

# Obsessed with Muslim Women's Rights? A Computational Text Analysis of U.S. News Coverage

Rochelle Terman\*  
University of California, Berkeley  
Department of Political Science

September 27, 2015

Replication Materials here: <https://github.com/rochelleterman/worlds-women>

**PRELIMINARY DRAFT: PLEASE DO NOT CIRCULATE OR CITE WITHOUT  
PERMISSION**

## Abstract

Are the U.S. media obsessed with Muslim women's rights? Since the terrorist attacks of September 11, a vast literature has employed the theory of "gendered orientalism" to critique American media coverage of women in Muslim and Middle Eastern societies. This paper tests those claims by conducting the first large-scale, systematic, and comparative analysis of U.S. news coverage of women abroad. I examine two proposed mechanisms derived from the theory of gendered orientalism. First, U.S. news coverage of women abroad is driven by confirmation bias; gender oppression in Muslim societies is noticed more and considered news-worthy, while women's oppression in Western societies tends to be ignored. Second, U.S. news media write differently about women in Muslim and Middle Eastern nations than they do about women in other societies. Specifically, stories about the former exhibit a more concentrated discussion about women's rights and gender equality (or lack thereof), even for countries with relatively good records of women's rights. Using novel computational text analytic methods, I test these two hypotheses with new data from 35 years of *New York Times* and *Washington Post* reporting. The results suggest that U.S. news media are more likely to address women's rights abuses if those abuses occur in Muslim or Middle Eastern nations, regardless of the state of gender equality on the ground.

---

\*Corresponding author, [rterman@berkeley.edu](mailto:rterman@berkeley.edu).

Are the U.S. media obsessed with Muslim women’s rights? Since the terrorist attacks of September 11, numerous scholars have employed the theory of “gendered orientalism” to critique American media coverage of women in Muslim and Middle Eastern societies. Popular media portrayals, these scholars claim, are riddled with inaccurate and offensive stereotypes concerning gender relations in Muslim and Middle Eastern societies. By painting these societies as uniquely or particularly misogynistic, U.S. media produce a narrative by which Western feminists must “save” Muslim women from Muslim men, which in turn justify political intervention and even war. While this argument serves as the bedrock for a vast literature spanning many disciplines, it has yet to be demonstrated empirically against a large dataset.

This paper tests the claims inherent in “gendered orientalism” by conducting the first large-scale, systematic, and comparative analysis of U.S. news coverage of women abroad. I first explicate two testable hypotheses derived from the theory. The first argues that U.S. news coverage of women abroad is driven by confirmation bias. Journalists are more likely to report on women living in Muslim and Middle Eastern countries if their rights are violated, but will report on women in other societies when their rights are respected. I call this the *confirmation bias hypothesis*. Second, U.S. news media will tend to frame reporting about women in Muslim or Middle Eastern countries around the specific issue of women’s rights and gender equality (or lack thereof). The coverage is biased insofar as it emphasizes this issue over other topics when discussing a Muslim or Middle Eastern country, regardless of the actual material status of women’s rights in these places. I call this the *reduction hypothesis*.

I test these two hypotheses with new data from 35 years of *New York Times* and *Washington Post* reporting, combined with novel computational text analytic methods that enable a systematic comparison of both the quantity and substance of media coverage. The analysis provides strong support for both hypotheses, although in some cases the magnitude of the bias is quite small. Taken together, the theory, research design, and findings provide a robust empirical basis for a ubiquitous claim, while raising new questions for scholars studying gender, media, and transnational advocacy.

## The Theory of Gendered Orientalism

Since Edward Said's influential *Orientalism*, and especially after 9/11, a large literature has developed critiquing American media coverage of women in Muslim and Middle Eastern societies. While the literature spans multiple disciplines, theoretical approaches, and empirical territory, scholars converge on three modal claims. First, American media discourse is purportedly obsessed with Muslim women's oppression, for which the veil is the ultimate symbol and case in point (Ahmad 2009; Macdonald 2006). Popular media outlets portray Middle Eastern and Muslim societies as uniquely or particularly misogynistic, especially compared to Western countries (Kumar 2012; Razack 2004). This misogyny is ascribed to Islam and/or Arab culture, which is said to justify or require female subjugation (Ahmed 1992; Bahramitash 2005; Razack 2004; Volpp 2000). Not only is this narrative simplistic and sensationalist, it conflicts with the reality of women's lives insofar as it inaccurately portrays the degree and cause of Muslim women's suffering (Abu-Lughod 2013; Razack 2008). Furthermore, it denies women's agency by reduces their lives to a totalizing oppression (Bilge 2010; Mahmood 2011; Scott 2009), while demonizing Muslim, Arab, and Middle Eastern men as inherently barbaric and cruel (Bhattacharyya 2008; Puar 2007).

Second, American media discourse tend to compare the lives of Muslim women to those of Western women, who are portrayed, by contrast, as liberated and free by comparison (Yegenoglu 1998). Together, this dichotomy justifies a rescue mission by which Western feminists must "save" Muslim women from their oppressive religion (i.e. Islam), culture, or traditions (Abu-Lughod 2002). The "savior" narrative has been heavily scrutinized by feminist and critical scholars, who denounce it as paternalistic and imperialist (Abu-Lughod 2002; Cooke 2002; Mohanty 1988).

Third, the need to "save" Muslim women, bolstered by American media portrayals, is often used to justify undesirable political imperatives, including military intervention, repressive or exploitative policies, and cultural imperialism (Abu-Lughod 2010; Maira 2009; Razack 2008). In many cases, the plight of women is not the central concern; rather, it is used cynically as a way to bolster public support for international aggression (Stabile and Kumar 2005). The increased coverage of Afghan women post 9/11 is an oft-cited case in point (Cloud 2004; Fowler 2013; Hirschkind and Mahmood 2002; Klaus and Kassel 2005; Shepherd 2006; Stabile and Kumar 2005), but scholars have also

looked to historical cases in which the veil { footbinding (Teng 1996), female genital mutilation (Wade 2009), and sati (Mani 1987) were used to legitimize colonialism. In sum, American media coverage of Muslim women is largely determined by political imperatives, and the distribution of attention on women’s issues mirrors the contours of American geopolitical interests and foreign policy.

Together, these three claims make up a general argument that I call “gendered orientalism,” and provide the foundation for a sprawling literature and even a number of subfields (Abu-Lughod 2001; Charrad 2011). But while rising to the level of common sense in some disciplines, the argument is treated with suspicion in others, perhaps due to the literature’s general prioritization of theoretical innovation over empirical findings. This is not to say that empirical evidence is entirely absent. Indeed, qualitative research on the topic has generated a number of insightful case studies, especially on the post 9-11 Afghan case, while a few medium-N content analyses have explored the phenomenon in greater scope. These works notwithstanding, we have yet to see an empirical analysis that is able to test these claims against a large dataset.

Such a study is challenging due to the practical difficulties involved in creating a large-scale, systematic, and comparative analysis of U.S. media coverage of Muslim and Middle Eastern women. Fortunately, newly available data sources and methodological tools in the realm of computational text analysis have opened the door to novel empirical possibilities. By virtue of their methodological limitations, such techniques could never replace careful qualitative work or rich theoretical analyses. But they would provide another lens through which we could evaluate an argument that has become so essential to such a large body of work.

## **Hypotheses**

It is not immediately clear what scholars of gendered orientalism mean when they say that the U.S. media are obsessed with Muslim women’s oppression. What mechanisms can we infer from the theory that we can then test empirically? One possibility involves the relationship between stereotype and confirmation bias. Few scholars would deny the existence of sexism or women’s oppression in Middle Eastern or Muslim societies. But they would argue that women’s oppression

in Muslim countries tend to be noticed more often than women’s oppression in Western countries, which are elided under the latent assumption of gender equality in these societies. In this way, American media are influenced by the stereotype that Muslim and Middle Eastern countries are uniquely or especially sexist, and perpetuate that stereotype via confirmation bias.

This leads to a testable hypothesis: women’s oppression in Muslim societies is considered more news-worthy, while women’s oppression in Western societies is ignored. In other words, there is an inverse relationship between women’s status and the quantity of news coverage about women in Muslim versus non-Muslim countries. Among Middle Eastern and Muslim-majority countries, American news media will tend to report more stories about women the greater inequality they face. But in other countries, women are more likely to be represented when they have relatively high political, social, and economic status. I call this the *confirmation bias hypothesis*.

**Hypothesis 1** *Women’s status is negatively correlated with the quantity of U.S. news coverage of Muslim and Middle Eastern countries, but positively correlated in other cases.*

While the first proposed mechanism has to do with quantity of coverage, a second possible mechanism centers around the tone or framing of such coverage. Not only is Muslim women’s oppression reported on more often in the American press; their entire lives are reduced to this supposed oppression or inequality (Ryan 2011; Abu-Lughod 2013). Western women, on the other hand, are reported in higher dimensions and with more complexity, according to the gendered orientalism argument.

This, too, generates a testable hypothesis. News coverage of women can take on a variety of content, from rights and equality to sports, fashion, travel, etc. If we believe the gender orientalist argument, however, we would expect coverage of Muslim or Middle Eastern women to feature a more concentrated discussion of women’s rights and gender (in)equality compared to coverage of women in non-Muslim or European countries. We would also expect the proportion of coverage dedicated to women’s rights to be higher regardless of women’s actual status in these countries. In other words, if we compare a Muslim country and a non-Muslim country with equivalent records of women’s rights, U.S. news reporting will focus more on women’s rights and equality in the Muslim

case. I label this the *reduction hypothesis*, since it claims that women in some countries are reduced to their (lack of) rights.

**Hypothesis 2** *U.S. news coverage of women in Muslim and Middle Eastern countries will focus more on women’s rights and gender equality, relative to other topics, compared to non-Muslim and non-Middle Eastern countries, even when controlling for women’s political, social, and economic status.*

## Data

The primary data used in this study consists of all articles about women in non-U.S. countries, published in the *New York Times* and *Washington Post*, 1980–2014. Clearly, the inferences drawn from this data cannot be straightforwardly applied to American media writ large. With that reservation, however, there are still reasons to value this sample. In addition to being available to researchers, these two outlets are often considered “papers of record,” i.e. the most prominent, accurate, and influential among U.S. news media. These sources are especially prominent among the foreign policy establishment due to their reputation for reporting of international affairs. Other media outlets, including print and television news, rely on the *New York Times* and *Washington Post* for their reporting (Schraeder and Endless 1998). Finally, the 35-year sample includes enough variation in geopolitical affairs to validly test the hypotheses raised above.

Using the LexisNexis database, I downloaded all articles containing the subject term “women” from these two outlets during the specified time period. Subject terms are derived from LexisNexis’s SmartIndexing technology, which applies controlled vocabulary terms for different taxonomies such as subject, geographic region, language, etc. In addition to subject, documents are assigned country terms along with a relevance score that calculates how important or salient each country is to a document. Scores of 85 percent or higher indicate a major term. I assign each article to a single country using its most salient country term, if that term has a relevance score of 85 percent or higher. Articles with missing major country terms were discarded.<sup>1</sup> Because this study explores

---

<sup>1</sup> Some articles contained more than one major country term; in these cases, I took the term with the highest relevance score. These cases accounted for only 9 per cent of the corpus.

how U.S. media represent women abroad, I discarded all articles that were primarily about the United States. The final sample includes 4522 documents: 3720 from the *New York Times* and 802 from the *Washington Post*.

These data were then combined into a country-year data set, with each document assigned to an observation based on the year in which it was published and the country it most concerned. The country-year data set includes all current and historic U.N. states, plus Palestine but excluding the United States, for a total of 198 countries and 6226 observations.

County-years were assigned a regional classification loosely based on Hafner-Burton and Ron's (2013) six regional groupings: *Powerful West* (West) with 28 countries; *Asia* (Asia) with 33 countries, including Pakistan; *Latin America* (LA) with 33 countries; the *Middle East and North Africa* (MENA) with 22 countries, including Afghanistan; *Sub-Saharan Africa* (Africa), with 46 countries; and the *Eastern Europe and Central Asia* (EECA) with 31 countries. These regional groupings generally conform to the United Nations' regional classification, with some exceptions. First, former countries are assigned to regions based on where their current territorial manifestations are classified. Second, due to ambiguity surrounding whether Pakistan and Afghanistan are part of the Middle East or Asia, I decided to code these countries based on their assignment in most U.S. higher education area studies programs, with Afghanistan going in MENA and Pakistan in Asia; these assignments generally reflect the location of these nations in U.S. popular consciousness. Finally, the Powerful West is a region that Hafner-Burton and Ron include in their study and that I agree is important given the theoretical argument we wish to test. This region includes advanced industrialized countries of North American and Western Europe, along with two highly developed Asian countries – Australia and New Zealand.

Figure 1 shows the number of documents per region over time. Note that these numbers are not normalized for the total number of articles produced by the *New York Times* or *Washington Post* during this time period. Thus it is difficult to ascertain whether an increase in the number of documents signifies an increase in the media's interest in the women living in these areas, or, alternatively, an increase in the coverage of these regions generally. After reading the corpus, however, some inferences can be made. For instance, we see distinct spikes in the Africa and Asia

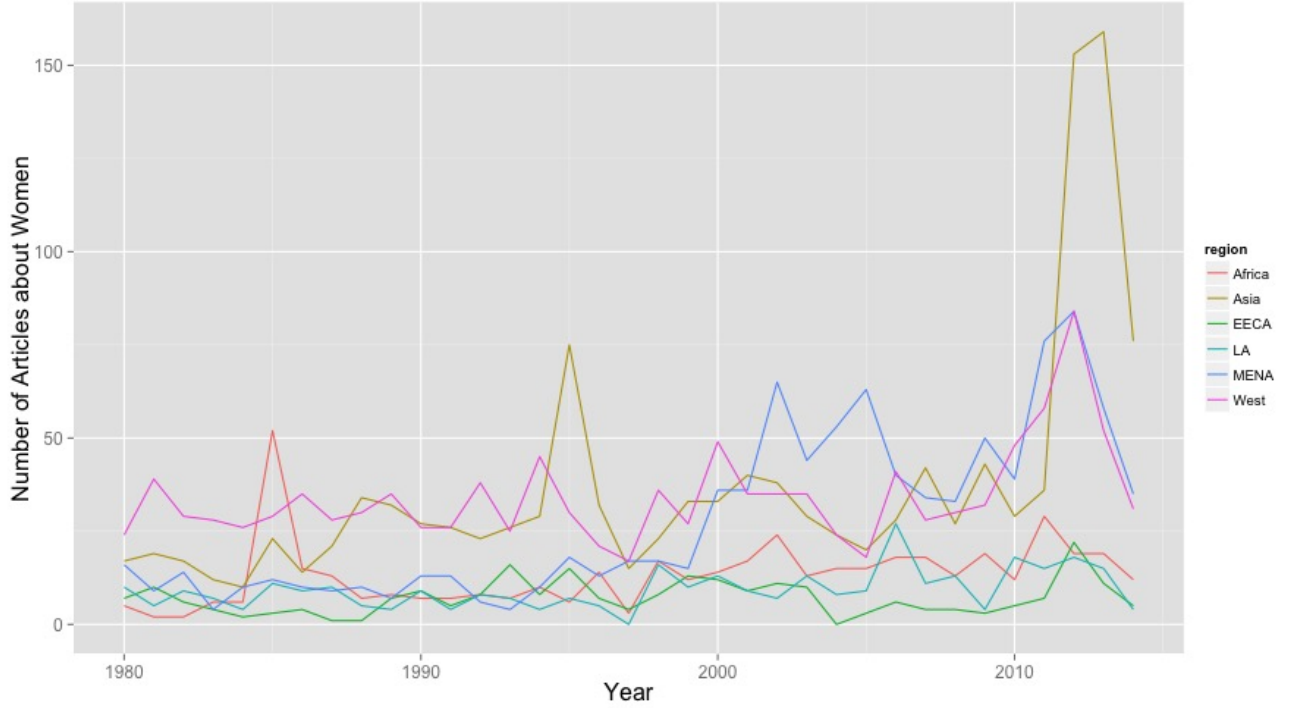


Figure 1: Number of Articles per Region over Time

regions corresponding to the UN's World Conferences on Women in Nairobi (1985) and Beijing (1990), which were heavily covered in these papers. Reporting on the Delhi rape case that occurred on 16 December 2012 contributes to the significant spike in the Asia region during 2012-2013. But while we do see a significant increase in coverage of women in the MENA region after 2001, these number must be weighed against the overall increase in reporting about this region following the wars in Afghanistan and Iraq.

## Modeling Hypothesis 1

The first hypotheses pertains to the likelihood of U.S. news coverage on women. Here, the dependent variable (likelihood of U.S. news coverage of women) is operationalized in two ways. The first is a simple binary (*Reported (Binary)*) indicating whether a country-year observation featured at least one article in the sample (true in 1451 cases). The second measure is a count indicating the total number of articles about women for that observation (*Reported (Count)*).

Hypothesis 1 claims that the effect of women's rights on the likelihood of coverage is conditional



on whether the unit of observation is a Muslim or Middle East country. Thus an interaction term is necessary for the model.<sup>2</sup> The mediating variable (whether the observation is a Muslim or Middle Eastern country) is operationalized in three ways: *Percentage Muslim* captures the Muslim percentage of a population according to research by the Pew Research Center. *Muslim Majority* is a binary indicating whether the *Percentage Muslim* is 50 per cent or above. *MENA* is a binary indicating whether a country is included in the Middle East and North Africa regional classification described above. I estimate models with all three measures in order to compare potential causal mechanisms.

Estimating the material condition of women’s rights and gender equality is problematic due to the conceptual difficulties surrounding this term (Peksen 2011). While recognizing the limitations of such a measure, I rely on the popular Cingranelli-Richards Rights Index (CIRI), which culls data from the U.S. State Department’s annual human rights country reports. This is appropriate for the task at hand, because we would expect that U.S. news media follow a commensurate understanding of “women’s rights” with that used by the U.S. State Department. CIRI offers three variables capturing the notion of women’s rights as they are effected in law and practice: *Women’s Economic Rights*, *Women’s Political Rights*, and *Women’s Social Rights*.<sup>3</sup> Each variable is an ordinal variable ranging from 0 (indicating that women’s rights were not guaranteed by law during a given year) to 3 (indicating that women’s rights were guaranteed in both law and practice).<sup>4</sup>

---

<sup>2</sup>Clearly, respect for women’s rights is itself affected by whether the observation is a Muslim or MENA country. However, tests using variable inflation factors indicate that collinearity was not a problem in the models, and, furthermore, the results are robust across a number of specifications.

<sup>3</sup> The *Women’s Political Rights* variable include the following rights: “The right to vote; the right to run for political office; the right to hold elected and appointed government positions; the right to join political parties; the right to petition government officials.” *Women’s Social Rights* includes: “The right to equal inheritance; the right to enter into marriage on a basis of equality with men; the right to travel abroad; the right to obtain a passport; the right to confer citizenship to children or a husband; the right to initiate a divorce; the right to own, acquire, manage, and retain property brought into marriage; the right to participate in social, cultural, and community activities; the right to an education; the freedom to choose a residence/domicile; freedom from female genital mutilation of children and adults without their consent; freedom from forced sterilization. *Women’s Economic Rights* is coded based on the following: “Equal pay for equal work; free choice of profession or employment without the need to obtain a husband or male relative’s consent; the right to gainful employment without the need to obtain a husband or male relative’s consent; equality in hiring and promotion practices; job security (maternity leave, unemployment benefits, no arbitrary firing or layoffs, etc.); non-discrimination by employers; the right to be free from sexual harassment in the workplace; the right to work at night; the right to work in occupations classified as dangerous; the right to work in the military and police force” (Cingranelli 2012).

<sup>4</sup> The *Women’s Political Rights* and *Women’s Economic Rights* variables are only available to 2011. The *Women’s Social Rights* variable is only available to 2004.

The composite variable *Women's Rights Index* estimates the overall situation of women's rights by taking the mean of these three indicators.<sup>5</sup>

I include a number of controls that may affect the likelihood of coverage. First, a straightforward alternative explanation suggests that reporting about women is proportional to general news coverage. The variable *Country Reports* records the number of articles that appear in the *New York Times* relating to a particular country-year, including those that are unrelated to the subject "women." We would expect coverage about women to be highly correlated with overall coverage for a given country-year.

On the other hand, articles about women may exhibit special features that make it different from general reporting. For instance, journalists may treat stories about women as "softer" news, requiring more personal interviews and field research than "hard" news items. One implication is that journalists may find it especially difficult to report on women in authoritarian countries, which tend to restrict freedoms of speech, assembly, and the press. To account for this possibility, I include a *Democracy* variable from the Polity IV dataset's Polity2 index, which is constructed by subtracting the 10-point autocracy index that measures the autocratic features from the 10-point democracy index that identifies the democratic characteristics of a polity (Marshall and Jaggers 2012).<sup>6</sup> Therefore, *Democracy* ranges from -10 (most autocratic) to +10 (most democratic).

Similarly, journalists may find it difficult to report on countries that are mired in domestic turmoil and violence. Thus I include a variable *Instability* culled from the Banks Cross-National Time Series Data Archive composite index of political instability, which encompasses multiple indicators including riots, antigovernment protests, guerrilla attacks, general strikes, purges, government crises, and assassinations. Higher values on this scale denote greater levels of political unrest and violence.

Finally, I include controls for *GDP per capita* (logged) using World Bank Development Indicators and *Population* (logged) using data from the United Nations. The rationale is that journalists find it easier to report about women in rich, populous countries, where it is easier to conduct field research and/or conduct interviews.

---

<sup>5</sup> The results are robust when using the three individual indices in place of the composite variable. See Appendix

<sup>6</sup> This data is only available to 2013.

I use statistical models that account for the cross-national time-series structure of the data. Because the panel data are highly correlated, I use generalized estimating equations (Zorn 2001). When modeling the dependent variable using the *Reporting* binary, I use a probit regression. When modeling the dependent variable using *Reported (Count)*, I use a negative binomial regression since this variable consists of over-dispersed counts. Note that a tobit is inappropriate here because coverage cannot assume negative values (Sigelman and Zeng 1999). To deal with heteroskedasticity, all estimates use Huber-White corrected robust standard errors clustered on country. Time-variant independent and control variables are lagged by one year to mitigate simultaneity issues and lessen any incorrect direction of inference. I also include a lagged-dependent variable to correct for serial correlation (Wooldridge 2010). The results are summarized in Tables 1 and 2.

The models provide strong support for Hypothesis 1. All three interaction terms are statistically significant and negative, suggesting that the effect of *Women’s Rights Index* varies across Muslim (MENA) and non-Muslim (non-MENA) countries. To help interpret these results, Figure 2 visualizes the marginal effect of *Women’s Rights* on *Reported (Count)* for each of the three Muslim/MENA-related variables.<sup>7</sup> The results suggest a double standard when it comes to how the U.S. news media represent women around the world. Not only are women in MENA and Muslim countries represented more often and in greater quantities, they garner special attention if their rights are violated. But the relationship is inverse for other nations, where better women’s rights records correlate with higher likelihood of coverage.

The findings are robust across a number of specifications. First, I ran alternative models replacing the *Women’s Rights Index* composite variable with individual scores representing *Women’s Political Rights*, *Women’s Social Rights*, and *Women’s Economic Rights* respectively. Second, to address potential bias stemming from missing data on covariates, I model effects using two imputation methods: a “nearest value” method, whereby I impute covariate data using the last reported value at the country level; and a multiple imputation method as described by King et al (2001).<sup>8</sup> Finally, I estimated models without the lagged dependent variable, and removing Israel from the

---

<sup>7</sup>Results are substantively identical for probit model on the *Reported (Binary)* DV.

<sup>8</sup> I use the *Amelia* package in *R* to estimate missing data using multiple imputation (Honaker, King and Blackwell 2011).

Table 1: Probit Analysis of U.S. News Coverage about Women Abroad

	<b>Dependent Variable: <i>Reported (Binary)</i></b>		
	<b>Model 1</b>	<b>Model 2</b>	<b>Model 3</b>
Lagged DV	0.496*** (0.061)	0.495*** (0.061)	0.495*** (0.061)
Country Reports	0.002*** (0.0003)	0.002*** (0.0003)	0.002*** (0.0003)
Women's Rights Index	0.078 (0.066)	0.087 (0.064)	0.102 (0.069)
Muslim Majority	0.425* (0.167)		
MENA		0.552** (0.183)	
Muslim Percentage			0.512** (0.185)
Democracy	0.006 (0.005)	0.011* (0.005)	0.007 (0.005)
Instability	−0.00001 (0.00002)	−0.00001 (0.00002)	−0.00001 (0.00002)
Population	0.381*** (0.023)	0.379*** (0.023)	0.376*** (0.023)
GDP per capita	0.134*** (0.024)	0.119*** (0.025)	0.131*** (0.024)
Women's Rights x Muslim Majority	−0.368** (0.127)		
Women's Rights x MENA		−0.342* (0.142)	
Women's Rights x Muslim Percentage			−0.392** (0.138)
Constant	−8.317*** (0.443)	−8.212*** (0.442)	−8.265*** (0.441)
N	3946	3962	3946
Log Likelihood	−1648.063	−1653.581	−1648.084
AIC	3316.127	3327.162	3316.168

\*\*\*p < .001; \*\*p < .01; \*p < .05

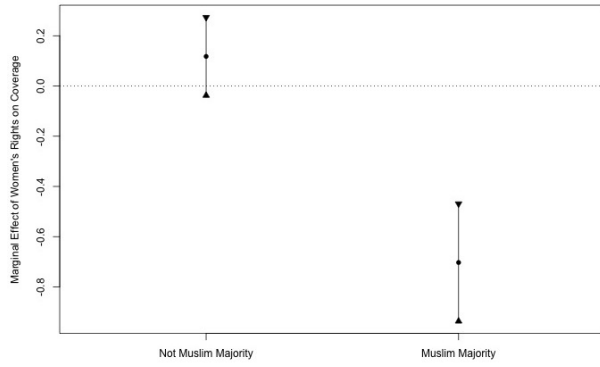
Robust standard errors clustered on country appear in parentheses.

Table 2: Negative Binomial Analysis of U.S. News Coverage about Women Abroad

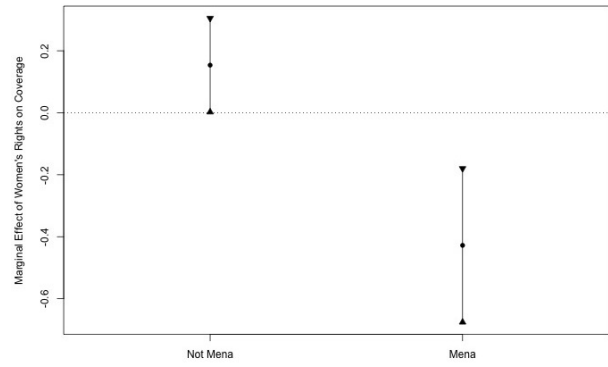
	<b>Dependent Variable: <i>Reported (Count)</i></b>		
	<b>Model 1</b>	<b>Model 2</b>	<b>Model 3</b>
Lagged DV	0.147*** (0.013)	0.149*** (0.013)	0.149*** (0.013)
Country Reports	0.001*** (0.0002)	0.001*** (0.0001)	0.001*** (0.0002)
Women's Rights Index	0.115 (0.096)	0.154 (0.093)	0.153 (0.103)
Muslim Majority	1.042** (0.320)		
MENA		1.298*** (0.340)	
Muslim Percentage			1.166*** (0.352)
Democracy	0.005 (0.010)	0.018 (0.010)	0.007 (0.010)
Instability	-0.00001 (0.00002)	-0.00001 (0.00002)	-0.00000 (0.00002)
Population	0.529*** (0.026)	0.539*** (0.027)	0.521*** (0.026)
GDP per capita	0.209*** (0.036)	0.169*** (0.037)	0.209*** (0.036)
Women's Rights x Muslim Majority	-0.816*** (0.205)		
Women's Rights x MENA		-0.581** (0.215)	
Women's Rights x Muslim Percentage			-0.825*** (0.218)
Constant	-11.631*** (0.622)	-11.641*** (0.632)	-11.579*** (0.629)
N	3946	3962	3946
Log Likelihood	-3569.768	-3566.883	-3569.499
$\theta$	0.905*** (0.061)	0.933*** (0.064)	0.907*** (0.061)
AIC	7159.537	7153.767	7158.999

\*\*\*p < .001; \*\*p < .01; \*p < .05

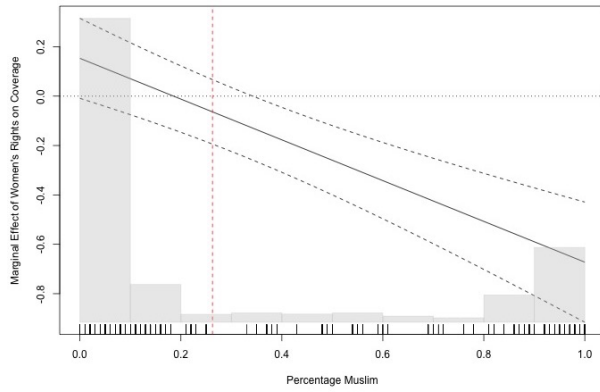
Robust standard errors clustered on country appear in parentheses.



(a) Interaction with *Muslim Majority*



(b) Interaction with *MENA*



(c) Interaction with *Muslim Percentage*

Figure 2: Marginal Effects of *Women's Rights Index* on *Reported (Count)*

sample if MENA countries. The results with substantively equivalent across all models.<sup>9</sup>

## Measuring Substantive Focus

While the above findings pertain to the quantity of coverage about women abroad, the second hypotheses pertains to the quality of coverage. What, precisely, is being reported in these articles? How does the substance of news vary depending on the society being discussed? Articles about women can take a variety of content, from elections and protests to sports and fashion. But gendered orientalism claims that U.S. media coverage of women in Muslim and Middle Eastern countries is obsessed about one issue in particular: women’s rights and gender equality. This obsession, scholars claim, is not present for other places, .

Testing this hypothesis requires information on the distribution of themes or topics in the corpus. Fortunately, recent advances in computational tools for the analysis of textual data enable new tools used to categorize and compare texts on a large scale (Grimmer and Stewart 2013). Among the most promising tools for social scientists and humanists, and the chief method used in this paper, is the probabilistic topic model, an algorithm used to code the content of a corpus of texts into substantively meaningful categories, or “topics,” using the statistical correlations between words in a corpus (Mohr and Bogdanov 2013).<sup>10</sup>

For the purposes of this study, topic modeling holds a number of advantages over other methods given the outcome of interest. The main benefit of this method lies in its ability to infer and analyze substantively meaningful categories (topics) with minimal assumptions and expense (Quinn et al. 2010). Unlike human-coder approaches, an automated topic model estimates topics from the observed data without assuming the substance, division, or keywords of topics beforehand. Thus it ameliorates the potential for confirmation bias. It is also fully replicable, because it is fully automated, which is an important validity concern for content analysis (Neuendorf 2011).

---

<sup>9</sup> All models are included in the Appendix.

<sup>10</sup> As a demonstration, imagine we have two sets of documents, one about food and the other about sports. We would expect that, given a document is about a particular topic, certain words would appear more or less frequently. “Sports” and “athlete” would appear more often in documents about sports, while “food” and “ingredient” would appear in documents about food, and “is” would appear equally in both. A topic model captures this intuition in a statistical framework, discovering abstract topics that might be thought of as a constellation of words that tend to come up in conversation (Mohr and Bogdanov 2013).

One alternative workflow is to categorize each document based on whether or not it pertains to women’s rights as a whole, and then calculate the proportion of articles in the “rights” category for each country-year. Unfortunately, such a blunt metric flattens important dimensions of variation. Most articles about women have at least one mention of women’s rights and gender equality, but differ in the degree to which they emphasize this theme. The gendered orientalist argument claims that for Muslim and/or MENA countries, *every* story, whether about politics or sports or literature, is framed as a women’s rights issue. A mixed-membership topic model estimates the outcome of interest more directly, because it represents texts as a distribution over many topics, not just one category. This allows us to compare how one document compares to another in terms of its *proportion* – not just presence – of a topic. However, as a robustness check, I also apply document-level labels indicating whether an article (as a whole) pertains to women’s rights and gender equality. This provides an alternative measure of the main outcome variable to be used in the models described below.<sup>11</sup>

## Data Preparation

To estimate the topic model, the corpus was first preprocessed following the standard recipe for automated text analysis. First, I removed capitalization, numbers, and punctuation. I then removed stop words, or those words that are extremely common but unrelated to the research topic, such as “and,” “or,” “the,” etc. I then applied a stemming algorithm to the corpus, which reduces words to their stem or root, using the Porter Snowball II stemmer for English, widely used in many automated text analysis applications (Porter 2001; Willett 2006). I also removed sparse terms by discarding all words used in only one or two documents out of the entire corpus.

In a less common step, I removed named entities such as the names of specific people, locations, and organizations, from the text of the articles. The reasoning behind this decision is that including terms that are specific to a particular region – such as the names of countries or political leaders –

---

<sup>11</sup>I use a computational strategy to apply boolean labels to documents. If a document contained the word ‘right’ (including the plural ‘rights’), ‘equal’, ‘sexist’, or ‘sexism’, it was labeled as pertaining to women’s rights. Approximately 52.6 per cent of documents were classified as ‘true’ on this boolean label. For each country-year observation, I summed all country-year documents containing the women’s rights label, and divided this count by the total number of articles for that country year. This offers a similar fractional variable to the *Women’s Rights Focus* variable used in the main analyses, described below.



would potentially bias the model to infer topics that are essentially regional categories instead of thematic ones. Because I am interested in the relationship between substantive topics and region, I chose to remove such terms, which I identified using Stanford’s Named Entity Recognizer as well as my own dictionary of nationalities (Finkel, Grenager and Manning 2005).

Once pre-processing was completed, the remaining words were transformed into a vector containing the count of each unique word in a document, disregarding information such as the order in which the words appear. The vectors are combined to construct a document-term matrix (DTM), where each row represents a document, each column represents a unique word, and each cell is the count of a unique word in a particular document. The final corpus had 4522 documents, 15,300 unique words and 1,047,653 total words. This DTM is the primary input for the structural topic model.

## Model Selection and Topics

To identify and explore the thematic topics in U.S. news media reporting about women abroad, I use the Structural Topic Model, developed by Roberts et al. (Roberts et al. 2013) with the goal to help “understand relationships between metadata and topics in their text corpus” (Lucas et al. 2015, p. 2).<sup>12</sup> The Structural Topic Model (STM) extends the popular topic modeling tool called Latent Dirichlet Allocation (LDA) by incorporating document-level metadata into the analysis as covariates. This allows us to measure systematic changes in topical prevalence according to changes in metadata, similar to a regression framework (Roberts et al. 2014, p. 5).

When estimating an STM, the analyst must make a number of important decisions pertaining to model selection. First, she must specify the number of topics ( $K$ ) to be estimated. There is no one answer for this decision (Roberts et al. 2014). The choice of  $K$  depends on a number of factors, including the nature of the documents under study and the level of granularity the analyst wishes to capture. In addition, STM depends on a multi-modal objective function, meaning that models with the same parameters (that is, the same data, covariates, and number of topic) may produce different results – that is, different topics – depending on the starting seed value.

---

<sup>12</sup>I use the *R* package *stm* to estimate the model (Roberts, Stewart and Tingley 2014).

Generally speaking, the best topic model is the one that generates the highest quality topics, and a higher quality topic is one that displays a higher degree of semantic interpretability. Two measures exist to evaluate interpretability quantitatively: cohesiveness and exclusivity. A topic is cohesive if its top words are likely to co-occur within documents. It is exclusive if top words are unique to that topic, or unlikely to appear in the top words for another topic (Roberts et al. 2014, p. 6). Oftentimes scholars will generate a number of candidate models and then discard those with low degrees of exclusivity and cohesiveness given a set  $K$  (Roberts et al. 2014)..

Even with these tools, however, selecting a model usually requires human judgment to evaluate the semantic meaning of topics (Mohr and Bogdanov 2013). For instance, we can try to summarize a topic by seeing which words have the highest probability of appearing in a topic. But because LDA topics are a distribution over all words, this technique usually prioritizes words that have high frequency in the entire corpus, but may not be semantically informative (i.e. “women” in this case). Another technique is to identify words that are exclusive to that topic, such as simplified frequency-exclusivity scoring (FREX) (Lucas et al. 2015, p. 5). A third technique is to see which document is most representative of a topic and read it to infer the topic’s content.

To address model specification issues, I estimated over 50 models varying the number of topics and starting seed values, and relied on the techniques described above to choose the highest quality model and apply semantic labels to each topic in that model. I did not look at corpus or region-level topic distributions before choosing a model to prevent bias. The final model estimated 15 topics by regressing topic prevalence on region and year covariates. Table 3 gives a summary of those topics, including each topic’s semantic label (applied by me), as well as top words calculated by frequency and FREX metrics. The words reported are stemmed, as explained above. There is a clearly discernible topic corresponding to “Women’s Rights & Gender Equality” in this corpus, as inferred by the words *right*, *equal*, *discrimin*, etc. <sup>13</sup>

---

<sup>13</sup> Applying semantic labels is more of an art than a science, because topics are merely statistical distributions of words, not hermeneutic concepts. Thus the “Politics” topic should not be considered all-encompassing of every aspect of political life; rather, it is a heuristic shortcut for a constellation of words that relate to specifically electoral politics (e.g. *polit*, *presid*, *govern*, *elect*, *minist*, *parti*). Still, some topics are more easily interpretable than others. The “Sports” topic, for instance, is fairly straightforward. The “Human Interest” topic, on the other hand, requires more thought because its top frequency words are ubiquitous and vague (*said*, *say*, *like*, *one*, *just*, *peopl*, *women*). More enlightening are the “Human Interest” topic’s top FREX words (*think*, *feel*, *know*, *thing*, *realli*, *see*, *someth*), which suggest individual perspectives over facts and events. Reading the most representative documents, i.e. those

Table 3: Summary of Topic Labels

Labels	Probability	FREX
1 Cancer	women, said, cancer, studi, breast, research, drug	cancer, implant, hormon, tumor, virus, breast, pill
2 Reproductive Health	abort, health, women, countri, children, birth, year	abort, birth, health, matern, aid, childbirth, pregnant
3 Religion	said, islam, religi, women, church, countri, ian	islam, religi, secular, circumcis, fundamentalist, koran, religion
4 Business & Work	women, said, percent, work, compani, year, job	ese, compani, corpor, board, employ, bank, job
5 Marriage & Family	said, famili, husband, girl, mother, year, children	father, husband, daughter, son, villag, bride, sister
6 Arts	art, work, show, artist, page, perform, photograph	museum, sculptur, artist, painter, art, galleri, exhibit
7 Migration	said, worker, citi, work, year, prostitut, govern	factori, brothel, prostitut, traffick, tourist, worker, migrant
8 Gender-Based Violence	said, rape, polic, case, report, sexual, violenc	rape, crime, prosecutor, crimin, prosecut, sentenc, victim
9 War & Combat	said, women, war, militari, kill, attack, soldier	soldier, troop, militari, combat, bomb, command, gun
10 Literature	book, film, write, stori, life, woman, love	novel, reader, film, charact, write, edit, literari
11 Human Interest	like, say, one, women, can, just, peopl	think, see, thing, feel, someth, realli, just
12 Women’s Rights & Gender Equality	women, right, law, men, said, issu, equal	law, equal, discrimin, constitut, legal, bill, divorc
13 Politics	said, polit, presid, govern, minist, elect, parti	elect, deleg, vote, voter, confer, minist, prime
14 Sports	team, women, game, said, world, play, olymp	game, olymp, sport, player, soccer, coach, athlet
15 Fashion	women, dress, fashion, wear, design, cloth, said	jacket, coutur, kimono, fashion, shirt, silk, skirt

## Comparing Coverage across Region

As a reminder, each document is represented as a mixture of these topics. For example, a hypothetical document devotes 54 percent of its content to “Business & Work,” 14 percent of its content to “Women’s Rights & Gender Equality,” 7 percent of its content to “Marriage & Family,” etc. Knowing this, we can calculate a number of summary statistics for topic distributions across the corpus. Figure 3 gives a corpus-level summary of topic distributions. The most common topics are “Women’s Rights & Gender Equality,” “Human Interest,” and “Politics,” with the average document devoting about 10-12 percent of its content to each of these topics. “Arts” is the least common topic, accounting for about 4 percent of the corpus’ content.

Coverage of these topics is unevenly distributed across regions. Table 4 tabulates the proportion of each topic (using the number of words assigned to it in the model) devoted to each region. Almost a third of American news coverage about “Women’s Rights & Gender Equality” abroad is dedicated to countries in the Middle East. These statistics, it should be noted, are not normalized for the total number of articles coming from each region. But even when we account for the population of articles, we see that regions vary greatly in their average content.

To get a better sense of this, STM allows us to plot the relationship between topical prevalence and metadata in a regression-like framework. Specifically, the model estimates the expected proportion of an unseen document devoted to a topic as a function of the region it is about and the year in which it was published. Holding time constant, a number of topics vary significantly in their expected proportions, depending on the region covered. Figure 4 visualizes these findings for a number of topics.

As the graphs show, if we came across an unseen article reporting about a MENA country, we would expect approximately 13 percent of its content to be devoted to “Women’s Rights & Gender Equality,” with a confidence interval of a little over 1 per cent. But if that article was about a Western country – even if it was published in the same year – we would expect less than 8 percent of its content devoted to ‘Women’s Rights & Gender Equality,’ In other words, reporting

---

with the highest percentage of its content devoted to the topic, offers the most insight. The most representative documents are interviews, profiles, inspiring stories of remarkable individuals, and so on. Taken together, I applied the label “Human Interest” because the topic appears to capture human interest reporting.

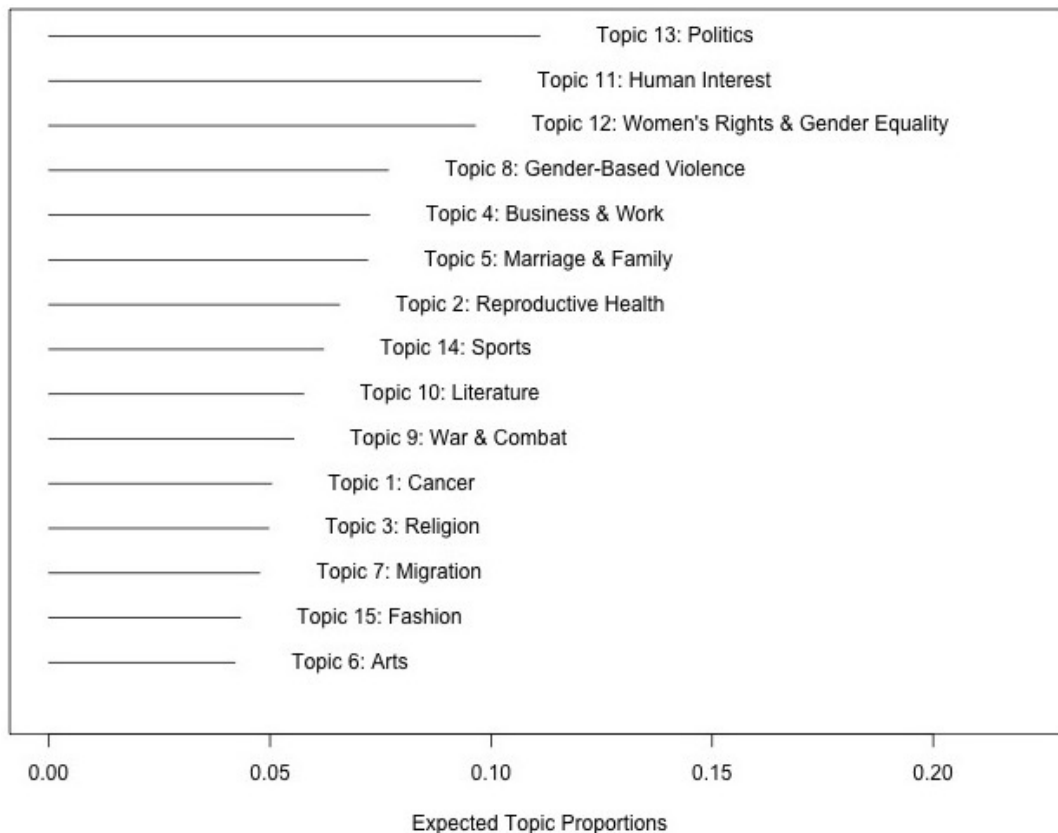


Figure 3: Corpus Summary

about women in MENA countries devotes 1.75 times the coverage to “Women’s Rights & Gender Equality,” compared to their counterparts in the West, and more than three times the attention to “Religion.”

## Modeling Hypothesis 2

The reader may find these results unsurprising, given the varying situation of women’s rights around the world. In other words, U.S. coverage about women’s rights may focus more on MENA and Muslim-majority countries because those are the societies that are least hospitable to women. Hypothesis 2 of the gendered orientalist argument, however, argues there is a bias, even when account for realities on the ground.

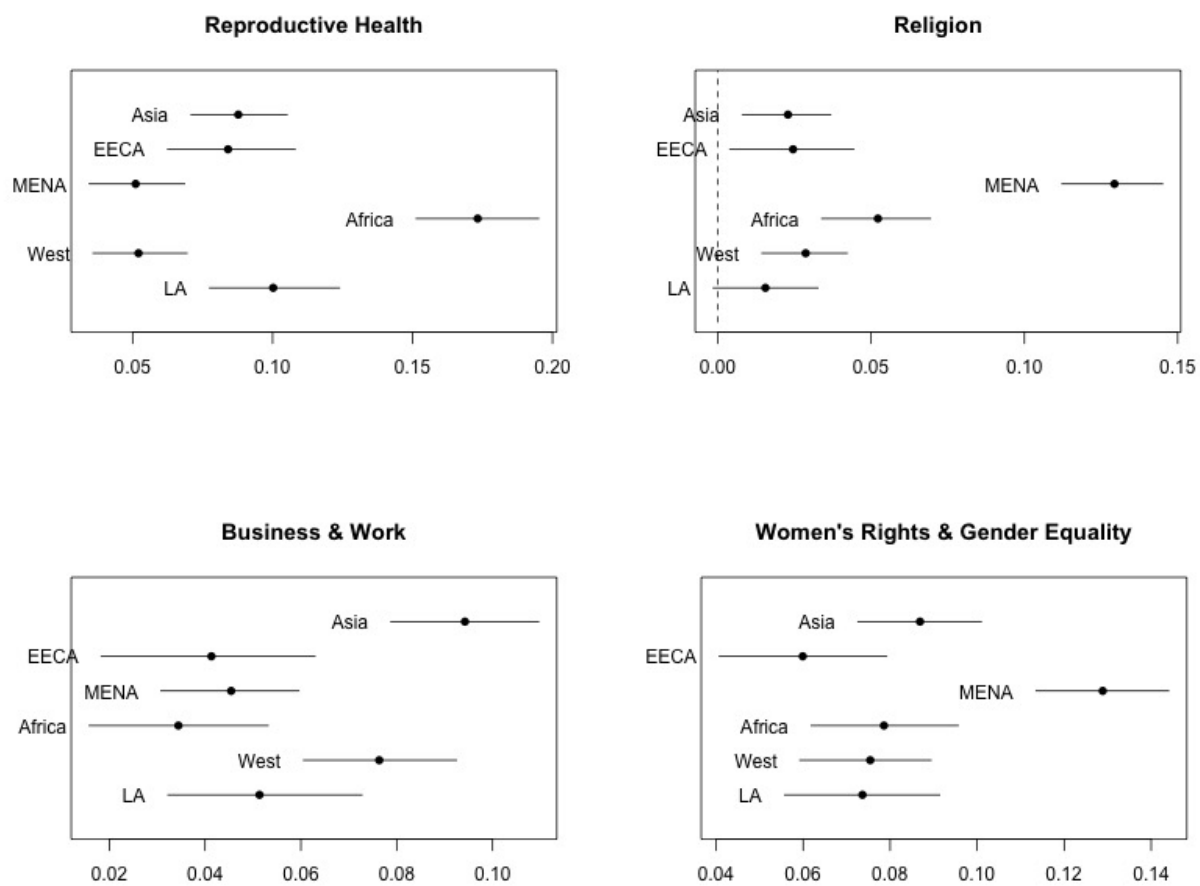


Figure 4: Expected Document Proportions of 4 Topics

Table 4: Topical Coverage across Region

	Africa	Asia	EECA	LA	MENA	West	Total
Cancer	12.64	17.46	2.84	5.04	6.84	55.17	100
Reproductive Health	26.02	32.05	6.31	9.65	11.98	13.99	100
Religion	10.65	13.14	2.78	2.09	54.89	16.46	100
Business & Work	5.49	40.16	3.29	5.47	14.82	30.78	100
Marriage & Family	13.70	32.87	5.22	7.32	21.34	19.54	100
Arts	7.12	25.79	7.24	8.04	11.58	40.22	100
Migration	9.17	39.96	9.00	10.20	16.99	14.68	100
Gender-Based Violence	9.04	40.40	6.40	10.12	19.83	14.21	100
War & Combat	7.96	19.84	8.59	6.08	45.41	12.13	100
Literature	6.69	25.13	8.14	6.99	13.80	39.24	100
Human Interest	10.12	29.64	5.46	6.36	21.37	27.05	100
Women’s Rights & Gender Equality	8.80	27.97	3.89	5.82	31.17	22.36	100
Politics	11.76	26.00	5.56	8.24	25.19	23.24	100
Sports	3.97	19.63	5.66	9.61	7.73	53.40	100
Fashion	8.26	27.07	4.53	7.01	12.79	40.33	100

The dependent variable in Hypothesis 2 is the percentage of coverage devoted to women’s rights for a particular country-year (*Rights Focus*). We expect this percentage to be higher for Muslim and Middle Eastern countries, even when controlling for women’s material status. I operationalized the outcome variable by taking the average proportion of articles assigned to the topic “Women’s Rights & Gender Equality,” weighted by number of words in each article. In other words, I sum the number of words about “Women’s Rights & Gender Equality” and divide it by the total number of words across all articles in that country-year. This gives us an estimate of the degree to which these newspapers focused on this topic relative to others for each observation, ranging from 0 to 1.

The dependent variable *Rights Focus* is then regressed onto two main explanatory variables in order to test the hypothesis described above. The first is *Women’s Rights Index*, measuring respect for women’s political, social, and economic rights, using the same CIRI indicator described above. The second is whether the observation represents a Muslim or Middle Eastern country, again using the same variable previously described: the fractional *Percentage Muslim* ranging from 0 to 1, the dichotomous *Muslim Majority*, and the dichotomous *MENA* variables.

I also include two controls that may affect the amount of rights language in reporting. First, coverage of women’s rights may be driven by the general state of human rights protections in certain countries. For instance, the poorer a country’s rights protections, the more coverage it may receive on its rights situation in general, including women’s rights. For this reason, I include a *Democracy* variable, described above. I also include a measure of general human rights protections, the *Physical Integrity Rights* index, also from the CIRI dataset,<sup>14</sup>

Because country-years had to contain at least one article to be included in the sample ( $n = 1451$ ), I use a two-step heckit model to account for potential selection effects. The selection equation is identical to the model presented in Table 1, where the dependent variable is the *Reported (Binary)* variable, indicating whether a country-year contained any articles in the data set. Conditional on this value being 1, an OLS model was estimated regressing *Rights Focus* on the four explanatory variables. I also include a one-year lagged dependent variable into the model, based on the reasoning that journalists maintain their thematic focus for a particular country from year to year. As with the previous models, I lag time-variant explanatory variables by one year and use Huber-White corrected robust standard errors clustered on country.

In all models, *Women’s Rights Index* is statistically significant and negative, indicating that U.S. news media will focus their coverage of women around “Women’s Rights & Gender Equality” in those areas with poor respect for women’s rights. This is hardly surprising. But even when we control for the material status of gender equality on the ground, we find that the coefficients on the *Muslim Majority*, *MENA* and *Muslim Percentage* variables are statistically significant and positive. In other words, U.S. news media talk more about “Women’s Rights & Gender Equality” if the country being covered is in the MENA region or has a larger Muslim population, regardless of women’s status in these societies. This finding supports Hypothesis 2, which states that women from Muslim and/or MENA countries are represented narrowly in U.S. news media, characterized largely

---

<sup>14</sup>The variable is composed of the four integrity rights variables, including disappearance, extra-judicial killing, political imprisonment, and torture. It is a nine-point scale that ranges from a minimum of zero to a maximum of eight, where zero indicates no respect for physical integrity rights and eight indicates full respect for those rights (Cingranelli 2012). The data is available to 2011. It should be noted that there are alternative measures for the state of human rights protections used in the literature, such as the CIRI’s Empowerment Rights Index, and the Political Terror Scale measures. I chose the Physical Integrity Index in the models discussed below, but I also estimated models using these two alternative measures, with the same substantive results.



Table 5: Two-Step Analysis of *Rights Focus* in U.S. News Coverage about Women Abroad

	<b>Dependent Variable: <i>Rights Focus</i></b>		
	<b>Model 1</b>	<b>Model 2</b>	<b>Model 3</b>
Intercept	0.096*** (0.012)	0.102*** (0.012)	0.094*** (0.013)
Lagged DV	0.101* (0.047)	0.107* (0.047)	0.103* (0.047)
Women's Rights Index	-0.018** (0.007)	-0.021** (0.007)	-0.018** (0.007)
Muslim Majority	0.033** (0.010)		
MENA		0.026* (0.010)	
Muslim Percentage			0.034** (0.012)
Democracy	-0.0005 (0.001)	-0.001 (0.001)	-0.001 (0.001)
Physical Integrity Rights	0.002 (0.002)	0.002 (0.002)	0.002 (0.002)
Inverse Mills Ratio	0.001 (0.009)	0.003 (0.009)	0.0001 (0.009)
N	656	656	656
R-squared	0.555	0.552	0.554
Adj. R-squared	0.550	0.547	0.549
Residual Std. Error (df = 649)	0.084	0.084	0.084
F Statistic (df = 7; 649)	115.565***	114.384***	115.032***

\*\*\*p < .001; \*\*p < .01; \*p < .05

Robust standard errors clustered on country appear in parentheses.

by their subordination, whereas women from other societies are portrayed in greater complexity. On the other hand, the magnitude of these coefficients are small, ranging from 0.026 to 0.034. So while U.S. media will focus more on women’s rights when it comes to Muslim and MENA countries, the bias amounts to roughly 3 per cent more coverage about this issue (in terms of topic proportions).

As in the previous section, results are robust to a range of alternative specifications. First, while the addition of a lagged DV helps correct for serial correlation, it shrinks the sample size considerably, because observations must contain at least one article at year  $t$  and  $t - 1$  to be included. As a robustness check, I removed this lagged DV, thus allowing for a larger sample. Second, I replaced the composite *Women’s Rights* variable with the three individual indicators representing Women’s Political Rights, Women’s Social Rights, and Women’s Economic Rights. Third, I estimate models using an alternative measure of the dependent variable *Rights Focus*, which is independent of the topic model.<sup>15</sup> Finally, I estimated one-step models using OLS and fractional logit. In all models, the *Muslim Majority*, *MENA* and *Muslim Percentage* variables were statistically significant and positive.<sup>16</sup>

## Conclusions

No society is immune from gender discrimination and oppression. But this paper demonstrates that the representation of women – and their rights – are unevenly portrayed in U.S. news reporting. First, I put forth a *confirmation bias hypothesis*, arguing that misogyny in Muslim and MENA countries are considered more newsworthy than in other places. Not only is there a bias in terms of quantity of coverage, but in the substance and framing as well. In the *reduction hypothesis*, women from Muslim and MENA societies are more likely to have their experience reduced to one facet – women’s rights and gender equality – in contrast to their counterparts in the rest of the world, who are represented in higher dimensions. Together, these findings provide strong support for the overall claim that U.S. news media will only “care” about women’s rights abuses if those abuses occur in Muslim or MENA nations.

---

<sup>15</sup>This alternative is a simple boolean label that marks a document as pertaining to women’s rights if it contains any word in a set of rights-related words. See note 10

<sup>16</sup>Reports of all alternative models are included in the appendix.

These findings are important because, as many scholars have argued, mass media is a key force in political and cultural change. The vast majority of people look to the media to make sense of events, especially in places where they have limited real-life exposure. News reporting, in particular, can serve an agenda-setting function (McCombs and Shaw 1972), with material consequences in realms such as humanitarian aid and foreign policy, referred to as the so-called “CNN effect” (Robinson 1999, 2001; Baum and Potter 2008). Media also shapes public opinion (Gamson and Modigliani 1989; Kinder 1998), especially on Muslims and Islam (Bail 2012; McAlister 2001; Said 2008).

The media’s obsession with Muslim women’s rights has two implications which are especially pertinent. First, while women from Muslim and Middle Eastern countries are front and center on the agenda, oppression in other societies is ignored. As with other humanitarian issues, the United States has a limited attention span when it to global women’s rights. Insofar as media attention drives awareness and resources, women from non-Muslim countries in Africa, Asia, Latin America and Europe lose out in this scenario, even if they suffer more egregiously.

Second, an obsession with Muslim women’s rights may, ironically, have counterproductive consequences for the goal of gender equality in these societies. Considering the already volatile environment surrounding Islam in the American public sphere, a disproportionate focus on Muslim women’s oppression is likely to be met with suspicion and incredulity among Muslim men and women alike. This is especially likely when the media’s diagnoses of sexism in Muslim societies point overwhelmingly to Islam.<sup>17</sup> Tired of feeling singled out, Muslims both at home and abroad may learn to equate feminist criticism with imperialism and Islamophobia, thus undermining even local initiatives for gender equality (Terman 2015).

Yet, a number of questions remain. First, due to the limited sample, we do not know whether and to what degree these biases vary across outlet or platform. Some scholars of “gendered orientalist” argue that conservative and right-wing factions are the worst offenders, whereas others insist that the stereotypes surrounding Islam and Muslim women are ubiquitous amongst even progressive crowds (Kumar 2012). Using similar techniques to the ones presented here, future research could examine these trends in liberal vs. conservative media. Likewise, scholars could compare coverage

---

<sup>17</sup>As illustrated in Figure 4 and Table 4, the topic of “Religion” – even more than “Women’s Rights & Gender Equality” – was disproportionately present in coverage about MENA countries.

in news outlets vs. entertainment, as well as media outside the United States.

Second, the precise mechanisms driving these trends – i.e. confirmation bias and reduction – remain unclear. What makes journalists write about women, or about Muslim / MENA societies, the way they do? Scholars writing in the gendered orientalist tradition differ in their accounts. Some writers attribute these biases to a cynical ploy by politicians, who strategically purpose women’s rights as a way to increase public support for the “War on Terror” (Stabile and Kumar 2005). Others argue that the problem goes much deeper, to longstanding and largely unconscious ideas about freedom, religion, and politics that are constitutive of Western modernity (Massad 2015). Throughout this literature, however, the empirical processes driving these phenomena remain imprecise. In different realms, social movement scholars have discovered important findings on the relationship between civil society organizations and media frames. Chris Bail, for instance, shows how anti-Muslim organizations originally occupied discursive niches but were amplified by mass media on account of their emotional energy, eventually drifting from the fringe of the discursive field into the mainstream (2012). Scholars interested in the positivist face of gendered orientalism argument could make similar inquiries into the ecological and organizational dynamics of media attention of women abroad.

Finally, while essay has largely neglected questions relating to *change over time*, the temporal aspect of this phenomenon remains important. In the literature on gendered orientalism, 9/11 occupies a crucial moment for American discourse on Islam, ushering in a new (or newly reinvigorated) obsession with Muslim women and their oppression. Indeed, the data confirm that from 2002 to 2006 more articles were published about women in MENA countries than in any other region (see Figure 1). But those numbers were dwarfed by the recent upswing in coverage about women in Asia, prompted in part to the Delhi gang rape case that occurred in late 2012 but extending to widespread coverage about India and China for several years. This raises new questions as to how women’s representations change (or don’t) in the face of global events.

## References

- Abu-Lughod, Lila. 2001. "'Orientalism' and Middle East Feminist Studies." *Feminist Studies* 27(1).
- Abu-Lughod, Lila. 2002. "Do Muslim women really need saving? Anthropological reflections on cultural relativism and its others." *American anthropologist* 104(3):783–790.
- Abu-Lughod, Lila. 2010. "The Active Social Life of 'Muslim Women's Rights': A Plea for Ethnography, not Polemic, with Cases from Egypt and Palestine." *Journal of Middle East Women's Studies* 6(1):1–45.
- Abu-Lughod, Lila. 2013. *Do Muslim women Need Saving?* Harvard University Press.
- Ahmad, Dohra. 2009. "Not Yet Beyond the Veil Muslim Women in American Popular Literature." *Social Text* 27(2 99):105–131.
- Ahmed, Leila. 1992. *Women and Gender in Islam: Historical Roots of a Modern Debate*. Yale University Press.
- Bahramitash, Roksana. 2005. "The War on Terror, Feminist Orientalism and Orientalist Ceminism: Case Studies of Two North American Bestsellers." *Critique: Critical Middle Eastern Studies* 14(2):221–235.
- Bail, Christopher A. 2012. "The Fringe Effect Civil Society Organizations and the Evolution of Media Discourse about Islam since the September 11th Attacks." *American Sociological Review* 77(6):855–879.
- Baum, Matthew A. and Philip B. K. Potter. 2008. "The Relationships Between Mass Media, Public Opinion, and Foreign Policy: Toward a Theoretical Synthesis." *Annual Review of Political Science* 11(1):39–65.
- URL:** <http://dx.doi.org/10.1146/annurev.polisci.11.060406.214132>
- Bhattacharyya, Gargi S. 2008. *Dangerous Brown Men: Exploiting Sex, Violence and Feminism in the 'War on Terror'*. Zed Books.

- Bilge, Sirma. 2010. "Beyond Subordination vs. Resistance: An Intersectional Approach to the Agency of veiled Muslim Women." *Journal of intercultural studies* 31(1):9–28.
- Charrad, Mounira M. 2011. "Gender in the Middle East: Islam, state, agency." *Annual Review of Sociology* 37:417–437.
- Cingranelli, David. 2012. "The Cingranelli-Richards (CIRI) human rights data project coding manual."
- Cloud, Dana L. 2004. "'To Veil the Threat of Terror': Afghan Women and the <Clash of Civilizations> in the Imagery of the US War on Terrorism." *Quarterly Journal of Speech* 90(3):285–306.
- Cooke, Miriam. 2002. "Saving Brown Women." *Signs: Journal of Women in Culture and Society* 28(1):468–470.
- Finkel, Jenny Rose, Trond Grenager and Christopher Manning. 2005. Incorporating non-local information into information extraction systems by gibbs sampling. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*. Association for Computational Linguistics pp. 363–370.
- Fowler, Corinne. 2013. "Journalists in feminist clothing: Men and women reporting Afghan women during Operation Enduring Freedom, 2001." *Journal of International Women's Studies* 8(2):4–19.
- Gamson, William A and Andre Modigliani. 1989. "Media discourse and public opinion on nuclear power: A constructionist approach." *American journal of sociology* pp. 1–37.
- Grimmer, Justin and Brandon M Stewart. 2013. "Text as data: The promise and pitfalls of automatic content analysis methods for political texts." *Political Analysis* .
- Hafner-Burton, Emilie and James Ron. 2013. "The Latin Bias: Regions, the Anglo-American Media, and Human Rights." *International Studies Quarterly* 57(3):474–491.
- Hirschkind, Charles and Saba Mahmood. 2002. "Feminism, the Taliban, and Politics of Counter-insurgency." *Anthropological Quarterly* 75(2):339–354.

- Honaker, James, Gary King and Matthew Blackwell. 2011. "Amelia II: A program for missing data." *Journal of Statistical Software* 45(7):1–47.
- Kinder, Donald R. 1998. "Communication and Opinion." *Annual Review of Political Science* 1(1):167–197.  
**URL:** <http://dx.doi.org/10.1146/annurev.polisci.1.1.167>
- King, Gary, James Honaker, Anne Joseph and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review* 95(01):49–69.
- Klaus, Elisabeth and Susanne Kassel. 2005. "The Veil as a Means of Legitimization An Analysis of the Interconnectedness of Gender, Media and War." *Journalism* 6(3):335–355.
- Kumar, Deepa. 2012. *Islamophobia and the Politics of Empire*. Haymarket Books.
- Lucas, Christopher, Richard A Nielsen, Margaret E Roberts, Brandon M Stewart, Alex Storer and Dustin Tingley. 2015. "Computer-Assisted Text Analysis for Comparative Politics." *Political Analysis* p. mpu019.
- Macdonald, Myra. 2006. "Muslim Women and the Veil: Problems of Image and Voice in Media Representations." *Feminist Media Studies* 6(1):7–23.
- Mahmood, Saba. 2011. *Politics of Piety: The Islamic Revival and the Feminist Subject*. Princeton University Press.
- Maira, Sunaina. 2009. "'Good' and 'Bad' Muslim Citizens: Feminists, Terrorists, and US Orientalisms." *Feminist Studies* pp. 631–656.
- Mani, Lata. 1987. "Contentious tTaditions: The Debate on Sati in Colonial India." *Cultural Critique* pp. 119–156.
- Marshall, Monty G and K Jagers. 2012. "Polity IV Project: Regime authority characteristics and transitions 1800–2011."

- Massad, Joseph A. 2015. *Islam in Liberalism*. University of Chicago Press.
- McAlister, Melani. 2001. *Epic Encounters: Culture, Media, and US Interests in the Middle East, 1945-2000*. Vol. 6 Univ of California Press.
- McCombs, Maxwell E. and Donald L. Shaw. 1972. "The Agenda-setting Function of Mass Media." *Public opinion quarterly* pp. 176–187.
- Mohanty, Chandra Talpade. 1988. "Under Western Eyes: Feminist Scholarship and Colonial Discourses." *Feminist review* pp. 61–88.
- Mohr, John W. and Petko Bogdanov. 2013. "Introduction–Topic Models: What They Are and Why They Matter." *Poetics* 41(6):545–569.
- Neuendorf, Kimberly A. 2011. "Content Analysis: A Methodological Primer for Gender Research." *Sex Roles* 64(3-4):276–289.
- Peksen, Dursun. 2011. "Foreign Military Intervention and Women's Rights." *Journal of Peace Research* 48(4):455–468.
- Porter, Martin F. 2001. *Snowball: A language for stemming algorithms*.  
**URL:** <http://snowball.tartarus.org/texts/introduction.html>
- Puar, Jasbir. 2007. *Terrorist assemblages: Homonationalism in queer times*. Duke University Press.
- Quinn, Kevin M, Burt L Monroe, Michael Colaresi, Michael H Crespin and Dragomir R Radev. 2010. "How to analyze political attention with minimal assumptions and costs." *American Journal of Political Science* 54(1):209–228.
- Razack, Sherene. 2008. *Casting Out: The Eviction of Muslims from Western Law and Politics*. University of Toronto Press.
- Razack, Sherene H. 2004. "Imperilled Muslim Women, Dangerous Muslim Men and Civilised Europeans: Legal and Social Responses to Forced Marriages." *Feminist legal studies* 12(2):129–174.



- Roberts, Margaret E, Brandon M Stewart and Dustin Tingley. 2014. “stm: R Package for Structural Topic Models.”.
- Roberts, Margaret E, Brandon M Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson and David G Rand. 2014. “Structural Topic Models for Open-Ended Survey Responses.” *American Journal of Political Science* 58(4):1064–1082.
- Roberts, Margaret E, Brandon M Stewart, Dustin Tingley, Edoardo M Airolti et al. 2013. The Structural Topic Model and Applied Social Science. In *Advances in Neural Information Processing Systems Workshop on Topic Models: Computation, Application, and Evaluation*.
- Robinson, Piers. 1999. “The CNN Effect: Can the News Media Drive Foreign Policy?” *Review of international studies* 25(02):301–309.
- Robinson, Piers. 2001. “Theorizing the Influence of Media on World Politics Models of Media Influence on Foreign Policy.” *European Journal of Communication* 16(4):523–544.
- Ryan, Louise. 2011. “Muslim Women Negotiating Collective Stigmatisation: ‘We’re Just Normal People’.” *Sociology* .
- Said, Edward W. 2008. *Covering Islam: How the Media and the Experts Determine How We See the Rest of the World (Fully Revised Edition)*. Random House.
- Schraeder, Peter J. and Brian Endless. 1998. “The Media and Africa: The Portrayal of Africa in the ‘New York Times’ (1955-1995).” *Issue: A Journal of Opinion* pp. 29–35.
- Scott, Joan Wallach. 2009. *The Politics of the Veil*. Princeton University Press.
- Shepherd, Laura J. 2006. “Veiled References: Constructions of gender in the Bush Administration Discourse on the Attacks on Afghanistan Post-9/11.” *International Feminist Journal of Politics* 8(1):19–41.
- Sigelman, Lee and Langche Zeng. 1999. “Analyzing Censored and Sample-selected Data with Tobit and Heckit Models.” *Political Analysis* 8(2):167–182.

- Stabile, Carol A. and Deepa Kumar. 2005. "Unveiling Imperialism: Media, Gender and the War on Afghanistan." *Media, Culture & Society* 27(5):765–782.
- Teng, Jinhua Emma. 1996. "The Construction of the 'Traditional Chinese Woman' in the Western Academy: A Critical Review." *Signs* pp. 115–151.
- Terman, Rochelle. 2015. "Islamophobia, Feminism and the Politics of Critique." *Theory, Culture & Society* p. 0263276415590236.
- Volpp, Leti. 2000. "Blaming Culture for Bad Behavior." *Yale JL & Human.* 12:89.
- Wade, Lisa. 2009. "Defining Gendered Oppression in U.S. Newspapers: The Strategic Value of 'Female Genital Mutilation'." *Gender & Society* 23(3):293–314.
- Willett, Peter. 2006. "The Porter Stemming Algorithm: Then and Now." *Program* 40(3):219–223.
- Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel data*. MIT press.
- Yegenoglu, Meyda. 1998. *Colonial Fantasies: Towards a Feminist Reading of Orientalism*. Cambridge University Press.
- Zorn, Christopher JW. 2001. "Generalized Estimating Equation Models for Correlated Data: A Review with Applications." *American Journal of Political Science* pp. 470–490.