Supplemental Information

for

Women's Authority in Patriarchal Social Movements: The Case of Female Salafi Preachers

Richard A. Nielsen[*]

February 20, 2019

Abstract

This online appendix provides supplemental information about some of the analysis in "Women's Authority in Patriarchal Social Movements: The Case of Female Salafi Preachers."

Contents

A Historical Website Data Shows When Female Preachers were Added

Section 4 of the paper gives information about the website that serves as my primary data source. Because of space constraints, I did not have room to include this supporting figure in the text.

This figure shows the number of male and female preachers with writings on saaid.net over time. I gather the data by from website captures on various dates between 2002 and 2018 by the Internet Archive, at archive.org. The figure shows that women are first added in small numbers to the website in 2004, three years after its founding. The number of women preachers on the website rises until 2013, when it plateaus at roughly 40, with only a handful of women joining the website in the last five years.
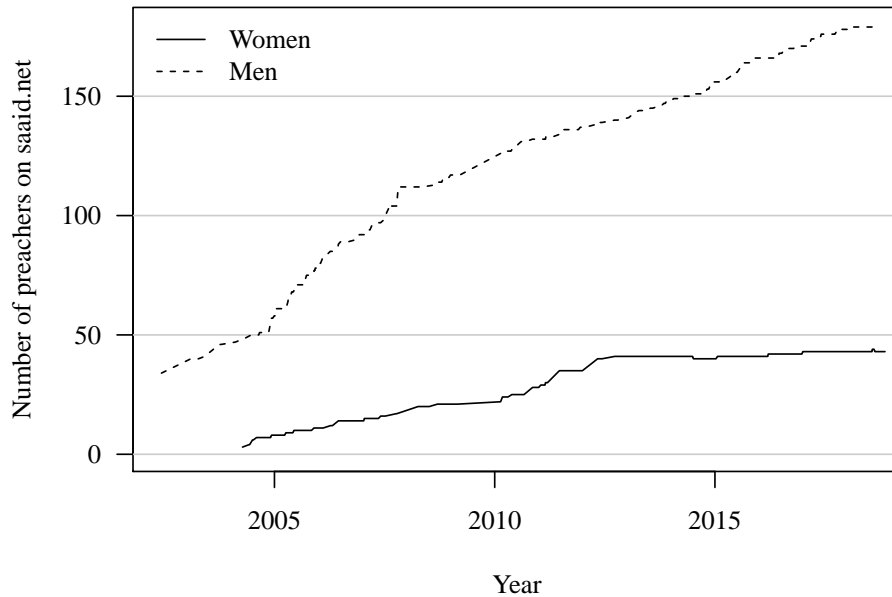

Figure 1: *Male and Female preachers on saaid.net over time*

This figure shows the number of male preachers (dashed line) and female preachers (solid line) on the Salafi-oriented website saaid.net over time. Data come from the Internet Archive, at archive.org.

B Varying the number of topics in the structural topic model

Topic models, such as the structural topic model (STM) I use in the paper, require that the analysis select a fixed number of topics to estimate. There is no single best number of topics, for the same reason that there is no correct answer to the question "how many topics are in this paper?" Estimating more topics will result in more fine-grained distinctions while estimating fewer gives broader topics.

In the paper, I estimated an STM with 10 topics because (a) the model provided insight, (b) the model gave similar results as models with about the same number of topics, (c) because I wanted the broad topics that a 10-topic model would uncover rather than the narrow topics that, say, a 100-topic model would uncover, and (d) 10 is a nice round number. An odd number like 13 might raise suspicion that I'm cherry-picking.

From this 10-topic model, I focused on the fact that women write about *Hadith* and *Salawat* far less frequently than men. To show that this conclusion is not overly dependent on my choice of the number of topics, I re-estimated the same STM specification with 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 75, and 100 topics.
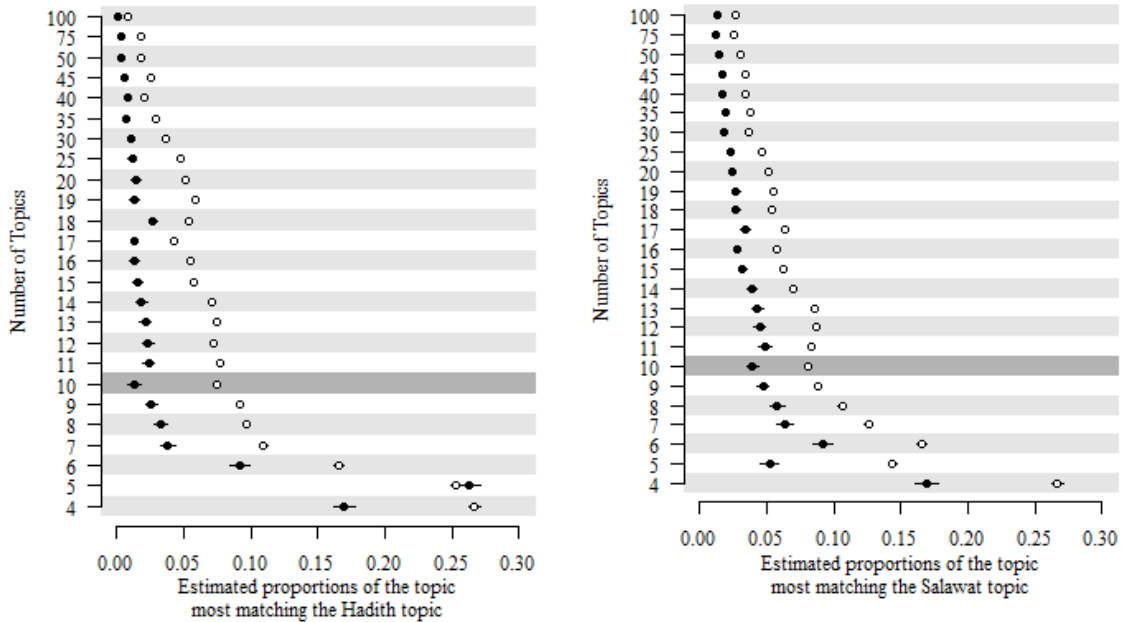

Figure 2: *Hadith and Salawat topics are largely invariant to the number of topics in the STM*

This figure shows how the estimates of topic proportions for the hadith topic (right) and salawat topic (left) change as I vary the number of topics in the STM. Each row shows the coefficients from a topic model with the number of topics indicated by the label. The black discs are the estimated proportion of women's writing devoted to each topic (with lines indicating 95% confidence intervals). The white discs are the estimated proportion of men's writing devoted to each topic. The consistent gap between white discs and black discs shows that men use hadith and salawat more than women.

Examining the results of these 24 models one by one, I find that my conclusions about the *Hadith* and *Salawat* topics are basically invariant to the number of topics. With six topics or more, the words I associate with Hadith and Salawat appear in their own topics and remain in their own topics regardless of how many other topics are in the model.[1] In all cases, the correlation of these topics with author gender strongly supports my claim that women use hadith citations far less frequently than men.

With only four or five topics, the words associated with Hadith and Salawat are lumped together in a single topic with many other themes that are only loosely related. In particular, the 5-5 topic STM suggests that there is no large difference between the men and women for the topic that contains the terms I associate with hadith. However, this does not contradict the findings of the paper. I find that the 5-topic model

---

[1] Aligning topics across runs of a topic model is difficult because topics are distributions over words, but these are not identical from model to model, even if a human would interpret both models to contain the same topics. To identify the hadith and salawat topics across runs, I selected the words that I used most as I was labeling the topics. For the salawat topic, these are the words *ṣalā* and *salam* from "Peace be upon him." For the hadith topic, these words are *hadith*, *rawah* (he narrated), *Hurayra* (a hadith compiler), *Tarmidhi* (a hadith compiler), *Bukhari* (a hadith compiler), *Dawud* (a hadith compiler), *ṣaḥīḍ* (Bukhari's hadith collection), and *da'if* (weak).

4

produces overly broad topics. The words for Hadith are lumped together with other words that are unrelated. A 5-topic model might be useful for other purposes, but it is not useful for isolating words related to hadith citation. Models with 6 or more topics do reliably isolate these words and confirm the result.

C   Stemming and Topic Model Insensitivity to Stemming Choices

For my statistical analysis text analysis, I use routine text analysis pre-processing procedures, adapted for Arabic. The most consequential part of this pre-processing is stemming, which reduces the variety of the texts by combining closely related words with the same morphological stem. For example, an analyst working with English-language text might stem the words "fights," "fought," and "fighting" to the stem "fight."

I stem each of the documents, using a stemmer based on the "Light 10" stemmer for Arabic (**?**). The stemmer is "light," meaning that it only removes common prespecified prefixes and suffixes, along with a list of function words (to, from, for, by). In practice, there are about 20 prefixes/suffixes that account from 95% of the stemming. For example, the definite al- (alif lam) is the most commonly removed prefix. Although infixing is a feature of Arabic, this stemmer does not attempt to remove infixes (because it is more error-prone) and does not recover the triliteral root of each Arabic word (because triliteral roots have broader meanings than I want). Thus, "a book" ("*kitāb*"), "the book" ("*al-kitāb*"), and "her book" ("*kitābuhā*"), are all stemmed to be "book" ("*kitāb*"), but the plural "books" ("*kutub*") is not. Arabic speakers will recognize that this is less than ideal, but Arabic stemming technology is far less developed than English-language stemming.

When an irregular plural is not stemmed to be singular, this does not mean that the information falls out of the dataset. Instead, the irregular plural is retained in the text analysis unchanged. With large amounts of text data like I have in this paper, this does not present a problem for the topic model. The model correctly identifies that "book" and "books" belong in the same topic because they are surrounded by similar words.

One reasonable concern is that becuase men and women use different words, the stemming affects them differentially. Examining the data, I find that the stemming of gendered suffixes follows predictable patters: men use more male-oriented suffixes and women use more female-oriented suffixes. The differences are fairly small, but definitely present. To be clear, the stemmer operates the same on all text. It is the differing word choices of men and women that lead to differences in the rates of particular suffixes. Notably, the rates of the gender-neutral definite article (al-) are identical for men and women.

To be sure that the stemming isn't affecting the result too much, I estimated two additional topic models. In the first, I used an unstemmed version of the corpus. In the second, I stem the corpus, but add the stemmed prefixes and suffixes back into the analysis as their own words. The results of these additional models support my conclusions. The size and key words for some of the substantive topics change somewhat, but there are no new "mssing" topics out of the blue. My analysis focuses most on topics related to form (hadith citation and salawat). These remain effectively identical, regardless of stemming.

D   Statistical models showing the hadith citation gap

Section 4 of the paper presents evidence that men use citations more than women. The following material was originally an extension of that section, but I moved the additional analysis here to make way for other material while adhering to journal length requirements.

To maintain the flow of the argument, I reproduce some text and a figure from the article, and then continue where the paper left off.

"The hadith citation gap between Salafi men and women is not due to confounding factors such as document topic or the author's scholarly training. It may be that writing on certain topics typically involves fewer hadith citations, and that these topics are disproportionately favored by women. To test this, I identify the main substantive topic of each document. Then, I collect all documents with the same main topic, divide them by author gender, and calculate the rate of citation to eight prominent hadith collections: al-Bukhari, Ibn Taymiyya, al-Bayhaqi, Abu Hurayra, al-Tabarani, Abu Dawud, Sahih Muslim, and al-Tarmidhi. Figure 3 displays the results graphically. Topics are plotted separately for men (white discs) and women (black discs), with the size of the disc indicating the proportion of documents by men or women respectively associated with each main topic. Topics are plotted on the x-axis according to the difference in topic proportions for women and men as estimated by the structural topic model above. Thus, topics listed to the right of zero in the figure are more prevalent in women's writing, while topics to the left of zero are more prevalent in men's writing. The y-axis indicates the percentage of documents about each main topic that make at least one reference to one of the eight hadith compilations just listed" (main text, page 19).
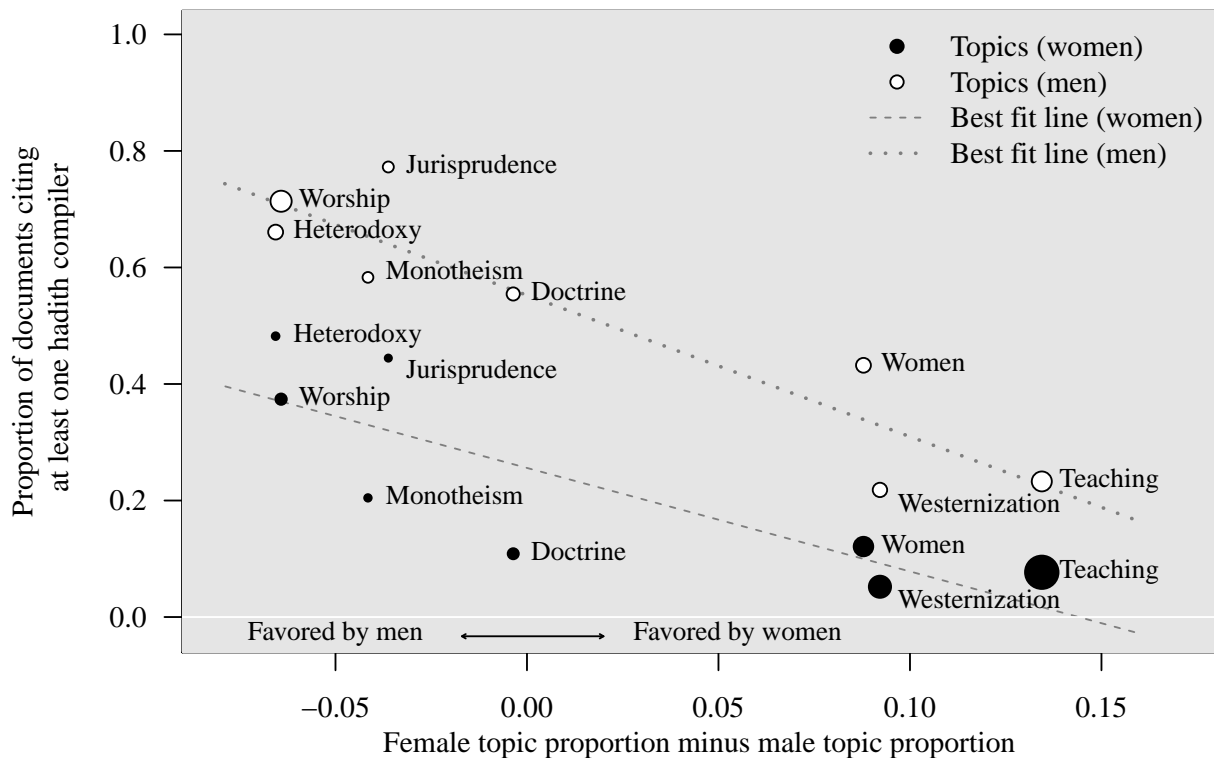

Figure 3: *Women cite hadith less frequently than men when writing on the same topics*

A reproduction of Figure 2 from the main text. "This figure shows that women are less likely to cite hadith than men, even when writing on the same topics. The black discs (women) and white discs (men) correspond to the eight substantive topics estimated in the main paper (see Figure 1). The size of the discs indicates the prevalence of each topic in writings by each gender. The x-axis location of the discs shows the difference in topic use by gender. Topics used more by women are right of zero and topics used more by men are left of zero. The y-axis shows the proportion of documents assigned to each main topic that cite at least one hadith compiler. The dashed lines show linear trends for men and women. The consistent gap between these trend lines shows that women cite hadith less often than men" (main text, page 20).

"The results show that hadith citations vary quite strongly with topics, but that this does not account for the difference in hadith citation rates by men and women. Figure 3 shows a persistent gap in the number of hadith citations by men and women, regardless of topic. This can be seen by examining the consistent size of the citation gap for any single topic, or by examining the roughly parallel slopes of the best fit lines for men and women. On average, men are twice as likely to cite a hadith compilation as women, regardless of whether the topic tends to have many hadith citations or few. The constant size of the gap is interesting because it reveals that at least some women are willing and able to use hadith citations, but use them far less often than men, even when the subject calls for them" (main text, pages 19–20).

Two more alternative arguments deserve attention: that men and women write in different genres, and that women may have less scholarly training for deploying hadith to support arguments. Conditioning on these factors requires a more complex statistical approach, so I estimate a series of regressions predicting hadith use in each document as a function of the author's gender, controlling for the main topic, scholarly credentials, and a proxy measure for genre.

I represent the main topic of each document as a categorical variable. I collect information on the scholarly credentials of each author from the saaid.net website. If the gap in hadith use is due to the relatively poor training of female Salafi preachers, then I expect the gap to disappear when I limit the sample to authors with PhDs in the Islamic sciences. I also include an interaction between the indicator for author gender and the indicator for whether an author has a PhD to test whether PhDs matter for one gender more than the other.

Measuring document genre is more difficult. I do not want to inadvertently include citation frequency in my coding of genre because then genre would be mechanically correlated with the outcome variable. Instead, I use document length as a proxy for genre. Women's documents are generally shorter than men's: only 10 percent of women's documents are over 1,172 words, while 37 percent of men's documents are longer than this threshold. Comparing documents of similar length makes it more likely that I am comparing citation rates in documents that are otherwise comparable styles of writing.

I present four statistical models in Table 1. In all four models, the outcome variable is a binary variable indicating whether a document cites at least one of the major hadith compilers (as in Figure 3). I estimate this outcome using a linear probability model for ease of interpretation, but logistic regression gives similar results. The first model predicts the probability of citing at least one hadith compiler as a function of author gender with no additional covariates. The second model includes the covariates linearly. The third and fourth models subset the data to account for possible non-linearities in the covariates, akin to a matching approach. The sample in model three is restricted to documents shorter than 1,172 words by authors with PhDs on topics preferred by women, while model four is the same but for topics preferred by men.

The results in Table 1 show that women consistently cite the hadith tradition less frequently than men, even when accounting for each document length, topic, and the credentials of authors. Model one shows that 50% of documents by men cite at least one hadith compiler, compared to only 11% of documents by women. Models two through four show that this gap is not fully explained by differences men's and women's preferred topics, document lengths, or scholarly training.

According to model 2, a document with the main topic of *Doctrine* will contain a citation to one of the major hadith compilers 44% of the time, but a similar document by a woman will contain such a citation only 26% of the time. The other covariates in model two predict citations to hadith compilers. When the author of a document has a PhD, the probability that their documents will cite hadith compilers is 5 percentage points higher for both men and women. As shown in Figure 3, topics favored by men are associated with more hadith citation while topics favored by women are associated with fewer citations. And as expected, longer documents are more likely to cite hadith compilers. Still, conditioning on these factors does not eliminate

7

Outcome: Binary variable indicating whether a document cites at least one of the prominent hadith compilers.

	(1) No covariates	(2) All covariates	(3) Subsample: Topics preferred by women	(4) Subsample: Topics preferred by men
Intercept	0.50*	0.47*	0.19*	0.57*
	0.004	0.01	0.008	0.014
Female Preacher	-0.39*	-0.21*	-0.16*	-0.21*
	0.009	0.01	0.013	0.075
Has PhD		0.057*	yes	yes
		0.007		
Female × Has Phd		-0.008		
		0.017		
Main Topic: Worship		0.2*		✓
		0.012		
Main Topic: Jurisprudence		0.25*		✓
		0.016		
Main Topic: Heterodoxy		0.15*		✓
		0.013		
Main Topic: Monotheism		0.058*		✓
		0.016		
Main Topic: Teaching		-0.25*	✓	
		0.011		
Main Topic: Westernization		-0.3*	✓	
		0.013		
Main Topic: Women		-0.09*	✓	
		0.013		
Document length		0.003*	< 1,172 words	< 1,172 words
		0.0001		
N documents	21,324	21,324	2,489	1,259
N documents by women	3,470	3,470	1,063	45
N documents by men	17,854	17,854	1,426	1,214
R^2	0.08	0.26	0.06	0.01
Sample	All documents	All documents	Documents on women, teaching, or westernization shorter than 1,172 words by authors with a PhD.	Documents on worship, jurisprudence, heterodoxy, or monotheism shorter than 1,172 words by authors with a PhD.

Table 1: *Regression models showing that men cite hadith compilers more than women.*

Four linear probability models showing that men cite hadith compilers more than women. Model 1 is estimated on the full data set of documents and includes no covariates. Model 2 is estimated on the full data set and includes covariates with "Doctrine" as the baseline main topic. Model 3 is estimated on the subset of documents that are shorter than 1,172 words, on topics favored by women, and by an author with a PhD. Model 3 is estimated on the subset of documents that are shorter than 1,172 words, on topics favored by men, and by an author with a PhD. * indicates $p < 0.05$.

the citation gap between men and women.

Models three and four subset the data to documents by authors with PhDs, all in the length-range that women seem to prefer (less than 1,172 words). If any set of women are able to cite hadith, it should be those with PhD's in the Islamic sciences. The gap still remains in this subset, whether I look at documents on topics favored by women (model 3) or favored by men (model 4).

E   An additional example of identity authority

Section 4 of the paper quotes from a document by Dr. Jawāhir bint ʿAbd al-ʿAzīz āl al-Shaykh to illustrate how women invoke identity authority in their writing. There is not space for more examples in the text of the paper, but I provide another here:

My second example of a woman invoking identity authority appears in a document by Hind ʿUmar, titled "Sally al-Faransiyya. . . in the Forum of the Saudi Women."[2] This document provides commentary on a comment made at a Saudi women's forum by a woman named Sally al-Faransiyya who reportedly spoke out against some attendees who were calling for attention to issues of women's rights in Saudi Arabia. She starts off by suggesting that those calling for rights are misguided: "I don't know if the Saudi women are fools? Or if they have been deceived. They talk of stolen rights?" She then gets personal. "I am a Muslim woman, of Tunisian origin, with a French name Sally, and an Arabic name Sayyida." She describes how she gained French citizenship because of her husband, but then left it all behind.

> But in spite of all the open horizons in front of us to live as citizens in France, obtaining all of the rights that any French citizen has, I left all that and moved to Saudi Arabia!! It was not because I couldn't find opportunities for work, or for the financial benefits,. . . but it was simply that we decided to live in Saudi Arabia because it was better for our daughters. Yes, it was better for our daughters.

Hind ʿUmar then comments at length on the foolishness of fixating on whether women can drive when Saudi Arabia provides so much for women in other respects. She concludes, "Praise God three times that I was born in a country that governs by God's law in the details of women's issues." The overall intention of the article is clear: Saudi women should be grateful for what they have no matter how restrictive Saudi society is to women. ʿUmar's support for this claim is based on identity. Sally al-Faransiyya, and by extension Hind ʿUmar, are women who have seen the so-called rights offered by the West and, as women, they can definitively say that they are not good.

F   List of personal pronouns used to identify womens' identity authority

The gender-neutral pronouns I use as linguistic traces of women's identity authority are *ana* (I), *anta* (you), *naḥnu* (we/us), *anī* (that I), *anak* (that you), *ankum* (that you), and *anna* (that we).

The gendered pronouns are *huwwa* (he/him), *hiyya* (she/her), *anha* (that she), *hum* (them (plural)), *huma* (them (dual)), *lahum* (to them), *minhum* (from them), *wahum* (and them), *anhu* (that he), *anhum* (that they), *banhum* (than they), *anhuma* (that they (dual)).

This is not an exhaustive list of all possible pronoun constructions in Arabic. Rather, it covers those commonly occurring in the corpus.

---

[2]https://saaid.net/daeyat/hindamer/16.htm, accessed 3/13/2017.