a random component. The *intrinsic utility* from playing action $i \in A$ is denoted by u_i, where $u_L < u_M < u_H$. The social component of an agent's utility represents the pressure to conform to others' actions, including shunning and ostracism of nonconformists, as documented for example by Shell-Duncan et al. (2011).[12] We shall assume that the *social utility* of an agent playing action i in state p takes the linear form:

$$-\sum_{j\in A} s_{ji}p_j, \tag{1}$$

where s_{ji} represents how much pressure an agent choosing j exerts on someone choosing i.[13] We assume $s_{ii} = 0$ for all i and $s_{ij} > 0$ for all $i \neq j$. Unlike much of the literature, we do not assume social pressure is symmetric between pairs of actions; that is, we allow $s_{ij} \neq s_{ji}$. This requires a different analysis from the symmetric case (under symmetry, the analysis is simplified by the fact that the underlying game is a potential game).

Note that although the model is presented as if agents observe p (for example, in the FGC application they would know the cutting status of members of their community), this is not necessary. Instead, agents could receive an imperfect signal about the actions of others and the results would go through – in fact, such imperfect information could be a source of the individual level heterogeneity we assume below.

When an agent updates at time t, her utility for action i is subject to a shock ε_{it}. Her utility is then

$$u_i - \sum_{j\in A} s_{ji}p_{jt} + \varepsilon_{it}. \tag{2}$$

Because time is continuous, the probability that two agents update at the same time is zero. As is common in the discrete-choice literature, we assume that shocks are i.i.d. and extreme-value distributed (Blume, 1993; McKelvey and Palfrey, 1995; Brock and Durlauf, 2001). This assumption is analytically convenient and consistent with our empirical results. Since an extreme-value random variable has zero mean, the expected utility of an agent choosing action i in state p is

$$v_i(p) = u_i - \sum_{j\in A} s_{ji}p_j. \tag{3}$$

It is a standard result that extreme-value shocks lead to logit choice probabilities.[14]

12 In fieldwork on FGC in Senegal and the Gambia, Shell-Duncan et al. (2011) find that "those who are not circumcised are contemptuously insulted by being labeled *solima*, meaning not only uncircumcised, but also rude, ignorant, immature, uncivilized, and unclean; women who are *solima* are told they know nothing, and are harassed and excluded by women for not knowing how to behave properly. [...] Uncircumcised women are excluded from participating in, or even being present for, some activities (most commonly listed were circumcision ceremonies and wedding ceremonies)" (pp. 7–8).

13 Other specifications exist in the literature (e.g., Brock and Durlauf (2001) consider the square of the difference between an individual action and the average action).

14 See, e.g., Sandholm 2010, p. 194.

Thus, the probability of an agent choosing action i in state p is

$$\sigma_i(p) = \frac{e^{\beta v_i(p)}}{\sum_{j \in A} e^{\beta v_j(p)}}, \tag{4}$$

where $\beta > 0$ captures agents' *precision.* If β is small, agents' choices are almost random; if β is large, agents' choices are close to pure best responses. Let $\sigma(p) = (\sigma_L(p), \sigma_M(p), \sigma_H(p))$. One can think of p as a *stock* and $\sigma(p)$ as a *flow.*

Shocks incorporate the idea that people are heterogeneous with respect to their intrinsic utilities u_i for different actions, and also with respect to the social pressure coefficients s_{ij}. Because of this heterogeneity, at any given point in time, the prevailing norm will be followed by most people but not by everyone. When the population is large, the expected direction of motion is given by $\sigma(p)$, which is the flow conditional on the current stock p.

The function σ determines a dynamical system on Δ defined by $\dot{p} = \sigma(p) - p$. A *path* is a function $p : \mathbb{R}_+ \to \Delta$. A path p is a *solution* if it is continuous and $\dot{p} = \sigma(p) - p$ almost everywhere. Since σ is Lipschitz continuous, by the Picard–Lindelöf theorem the system admits a unique solution path from any initial state.

It will be insightful to consider the limiting best-response case as $\beta \to \infty$. Let $B(p)$ denote the set of best responses in state p; that is,

$$B(p) = \arg\max_{i \in A} v_i(p). \tag{5}$$

Define the best-response revision protocol

$$\sigma_i^*(p) = \frac{\mathbf{1}_{B(p)}(i)}{|B(p)|} \tag{6}$$

That is, $\sigma_i^*(p)$ denotes the probability that action i is chosen when agents play best response, assuming agents randomize when there are multiple best responses. Note that if the precision β is large, $\sigma_i(p)$ will be close to $\sigma_i^*(p)$. (Formally, viewing $\sigma(p) = \sigma(p, \beta)$ as a function of β as well as p, we have $\sigma(p, \beta) \to \sigma^*(p)$ as $\beta \to \infty$.) Note also that, unlike σ, σ^* is discontinuous at states where two or more actions are best responses. Nevertheless σ^* is continuous almost everywhere, which is enough to ensure that solution paths exist.

Finally, the assumption that the population is infinite is simply for convenience. The case in which the population is finite but large can be addressed using standard methods from stochastic approximation theory. In particular, over a given finite time interval, provided the population is sufficiently large, the finite-population process will stay close to the corresponding infinite-population process with high probability (Benaim and Weibull, 2003).

2.2 Analysis

We are interested in analyzing the dynamics generated by the above model. In particular, we will see that a great deal of intuition can be garnered from considering the limiting best-response case.

First, we define what it means for M to be a stepping stone. Intuitively, we say that the intermediate action is a stepping stone if, starting at the L-norm, updating agents initially mostly choose M, but later mostly choose H.

Definition 1. M is a *stepping stone* if, for every solution path $p(t)$ such that $p(0) = p^L$, (i) there is some $t' > 0$ such that $\sigma_M(p(t)) > \sigma_H(p(t))$ on $[0, t')$ and $\sigma_M(p(t)) < \sigma_H(p(t))$ on (t', ∞), and (ii) there is some $t'' > 0$ such that $p_H(t) > 1/2$ for all $t > t''$.

We say that t' in the above is the *inflection time*, and $p(t')$ is the *inflection point*. Note that the inflection point occurs when the expected payoffs from H and M are equal.

The dynamics of the process can be illustrated using phase diagrams. Figure 2 illustrates the four possible dynamics starting from p_L.[15] The arrows show the direction of travel and the filled circles show the stable states of the process. Figures 2a and 2b capture the key trade-off of intermediate actions: on the one hand, M could allow a transition from p^L to a state close to p^H; on the other hand, M could be absorbing. Figures 2c and 2d illustrate the remaining two possible cases, which are less interesting from the point of view of our analysis.

In what follows we provide conditions for M to be a stepping stone under the best-response dynamics. This case is particularly tractable and allows us to obtain clean, intuitive results. Note that the general dynamics will be close to the best-response dynamics whenever the precision β is sufficiently large.

Theorem 1. *Under the best-response dynamics, M is a stepping stone if and only if*

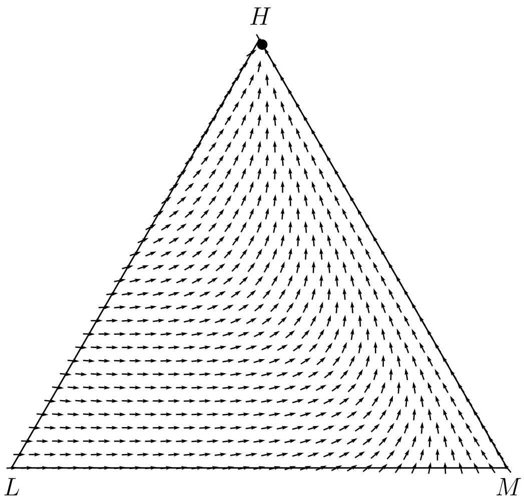
$$\frac{s_{LH} - s_{LM}}{u_H - u_M} > 1, \tag{7a}$$
$$\frac{s_{LM}}{u_M - u_L} \leq 1, \text{ and} \tag{7b}$$
$$\frac{s_{MH}}{u_H - u_M} \leq 1. \tag{7c}$$

Moreover, if M is a stepping stone under the best-response dynamics, then its inflection point is $(1 - q^*, q^*, 0)$, where

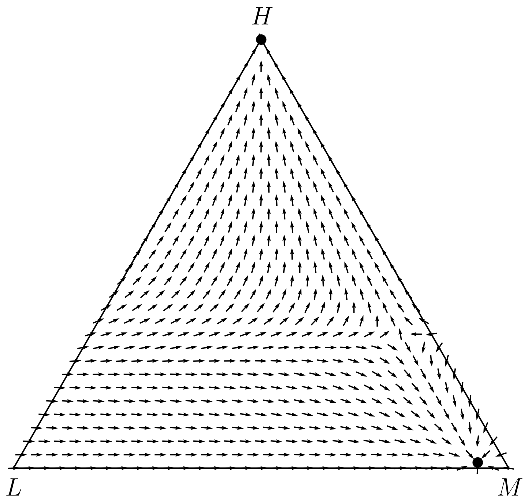
$$q^* = \frac{u_M - u_H + s_{LH} - s_{LM}}{s_{LH} - s_{LM} - s_{MH}}. \tag{8}$$

15 In each case $u_L = 0$, $u_M = 0.5$, $u_H = 1$, $\beta = 3$, and $s_{ij} = s_{ji}$ for each $i, j \in A$. In figure 2a, $s_{LM} = 0.5$, $s_{MH} = 1$, and $s_{LH} = 2$. In figure 2b, $s_{LM} = 0.5$, $s_{MH} = 2$, and $s_{LH} = 2$. In figure 2c, $s_{LM} = 1$, $s_{MH} = 1$, and $s_{LH} = 1$. In figure 2d, $s_{LM} = 2$, $s_{MH} = 2$, and $s_{LH} = 3$.

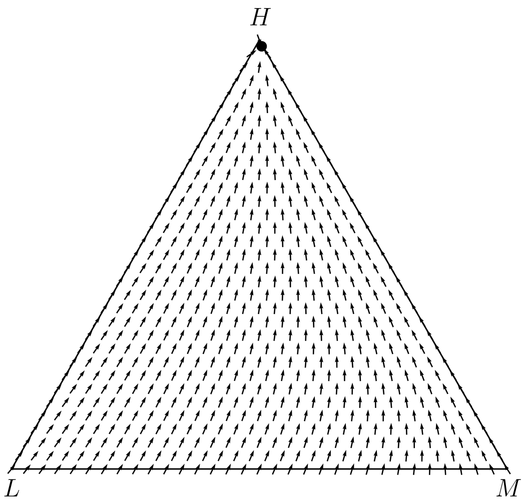
Figure 2: Dynamics starting from p^L.



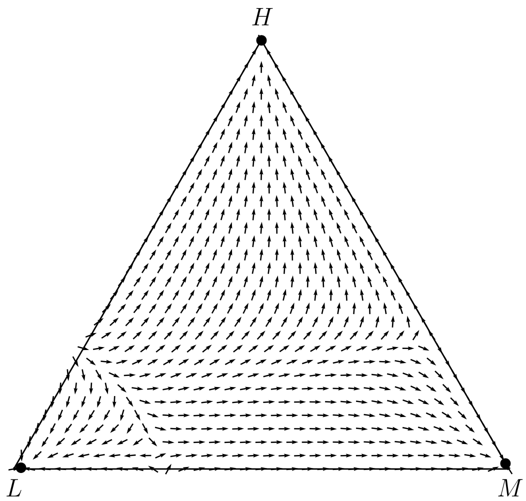
(a) M is a stepping stone.



(b) M is absorbing.



(c) Direct transition to H.



(d) L is absorbing.

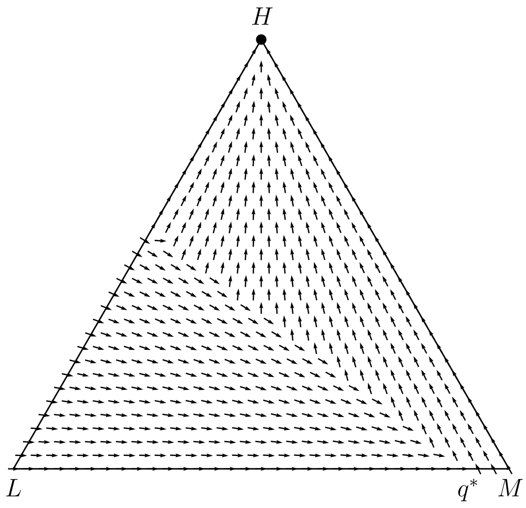
Proof. See the Appendix. □

Intuitively, inequality (7a) means that the social sanctions for going directly from L to H rather than from L to M are high relative to the intrinsic-utility gains, so that M is preferred to H at p^L. Inequality (7b) means that the sanctions for going from L to M are lower than the intrinsic-utility gains, so that M is preferred to L at p^L; similarly, inequality (7c) means that H is preferred to M at p^M. In other words, M must be a good *social substitute* for both L and H. A way to see this directly is to note that a necessary condition for expression (7) to hold is the *reverse triangle inequality*:

Corollary 1 (Reverse triangle inequality)**.** *If expression (7) holds, then*

$$s_{LH} > s_{LM} + s_{MH}. \tag{9}$$

Figure 3: q^* under the best-response process.



To see why, note that inequalities (7a) and (7c) imply $s_{LH} - s_{LM} > u_H - u_M$ and $u_H - u_M \geq s_{MH}$, which together imply inequality (9). One can thus think of stepping stones as requiring a form of convexity of sanctions, which does not depend on the intrinsic utilities. Note that expression (7) also implies monotonicity of sanctions, in the sense that $s_{LH} > s_{LM}$.

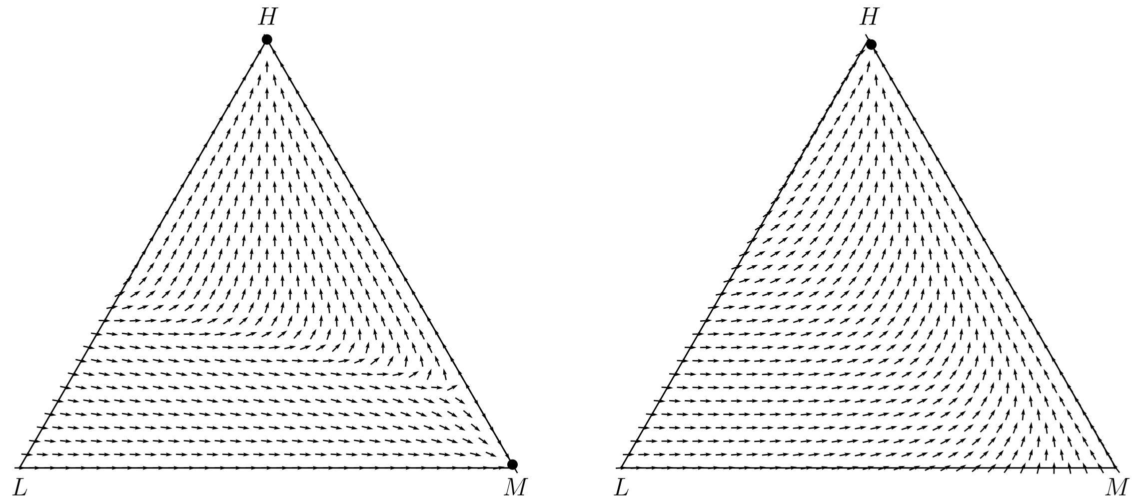
The second part of theorem 1 has two important implications for the process dynamics. First, a stepping-stone transition may take longer than a direct transition to the high-utility norm. Second, the process may come quite close to the intermediate norm before moving towards the good action; that is, the fact that few agents are choosing the good action does not imply that the intermediate norm is absorbing. Figure 3 illustrates.[16] Note that the inflection point $p^* = (1 - q^*, q^*, 0)$ is the unique state such that $p^*_H = 0$ and $v_H(p^*) = v_M(p^*)$.

If the precision β is large but finite, then the behavior of the general dynamics will be close to the behavior of the best-response dynamics. Under heterogeneity, the inflection point will not be $(1 - q^*, q^*, 0)$, but will be another state that equalizes the expected utilities of H and M, and thus equalizes the flows into H and M. In other words, the states that satisfy $v_H(p) = v_M(p)$ form a boundary that the path must cross for M to be a stepping stone. Note that this boundary does not depend on β (since the condition $v_H(p) = v_M(p)$ does not depend on β).

If β is not large, then expression (7) is neither necessary nor sufficient for M to be a stepping stone; that is, M can be a stepping stone if expression (7) fails, and it can fail to be a stepping stone if expression (7) holds. To see why it is not necessary, consider figure 4. The parameters are the same in both figures 4a and 4b, except that β is smaller

16 Intrinsic utilities are $u_L = 0$, $u_M = 0.5$, $u_H = 1$; social pressure parameters are $s_{LM} = 0.4$, $s_{MH} = 0.4$, and $s_{LH} = 2$, and satisfy $s_{ij} = s_{ji}$ for each $i, j \in A$.

Figure 4: Expression (7) is not necessary for a stepping stone with enough heterogeneity.



(a) M is not a stepping stone when β is large. (b) M is a stepping stone when β is small.

in figure 4b. Moreover, the parameters are such that expression (7) fails.[17] However, M is still a stepping stone in figure 4b. Intuitively, in the presence of heterogeneity, a sufficient proportion of agents choose H in the neighborhood of p^M. Thus, the process escapes the basin of attraction of p^M even though expression (7) is not satisfied. To see why expression (7) is not sufficient for M to be a stepping stone when β is small, suppose β is close to 0. Then $\sigma(p) \approx (1/3, 1/3, 1/3)$ for any state p. Intuitively, the process will tend towards a state close to $(1/3, 1/3, 1/3)$ from any initial state, so M is not a stepping stone even if expression (7) holds.

Another implication of the model is that an increase in the proportion of agents choosing the good action does not necessarily imply that the intermediate action is a stepping stone. In fact, when M is absorbing, p_H may increase at first but then subsequently decrease. Figure 5 illustrates.[18] Intuitively, starting at p^L, as agents switch to M and H, the expected utilities of both M and H increase. However, the utility of M increases faster than the utility of H. Eventually, M becomes much more attractive than H and agents start to switch from H to M.

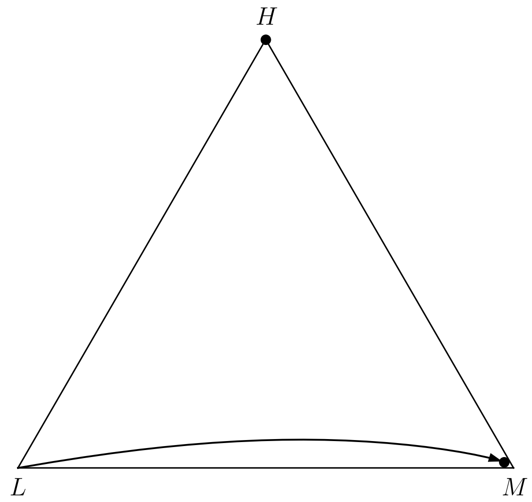
2.3 Discussion

In the interest of analytical tractability the preceding model makes several simplifying assumptions that are fairly standard in the social interactions literature, but that may or may not be empirically valid. One is the assumption that everyone knows the current

17 Specifically, intrinsic utilities are $u_L = 0$, $u_M = 0.5$, $u_H = 1$; social pressure parameters are $s_{LM} = 0.6$, $s_{MH} = 1$, and $s_{LH} = 2$, and satisfy $s_{ij} = s_{ji}$ for each $i, j \in A$. In figure 4a, $\beta = 10$; in figure 4b, $\beta = 3$.

18 The parameters are $u_L = 0$, $u_M = 0.5$, $u_H = 1$, $\beta = 3$, $s_{LM} = 1$, $s_{MH} = 2$, and $s_{LH} = 2$, and $s_{ij} = s_{ji}$ for each $i, j \in A$.

Figure 5: p_H can increase even when M is absorbing.



prevalence of each action in society, and can therefore accurately predict social sanctions. This assumption is most plausible in small communities; in larger communities it may be more reasonable to assume agents only have partial information, say based on the actions of a subset of the community.

However, it is not necessary to assume that p is fully observable for our results to hold. We could assume that agents act on the basis of an imperfect signal about the actions of others – indeed this could be one of the sources of individual-level heterogeneity in the model. More generally, heterogeneity incorporates a variety of factors that translate into idiosyncratic choices by individuals that result from differences in their intrinsic utilities, their susceptibility to social pressure, their perception of the current state of the system, and so forth.

Another assumption concerns the ordering of the intrinsic utilities of the various actions. In the context of our application, we assume that the intrinsic utility of Uncut is higher than that of Sunna, which is higher in turn than that of Pharaonic. This is plausible if intrinsic utilities reflect, for example, health costs.[19] However, intrinsic utilities could also reflect other considerations and some of these considerations may reverse the ranking. For instance, some individuals might prefer Sunna to Uncut for religious reasons, regardless of what other agents chose. If that motive dominates the health motive, the intrinsic utility of Sunna might be higher than the intrinsic utility of Uncut. In that case, Sunna would necessarily be absorbing, and could not serve as a stepping stone. In the theory we maintain the assumption $u_L < u_M < u_H$ to study the case in which, at least in principle, stepping stones are possible. However, it is important to note that our empirical analysis is not dependent on this assumed mapping.

19 Table A.1 in Online Appendix A shows that, based on our survey data, respondents associate significantly higher likelihood of health complications to Pharaonic than to Sunna.

3 Empirical application: FGC in Somalia

In this section we apply the model to the norm of female genital cutting (FGC) using originally collected data from Somalia.

3.1 Context

We start by giving some background on FGC in general and on FGC in Somalia in particular. FGC is the practice of cutting or removing part of the external female genitalia for non-medical reasons. An estimated 200 million women are cut worldwide (UNICEF, 2022) and every year, 3 million female infants and children are at risk of undergoing FGC (Spisma et al., 2012). FGC is a harmful practice, as it leads to serious health consequences both at the time of cutting (e.g., excessive bleeding and increased mortality) and in the long run (e.g., birth-related complications).Also, given that FGC is generally performed on young girls without their informed consent, it is considered a human rights violation (WHO, 2018). The practice is geographically dispersed, being prevalent in 29 African and Middle Eastern countries (Camilotti, 2016). In some of these countries FGC is almost universal: the share of cut women is estimated to be 98% in Somalia, 96% in Guinea, 93% in Djibouti, and 91% in Egypt and Sierra Leone (Yoder et al., 2013).

The WHO distinguishes three main types of female circumcision: *type I* is the partial or total removal of the clitoris and/or the prepuce (clitoridectomy); *type II* is the partial or total removal of the clitoris and the labia minora, with or without excision of the labia majora (excision); *type III*, also known as infibulation, is the narrowing of the vaginal orifice with the creation of a covering seal by cutting and apposing the labia minora and/or the labia majora, sometimes through stitching (WHO, 2018).

We will focus on this difference in Somalia, where FGC is divided into two broad categories: "Sunna" (types I–II) and "Pharaonic" (type III). Historically, Pharaonic circumcision was the dominant type in Somalia, practiced almost universally (Abdalla, 1982). Sunna was introduced in Somalia much later than the Pharaonic type.[20] In 1984, the Inter-African Committee on Harmful Traditional Practices Affecting the Health of Women and Children was established and in the 1990s pushes for abandonment gained momentum. Anecdotal evidence suggests that there was "a shift from infibulation to Sunna [...] as a result of FGC campaigns that have been emphasizing health effects of infibulation" (MOLSA, 2009). This shift was facilitated by religious leaders who started opposing Pharaonic as harmful and non-Islamic, but supported Sunna as in line with

20 Historical accounts suggest that the intermediate type of cutting was first introduced in Sudan, as a result of the prohibition of infibulation by the British colonial authorities in 1945 (El Dareer, 1982; Slack, 1988).

Islam (Newell-Jones, 2016).[21]

3.2 Data

We use data from a household survey that we conducted in 141 communities in the Somaliland and Puntland regions of Somalia between January and May 2020. The communities are located in the districts of Badhan, Buraou, Erigabo, Galdagob, Galkayo and Hargeisa. Figure A.1 shows the locations of these communities. The survey collected information on a sample of 4,130 individuals – on average 29 per community – roughly equally split by gender (2,040 men and 2,090 women). The respondents were sampled from the list of participants in community meetings that were conducted as part of an ongoing project (Gulesci et al., 2020).[22]

The survey collected information on household demographics, socioeconomic status, and gender norms. In particular, we elicited information on respondents' own attitudes as well as their perceptions of community members' attitudes towards different types of FGC. We also collected the history of FGC types in the respondents' family, a list of those among their daughters that were cut (or were intended to be cut) and those that were uncut (and were intended to remain so). For all cut daughters, we asked what type of FGC was performed on them. Finally, we elicited expectations about community reactions to deviations from FGC, to understand the degree of social pressure that may be faced by people deciding not to comply with local norms.

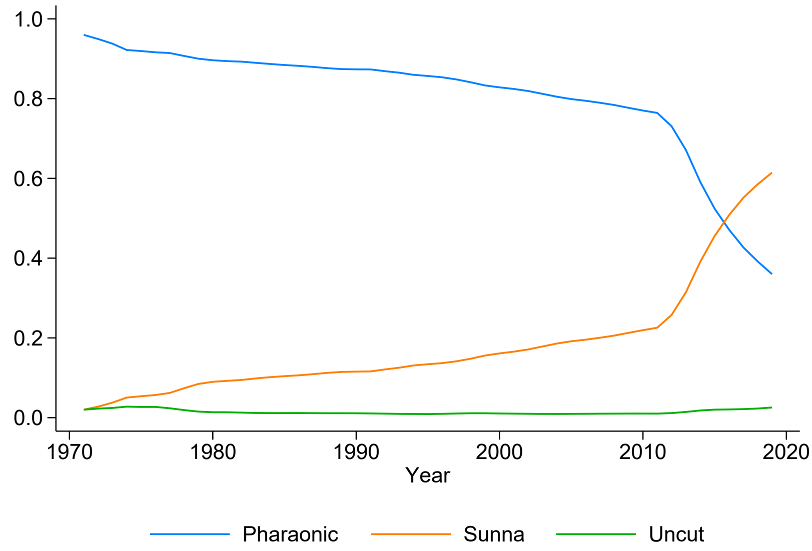
In our sample, among women and girls aged 10 to 60, 38% are Pharaonic cut, 59% are Sunna cut, and 3% are uncut. Online Appendix Figure A.2 shows that the choice of Sunna vs. Pharaonic circumcision is not an ethnic marker: in fact, cutting rates are similar across different sub-clans (panel (a)). On the other hand, there is significant variation across communities (panel (b)). Table A.2 presents summary statistics on the age at cutting and the relevant decision makers by type of FGC. We see that, on average, girls were cut at the age of 9 for Pharaonic and 8.3 for Sunna. The decision to cut is most often taken by the mother: mothers are reported as making the decision about circumcision for 72% of girls who are Pharaonic cut and 85% percent of the girls who are Sunna cut.

Due to the sensitive nature of the questions related to FGC status, one may worry about reporting bias. In order to assess this, we followed the approach in Dhar et al. (2022) and collected information on respondent's propensity to give socially desirable

21 More recently, a religious Fatwa issued in 2018 in Somaliland banned the practice of Pharaonic circumcision, but not Sunna.

22 These meetings were organized in collaboration with the NGO Save the Children and were facilitated by local personnel trained in interpersonal communication on sensitive topics. When participants were recruited to take part in the meetings, the latter were advertised as part of a research effort aimed towards better understanding local conditions and community life.

Figure 6: Stocks, or empirical approximation of the state, over time



Source: Authors' calculations on originally collected data.

answers, creating an index using the scale developed by Crowne and Marlowe (1960). In Table A.3 we compare the type of FGC reported by respondents with high and low social desirability, as measured by the index. We do not find any evidence of differential reporting in types of FGC, suggesting that reporting bias in FGC status due to social desirability may not be a significant concern in this context. This is likely due to the fact that both types of cutting are widespread, and the overall cutting rate is 98%.

We now describe how we map the data to the model. Our data record three cutting statuses: Pharaonic, Sunna, and Uncut. We define the *stock* in a given year to be the proportion of girls and women aged 10 or more who have a given cutting status.[23] We count as cut girls aged 10 or more who are not cut but whose families plan to have them cut.[24] Figure 6 shows the evolution of the stocks over time. To reduce noise, we work with three-period moving averages of the stocks and flows throughout the analysis.[25]

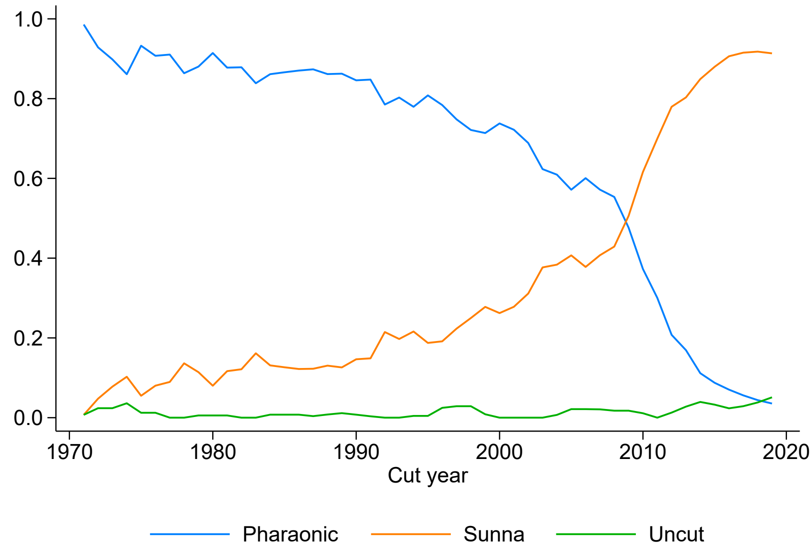
We define the *flow* in a given year to be the proportion of girls who turn 10 that year and who have a given cutting status. Again, we count as cut girls aged 10 who are uncut but whose families plan to have them cut. Figure 7 shows the evolution of the flows over

23 Age 10 is a natural benchmark given that the average age at cutting in our sample is 9 (see Appendix Table A.1). We have explored robustness to alternative age cutoffs (specifically, 9, 11 and 12) and our results are qualitatively unchanged.

24 This reflects the fact that, in the rural context we study, intentions to cut may be discussed and be known by others and will influence the choice of those who have not yet made a decision. In Online Appendix B we present an alternative method of computing stocks and flows and we show that all our results are robust to using this alternative method.

25 We consider cohorts of women born after 1960 because the number of observations for earlier cohorts is too small to yield reliable estimates.

Figure 7: Flows, or empirical approximation of the choice probabilities, over time



Source: Authors' calculations on originally collected data.

time.

Figure 7 shows that, until the early 1990s, virtually all girls aged 10 were cut and, moreover, were Pharaonic-cut. Starting from the mid-1990s, the rate of Sunna-cut women has been steadily rising and circumcised girls aged 10 in 2010 or after are more likely to be Sunna-cut than Pharaonic-cut. In recent years, Sunna has replaced Pharaonic circumcision as the dominant choice for younger generations. The pattern is very similar if we use the most recent nationally representative dataset with information on FGC for Somalia, the 2011 Multiple Indicator Cluster Survey (see Appendix Figure A.3).[26]

While many factors may have contributed to this transition, two (not mutually exclusive) are worth mentioning. The first is model dynamics: the pattern of acceleration in the 1990s could arise endogenously as a result of the gradual accumulation of Sunna. The second is exogenous interventions that may have facilitated the abandonment of Pharaonic. As discussed in section 3.1, since the late 1980s human rights campaigns made a strong push against Pharaonic circumcision and religious leaders started emphasizing that this type of FGC was not a requirement of Islam. In terms of our model, these could be represented as changes in s_{LM}, u_L, and u_M that would lead to M taking over

26 We pooled data from the 2011 Multiple Indicator Cluster Survey (MICS) for Somaliland and Northeast Somalia. Each of these are representative household surveys, with the Somaliland MICS covering 5,865 women aged 15–49 and their children, while the Northeast Somalia MICS entails information on 5,492 women aged 15–49 and their children. As Appendix Figure A.3 shows, the share of Pharaonic (Sunna) cut women gradually declines (rises) over time, with girls circumcised in the mid-2000s or after being more likely to be Sunna-cut as opposed to Pharaonic. Since the MICS data are from 2011, we are not able to replicate the corresponding trends post-2011 with these data.

L.[27]

Finally, figure 7 also shows a small but noticeable upward trend in Uncut in recent years. Our analysis in figure 5 shows that this doesn't necessarily mean the intermediate action is a stepping stone. We address this question in the next subsection.

3.3 Tests for stepping stone

The mapping from stocks to flows, $\sigma(p)$, is given by the logit function in equation (4). A convenient feature of this relationship is that the log of the ratio $\sigma_i(p)/\sigma_j(p)$ is a linear function of the difference in expected payoffs between action i and action j. In the present case, this reduces to the expression

$$\ln\left(\frac{\sigma_i(p)}{\sigma_j(p)}\right) = \beta\left(u_i - u_j - \sum_{k\in A}(s_{ki} - s_{kj})p_k\right) \tag{10}$$

which can be estimated via OLS. Note that this representation stems from our assumptions of (i) logit choice, (ii) separability between intrinsic and social utility, and (iii) linear social utility.

Appendix Table A.4 reports the estimates of eq. (10) for different i and j. In particular, columns 1 and 2 show the estimated coefficients from regressing $\ln(\sigma_M/\sigma_L)$ and $\ln(\sigma_M/\sigma_H)$, respectively, on p_M and p_H. (Recall that $p_L + p_M + p_H = 1$, hence p_L is omitted). Both regressions show a good fit, suggesting that the data are well approximated by the model.

We present three tests to assess whether an intermediate action is a stepping stone. These tests are derived from the general formulation of our model and can be applied to a variety of settings where one may be interested in assessing whether an intermediate action may become absorbing or not. We will empirically illustrate them for the case of FGC in Somalia, and we will see that they suggest that Sunna may be absorbing.

3.3.1 Reverse-triangle-inequality test

Recall that the reverse triangle inequality in expression (9) is a necessary condition for M to be a stepping stone, provided heterogeneity is small. Under the model, expanding

27 Given the nature of our data, we cannot fully isolate the contribution of endogenous bottom-up factors and that of exogenous factors, such as interventions by social leaders (Acemoglu and Jackson, 2015). However, we can check whether the parameters changed substantially in the 1990s by replicating the analysis using only cohorts that were of cutting age from 1990 onwards. The results in the two cases are qualitatively similar, i.e., Sunna being absorbing, suggesting that the dynamics did not change greatly as a result of the interventions (see Online Appendix C). Since the estimates in the truncated sample are considerably noisier, we use the entire sample in our main analysis.

eq. (10) for $\ln(\sigma_H/\sigma_M)$, we have

$$\ln\left(\frac{\sigma_H(p)}{\sigma_M(p)}\right) = \alpha + \gamma p_H + \lambda(s_{LM} + s_{MH} - s_{LH})p_L, \tag{11}$$

where $\alpha \in \mathbb{R}$, $\gamma > 0$, and $\lambda > 0$ are constants that depend on the intrinsic utilities, the social pressure parameters, and β. A useful feature of the above equation is that it allows us to perform a test of the reverse triangle inequality without requiring to have empirical estimates of the sanction parameters s_{ij}: in fact, equation (11) shows that the reverse triangle inequality holds if and only if the coefficient on p_L in equation (11) is negative. Note that this tests depends on the same assumptions that yield eq. (10), namely (*i*) logit choice, (*ii*) separability between intrinsic and social utility, and (*iii*) linear social utility.

Table 1: Relationship between $\ln(\sigma_H/\sigma_M)$ and state.

	Log-ratio Uncut flow to Sunna flow				
Uncut stock	109.843***	35.200***	36.608**	71.738**	73.173**
	(18.484)	(7.532)	(13.986)	(8.907)	(14.275)
Pharaonic stock	3.689***	5.128***	5.752***	4.814**	5.259***
	(0.577)	(0.661)	(0.644)	(0.615)	(0.051)
Constant	-7.365***	-6.756***	-7.227***	-7.554***	-8.284***
	(0.641)	(0.561)	(0.906)	(0.595)	(0.192)
Sample	Year	Year-District	Year-District	Year-Clan	Year-Clan
District F.E.	No	No	Yes	No	No
Clan F.E.	No	No	No	No	Yes
Adjusted R^2	0.616	0.561	0.631	0.458	0.536
Observations	38	60	60	58	58

Notes: Observations are years (column 1), year-district (columns 2–3), and year-clan (columns 4–5). Uncut stock is the share of uncut women aged 10 or more in a given year who were also planned to remain uncut. Pharaonic stock is the share of Pharaonic cut women aged 10 or more in a given year. Uncut flow is the share of girls who turned 10 in a given year and were uncut and planned to remain uncut. Sunna flow is the share of girls who turned 10 in a given year and were Sunna cut at some point in their lives. Stocks and flows are 3-year moving averages for each year, from 1971 to 2019. They comprise female respondents and all respondents' daughters aged 0 to 18.

Table 1 shows the results of a regression of $\ln(\sigma_H/\sigma_M)$ on p_H and p_L. In column 1 we exploit variation across years, in columns 2-3 across district-years, and in columns 4-5 across clan-years. The specifications in column 3 and 5 also include district and clan fixed effects, respectively.[28] It can be seen from the table that the coefficient on p_L is positive and significant in all specifications, suggesting that the reverse triangle inequality fails to hold in our data. Without heterogeneity, failure of the reverse triangle inequality is sufficient to conclude that Sunna is absorbing. Of course, in practice, there can be

28 We cannot exploit variation at the village-year level because we would have too few observations for each cohort.

substantial heterogeneity in the data. Thus the above test should be viewed as indicative rather than conclusive. In contrast, the next test holds for any degree of heterogeneity in the agents' responses.

3.3.2 Extrapolation test

The following test involves simulating – that is, extrapolating – the process based on historical data. Observe that the path of the process is determined by the mapping from states p to flows $\sigma(p)$. Model parameters only matter inasmuch as they determine $\sigma(p)$. Moreover, this mapping can be estimated from the data. Under the model, we have

$$\ln\left(\frac{\sigma_H(p)}{\sigma_M(p)}\right) = \alpha + \gamma_H p_H + \gamma_L p_L, \tag{12}$$

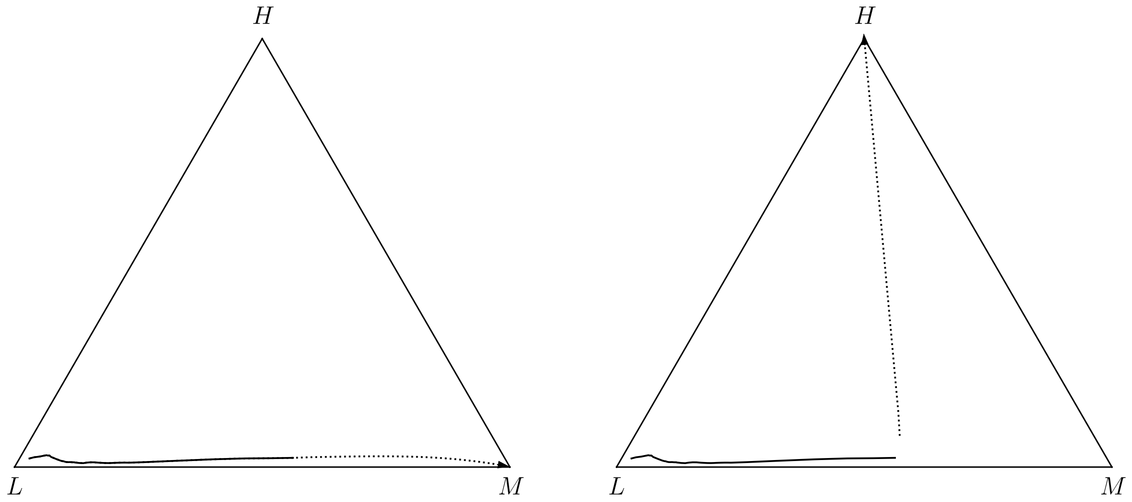
$$\ln\left(\frac{\sigma_M(p)}{\sigma_L(p)}\right) = \alpha' + \gamma'_H p_H + \gamma'_L p_L, \tag{13}$$

where α, γ_H, γ_L, α', γ'_H, and γ'_L are constants that depend on the intrinsic utilities, social sanction parameters, and β. The relationships in equations (12) and (13) can be estimated via OLS. Moreover, since $\sigma_L(p) + \sigma_M(p) + \sigma_H(p) = 1$ for all states p, equations (12) and (13) are sufficient to infer $\sigma_L(p)$, $\sigma_M(p)$, and $\sigma_H(p)$.

Once the mapping $\sigma(p)$ is estimated, one can simulate the process starting at the most recent state. The results of this simulation are shown in figure 8a, with the full line showing historical data, and the dotted line showing simulated data. These results suggest that, in the case of FGC in Somalia, Sunna is absorbing. Note also that the simulation suggests that the modest recent increase in Uncut will continue for some time before being reversed. This type of behavior is consistent with the model (recall the illustration previously shown in figure 5).

It bears emphasizing that this extrapolation approach is very general and is valid whether or not the log ratios of the flows are linear in the stock variables. All that is required is that it is possible to estimate the map $\sigma(p)$. In our case, eq. (10) means that this can be done via OLS, but in general other estimation techniques could be used. This approach can also be used to forecast the impact of exogenous shifts in the state, since we can simulate the process starting from any state. For example, suppose a targeted policy led to 5% of agents switching from Sunna to Uncut. Figure 8b shows the simulated path after such a shift. The results in this figure suggest that, at the juncture at which the Somali communities are now, a small shift in the state may be sufficient to prevent Sunna from becoming absorbing.

Figure 8: Extrapolation test and exogenous shift in the state.



(a) Extrapolation test. (b) Exogenous shift in the state.

3.3.3 Equal-flow test

The two preceding tests rely on an econometric estimation of the mapping from states p to flows $\sigma(p)$. The logit choice structure made this particularly straightforward. In other contexts, it may be more challenging to estimate the mapping. In this section we present a simple test that does not depend on the functional form specification of the model.

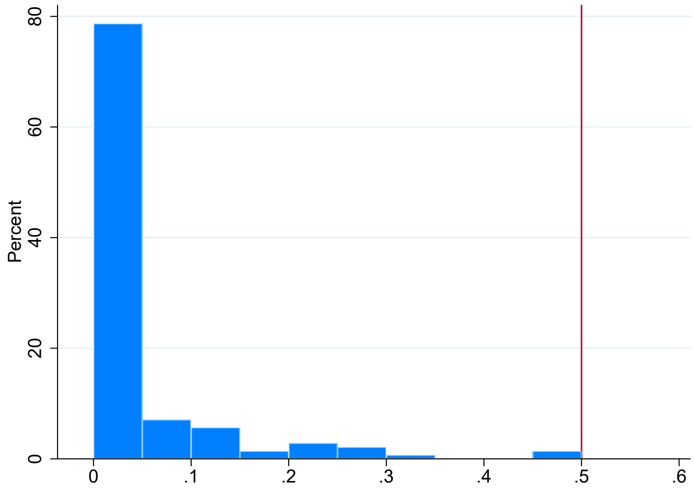
If M is a stepping stone, then by definition the process must eventually reach an inflection point, that is, a state where the flow into H becomes at least as large as the flow into M. If the flows are a continuous function of the state, this inflection point occurs where the flows into H and M are equal.[29] The equal-flow state may occur well before the *stocks* are equal; indeed, under the best-response dynamics it occurs while $p_H = 0$. Reaching an equal-flow state is not in general a sufficient condition for M to be a stepping stone. However, reaching a state in which $\sigma_H \geq 1/2$ is sufficient under a broad class of models, including the logit and uniform error models.

On the other hand, not reaching an equal-flow state by a certain date does not imply that such a state will never be reached. The fact that at a given point in time the equal-flow test is not satisfied means that, up to that point, the path of the process is consistent with M being absorbing. It is still possible that, at a later stage, the transition may occur.

What does the equal-flow test tell us in the context of FGC in Somalia? Figure 9 plots the frequency of $\sigma_H/(\sigma_H+\sigma_M)$ values across communities (we take the average over the period 2010–2020 to reduce variability). The equal-flow test is met in a community

29 In a utility-based model, this is where the utilities of H and M are equal.

Figure 9: Frequency of $\sigma_H/(\sigma_H + \sigma_M)$ values across communities.



Notes: Observations are communities. Uncut flow (σ_H) is the share of girls who turned 10 in a given year and were uncut and planned to remain uncut. Sunna flow (σ_M) is the share of girls who turned 10 in a given year and were Sunna cut at some point in their lives. Flows for each community are computed pooling women and girls born between 2000 and 2010, who were of cutting age between 2010 and 2020.

if and only if $\sigma_H/(\sigma_H + \sigma_M) \geq 1/2$. The figure shows that the distribution of this ratio lies entirely to the left of 0.5, i.e., there is not a single community that has achieved an equal flow to date. The results of this third test are therefore consistent with those of the previous two tests: the observed path in Somalia to date is not consistent with the start of a stepping-stone transition from Sunna to Uncut.

3.3.4 Discussion

Consistent with our model, the above tests rest upon the assumption that the parameters governing the process do not change. To test this assumption, one would need data on the parameters u_i and s_{ij} over time – something that is unfortunately unavailable. We assess the empirical relevance of this potential limitation in two ways. First, in Online Appendix C we repeat our analysis using only the sample 1990–2020. The three tests for stepping stone yield similar results (i.e., Sunna absorbing), suggesting that, even if the parameters had changed over time, they did not do so in a way that significantly alters our conclusions.

Second, in Online Appendix D we consider an additional way to shed light on the assumption of parameter stability by exploiting variation *across cohorts* in a measure of s_{ij} constructed from survey responses to a set of original questions that we designed for this purpose. The underlying idea is the following. If the parameters varied over time, one would expect younger respondents to give very different answers from older ones when asked how bad it is for people to choose action i instead of action j. Table D.1

shows that this is not the case, i.e., our empirical proxies for s_{ij} are uncorrelated with age. While our proxies may be imperfect and this result is obviously not conclusive, it provides supportive evidence for the assumption that parameters did not vary significantly over time.

Obviously, what will happen over the coming years will also depend on factors external to the model. For example, interventions that increase the perceived attractiveness of Uncut relative to Sunna, either by changing intrinsic utilities or decreasing the force of sanctions, may alter our conclusion that Sunna is likely to be absorbing.

4 Conclusions and policy implications

This paper argues that intermediate actions can be crucial for understanding the dynamics of social norms. We propose a model of harmful social norms in which agents choose between three actions and are motivated by intrinsic utility, social pressure, and individual-level heterogeneity. We note that intermediate actions may be beneficial if they allow a *stepping-stone transition*, but may also be counterproductive if the intermediate action is *absorbing*. We provide intuitive conditions under which a stepping-stone transition occurs. The key parameters governing the transition are the social sanctions imposed on those who abandon the prevailing option in favor of another, relative to the difference in intrinsic utilities of the options. Intuitively, stepping stones occur when the intermediate action is a good "social substitute" for both the low-payoff and the high-payoff action.

We develop an econometric approach to testing whether such a transition will occur, and apply this approach to original data on FGC collected across 141 communities in Somalia. We find evidence that, on balance, based on the data available so far the intermediate action is unlikely to be a stepping stone in the communities we study.

Our theoretical and empirical analyses provide a framework for thinking about potential policy responses. Since the key condition for transitioning from Sunna to Uncut requires that sanctions are low relatively to perceived utility gains, policymakers may work on two fronts. First, they may try to reduce sanctions associated with the decision not to cut, for example by changing the narrative around the value of circumcision in the marriage market. This is what NGOs like Tostan have been proposing, for example through "public declarations" by community members that pledge not to cut their daughters and not to marry their sons to girls who are cut.[30] Second, policymakers may work on changing perceptions and knowledge of the health costs of FGC, e.g., through health information campaigns.[31] Clearly, the two approaches are not exclusive but rather com-

30 https://tostan.org
31 A recent example of such an intervention, in the context of Sierra Leone, is Corno and La Ferrara (2023).

plement each other. Moreover, we find that a shift of a relatively small proportion of agents may be sufficient to allow the process to tip towards Uncut. This suggests that interventions targeting a small proportion of agents may prove successful in eliminating FGC.

Although our empirical application concerns FGC in Somalia, the theoretical framework is general and can be applied to a variety of settings. An interesting historical example is the case of dueling. In the United Kingdom, dueling died out suddenly in the mid-nineteenth century (Banks, 2008), while in France it endured up to the turn of the century (Hopton, 2007, 323). In France, the form of dueling changed during this period. Swords became popular again, replacing pistols (Nye, 1993, 186), and, increasingly, the risks involved in dueling were made explicit: duels were announced in advance as being *au premier sang*, to serious wounds, or, rarely, *à la mort* (Hopton, 2007, 79). These changes had the effect of drastically reducing fatality rates. In our framework, these changes are consistent with a shift to an absorbing intermediate norm in the case of France, and a direct shift to the high-utility norm in the case of the United Kingdom.

A contemporary example is child marriage. Today, most countries outlaw marriage before adult age (typically 18 years). This approach has failed to eradicate child marriage in some countries, most notably in South Asia. One might ask whether reducing the legal age for marrying to, say, 16 and then gradually increasing it could prove more effective. Of course, as in the case of FGC, such a policy runs the risk of making the lower age the new norm. An interesting avenue for future research is to explore these implications in countries that have indeed changed the legal age of marriage at different points in time (e.g., Bangladesh).

Finally, another setting in which our model could be applied is that of cigarette smoking, particularly among communities where smoking is the majority norm. Recent years have seen the introduction of electronic cigarettes, and the share of consumers that have substituted tobacco with e-cigarettes has risen sharply. Will e-cigarettes ultimately lead people to quit smoking for good, or will they become the new norm? Again, as more data becomes available, it may be possible to explore transition patterns and give an answer to this question.

Appendix. Proof of theorem 1

Consistent with definition 1, we will say that a path $p : \mathbb{R}_+ \to \Delta$ is a *stepping-stone path* if (*i*) $p(0) = p^L$, (*ii*) there is some $t' > 0$ such that $\sigma_M(p(t)) > \sigma_H(p(t))$ on $[0, t')$ and $\sigma_M(p(t)) < \sigma_H(p(t))$ on (t', ∞), and (*iii*) there is some $t'' > 0$ such that $t \geq t''$ implies $p_H(t) > 1/2$. Let

$$q^* = \frac{u_M - u_H + s_{LH} - s_{LM}}{s_{LH} - s_{LM} - s_{MH}}. \tag{14}$$

We divide the theorem into two lemmas.

Lemma 1. *Suppose inequalities (7a) to (7c) in the main text hold. Then under* σ^**, starting at* p^L*, there exists a unique solution path* $p(t)$*. Moreover,* $p(t)$ *is a stepping-stone path with inflection point* $(1-q^*, q^*, 0)$*.*

Proof. Let $t^* = -\ln(1-q^*)$. We claim that the unique solution is

$$p_M(t) = \begin{cases} 1-e^{-t} & \text{if } t \leq t^* \\ \frac{q^*}{1-q^*}e^{-t} & \text{otherwise} \end{cases} \tag{15}$$

$$p_H(t) = \begin{cases} 0 & \text{if } t \leq t^* \\ 1-\frac{1}{1-q^*}e^{-t} & \text{otherwise.} \end{cases} \tag{16}$$

It is straightforward to verify that $\dot{p}(t) = \sigma^*(p(t)) - p(t)$ for all $t \neq t^*$. Note that $\lim_{t\to\infty} p_H(t) = 1$, so there exists some $t^{**} \in \mathbb{R}_+$ such that $p_H(t) \geq 1/2$ for all $t \geq t^{**}$.

To see that $p(t)$ is the unique solution, note that any solution must agree with $p(t)$ on $[0, t^*)$ because $\sigma^*(p)$ is continuous at any state in that interval. Continuity of the solution then implies any solution must agree with $p(t)$ on $[0, t^*]$.

Suppose that $p'(t)$ is another solution path on $[0,\infty)$ and let $t^{**} > t^*$ be such that $p'(t^{**}) \neq p(t^{**})$. First, suppose that H is a unique best response on (t^*,∞) under $p'(t)$. Write $q^{**} = (q_L^{**}, q_M^{**}, q_H^{**}) = p'(t^{**})$. Consider the set of solution paths on (t^*,∞) whose state at t^{**} is q^{**} and under which H is a unique best response on (t^*,∞). It is straightforward to check that $p''(t)$ defined by

$$p''_M(t) = q_M^{**}e^{t^{**}-t} \tag{17}$$
$$p''_H(t) = 1-(1-q_H^{**})e^{t^{**}-t} \tag{18}$$

is such a solution path. But note that σ is constant, and therefore continuous, on (t^*,∞) under the assumption that H is a unique best response. So there is a unique such solution, and therefore $p''_M(t)$ and $p'_M(t)$ agree on (t^*,∞). By continuity, they agree on $[t^*,\infty)$.

Now, by assumption, $p'(t^{**}) \neq p(t^{**})$. Suppose

$$p'_M(t^{**}) \neq p_M(t^{**}) \tag{19}$$
$$\iff q_M^{**} \neq \frac{q^*}{1-q^*}e^{-t^{**}} \tag{20}$$
$$\iff q_M^{**}e^{t^{**}-t^*} \neq \frac{q^*}{1-q^*}e^{-t^*} \tag{21}$$
$$\iff p'_M(t^*) \neq p_M(t^*), \tag{22}$$

which contradicts the fact that any solution on $[0,\infty)$ must agree with $p(t)$ on $[0,t^*]$. The cases $p'_H(t^{**}) \neq p_H(t^{**})$ and $p'_L(t^{**}) \neq p_L(t^{**})$ are analogous and are omitted. Hence $p'(t^{**}) \neq p(t^{**})$ leads to a contradiction, and therefore $p(t)$ and $p'(t)$ agree on $[0,\infty)$.

The general case where H is not a unique best response on (t^*, ∞) under $p'(t)$ is similar but cumbersome, and is omitted (results available from the authors).

We conclude that $p(t)$ is the unique solution. To complete the proof, note that $p(t)$ is clearly a stepping-stone path with inflection time t^* and inflection point $(1-q^*, q^*, 0)$. □

Lemma 2. *Suppose that at least one of inequalities (7a) to (7c) in the main text fails to hold. Then under the best-response dynamics, M is not a stepping stone.*

Proof. First, suppose inequality (7a) fails, so that $\frac{s_{LH}-s_{LM}}{u_H-u_M} \leq 1$. There are two subcases. First, suppose $H \in B(p^L)$. Then $B((1-q, 0, q)) = \{H\}$ for any $q \in (0, 1]$. This implies that there exists a solution path $p(t)$ such that $p_M(t) = 0$ for all $t \in \mathbb{R}_+$ and $\lim_{t\to\infty} p_H(t) = 1$. Second, suppose $H \notin B(p^L)$. Then $\frac{s_{LH}-s_{LM}}{u_H-u_M} \leq 1$ implies $M \notin B(p^L)$, and hence $B(p^L) = \{L\}$. This implies that there exists a unique solution $p(t)$, where $p(t) = p^L$ for all $t \in \mathbb{R}_+$. In each subcase, M is not a stepping stone.

Second, suppose inequality (7b) fails, so that $\frac{s_{LM}}{u_M-u_L} > 1$. There are two subcases. First, suppose $H \in B(p^L)$. Then, as in the previous paragraph, there exists a solution path $p(t)$ such that $p_M(t) = 0$ for all $t \in \mathbb{R}_+$ and $\lim_{t\to\infty} p_H(t) = 1$. Second, suppose $H \notin B(p^L)$. Then $\frac{s_{LM}}{u_M-u_L} > 1$ implies $B(p^L) = \{L\}$. As in the previous paragraph, there exists a unique solution $p(t)$, where $p(t) = p^L$ for all $t \in \mathbb{R}_+$. In each subcase, M is not a stepping stone.

Finally, suppose inequality (7c) fails, so that $\frac{s_{MH}}{u_H-u_M} > 1$. Again, there are two subcases. First, suppose $M \in B(p^L)$. Then $\frac{s_{MH}}{u_H-u_M} > 1$ implies that $B((1-q, 0, q)) = \{M\}$ for any $q \in (0, 1]$. This implies that there exists a solution path $p(t)$ such that $p_H(t) = 0$ for all $t \in \mathbb{R}_+$ and $\lim_{t\to\infty} p_M(t) = 1$. Second, suppose $M \notin B(p^L)$. If $H \in B(p^L)$, then there exists a solution path $p(t)$ such that $p_M(t) = 0$ for all $t \in \mathbb{R}_+$ and $\lim_{t\to\infty} p_H(t) = 1$. If $H \notin B(p^L)$, then $B(p^L) = \{L\}$ and there exists a unique solution $p(t)$, where $p(t) = p^L$ for all $t \in \mathbb{R}_+$. In each subcase, M is not a stepping stone. □

References

28 Too Many (2018). The law and FGM: an overview of 28 African countries.

Abdalla, R. H. D. (1982). *Sisters in Affliction: Circumcision and Infibulation of Women in Africa.* Connecticut: Lawrence Hill and Co.

Acemoglu, D. and M. O. Jackson (2015). History, expectations, and leadership in the evolution of social norms. *Review of Economic Studies 82*(2), 423–456.

Akerlof, G. A. (1997). Social distance and social decisions. *Econometrica 65*(5), 1005–1027.

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *Quarterly Journal of Economics 115*(3), 715–753.

Alesina, A., P. Giuliano, and N. Nunn (2013). On the origins of gender roles: Women and the plough. *Quarterly Journal of Economics 128*(2), 469–530.

Ambrus, A. and E. Field (2008). Early marriage, age of menarche and female schooling attainment in bangladesh. *Journal of Political Economy 116*(5), 881–930.

Anderson, S. (2007). The economics of dowry and brideprice. *Journal of Economic Perspectives 21*(4), 151–174.

Ashraf, N., N. Bau, N. Nunn, and A. Voena (2020). Bride price and female education. *Journal of Political Economy 128*(2), 591–641.

Banks, S. (2008). Killing with courtesy: The english duelist, 1785–1845. *Journal of British Studies 47*(3), 528–558.

Becker, A. (2022). On the origins of restricting women's promiscuity. Working paper.

Bellemare, M. F., 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.

Belloc, M. and S. Bowles (2013). The persistence of inferior cultural-institutional conventions. *American Economic Review 103*(3), 93–98.

Benaim, M. and J. W. Weibull (2003). Deterministic approximation of stochastic evolution in games. *Econometrica 71*(3), 873–903.

Bicchieri, C. (2005). *The grammar of society: The nature and dynamics of social norms.* Cambridge: Cambridge University Press.

Bicchieri, C. and A. Marini (2016). Female genital cutting: Fundamentals, social expectations and change. MPRA working paper No. 72927.

Blume, L. E. (1993). The statistical mechanics of strategic interaction. *Games and Economic Behavior 5*(3), 387–424.

Blume, L. E. (1995). The statistical mechanics of best-response strategy revision. *Games and Economic Behavior 11*(2), 111–145.

Bowles, S. (2004). *Microeconomics: Behavior, institutions, and evolution.* Princeton: Princeton University Press.

Bowles, S. (2006). Group competition, reproductive leveling, and the evolution of human altruism. *Science 314*.

Bowles, S. and J.-K. Choi (2013). Coevolution of farming and private property during the early holocene. *Proceedings of the National Academy of Sciences 110*(22), 8830–8835.

Bowles, S. and J.-K. Choi (2019). The neolithic agricultural revolution and the origin of private property. *Journal of Political Economy 127*(5).

Brock, W. A. and S. N. Durlauf (2001). Discrete choice with social interactions. *Review of Economic Studies 68*(2), 235–260.

Camilotti, G. (2016). Interventions to stop female genital cutting and the evolution of the custom: Evidence on age at cutting in senegal. *Journal of African Economies 25*(1), 133–158.

Carvalho, J.-P. (2013). Veiling. *Quarterly Journal of Economics 128*(1), 337–370.

Corno, L., N. Hildebrandt, and A. Voena (2020). Age of marriage, weather shocks, and the direction of marriage payments. *Econometrica 88*(3), 879–915.

Corno, L. and E. La Ferrara (2023). Norms replacement. a field experiment on ending female genital cutting. Working paper.

Corno, L., E. La Ferrara, and A. Voena (2020). Female genital cutting and the slave trade. Working paper.

Crowne, D. and D. Marlowe (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology 24*(4), 349–354.

Dhar, D., T. Jain, and S. Jayachandran (2022). Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in india. *American economic review 112*(3), 899–927.

Diop, N. J., M. Faye, A. Moreau, J. Cabral, H. Benga, F. Cissé, B. Mané, I. Baumgarten, and M. Melching (2004). The tostan program. evaluation of a community based education program in senegal.

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

Efferson, C., S. Vogt, and E. Fehr (2020). The promise and the peril of using social influence to reverse harmful traditions. *Nature Human Behaviour 4*(1), 55–68.

El Dareer, A. (1982). *Woman, Why do you Weep?* London, U.K.: Zed Press.

Ellison, G. (2000). Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. *Review of Economic Studies 67*(1), 17–45.

Ely, J. C. and M. Szydlowski (2020). Moving the goalposts. *Journal of Political Economy 128*(2), 468–506.

Fernández, R. (2013). Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review 103*(1), 472–500.

Fernández, R. and A. Fogli (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics 1*(1), 146–77.

García-Hombrados, J. and E. Salgado (2023). Legal bans, female genital cutting, and education: Evidence from senegal. *The World Bank Economic Review 37*(1), 74–92.

Goyal, S. (2012). *Connections: an introduction to the economics of networks.* Princeton: Princeton University Press.

Gulesci, S., E. La Ferrara, D. Smerdon, and M. Sulaiman (2020). Changing harmful norms through information and coordination: Experimental evidence from somalia. mimeo.

Hopton, R. (2007). *Pistols at dawn: A history of duelling.* London: Piatkus.

Jackson, M. O. (2008). *Social and economic networks.* Princeton: Princeton University Press.

Kandori, M., G. J. Mailath, and R. Rob (1993). Learning, mutation, and long run equilibria in games. *Econometrica 61*(1), 29–56.

Kimura, M. (1983). *The neutral theory of molecular evolution.* Cambridge: CUP.

Kimura, M. and G. H. Weiss (1964). The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics 49*(4), 561.

Kudo, Y. (2019). Political efforts to eliminate female genital cutting: Long-term effects on women's health and marriage in west africa. Working paper. Institute of Developing Economies (IDE-JETRO), Japan.

Lewis, D. (1969). *Convention: A philosophical study.* Cambridge, MA: Harvard University Press.

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

McFadden, D. (1974). The measurement of urban travel demand. *Journal of public economics 3*(4), 303–328.

McKelvey, R. D. and T. R. Palfrey (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior 10*(1), 6–38.

MOLSA (2009). National policy for the abandonment of Female Genital Mutilation (FGM). Technical report, Ministry of Labor & Social Affairs, Republic of Somaliland.

Newell-Jones, K. (2016). Empowering communities to collectively abandon fgm/c in somaliland: Baseline research report.

Newton, J. (2020). Conventions under heterogeneous behavioural rules. *Review of Economic Studies*.

Norman, T. W. (2009). Rapid evolution under inertia. *Games and Economic Behavior 66*(2), 865–879.

Novak, L. (2020). Persistent norms and tipping points: The case of female genital cutting. *Journal of Economic Behavior & Organization 177*, 433 – 474.

Nye, R. A. (1993). *Masculinity and male codes of honor in modern France.* Oxford: Oxford University Press.

Platteau, J.-P., G. Camilotti, and E. Auriol (2018). Eradicating women-hurting customs: What role for social engineering? In S. Anderson, L. Beaman, and J.-P. Platteau (Eds.), *Towards gender equity in development.* Oxford: Oxford University Press.

Samuelson, L. (1997). *Evolutionary games and equilibrium selection.* Cambridge, MA: MIT Press.

Sandholm, W. H. (2010). *Population games and evolutionary dynamics.* Cambridge, MA: MIT Press.

Schelling, T. C. (1978). *Micromotives and macrobehavior.* New York: W. W. Norton.

Shell-Duncan, B., K. Wander, Y. Hernlund, and A. Moreau (2011). Dynamics of change in the practice of female genital cutting in senegambia: testing predictions of social convention theory. *Social science & medicine 73*(8), 1275–1283.

Slack, A. (1988). Female circumcision: A critical appraisal. *Human Rights Quarterly 10*(4), 437–486.

Spisma, H. L., P. G. Chen, A. Ofori-Atta, U. O. Ilozumba, K. Karfo, and E. H. Bradley (2012). Female genital cutting: current practices and beliefs in western africa. *Bulletin of the World Health Organization 90*, 120–127F.

UNICEF (2008). Long-term evaluation of the tostan programme in senegal: Kolda, thiès and fatick regions. www.unicef.org/evaldatabase/index_59605.html.

UNICEF (2022). Eliminating female genital mutilation: Unicef 2021 annual report to the u.s. department of state. https://www.unicef.org/documents/eliminating-female-genital-mutilation.

Vega-Redondo, F. (1996). *Evolution, games, and economic behaviour.* Oxford: Oxford University Press.

Vogt, S., N. Zaid, H. Ahmed, E. Fehr, and C. Efferson (2016). Female circumcision: A critical appraisal. *Nature 538*, 506–509.

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.

Weibull, J. W. (1995). *Evolutionary game theory.* Cambridge, MA: MIT Press.

WHO (2018). Care of women and girls living with female genital mutilation: a clinical handbook. Technical report, World Health Organization, Geneva.

Yoder, P. S., S. Wang, and E. Johansen (2013). Estimates of female genital mutilation/cutting in 27 african countries and yemen. *Studies in Family Planning 44*(2), 189–204.

Young, H. P. (1993). The evolution of conventions. *Econometrica 61*(1), 57–84.

Young, H. P. (1998). *Individual strategy and social structure.* Princeton: Princeton University Press.

Young, H. P. (2015). The evolution of social norms. *Annual Review of Economics 7*(1), 359–387.

Online Appendix
A Stepping-Stone Approach to Norm Transitions

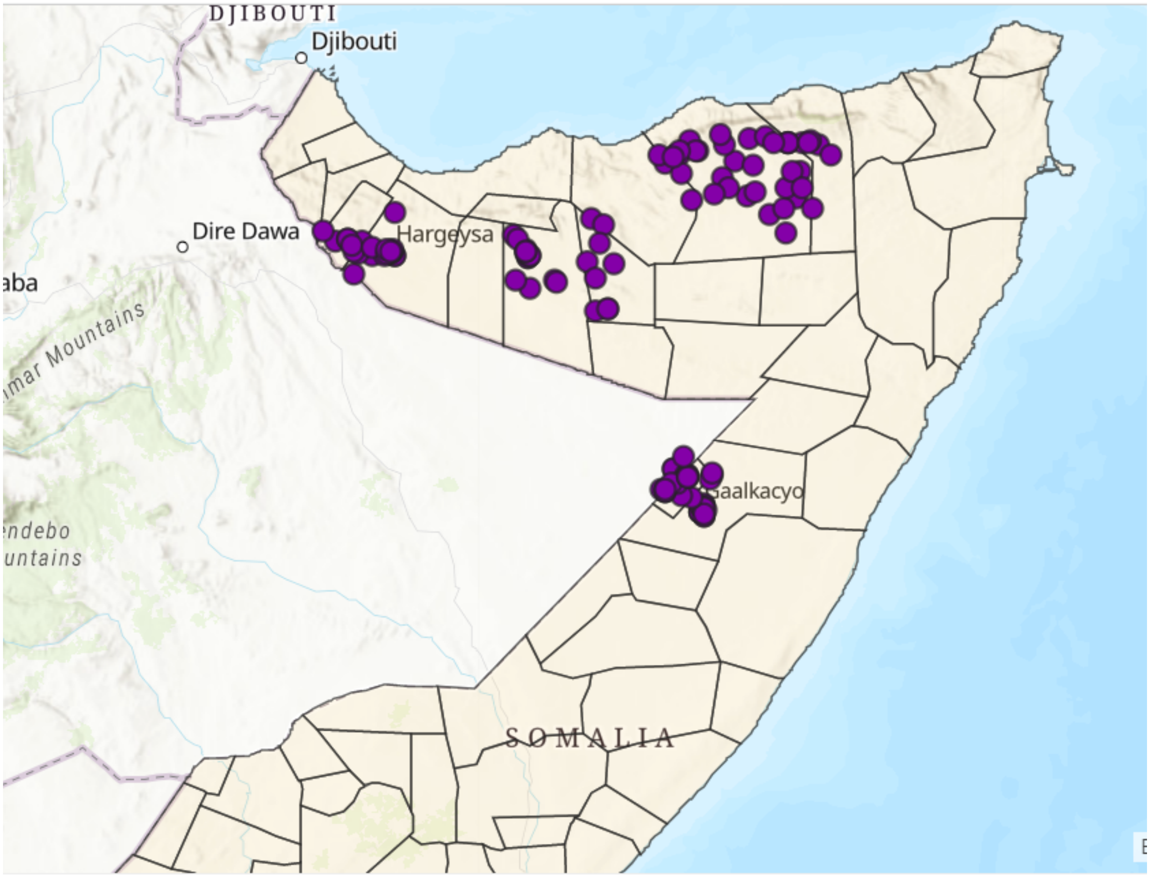
8 April 2023

Selim Gulesci[1] Sam Jindani[2] Eliana La Ferrara[3]
David Smerdon[4] Munshi Sulaiman[5] H. Peyton Young[6]

1 Trinity College Dublin; gulescis@tcd.ie.
2 University of Cambridge; sj608@cam.ac.uk.
3 Harvard University, CEPR, NBER and LEAP; elaferrara@hks.harvard.edu.
4 University of Queensland; d.smerdon@uq.edu.au.
5 BRAC; munshi.sulaiman@brac.net.
6 University of Oxford; peyton.young@economics.ox.ac.uk.

A Additional figures and tables

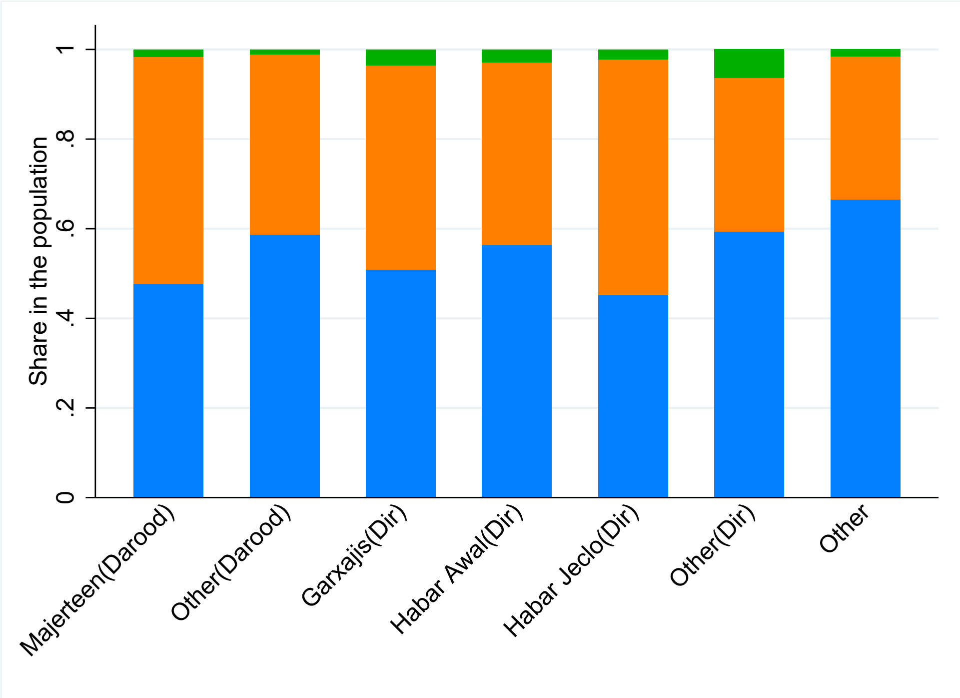
Figure A.1: Location of the Study



Notes: The map shows the location of communities included in the study, along with district boundaries of Somalia.

Figure A.2: Type of FGC

(a) By sub-clan



(b) By community

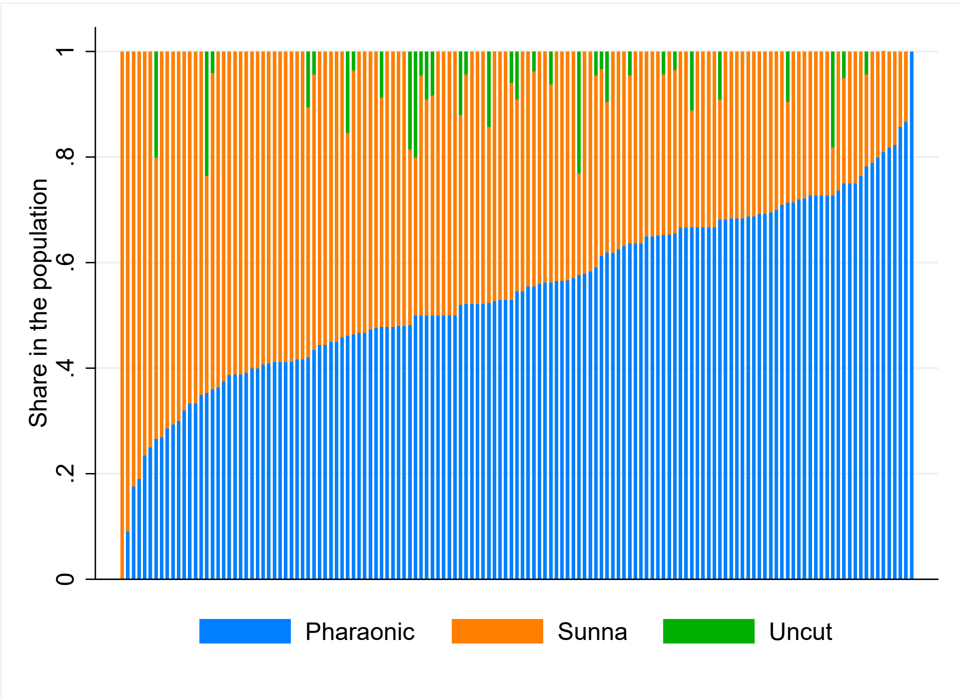


Table A.1: Health complications, by type of FGC

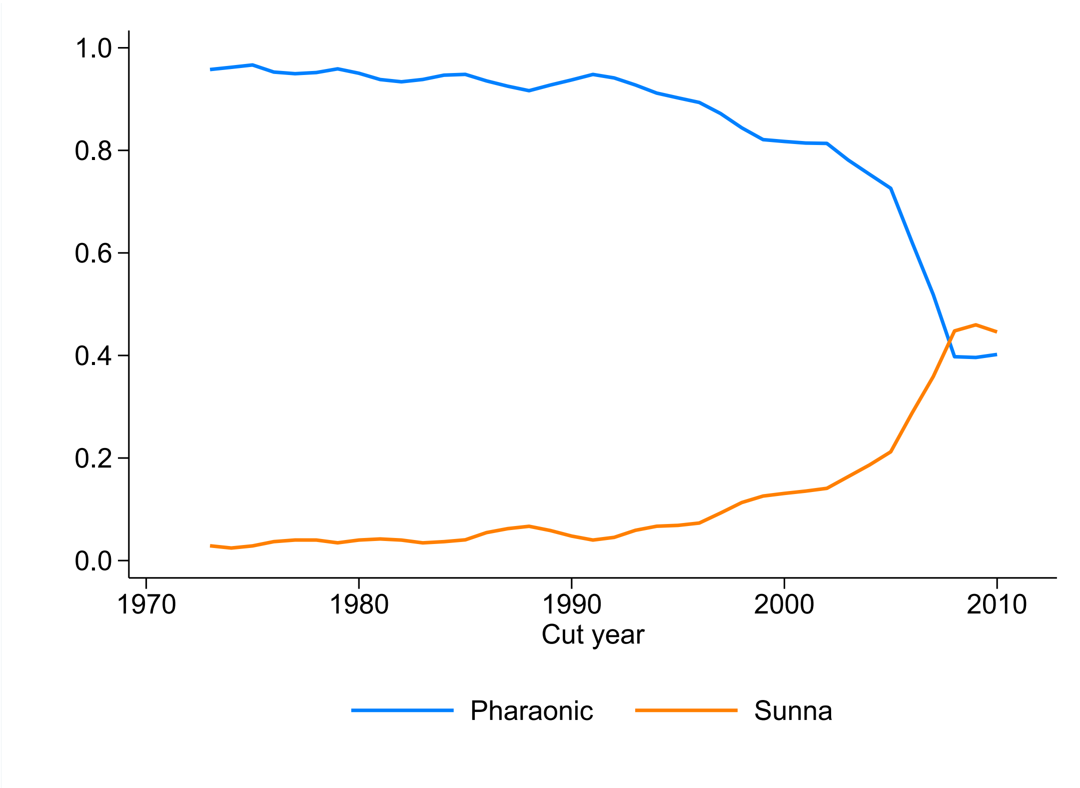
	(1) *Pharaonic* Mean	SD	(2) *Sunna* Mean	SD
Panel A: For female respondents				
Any health complication (Yes=1)	0.624	(0.484)	0.116	(0.320)
N	1,376		403	
Panel B: For female daughters				
Any health complication (Yes=1)	0.592	(0.492)	0.029	(0.169)
N	1,488		3,154	
Panel C: Perceived health complications				
Any complication	0.753	(0.432)	0.060	(0.238)
Infection	0.404	(0.491)	0.041	(0.199)
Bleeding	0.570	(0.495)	0.029	(0.169)
Difficulty in delivery	0.610	(0.488)	0.034	(0.182)
Reduction in sexual feeling	0.462	(0.499)	0.031	(0.172)
Difficulty in penetration	0.328	(0.470)	0.004	(0.064)
Other	0.014	(0.117)	0.004	(0.064)
Number of complications	2.388	(1.726)	0.159	(0.718)
N	939		3,638	

Table A.2: Summary statistics.

	(1) *Pharaonic* Mean	SD	(2) *Sunna* Mean	SD
Age at FGM	9.011	2.075	8.272	1.829
Decision to cut by (not mutually exclusive):				
Mother	0.721		0.855	
Father	0.520		0.543	
Grandmother	0.083		0.038	

Notes: Sample restricted to female respondents and respondents' daughters between 0 and 18 years who have been cut. "Age of FGC" is the age at which the individual was cut. "Decision to cut by" is the fraction of respondents reporting that the decision to cut was taken, by, respectively, the child's mother, father, or grandmother.

Figure A.3: Flows, or empirical approximation of the choice probabilities, over time (MICS 2011 data)



Source: Authors' calculations based on Somalia Multiple Indicator Cluster Survey 2011. The sample includes all respondents (all female household members aged 15-50) and their daughters aged 0 to 18. Pharaonic is the fraction of females who are reported as being circumcised with their genital area sewn closed, Sunna is the fraction of females who are circumcised without their genital area being sewn closed. Data points are three-year moving averages.

Table A.3: Type of cut reported, by social desirability bias

	(1)	(2)	(3)	(4)
	High social desirability	Low social desirability	Difference p-value	Normalized difference
Pharaonic cut	0.178	0.189	0.198	-0.021
	(0.382)	(0.392)		
Sunna cut	0.396	0.389	0.538	0.010
	(0.489)	(0.488)		
Uncut	0.418	0.418	0.967	0.001
	(0.493)	(0.493)		
Uncut nor planned to be	0.078	0.072	0.364	0.015
	(0.267)	(0.259)		

Notes: The sample includes all daughters aged 0-18 and all female respondents. "Pharaonic cut" is a dummy variable =1 if the girl was reported to be Pharaonic cut. "Sunna cut" is a dummy variable =1 if the girl was reported to be Sunna cut. "Uncut" is a dummy variable =1 if the girl was reported to be uncut. "Uncut nor planned to be" is a dummy variable =1 if the girl was reported to be uncut and planned not to be cut. Column 1 provides the mean and standard deviation of the relevant variables for respondents whose social desirability score is above the median; column 2 for below-median. Column 3 provides the p-value for the null hypothesis that the difference between columns 1 and 2 is equal to 0. Column 4 shows the normalized difference between columns 1 and 2.

Table A.4: Relationship between $\ln(\sigma_M/\sigma_L)$ and $\ln(\sigma_M/\sigma_H)$ and state.

	Log-ratio Sunna flow to Pharaonic flow	Log-ratio Sunna flow to Uncut flow
Sunna stock	11.845***	3.689***
	(0.669)	(0.577)
Uncut stock	-41.073***	-106.154***
	(15.259)	(18.099)
Constant	-2.397***	3.676***
	(0.171)	(0.244)
Adjusted R^2	0.917	0.616
Observations	49	38

Notes: Observations are years. The number of observations is smaller than the total number of years in our sample because the dependent variable is not defined for $\sigma_H = 0$ and there are 11 observations (years) for which this is the case. Sunna stock is the share of women aged 10 or more who are Sunna cut in a given year. Uncut stock is the share of women aged 10 or more who are uncut in a given year. Sunna flow is the share of girls who turned 10 in a given year and were Sunna cut at some point in their lives. Pharaonic flow is the share of girls who turned 10 in a given year and were Pharaonic cut at some point in their lives. Uncut flow is the share of girls who turned 10 in a given year, were uncut and planned to remain uncut. Stocks and flows are 3-year moving averages for each year, from 1971 to 2019. They comprise female respondents born after 1960 and all respondents' daughters aged 0 to 18.

B Alternative estimation of stocks and flows

Our survey data comprises information on the year of birth, the cutting year, and the type of cut of female respondents and of all respondents' daughters aged 0 to 18. In order to test whether Sunna is a stepping stone we need to compute the stocks p and flows σ for each action (Pharaonic, Sunna, Uncut) at different points in time. This involves two challenges. First, while we do observe the year of cutting for women who were cut, we do not know in which year it was decided that uncut women would remain uncut. This implies that it is hard to determine the flow of uncut women in any given year based on survey responses. Second, while the cutting year indicates when girls were cut, the timing of parents' decision might precede it.

To address these challenges, in the main body of the paper we assumed that all girls are cut (or decided to be uncut) at the age of ten. This is approximately one year after the average age of cutting in our data, hence we can assume that parents have decided whether to cut their daughters (and with which type of cut) by then.[7] To compute stocks and flows using this method, we ignore girls from their birth until age nine. Each girl enters the flow for their type of cut precisely *in* the year she turns ten and enters the stock of her type of cut *from* the year she turns ten. The rationale behind this method is that parents get one chance of choosing their daughters' cutting status –as in the model– and that when the time for deciding comes, they take into account other parents' choices for the same 'cohort' of girls (these choices may be actions already taken or intentions to cut/not cut, as each parent gets one opportunity to choose at the relevant age).

As an alternative to the method used in the main text, we could use girls' actual cutting ages. In this alternative method, we still need to set a cutoff age from which girls can be accounted in stocks and flows (e.g., girls below the cutoff age are too young to influence agents' decision of cutting).[8] Then if survey responses indicate that a girl was cut after the cutoff age, she is considered uncut until she is actually cut. In other words, such a girl would first enter the stock of uncut girls and then enter the flow of her type of cut once in the year in which she is cut according to the survey, and the stock of her type of cut from that year on. If a girl was cut *before* the cutoff age, she is considered to be cut at the cutoff age and she is not accounted in any stock or flow until she reaches this age.[9] While this method may seem more natural, as it reflects the actual cutting year of each girl, it has two shortcomings. First, by attributing different actions to a girls at different points in time (e.g., first uncut and then cut), it is as if parents got more than one chance of choosing their daughters' cutting status – which is not what we assume in the model. Second, in reality parents may take into account other parents' *intentions*, which means that considering uncut someone who will be cut, say, the following year (and whose parents have communicated that intention) may not reflect the information set of the decision makers.

Figures B.1 and B.2 reproduce Figures 6 and 7 in the main text, using this alternative approach. As one can see, the differences are minimal. Therefore it is not surprising that, when we implement the three tests for stepping stone proposed in the paper, we reach the same conclusion that Sunna is absorbing.

Table B.1 Shows the regression output corresponding to Table 2: the coefficient on Pharaonic stock is still positive and significant, hence the reverse triangle inequality fails. Figure B.3 shows the results of the extrapolation test and confirms that action M (i.e.,

7 We also applied the same method using alternative age cutoffs, e.g., 9, 11, or 12 and our results are robust.

8 The reason we set this cutoff is that we need to establish a time when it is decided that uncut girls will remain uncut.

9 Setting the cutoff age at 10 means that, if a girl was cut at age 9, she will not be accounted for until she turns 10 – and she will enter the stock of cut women when she turns 10. If a girl was cut at 11, she will enter the stock of uncut women when she is 10, and when she turns 11 she will move to the stock of cut women.

Sunna) is absorbing according to our simulations. Finally, Figure B.4 shows the results of the equal flow test (analogous to Figure 11 in the text). Again, the results are very similar to those in the text.

Figure B.1: Stocks, or empirical approximation of the state, over time (alternative estimation of stocks and flows).

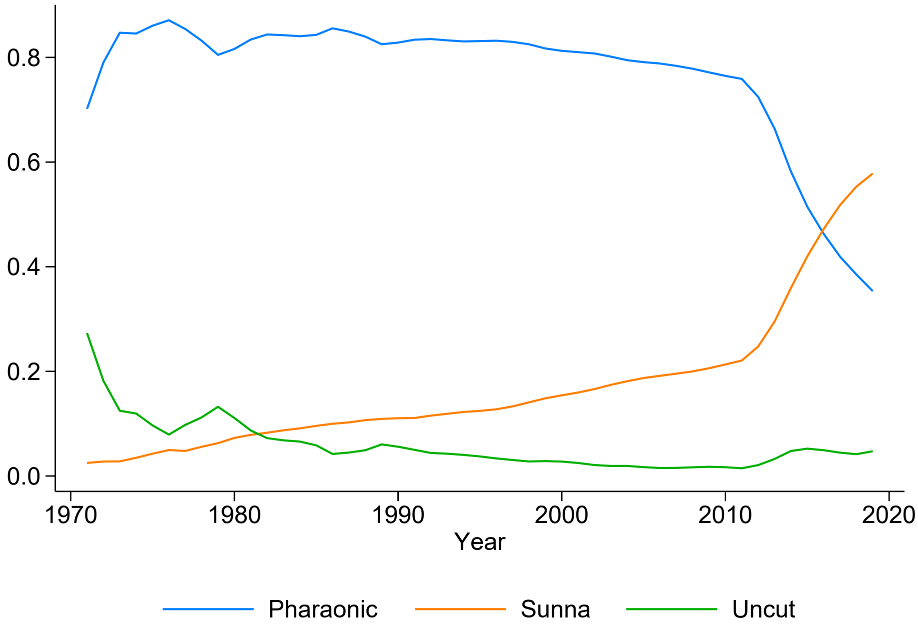


Figure B.2: Flows, or empirical approximation of the choice probabilities, over time (alternative estimation of stocks and flows).

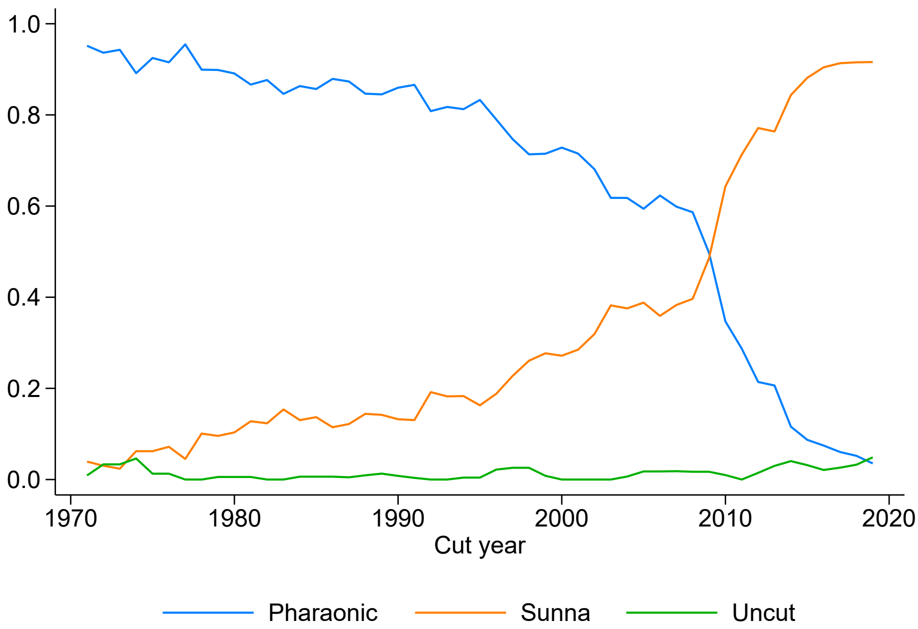


Figure B.3: Extrapolation test (alternative estimation of stocks and flows).

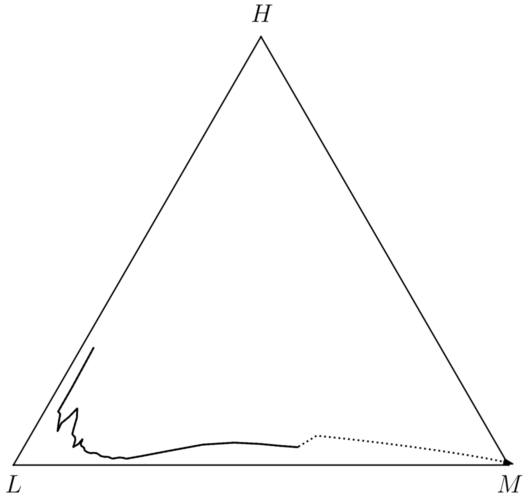


Table B.1: Relationship between $\ln(\sigma_H/\sigma_M)$ and state (alternative estimation of stocks and flows)

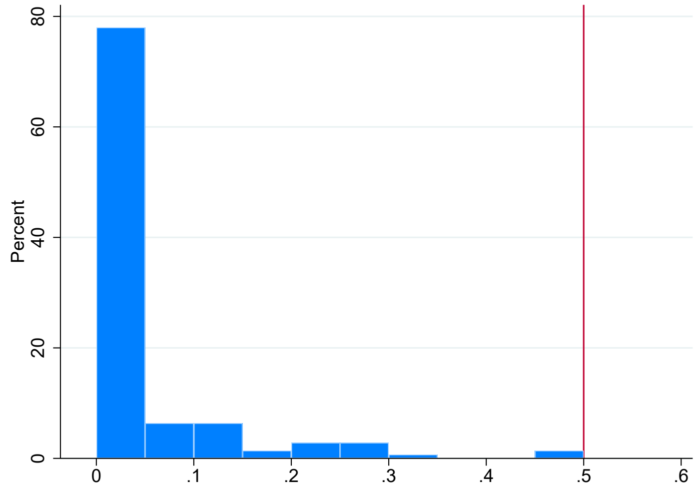
	Log-ratio Uncut flow to Sunna flow				
Uncut stock	13.013**	9.366	10.115	19.115**	19.072**
	(4.955)	(6.498)	(9.120)	(2.702)	(3.779)
Pharaonic stock	1.652**	5.577***	5.287***	3.278	3.788
	(0.688)	(0.804)	(1.074)	(2.265)	(2.499)
Constant	-4.812***	-6.723***	-5.835***	-6.074*	-6.610*
	(0.492)	(0.689)	(0.868)	(1.614)	(1.669)
Sample	Year	Year-District	Year-District	Year-Clan	Year-Clan
District F.E.	No	No	Yes	No	No
Clan F.E.	No	No	No	No	Yes
Adjusted R^2	0.457	0.449	0.565	0.489	0.521
Observations	38	59	59	60	60

Notes: Observations are years (column 1) and year-district (columns 2–3). Uncut stock is the share of uncut women aged 10 or more in a given year. Pharaonic stock is the share of Pharaonic cut women aged 10 or more in a given year. Uncut flow is the share of girls who were chosen to remain uncut in a given year. Sunna flow is the share of girls who were chosen to be Sunna cut in a given year. Stocks and flows are 3-year moving averages for each year, from 1971 to 2019. They comprise female respondents and all respondents' daughters aged 0 to 18.

C Robustness: sample restricted to 1990-2020

As discussed in Section 3.2, there is anecdotal evidence that societal changes in the 1990s affected the dynamics of FGC. In the context of our model, this may imply that a shift in parameters occurred in the 1990s. To check that our results are robust to this possibility, we conducted the analysis using only data from 1990 onward. We found

Figure B.4: Frequency of $\sigma_H/(\sigma_H+\sigma_M)$ values across communities (alternative estimation of stocks and flows).



that this did not impact the results and that our estimates remain similar in magnitude, despite being more noisy. Table C.1 shows that the reverse-triangle-inequality test still fails directionally, although the relevant coefficient is no longer significant. Figure C.3 shows that the extrapolation test still predicts that Sunna is absorbing.

Figure C.1: Stocks, or empirical approximation of the state, over time (restricted to 1990-2020).

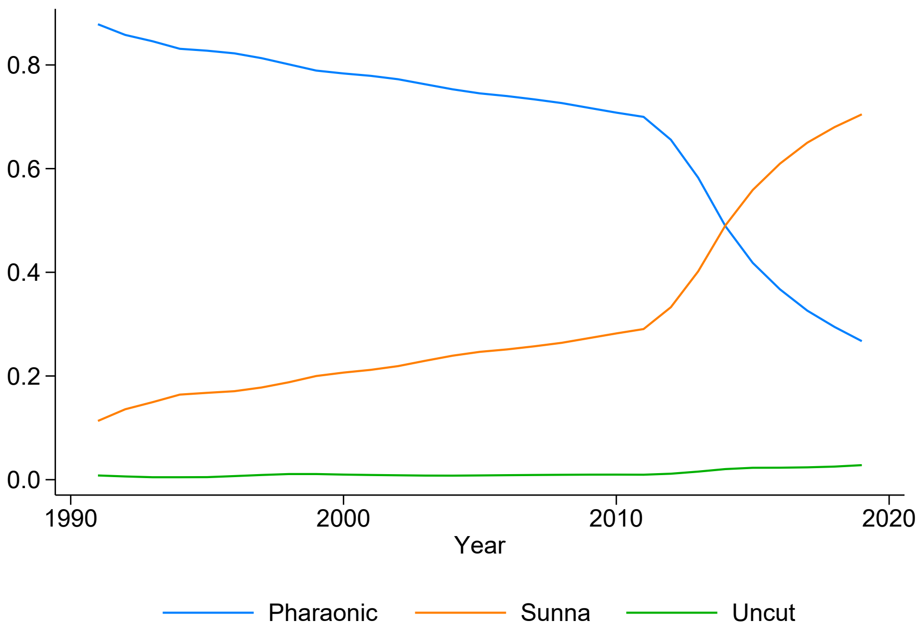


Figure C.2: Flows, or empirical approximation of the choice probabilities, over time (restricted to 1990-2020).

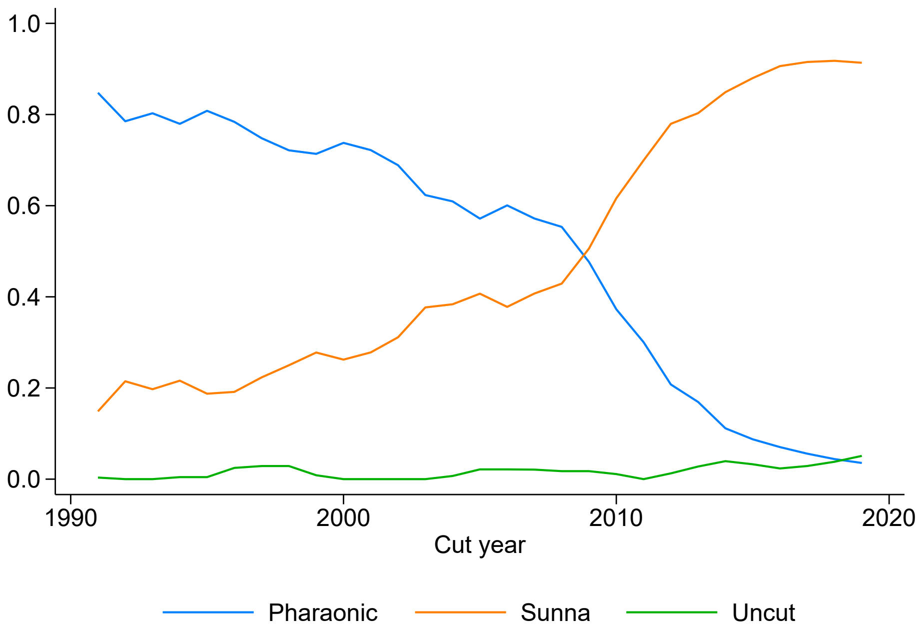


Figure C.3: Extrapolation test (restricted to 1990-2020).

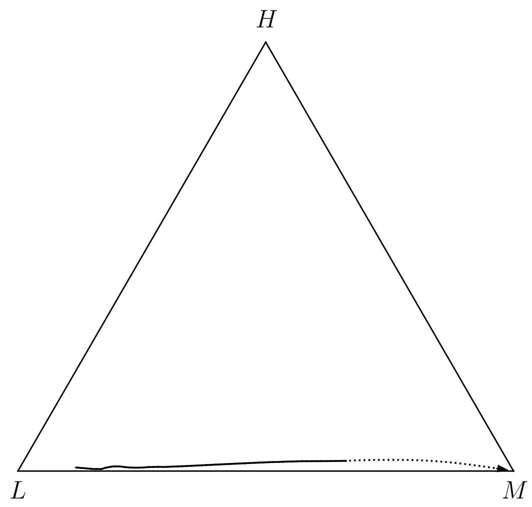


Table C.1: Relationship between $\ln(\sigma_H/\sigma_M)$ and state (alternative estimation of stocks and flows)

	Log-ratio Uncut flow to Sunna flow				
Uncut stock	146.497	35.164***	52.314***	90.880***	85.348**
	(97.650)	(4.111)	(6.498)	(1.150)	(9.606)
Pharaonic stock	5.481	4.368***	6.374***	5.688***	5.787***
	(3.866)	(0.595)	(0.709)	(0.247)	(0.063)
Constant	-8.735**	-5.984***	-8.410***	-8.101***	-8.272***
	(3.730)	(0.429)	(0.677)	(0.300)	(0.069)
Sample	Year	Year-District	Year-District	Year-Clan	Year-Clan
District F.E.	No	No	Yes	No	No
Clan F.E.	No	No	No	No	Yes
Adjusted R^2	0.081	0.564	0.711	0.556	0.584
Observations	22	48	48	42	42

Notes: Observations are years. Uncut stock is the share of uncut women aged 10 or more in a given year who were also planned to remain uncut. Pharaonic stock is the share of Pharaonic cut women aged 10 or more in a given year. Uncut flow is the share of girls who turned 10 in a given year and were uncut and planned to remain uncut. Sunna flow is the share of girls who turned 10 in a given year and were Sunna cut at some point in their lives. Stocks and flows are 3-year moving averages for each year, from 1991 to 2019. They comprise female respondents and all respondents' daughters aged 0 to 18.

D Parameter stability, survey-based proxies

An important assumption in our analysis is that the parameters are roughly stable over time. If instead they change drastically over time, then the dynamics produced may be different from those predicted by the model. While we cannot test this assumption directly, we can gain some insights using proxies elicited through our survey.

To obtain proxies for the social-pressure parameters, we asked respondents a series of questions specifically designed to elicit pairwise comparisons between Pharaonic, Sunna, and Uncut.

We presented each respondent with different situations where hypothetical parents have cut their daughter with a certain type of FGC, but their daughter-in-law may have a different type of FGC. We chose to frame this in the context of marriage choices because most of the literature on FGC highlights consequences in the marriage market as a potential cost for deviating from prevailing norms (see Wagner, 2015, for cross-country evidence). The idea is that, in each vignette, the hypothetical parents' choice about their daughter would reveal their own preferred action. The daughter-in-law represents someone these parents would also care about (e.g., in terms of reputation concerns), hence the comparison is made between two scenarios that both affect the hypothetical family.

For example, we asked: "Suppose a mother and father in your community chose Pharaonic circumcision for their daughter, but their son wants to marry a girl with Sunna.

How would these parents feel?" The possible answers were: "happy", "indifferent", or "unhappy". We repeated the same question for Sunna vs. Uncut and for Pharaonic vs. Uncut. Appendix Figure D.1 provides a visual summary of the responses. The rationale underlying these questions about hypothetical parents is not to ask respondents how *they themselves* would feel, but to elicit second-order beliefs about the attitudes of other community members. This is because it is other people's views that matter if we want to measure expected sanctions for noncompliance with local norms.[10]

Table D.1 reports the results of regressing of the proxies for social-pressure parameters on respondents' age. The estimates suggest that age is not correlated with beliefs about social pressure.

Table D.1: Correlation between proxies for s_{ij} and respondents' age

	s_{LH}		s_{LM}		s_{MH}	
	(1)	(2)	(3)	(4)	(5)	(6)
Respondent's age	0.000	-0.003	-0.000	0.005	-0.000	-0.006
	(0.005)	(0.003)	(0.005)	(0.004)	(0.006)	(0.004)
Respondent's age^2	0.000	0.000	0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	3385	3385	3437	3437	3339	3339
Outcome mean	0.693	0.693	0.537	0.537	0.553	0.553
Community F.E.s	No	Yes	No	Yes	No	Yes

Notes: The sample includes all respondents. Regressions in columns 2, 4 and 6 control for community fixed effects. Standard errors are clustered by respondent's age cohort.

10 Bicchieri (2005, p. 15 and ff.) emphasises that this is a key feature of social norms.

Figure D.1: Survey-based proxies for s_{ij} parameters

(a) Pharaonic v.s. Uncut ($\hat{s}_{LH}$)

(b) Pharaonic v.s. Sunna ($\hat{s}_{LM}$)

(c) Sunna v.s. Uncut ($\hat{s}_{MH}$)