Supplemental Appendix
A Stepping-Stone Approach to Norm Transitions

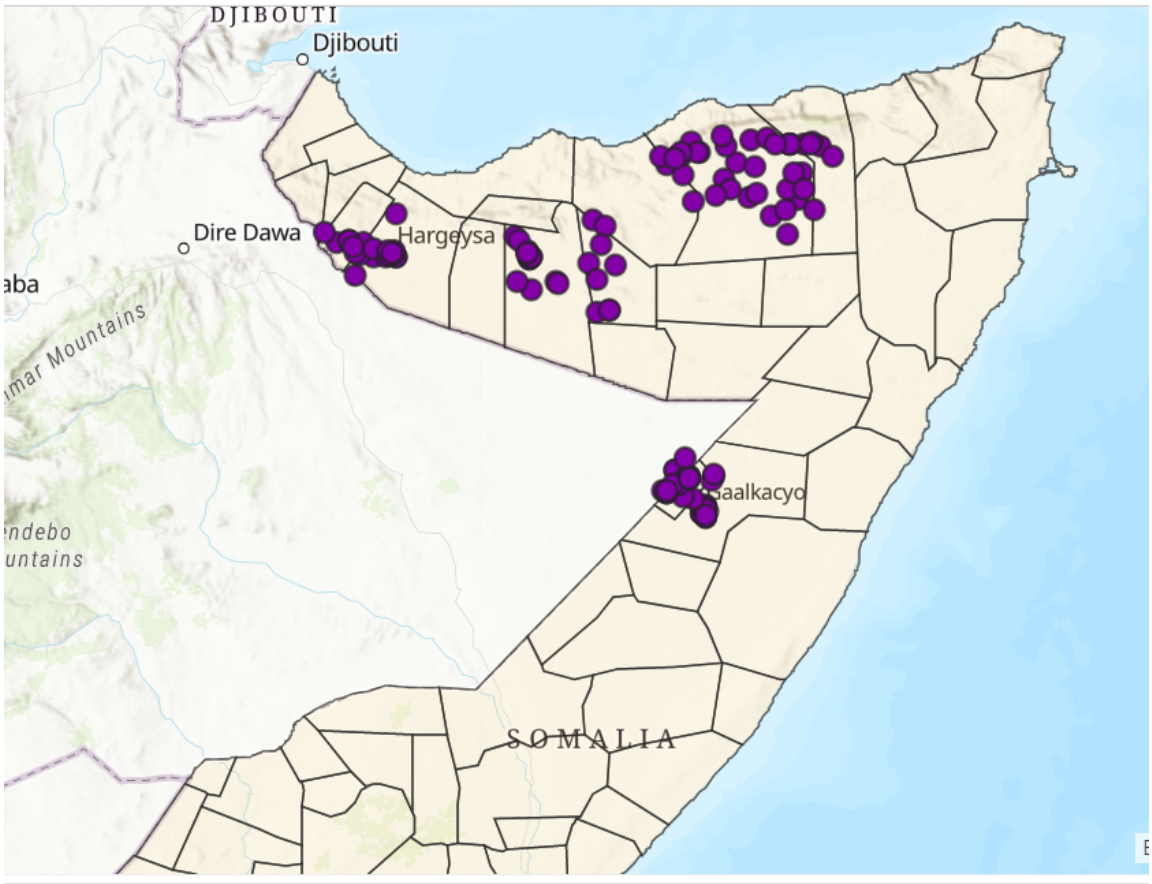
February 2025

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]BIGD, BRAC University; munshi.sulaiman@bracu.ac.bd.
[6]University of Oxford; peyton.young@economics.ox.ac.uk.

A Additional figures and tables

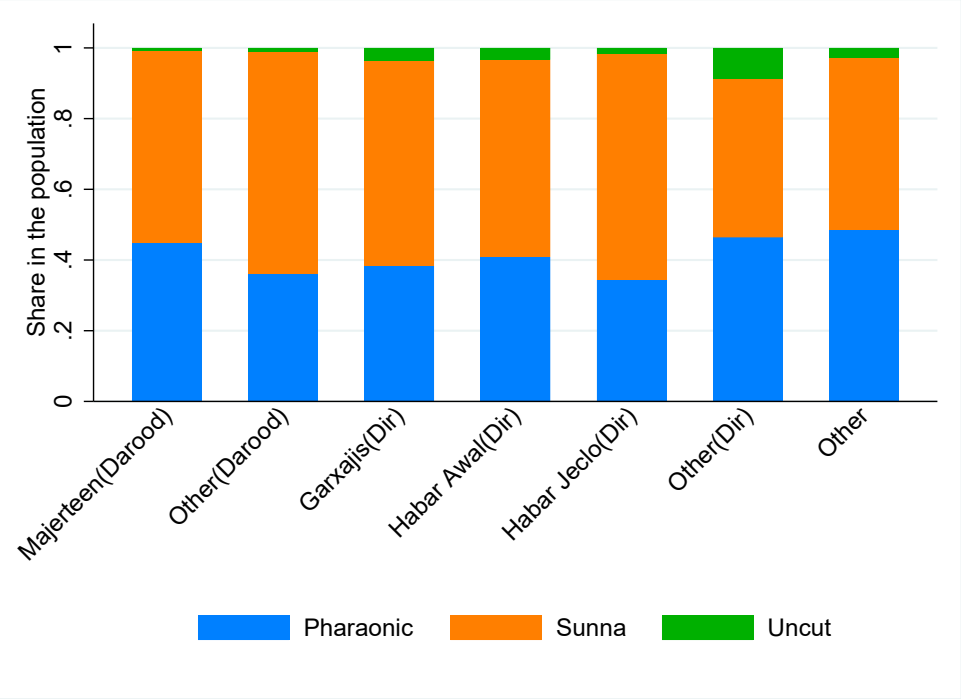
Figure O.1: Location of the Study



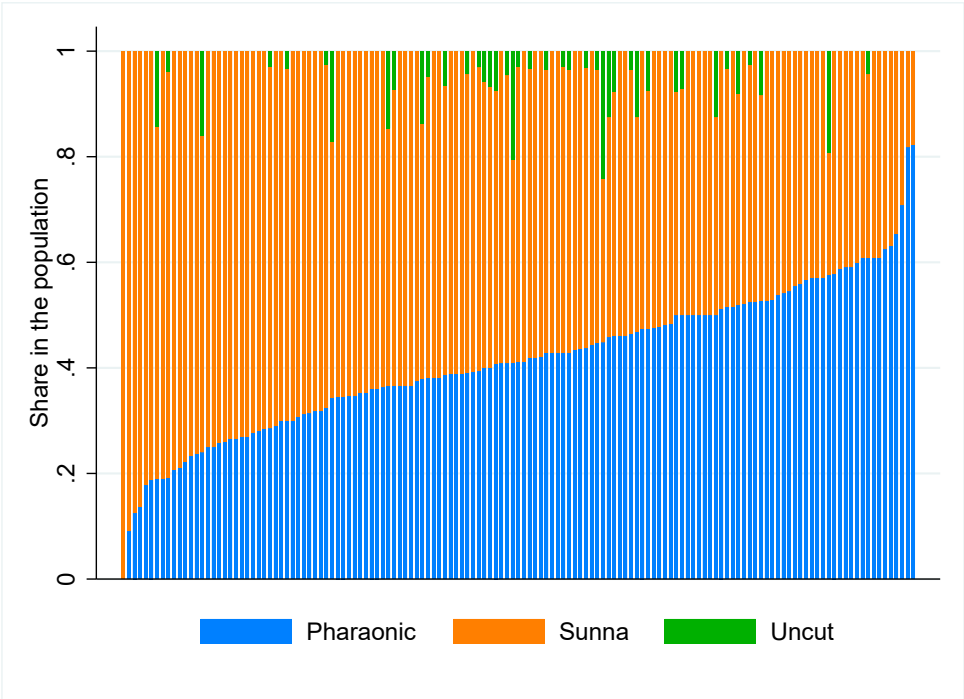
Notes: The map shows the location of communities included in the study, along with district boundaries of Somalia.

Figure O.2: Type of FGC

(a) By sub-clan

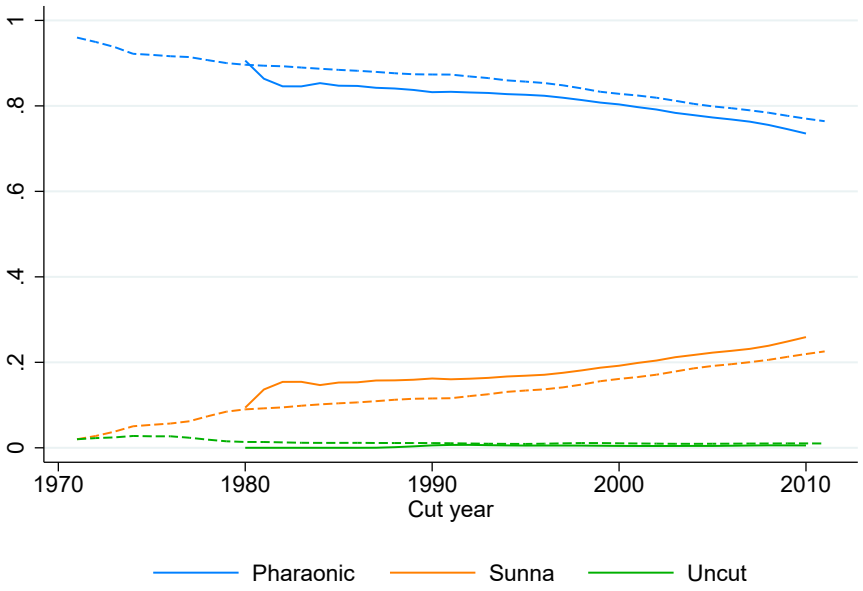


(b) By community

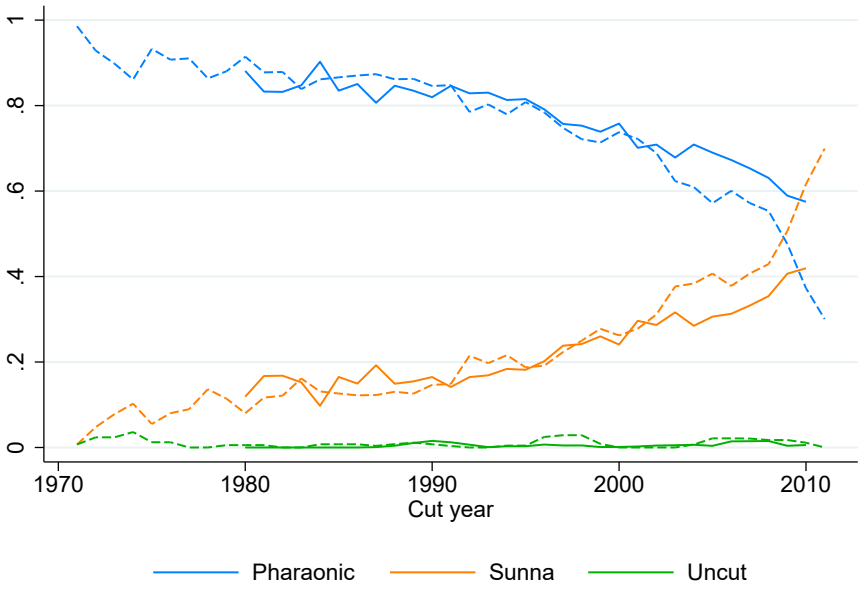


Notes: Sample is restricted to respondents and their daughters aged 12 or older.

Figure O.3: Stocks and flows in SHDS 2020 vs. our sample



(a) Stocks



(b) Flows

Notes: Stocks and flows for SHDS 2020 (Federal Government of Somalia, 2020) are represented in solid lines and include respondents aged 18 to 49. Stocks and flows for our sample of respondents (excluding daughters) are represented in dashed lines. For comparability, we restrict the SHDS data to non-urban (rural and nomadic) respondents from similar regions as our study sample [Woqooyi Galbeed (Somaliland: Hargeisa), Togdheer (Somaliland: Burao), Sanaag (Somaliland: Badhan, Erigavo), and Mudug (Puntland: Galkayo, Galdogob)]. Sample weighting is used for SHDS 2020 data.

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

	(1) *Pharaonic*		*(2)* *Sunna*	
	Mean	SD	Mean	SD
Panel A: For female respondents				
Any health complication (Yes=1)	0.638	(0.481)	0.115	(0.320)
N	1,282		395	
Panel B: For female daughters				
Any health complication (Yes=1)	0.284	(0.452)	0.017	(0.130)
N	203		2,743	
Panel C: Perceived health complications				
Any complication	0.754	(0.431)	0.060	(0.238)
Infection	0.403	(0.491)	0.041	(0.199)
Bleeding	0.570	(0.495)	0.029	(0.169)
Difficulty in delivery	0.611	(0.488)	0.034	(0.182)
Reduction in sexual feeling	0.462	(0.499)	0.031	(0.172)
Difficulty in penetration	0.329	(0.470)	0.004	(0.064)
Other	0.014	(0.117)	0.004	(0.064)
Number of complications	2.389	(1.724)	0.159	(0.719)
N	932		3,620	

Notes: In Panel A, the sample in column 1 (2) includes female respondents who reported that they were cut Pharaonic (Sunna). In Panel B, the sample in column 1 (2) includes respondents who reported that their daughter/s was/were cut Pharaonic (Sunna). In Panel A (B), "Any health complication" is a dummy variable =1 if the respondent (or respondent's daughter) experienced any health complications because of FGC. In Panel C, the sample in column 1 (2) includes respondents who reported that Pharaonic (Sunna) cut was practiced in their community. The variables are based on responses to the question "What are some of the consequences of Pharonic (Sunna) circumcision?". The respondents could report "No consequences" or report multiple options among "Infection", "Bleeding", "Difficulty in delivery", "Reduction in sexual feeling", "Difficulty in penetration", "Other". "Any complication" is a dummy variable =1 if the respondent reported any consequence, following variables are dummies =1 if they reported the respective consequence, "Number of complications" is the number of consequences reported by the respondents.

Table O.2: Summary statistics

	(1) *Pharaonic*		(2) *Sunna*	
	Mean	SD	Mean	SD
Age at FGM	9.014	2.078	8.279	1.810
Decision to cut by (not mutually exclusive):				
Mother	0.724	0.448	0.857	0.350
Father	0.522	0.501	0.543	0.498
Grandmother	0.084	0.278	0.037	0.189

Notes: Sample restricted to female respondents and respondents' daughters aged 0–18 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.

Table O.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	0.181	0.192	0.214	-0.020
	(0.385)	(0.394)		
Sunna cut	0.403	0.394	0.419	0.013
	(0.491)	(0.489)		
Uncut	0.416	0.413	0.855	0.003
	(0.493)	(0.493)		
Uncut nor planned to be	0.079	0.072	0.252	0.019
	(0.270)	(0.259)		

Notes:
The sample includes all daughters aged 0 to 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 O.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	11.886	3.698
	(0.672)	(0.579)
Uncut stock	-40.873	-106.023
	(15.244)	(18.058)
Constant	-2.404	3.673
	(0.171)	(0.244)
Adjusted R^2	0.917	0.615
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.[1] 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).[2] 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.[3] While this method may seem more natural, as it reflects the actual cutting year

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

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

[3] 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.

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 O.4 and O.5 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 O.5 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 O.6 shows the results of the extrapolation test and confirms that action M (i.e., Sunna) is absorbing according to our simulations. Finally, Figure O.7 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 O.4: Stocks, or empirical approximation of the state, over time (alternative estimation of stocks and flows)

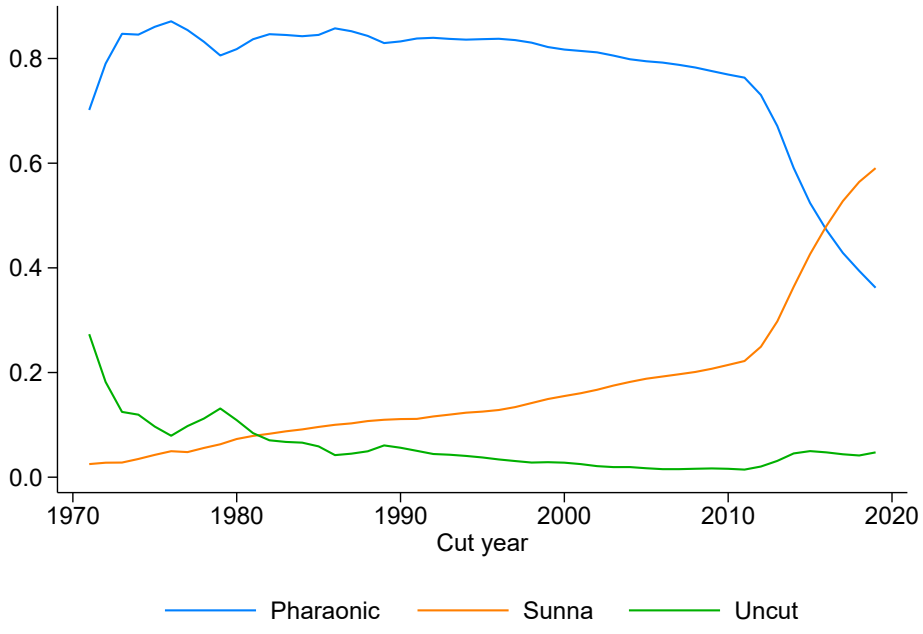


Figure O.5: Flows, or empirical approximation of the choice probabilities, over time (alternative estimation of stocks and flows)

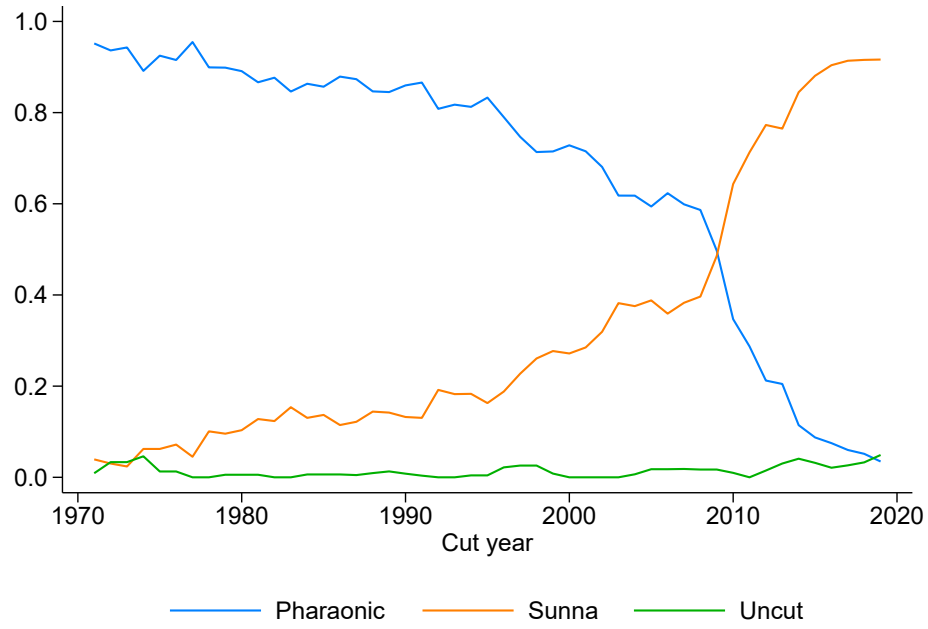


Figure O.6: Extrapolation test (alternative estimation of stocks and flows)

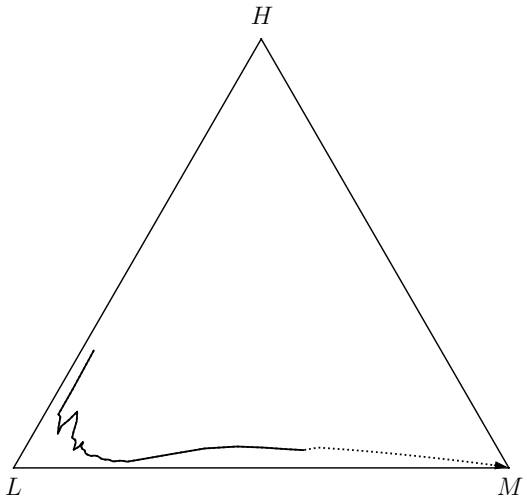


Table O.5: Relationship between $\ln(\sigma_H/\sigma_M)$ and state (alternative estimation of stocks and flows)

	Log-ratio Uncut flow to Sunna flow				
	(1)	(2)	(3)	(4)	(5)
Pharaonic stock	1.614	5.629	5.307	3.276	3.789
	(0.691)	(0.813)	(1.067)	(0.784)	(0.835)
	[0.728]	[0.795]	[1.087]	[0.784]	[0.861]
Uncut stock	13.125	9.971	10.844	19.619	19.483
	(5.036)	(7.153)	(10.250)	(4.473)	(4.655)
	[4.857]	[7.677]	[13.830]	[4.364]	[4.439]
Constant	-4.790	-6.811	-5.905	-6.101	-6.628
	(0.489)	(0.714)	(0.937)	(0.656)	(0.699)
	[0.530]	[0.723]	[1.195]	[0.655]	[0.723]
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	0.460	0.452	0.574	0.496	0.526
Observations	38	59	59	60	60

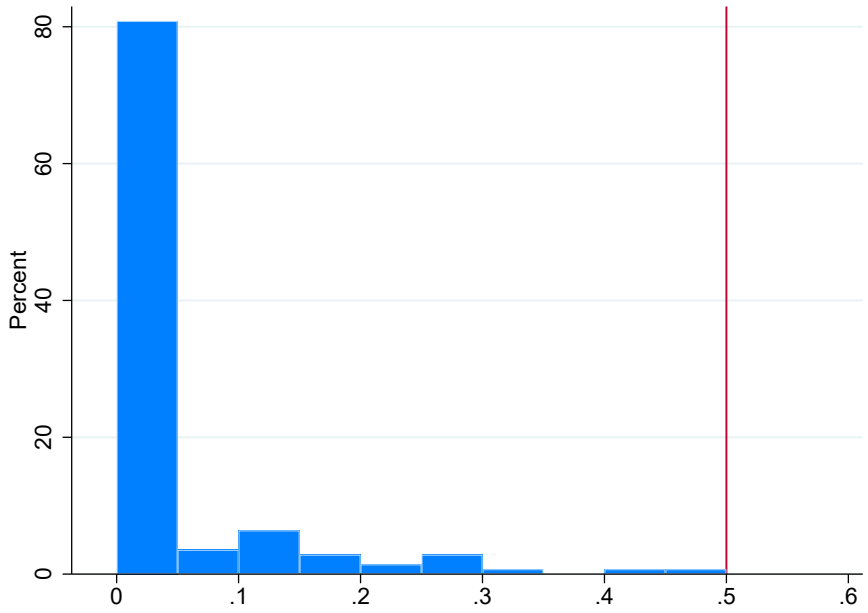
Notes: Observations are years (column 1), year-district (columns 2–3), and year-clan (columns 4–5). Uncut stock is the share of uncut women aged 10 or more in a given year. Pharaonic stock is the share of Pharaonic cut women aged 10 or more in a given year. Uncut flow is the share of girls who were chosen 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. OLS estimates with robust standard errors in parenthesis and bootstrapped standard errors in square brackets (calculated using 1,000 repetitions). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C Parameter stability, survey-based proxies

This section presents suggestive evidence on the stability of model parameters, based on survey-based proxies for the social pressure parameters.

To obtain proxies for s_{ij}, 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 had 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.

Figure O.7: Frequency of $\sigma_H/(\sigma_H+\sigma_M)$ values across communities (alternative estimation of stocks and flows)

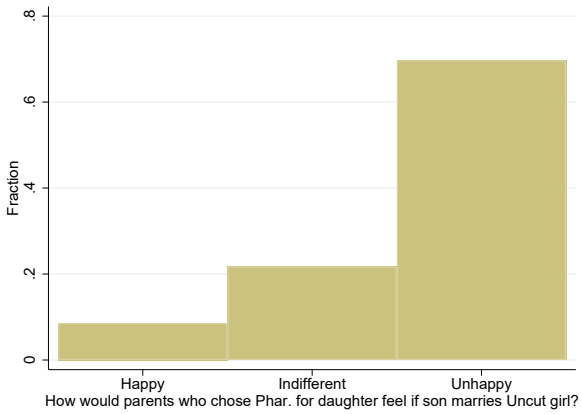


For example, we asked: "Suppose a mother and father in your community chose Pharaonic circumcision for their daughter, but their son wants to marry a girl with Sunna. How would these parents feel?" The possible answers were: "happy", "indifferent", or "unhappy". We repeated the same question for Sunna vs. Uncut and for Pharaonic vs. Uncut. Appendix Figure O.8 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.[4]

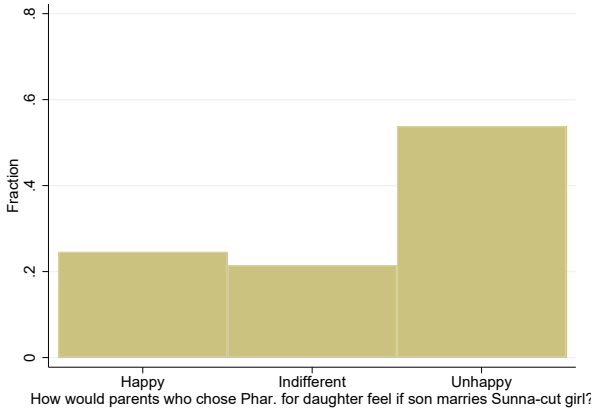
[4] Bicchieri (2005, p. 15 and ff.) emphasizes that this is a key feature of social norms.

Figure O.8: 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 ($\hat{s}_{MH}$)

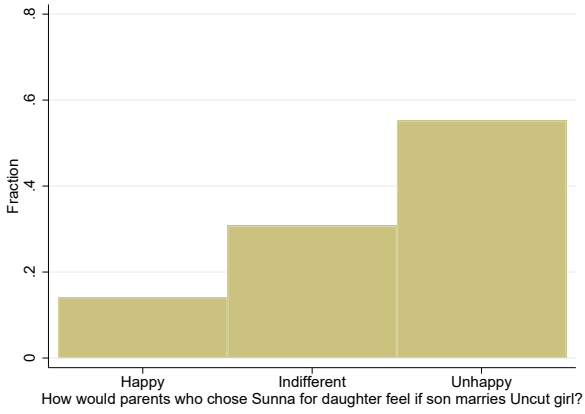


Table O.6 reports the results of regressing the proxies for social-pressure parameters on respondents' age. The estimates suggest that age is not robustly correlated with beliefs about social pressure. Even in cases where a correlation is detected (i.e., two of the linear specifications without controls), the magnitude of the coefficient is very small: the estimated value of 0.002 implies that a difference in age of 20 years would increase the probability that parents are unhappy by 4 percentage points, relative to a mean of 69 percentage points for s_{LH} and 55 percentage points for s_{MH}.

Table O.6: Correlation between proxies for s_{ij} and respondents' age

	s_{LH}			s_{LM}			s_{MH}		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Respondent's age	0.002	-0.002	-0.006	-0.000	0.005	0.006	0.002	-0.006	-0.006
	(0.001)	(0.006)	(0.005)	(0.001)	(0.006)	(0.007)	(0.001)	(0.007)	(0.005)
	[0.001]	[0.005]	[0.005]	[0.001]	[0.005]	[0.006]	[0.001]	[0.005]	[0.005]
Respondent's age sq. [(a)]		0.005	0.009		-0.007	-0.007		0.010	0.009
		(0.008)	(0.006)		(0.007)	(0.008)		(0.009)	(0.005)
		[0.006]	[0.006]		[0.007]	[0.007]		[0.007]	[0.007]
Observations	3289	3289	3289	3339	3339	3339	3245	3245	3245
Outcome mean	0.690	0.690	0.690	0.534	0.534	0.534	0.552	0.552	0.552
Controls	No	No	Yes	No	No	Yes	No	No	Yes
Adj. R^2	0.002	0.002	0.130	-0.000	-0.000	0.058	0.001	0.001	0.175

Notes: In parenthesis: standard errors clustered at the age level. Controls include district fixed effects, dummy =1 if respondent is female, dummy =1 if respondent has formal education, =1 if respondent did Koranic studies, number of children at baseline, wealth index, and subclan fixed effects, dummy = 1 if respondent is Pharaonic cut and =1 if cut status is missing. In brackets: bootstrapped standard errors with 500 iterations. In curly brackets: standard errors computed through jackknife. (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$)

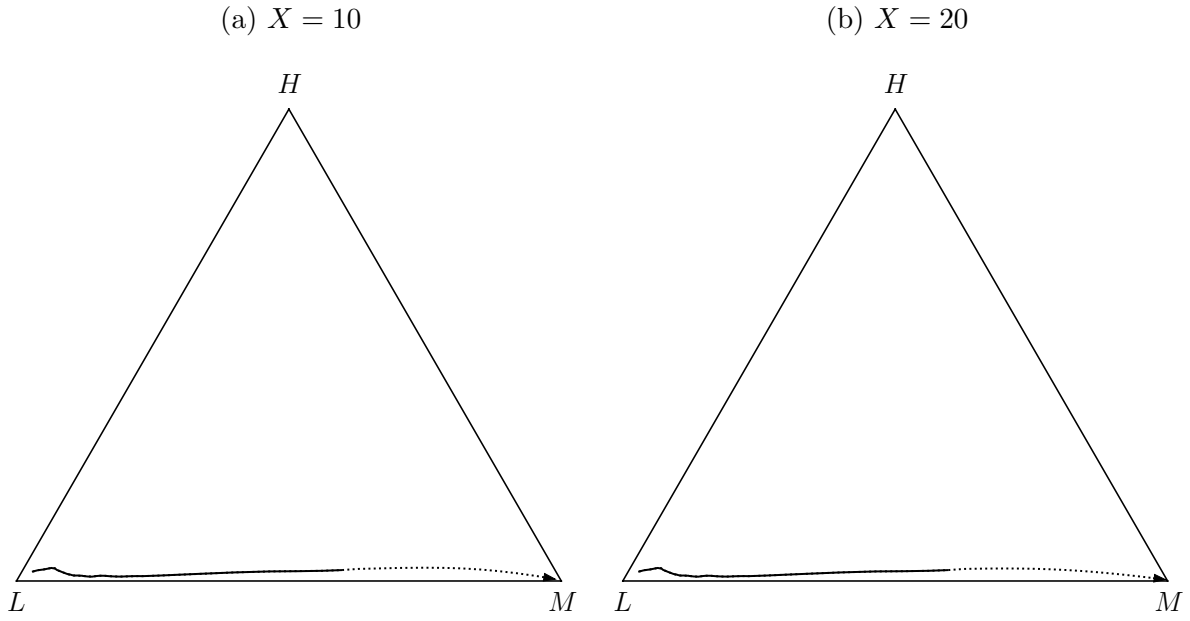
(a) Coefficient multiplied by 100.

D Social utility restricted to recent actions

In the main analysis, we assume that social utility depends on the actions of all other agents in the population. This includes actions that were taken decades ago, which may be unrealistic. For example, it could be the case that a woman who was cut with Pharaonic in the 1980s but chose Sunna for her daughter in the 2000s would not exert social pressure on others choosing Sunna today, yet our analysis assumes that she would.

To show that our empirical results do not depend on this assumption, we replicate the analysis assuming that agents making a choice in a given year only take into account actions taken not too long before. In particular, we construct 'rolling' stocks that only include choices made in the most recent X years. We then conduct the reverse-triangle-inequality and extrapolation tests using these rolling stocks. Our results are qualitatively unchanged for $X = 10, 20$: the analysis continues to predict that Sunna will not be a stepping stone. The two panels of table O.7 replicate the reverse-triangle-inequality test of table 1 for $X = 10, 20$. As in the main analysis, the coefficient on p_L is positive and significant in all specifications, suggesting that the reverse triangle inequality fails to hold. Similarly, figure O.9 replicates the extrapolation test of figure 8a for the same two time intervals and predicts that Sunna will be absorbing.

Figure O.9: Extrapolation test when social utility depends on recent actions



Notes: The solid line represents the path of stocks from observed data, hence they are the same in (a) and (b). The dotted line represents extrapolated based on rolling stocks for $X = 10$ in (a) and $X = 20$ in (b).

Table O.7: Reverse-triangle-inequality test when social utility depends on recent actions

(a) $X = 10$

	Log-ratio Uncut flow to Sunna flow				
	(1)	(2)	(3)	(4)	(5)
Pharaonic stock	3.438	3.604	4.038	3.961	3.995
	(0.477)	(0.347)	(0.268)	(0.182)	(0.033)
	[0.494]	[0.372]	[0.414]	[0.525]	[0.507]
Uncut stock	84.160	23.974	24.107	54.992	50.541
	(13.217)	(2.494)	(7.930)	(4.442)	(8.099)
	[13.785]	[2.606]	[7.982]	[8.114]	[8.719]
Constant	-6.544	-5.073	-5.243	-6.324	-6.418
	(0.500)	(0.328)	(0.776)	(0.147)	(0.134)
	[0.517]	[0.276]	[0.906]	[0.468]	[0.463]
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	0.626	0.650	0.707	0.611	0.616
Observations	38	60	60	58	58

(b) $X = 20$

	Log-ratio Uncut flow to Sunna flow				
	(1)	(2)	(3)	(4)	(5)
Pharaonic stock	3.408	3.983	4.471	4.045	4.277
	(0.492)	(0.433)	(0.340)	(0.399)	(0.056)
	[0.513]	[0.465]	[0.544]	[0.628]	[0.625]
Uncut stock	97.375	28.239	28.821	64.115	59.312
	(16.313)	(2.683)	(9.042)	(5.171)	(10.827)
	[16.982]	[3.779]	[10.067]	[11.254]	[10.828]
Constant	-6.855	-5.596	-5.913	-6.688	-7.030
	(0.565)	(0.353)	(0.583)	(0.340)	(0.195)
	[0.586]	[0.365]	[0.935]	[0.605]	[0.618]
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	0.606	0.587	0.640	0.524	0.561
Observations	38	60	60	58	58

References

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

Federal Government of Somalia. 2020. "The Somali Health and Demographic Survey 2020." *https://somalia.unfpa.org/en/publications/somali-health-and-demographic-survey-2020.*

Wagner, Natascha. 2015. "Female Genital Cutting and Long-Term Health Consequences – Nationally Representative Estimates across 13 Countries." *The Journal of Development Studies*, 51(3): 226–246.