# Online Appendix

# #No2Sectarianism:

## Experimental Approaches to Reducing Sectarian Hate Speech Online

Alexandra A. Siegel* and Vivienne Badaan†

April 2020

**Abstract**

We use an experiment across the Arab Twittersphere and a nationally representative survey experiment in Lebanon to evaluate what types of counter-speech interventions are most effective in reducing sectarian hate speech online. We explore whether and to what extent messages priming common national identity or common religious identity, with and without elite endorsements, decrease the use of hostile anti-outgroup language. We find that elite-endorsed messages that prime common religious identity are the most consistently effective in reducing the spread of sectarian hate speech. Our results provide suggestive evidence that religious elites may play an important role as social referents—alerting individuals to social norms of acceptable behavior. By randomly assigning counter-speech treatments to actual producers of online hate speech, and experimentally evaluating the effectiveness of these messages on a representative sample of citizens that might be incidentally exposed to such language, this work offers insights for researchers and policymakers on avenues for combating harmful rhetoric on and offline.

*Postdoctoral Fellow, Immigration Policy Lab, Encina Hall, 616 Serra Street, Stanford University, Stanford, CA 94305

†PhD. Candidate, Department of Psychology, New York University, 6 Washington Place, New York, NY 10003

# List of Figures



# List of Tables

# Twitter Experiment

## Measuring anti-Shia Tweet Content

Table A1: Anti-Shia Slurs

| Anti-Shia | Translation |
|---|---|
| الرافضة | Rejectionist |
| الروافض | Rejectionists |
| حزب الشيطان | Party of the Devil |
| حزب اللات | Party of Laat |
| مجوس | Majus |
| نصيري | Nusayri |
| صفوي | Safavid |

*These keywords were used to filter the initial Twitter dataset to include tweets that contained at least one derogatory reference to the Shia population.*

### Description of anti-Shia Slurs:

In the years following the escalation of the Syrian civil war, six main slurs have frequently been used to disparage Shia Muslims (Abdo 2015; Zelin and Smyth 2014): "Rejectionist" (Rafidha), "Party of the Devil"(Hizb al-Shaytan), "Party of Lat" (Hizb al-Laat), "Majus," "Followers of Nusayr" (Nusayri), and"Safavid" (Safawi). "Rejectionist" refers to Twelver Shiites, the largest of the Shia sects, and implies that they have rejected "true" Islam as they allegedly do not recognize Abu Bakr and his successors as having been legitimate rulers after the death of the Prophet Mohammad. "Party of the Devil" and "Party of Laat" are both used in reference to Hezbollah and its Shia followers. "Laat" alludes to the pre-Islamic Arabian goddess al-Laat, who was believed to be a daughter of God. This brands Hezbollah and its supporters as a group of polytheist non-believers. "Majus" is a derogatory term that references Zoroastrianism, implying that Shia Islam is nothing more than a deviant religion

of the past. "Nusrayri" or "Followers of Nusayr" is a reference to Abu Shuayb Muhammad Ibn Nusayr, the founder of the Alawite offshoot of Shia Islam during the eighth century. It implies that the Alawite religion is not divinely inspired as it follows a man, rather than God. Finally, "Safawi," which recalls the Safavid dynasty that ruled Persia from 1501 to 1736, is used to depict Shia ties to Iran. Sometimes the term is also used as a neologism of "Sahiyyu-Safawi" (Zionist-Safawi) to suggest that there is a conspiracy between Israel and Iran against Sunni Muslims.

## Descriptive Statistics

Table A2: Descriptive Statistics for Twitter Experiment

| | n | mean | median | sd | min | max |
|---|---|---|---|---|---|---|
| Anti-Shia Pre-Treatment Tweet Count (Days 1-30) | 957 | 10.55 | 3 | 25.04 | 0 | 350 |
| Anti-Shia Post-Treatment Tweet Count (Days 1-30) | 795 | 12.38 | 3 | 33.22 | 0 | 612 |
| Pre-Treatment Total Tweet Count (Days 1-30) | 957 | 247.58 | 111 | 378.47 | 0 | 3719 |
| Post-Treatment Total Tweet Count (Days 1-30) | 795 | 289.93 | 138 | 405.16 | 0 | 3124 |
| Anti-Shia Pre-Treatment Tweet Count (Days 15-30) | 957 | 4.93 | 1 | 11.18 | 0 | 117 |
| Anti-Shia Post-Treatment Tweet Count (Days 15-30) | 795 | 5.86 | 1 | 17.99 | 0 | 385 |
| Pre-Treatment Total Tweet Count (Days 15-30) | 957 | 127.79 | 54 | 200.54 | 0 | 2071 |
| Post-Treatment Total Tweet Count (Days 15-30) | 795 | 150.59 | 70 | 217.87 | 0 | 1775 |
| Anti-Shia Pre-Treatment Tweet Count (Days 8-14) | 957 | 2.82 | 1 | 7.47 | 0 | 104 |
| Anti-Shia Post-Treatment Tweet Count (Days 8-14) | 922 | 3.36 | 0 | 8.97 | 0 | 162 |
| Pre-Treatment Total Tweet Count (Days 8-14) | 957 | 60.46 | 24 | 99.84 | 0 | 861 |
| Post-Treatment Total Tweet Count (Days 8-14) | 921 | 63.95 | 28 | 97.29 | 0 | 906 |
| Anti-Shia Pre-Treatment Tweet Count (Days 2-7) | 957 | 2.02 | 0 | 6.42 | 0 | 104 |
| Anti-Shia Post-Treatment Tweet Count (Days 2-7) | 944 | 2 | 0 | 6.20 | 0 | 108 |
| Pre-Treatment Total Tweet Count (Days 2-7) | 957 | 44.63 | 18 | 71.62 | 0 | 612 |
| Post-Treatment Total Tweet Count (Days 2-7) | 944 | 47.98 | 20 | 77.17 | 0 | 657 |
| Anti-Shia Pre-Treatment Tweet Count (Day 1) | 957 | 0.78 | 0 | 2.15 | 0 | 38 |
| Anti-Shia Post-Treatment Tweet Count (Day 1) | 952 | 0.67 | 0 | 2.21 | 0 | 34 |
| Pre-Treatment Total Tweet Count (Day 1) | 957 | 14.69 | 5 | 24.55 | 0 | 193 |
| Post-Treatment Total Tweet Count (Day 1) | 952 | 15.06 | 5 | 24.38 | 0 | 186 |
| Followers Count | 957 | 846.76 | 245 | 1433.43 | 1 | 9558 |

## Full Sample Results (Main Specification)

Table A3: Effect of Treatment on Anti-Shia Tweet Count

| | *Change in Anti-Shia Tweet Count (Non-Overlapping Treatment Periods)* | | | |
|---|---|---|---|---|
| | Day | Week | Two Weeks | Month |
| Arab ID | 0.074 | 0.116 | −0.249 | −0.403 |
| | (0.236) | (0.574) | (0.811) | (1.811) |
| Religious ID | 0.006 | −0.421 | −0.221 | 1.376 |
| | (0.244) | (0.594) | (0.836) | (1.896) |
| Arab ID (Elite) | 0.243 | 0.072 | −0.557 | −0.461 |
| | (0.239) | (0.584) | (0.828) | (1.889) |
| Religious ID (Elite) | −0.948*** | −1.625** | −2.817*** | −6.405*** |
| | (0.238) | (0.580) | (0.816) | (1.867) |
| No ID | −0.108 | −0.147 | −0.787 | −1.657 |
| | (0.232) | (0.567) | (0.798) | (1.820) |
| Constant | 0.015 | 0.298 | 1.304* | 1.816 |
| | (0.176) | (0.431) | (0.612) | (1.404) |
| Observations | 952 | 944 | 922 | 795 |
| $R^2$ | 0.035 | 0.015 | 0.019 | 0.029 |
| Adjusted $R^2$ | 0.030 | 0.010 | 0.014 | 0.023 |
| Residual Std. Error | 2.033 (df = 946) | 4.934 (df = 938) | 6.839 (df = 916) | 14.245 (df = 789) |
| F Statistic | 6.922*** (df = 5; 946) | 2.892* (df = 5; 938) | 3.582** (df = 5; 916) | 4.726*** (df = 5; 789) |

*Note:* †p<0.1; *p<0.05; **p<0.01; ***p<01

*Estimates are from OLS models (difference in means). The dependent variable is the change in the anti-Shia tweet count in each post-treatment period. Standard Errors are in parentheses. Analysis is conducted on full sample.*

## Geographic Breakdown

It is also possible that the effects of our treatments might be different for Twitter users from different countries or regions. Examining the top 10 locations of users in our sample, we see that while the majority of users are located in the Gulf—and therefore share an identity with our sockpuppet Mohammed Ahmed—many users are located in other parts of the Arab World and may have had different responses to our treatments.

Table A4: Distribution of User Locations

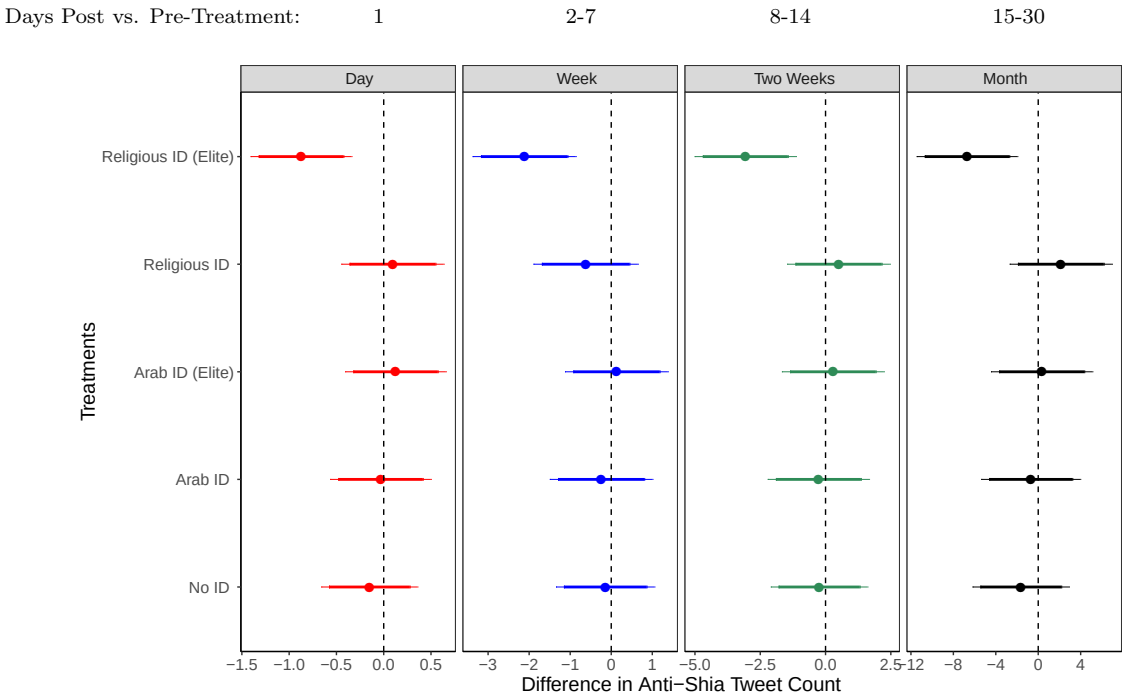
| Locations | Freq |
|---|---|
| Gulf Cooperation Council Countries | 630 |
| Yemen | 127 |
| Syria | 37 |
| Iraq | 23 |
| Lebanon | 11 |
| Egypt | 9 |
| US | 8 |
| Algeria | 7 |
| Turkey | 6 |
| Libya | 5 |

*Table A4 shows the top ten locations of subjects in our experiment. Location was determined first by analyzing tweet-level metadata and then by manual coding of each user's profile to ensure more accurate location identification.*

Restricting our analysis only to users in Gulf countries, we see slightly larger treatment effects than what we observe in the full sample. By contrast, when we restrict our data analysis to users from conflict zones–Yemen, Syria and Iraq—the next most popular locations in our sample after the Gulf countries, we do not see statistically significant treatment effects

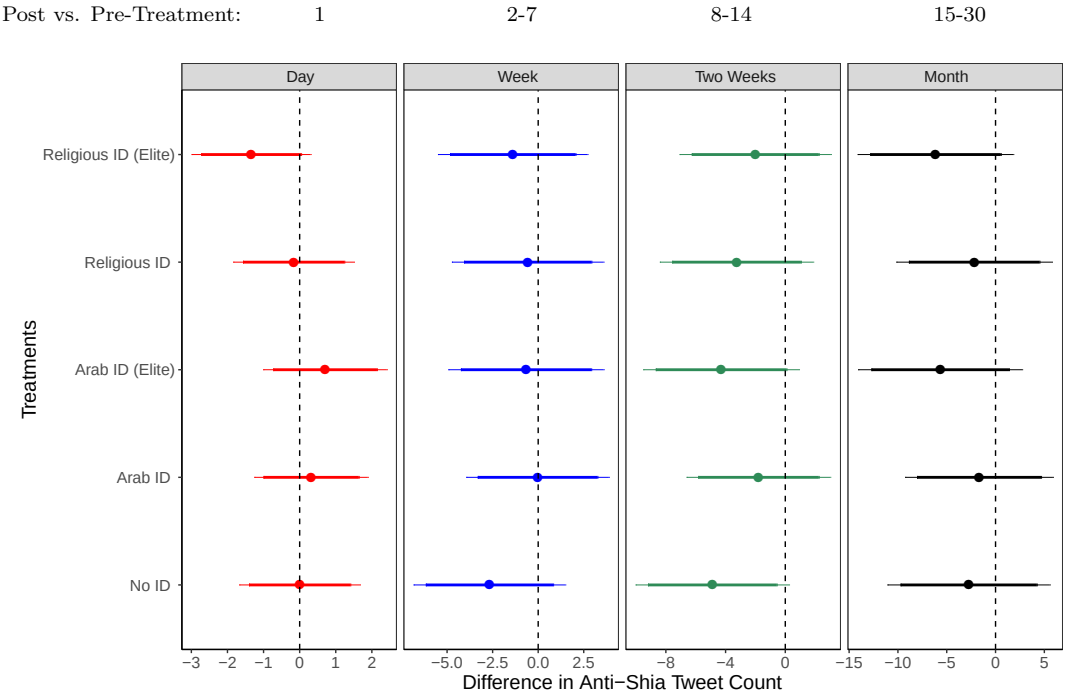
one day, week, or two weeks after treatment. These results are displayed in Figure A2. Perhaps, unsurprisingly, in locations where continuous exposure to violent sectarian conflict is commonplace, simply being criticized on Twitter is not sufficient to stop people from expressing anti-outgroup hatred on social media. Of course, it is important to note that due to the much smaller sample size of users located in conflict zones, these differences may also be driven by a lack of statistical power.

Figure A1: Effect of Treatment on Anti-Shia Tweet Count (Gulf Users)




*This coefficient plot shows the results of four OLS models, where the outcome variable is the daily, weekly, biweekly, or monthly count of anti-Shia tweets produced by subjects (Twitter users) from the Gulf. Error bars show 90% and 95% confidence intervals. Location of Twitter users was determined using account metadata and manual coding of each user profile. The full output is displayed in regression Table A5.*

Figure A2: Effect of Treatment on Anti-Shia Tweet Count
(Conflict Zone Users)




*This coefficient plot shows the results of four OLS models, where the outcome variable is the daily, weekly, biweekly, or monthly count of anti-Shia tweets produced by Twitter users from conflict zones (Yemen, Iraq, and Syria). Error bars show 90% and 95% confidence intervals. Location of Twitter users was determined using account metadata and manual coding of each user profile. The full output is displayed in regression Table A6.*

Table A5: Effect of Treatment on Anti-Shia Tweet Count
(Gulf Users)

| | *Change in Anti-Shia Tweet Count* | | | |
|---|---|---|---|---|
| | Day | Week | Two Weeks | Month |
| Arab ID | −0.034 | −0.246 | −0.273 | −0.706 |
| | (0.272) | (0.638) | (0.989) | (2.382) |
| Religious ID | 0.093 | −0.623 | 0.498 | 2.131 |
| | (0.276) | (0.648) | (1.003) | (2.458) |
| Arab ID (Elite) | 0.125 | 0.127 | 0.283 | 0.328 |
| | (0.272) | (0.639) | (0.996) | (2.439) |
| Religious ID (Elite) | −0.874** | −2.120*** | −3.071** | −6.711** |
| | (0.272) | (0.641) | (0.989) | (2.420) |
| No ID | −0.152 | −0.146 | −0.244 | −1.644 |
| | (0.259) | (0.612) | (0.944) | (2.318) |
| Constant | 0.024 | 0.463 | 1.051 | 1.500 |
| | (0.201) | (0.474) | (0.738) | (1.815) |
| Observations | 605 | 599 | 590 | 539 |
| $R^2$ | 0.033 | 0.030 | 0.032 | 0.032 |
| Adjusted $R^2$ | 0.025 | 0.022 | 0.024 | 0.023 |
| Residual Std. Error | 1.839 (df = 599) | 4.291 (df = 593) | 6.558 (df = 584) | 15.188 (df = 533) |
| F Statistic | 4.069** (df = 5; 599) | 3.725** (df = 5; 593) | 3.900** (df = 5; 584) | 3.481** (df = 5; 533) |

±p<0.01; *p<0.05; **p<0.01; ***p<0.001

*Estimates are from four OLS models, where the outcome variable is the daily, weekly, biweekly, or monthly change in the count of anti-Shia tweets produced by subjects (Twitter users) from the Gulf. Location of Twitter users was determined using account metadata and manual coding of each user profile.*

Table A6: Effect of Treatment on Anti-Shia Tweet Count
(Conflict Zone Users)

| | *Change in Anti-Shia Tweet Count* | | | |
|---|---|---|---|---|
| | Day | Week | Two Weeks | Month |
| Arab ID | 0.317 | −0.032 | −1.797 | −1.687 |
| | (0.802) | (1.995) | (2.454) | (3.861) |
| Religious ID | −0.167 | −0.572 | −3.264 | −2.186 |
| | (0.853) | (2.123) | (2.620) | (4.066) |
| Arab ID (Elite) | 0.704 | −0.665 | −4.296 | −5.670 |
| | (0.873) | (2.173) | (2.661) | (4.284) |
| Religious ID (Elite) | −1.344 | −1.395 | −2.011 | −6.150 |
| | (0.841) | (2.094) | (2.585) | (4.066) |
| No ID | −0.000 | −2.672 | −4.885$^{\dagger}$ | −2.753 |
| | (0.853) | (2.123) | (2.620) | (4.241) |
| Constant | 0.000 | 0.739 | 3.333$^{\dagger}$ | 4.579 |
| | (0.642) | (1.597) | (1.996) | (3.138) |
| Observations | 183 | 183 | 178 | 157 |
| $R^2$ | 0.043 | 0.015 | 0.028 | 0.024 |
| Adjusted $R^2$ | 0.016 | −0.013 | −0.0001 | −0.009 |
| Residual Std. Error | 3.077 (df = 177) | 7.659 (df = 177) | 9.145 (df = 172) | 13.679 (df = 151) |
| F Statistic | 1.578 (df = 5; 177) | 0.533 (df = 5; 177) | 0.998 (df = 5; 172) | 0.736 (df = 5; 151) |

$^{\pm}$p<0.01; *p<0.05; **p<0.01; ***p<0.001

*Estimates are from four OLS models, where the outcome variable is the daily, weekly, biweekly, or monthly change in the count of anti-Shia tweets produced by subjects (Twitter users) from Yemen, Iraq, or Syria. Location of Twitter users was determined using account metadata and manual coding of each user profile.*

# Attrition

Table A7: Effect of Treatment on Anti-Shia Tweet Count
(Excluding Suspended Accounts)

| | *Change in Anti-Shia Tweet Count (Non-Overlapping Treatment Periods)* | | | |
|---|---|---|---|---|
| | Day | Week | Two Weeks | Month |
| A Arab ID | −0.210 | −0.117 | −0.570 | −0.403 |
| | (0.264) | (0.664) | (0.917) | (1.811) |
| Religious ID | −0.058 | −0.337 | −0.147 | 1.376 |
| | (0.276) | (0.695) | (0.960) | (1.896) |
| Arab ID (Elite) | 0.185 | 0.048 | −0.791 | −0.461 |
| | (0.275) | (0.693) | (0.957) | (1.889) |
| Religious ID (Elite) | −1.116*** | −1.682* | −2.832** | −6.405*** |
| | (0.272) | (0.685) | (0.946) | (1.867) |
| No ID | −0.285 | −0.276 | −1.245 | −1.657 |
| | (0.265) | (0.668) | (0.922) | (1.820) |
| Constant | 0.146 | 0.369 | 1.563* | 1.816 |
| | (0.204) | (0.515) | (0.711) | (1.404) |
| Observations | 795 | 795 | 795 | 795 |
| $R^2$ | 0.038 | 0.013 | 0.016 | 0.029 |
| Adjusted $R^2$ | 0.032 | 0.006 | 0.010 | 0.023 |
| Residual Std. Error (df = 789) | 2.074 | 5.225 | 7.217 | 14.245 |
| F Statistic (df = 5; 789) | 6.316*** | $2.033^{dagger}$ | 2.635* | 4.726*** |

*Note:* $^{\dagger}$p<0.1; *p<0.05; **p<0.01; ***p<0.001

*Estimates are from OLS models. The dependent variable is the change in the anti-Shia tweet count in each post-treatment period. Standard Errors are in parentheses. Analysis is conducted on full sample excluding deleted or suspended accounts.*

Table A8: Effect of Treatment on Anti-Shia Tweet Count
(Suspended Accounts = 0)

| | *Change in Anti-Shia Tweet Count (Non-Overlapping Treatment Periods)* | | | |
|---|---|---|---|---|
| | Day | Week | Two Weeks | Month |
| Arab ID | 0.067 | 0.141 | −0.119 | 0.167 |
| | (0.235) | (0.567) | (0.782) | (1.530) |
| Religious ID | 0.005 | −0.393 | 0.001 | 1.487 |
| | (0.243) | (0.587) | (0.808) | (1.581) |
| Arab ID (Elite) | 0.229 | 0.086 | −0.379 | 0.124 |
| | (0.238) | (0.575) | (0.793) | (1.550) |
| Religious ID (Elite) | −0.942*** | −1.588** | −2.560** | −5.165*** |
| | (0.237) | (0.572) | (0.788) | (1.542) |
| No ID | −0.108 | −0.134 | −0.639 | −1.056 |
| | (0.231) | (0.559) | (0.770) | (1.505) |
| Constant | 0.015 | 0.271 | 1.060$^{dagger}$ | 0.744 |
| | (0.176) | (0.425) | (0.586) | (1.146) |
| Observations | 957 | 957 | 957 | 957 |
| $R^2$ | 0.035 | 0.015 | 0.018 | 0.025 |
| Adjusted $R^2$ | 0.030 | 0.010 | 0.012 | 0.020 |
| Residual Std. Error (df = 951) | 2.030 | 4.902 | 6.754 | 13.212 |
| F Statistic (df = 5; 951) | 6.813*** | 2.886* | 3.405** | 4.911*** |

*Note:* $^{\dagger}$p<0.1; *p<0.05; **p<0.01; ***p<0.001

*Estimates are from OLS models. The dependent variable is the change in the anti-Shia tweet count in each post-treatment period. Standard Errors are in parentheses. Analysis is conducted on full sample with suspended/deleted accounts treated as 0 after they drop out of the sample.*

## Results by Follower Count

Table A9: Effect of Treatment on Anti-Shia Tweet Count
(Users with < Median Followers)

| | *Dependent variable: Difference in Anti-Shia Tweet Count* | | | |
|---|---|---|---|---|
| | Day | Week | Two Weeks | Month |
| Arab ID | −0.271 | −0.061 | −0.727 | −0.838 |
| | (0.405) | (0.974) | (1.504) | (3.518) |
| Religious ID | −0.450 | −0.384 | −0.637 | 3.316 |
| | (0.411) | (0.990) | (1.527) | (3.614) |
| Arab ID (Elite) | 0.052 | −0.541 | −1.614 | −0.806 |
| | (0.398) | (0.960) | (1.490) | (3.527) |
| Religious ID (Elite) | −1.662*** | −2.086* | −3.630* | −8.441* |
| | (0.404) | (0.974) | (1.501) | (3.573) |
| No ID | −0.620 | −0.999 | −2.318 | −2.751 |
| | (0.392) | (0.946) | (1.466) | (3.494) |
| Constant | 0.357 | 0.691 | 1.840 | 2.238 |
| | (0.311) | (0.752) | (1.180) | (2.817) |
| Observations | 477 | 473 | 461 | 403 |
| $R^2$ | 0.058 | 0.016 | 0.020 | 0.037 |
| Adjusted $R^2$ | 0.048 | 0.005 | 0.009 | 0.025 |
| Residual Std. Error | 2.329 (df = 471) | 5.574 (df = 467) | 8.344 (df = 455) | 18.256 (df = 397) |
| F Statistic | 5.797*** (df = 5; 471) | 1.502 (df = 5; 467) | 1.834 (df = 5; 455) | 3.056* (df = 5; 397) |

*Note:* $^{\dagger}$p<0.1; *p<0.05; **p<0.01; ***p<0.001

*Estimates are from OLS models. The dependent variable is the change in the anti-Shia tweet count in each post-treatment period. Standard Errors are in parentheses. Analysis is limited to Twitter users who have equal to or less than the median number of Twitter followers.*

Table A10: Effect of Treatment on Anti-Shia Tweet Count
(Users with < 200 Followers)

| | *Dependent variable: Difference in Anti-Shia Tweet Count* | | | |
|---|---|---|---|---|
| | Day | Week | Two Weeks | Month |
| Arab ID | −0.350 | −0.126 | −0.654 | −1.126 |
| | (0.442) | (1.026) | (1.333) | (2.196) |
| Religious ID | −0.552 | −0.955 | −2.137 | −1.456 |
| | (0.457) | (1.059) | (1.370) | (2.265) |
| Arab ID (Elite) | −0.014 | −0.647 | −1.519 | −0.888 |
| | (0.436) | (1.014) | (1.324) | (2.201) |
| Religious ID (Elite) | −1.823*** | −2.285* | −3.828** | −9.278*** |
| | (0.448) | (1.043) | (1.350) | (2.251) |
| No ID | −0.831† | −1.141 | −2.080 | −3.779† |
| | (0.436) | (1.014) | (1.321) | (2.213) |
| Constant | 0.429 | 0.771 | 1.814† | 2.541 |
| | (0.345) | (0.803) | (1.063) | (1.776) |
| Observations | 425 | 421 | 410 | 363 |
| $R^2$ | 0.062 | 0.017 | 0.027 | 0.077 |
| Adjusted $R^2$ | 0.051 | 0.006 | 0.015 | 0.064 |
| Residual Std. Error | 2.415 (df = 419) | 5.564 (df = 415) | 6.969 (df = 404) | 10.803 (df = 357) |
| F Statistic | 5.529*** (df = 5; 419) | 1.478 (df = 5; 415) | 2.246* (df = 5; 404) | 5.942*** (df = 5; 357) |

*Note:* †p<0.1; *p<0.05; **p<0.01; ***p<0.001

*Estimates are from OLS models. The dependent variable is the change in the anti-Shia tweet count in each post-treatment period. Standard Errors are in parentheses. Analysis is limited to Twitter users who have 200 followers or fewer.*

Table A11: Effect of Treatment on Anti-Shia Tweet Count
(Users with < 300 Followers)

| | *Dependent variable: Difference in Anti-Shia Tweet Count* | | | |
|---|---|---|---|---|
| | Day | Week | Two Weeks | Month |
| Arab ID | −0.237 | −0.002 | −0.596 | −0.891 |
| | (0.377) | (0.920) | (1.412) | (3.299) |
| Religious ID | −0.294 | 0.116 | 0.095 | 3.400 |
| | (0.382) | (0.934) | (1.432) | (3.390) |
| Arab ID (Elite) | 0.074 | −0.468 | −1.436 | −0.836 |
| | (0.372) | (0.912) | (1.409) | (3.329) |
| Religious ID (Elite) | −1.565*** | −1.945* | −3.250* | −8.033* |
| | (0.378) | (0.927) | (1.418) | (3.371) |
| No ID | −0.528 | −0.845 | −2.011 | −2.676 |
| | (0.363) | (0.892) | (1.370) | (3.278) |
| Constant | 0.306 | 0.600 | 1.655 | 2.200 |
| | (0.288) | (0.708) | (1.103) | (2.645) |
| Observations | 511 | 505 | 493 | 429 |
| $R^2$ | 0.055 | 0.017 | 0.020 | 0.036 |
| Adjusted $R^2$ | 0.045 | 0.007 | 0.010 | 0.025 |
| Residual Std. Error | 2.265 (df = 505) | 5.485 (df = 499) | 8.177 (df = 487) | 17.742 (df = 423) |
| F Statistic | 5.847*** (df = 5; 505) | 1.683 (df = 5; 499) | 1.984† (df = 5; 487) | 3.194** (df = 5; 423) |

*Note:* †p<0.1; *p<0.05; **p<0.01; ***p<0.001

*Estimates are from OLS models. The dependent variable is the change in the anti-Shia tweet count in each post-treatment period. Standard Errors are in parentheses. Analysis is limited to Twitter users who have 300 followers or fewer.*

Table A12: Effect of Treatment on Anti-Shia Tweet Count
(Users with <350 Followers)

| | *Dependent variable: Difference in Anti-Shia Tweet Count* | | | |
|---|---|---|---|---|
| | Day | Week | Two Weeks | Month |
| Arab ID | −0.216 | 0.030 | −0.661 | −0.939 |
| | (0.371) | (0.969) | (1.358) | (3.126) |
| Religious ID | −0.167 | −0.519 | 0.093 | 2.596 |
| | (0.374) | (0.976) | (1.368) | (3.187) |
| Arab ID (Elite) | 0.089 | −0.425 | −1.459 | −0.824 |
| | (0.367) | (0.960) | (1.355) | (3.155) |
| Religious ID (Elite) | −1.242*** | −2.335* | −3.402* | −8.302** |
| | (0.365) | (0.954) | (1.335) | (3.133) |
| No ID | −0.500 | −0.554 | −1.720 | −2.204 |
| | (0.355) | (0.931) | (1.306) | (3.080) |
| Constant | 0.284 | 0.554 | 1.683 | 2.204 |
| | (0.280) | (0.737) | (1.048) | (2.478) |
| Observations | 544 | 538 | 526 | 458 |
| $R^2$ | 0.037 | 0.018 | 0.021 | 0.037 |
| Adjusted $R^2$ | 0.029 | 0.009 | 0.012 | 0.026 |
| Residual Std. Error | 2.295 (df = 538) | 5.939 (df = 532) | 8.114 (df = 520) | 17.349 (df = 452) |
| F Statistic | 4.186*** (df = 5; 538) | 1.973† (df = 5; 532) | 2.272* (df = 5; 520) | 3.472** (df = 5; 452) |

*Note:* †p<0.1; *p<0.05; **p<0.01; ***p<0.001

*Estimates are from OLS models. The dependent variable is the change in the anti-Shia tweet count in each post-treatment period. Standard Errors are in parentheses. Analysis is limited to Twitter users who have 350 followers or fewer.*

Table A13: Effect of Treatment on Anti-Shia Tweet Count
(Users with < 400 Followers)

| | *Dependent variable: Difference in Anti-Shia Tweet Count* | | | |
|---|---|---|---|---|
| | Day | Week | Two Weeks | Month |
| Arab ID | −0.138 | 0.180 | −1.088 | −1.114 |
| | (0.358) | (0.922) | (1.297) | (2.990) |
| Religious ID | 0.017 | −0.634 | −0.655 | 1.406 |
| | (0.360) | (0.929) | (1.306) | (3.053) |
| Arab ID (Elite) | 0.206 | −0.277 | −1.882 | −1.092 |
| | (0.355) | (0.918) | (1.300) | (3.038) |
| Religious ID (Elite) | −1.172** | −2.310* | −3.981** | −8.916** |
| | (0.354) | (0.916) | (1.287) | (3.031) |
| No ID | −0.426 | −0.541 | −2.279† | −2.486 |
| | (0.345) | (0.894) | (1.258) | (2.972) |
| Constant | 0.222 | 0.514 | 2.215* | 2.519 |
| | (0.270) | (0.700) | (0.998) | (2.359) |
| Observations | 569 | 563 | 551 | 480 |
| $R^2$ | 0.039 | 0.019 | 0.024 | 0.036 |
| Adjusted $R^2$ | 0.031 | 0.011 | 0.015 | 0.025 |
| Residual Std. Error | 2.288 (df = 563) | 5.855 (df = 557) | 8.043 (df = 545) | 17.334 (df = 474) |
| F Statistic | 4.582*** (df = 5; 563) | 2.215† (df = 5; 557) | 2.687* (df = 5; 545) | 3.500** (df = 5; 474) |

*Note:* †p<0.1; *p<0.05; **p<0.01; ***p<0.001

*Estimates are from OLS models. The dependent variable is the change in the anti-Shia tweet count in each post-treatment period. Standard Errors are in parentheses. Analysis is limited to Twitter users who have 400 followers or fewer.*

## Results by Network Type

Table A14: Effect of Treatment on Anti-Shia Tweet Count
(Low Anti-Shia Networks)

| | *Change in Anti-Shia Tweet Count* | | | |
|---|---|---|---|---|
| | Day | Week | 2 Weeks | Month |
| Arab ID | 0.014 | 0.166 | −0.200 | −0.701 |
| | (0.384) | (0.902) | (1.376) | (3.499) |
| Religious ID | −0.285 | 0.221 | 0.375 | 3.695 |
| | (0.390) | (0.916) | (1.390) | (3.583) |
| Arab ID (Elite) | 0.358 | −0.062 | −0.925 | −0.381 |
| | (0.387) | (0.911) | (1.393) | (3.616) |
| Religious ID (Elite) | −1.345*** | −1.947* | −3.334* | −8.641* |
| | (0.389) | (0.913) | (1.386) | (3.665) |
| No ID | −0.879* | −1.537† | −2.696* | −4.423 |
| | (0.379) | (0.894) | (1.355) | (3.525) |
| Constant | 0.172 | 0.404 | 1.296 | 2.048 |
| | (0.297) | (0.700) | (1.071) | (2.801) |
| Observations | 480 | 476 | 464 | 377 |
| $R^2$ | 0.064 | 0.028 | 0.032 | 0.044 |
| Adjusted $R^2$ | 0.054 | 0.017 | 0.021 | 0.031 |
| Residual Std. Error | 2.260 (df = 474) | 5.284 (df = 470) | 7.870 (df = 458) | 18.155 (df = 371) |
| F Statistic | 6.500*** (df = 5; 474) | 2.680* (df = 5; 470) | 2.997* (df = 5; 458) | 3.399** (df = 5; 371) |

*Note:* *p<0.1; **p<0.05; ***p<0.01

*Estimates are from OLS models. The dependent variable is the change in anti-Shia tweet count in each post-treatment period. Standard Errors are in parentheses. Analysis is limited to subjects (Twitter users) with below the median number of friends in their networks who tweet anti-Shia content.*

Table A15: Effect of Treatment on Anti-Shia Tweet Count
(High Anti-Shia Networks)

| | *Change in Anti-Shia Tweet Count* | | | |
|---|---|---|---|---|
| | Day | Week | 2 Weeks | Month |
| Arab ID | 0.095 | 0.037 | −0.298 | −0.181 |
| | (0.275) | (0.718) | (0.892) | (1.580) |
| Religious ID | 0.288 | −1.247 | −0.941 | −1.317 |
| | (0.292) | (0.760) | (0.944) | (1.697) |
| Arab ID (Elite) | 0.067 | 0.184 | −0.157 | −0.609 |
| | (0.281) | (0.735) | (0.920) | (1.663) |
| Religious ID (Elite) | −0.588* | −1.327† | −2.297* | −4.669** |
| | (0.276) | (0.722) | (0.897) | (1.602) |
| No ID | 0.640* | 1.233† | 1.167 | 0.813 |
| | (0.270) | (0.706) | (0.877) | (1.584) |
| Constant | −0.107 | 0.216 | 1.310* | 1.656 |
| | (0.199) | (0.522) | (0.653) | (1.190) |
| Observations | 472 | 468 | 458 | 418 |
| $R^2$ | 0.047 | 0.039 | 0.038 | 0.038 |
| Adjusted $R^2$ | 0.037 | 0.028 | 0.027 | 0.026 |
| Residual Std. Error | 1.727 (df = 466) | 4.489 (df = 462) | 5.501 (df = 452) | 9.294 (df = 412) |
| F Statistic | 4.572*** (df = 5; 466) | 3.720** (df = 5; 462) | 3.567** (df = 5; 452) | 3.241** (df = 5; 412) |

*Note:* *p<0.1; **p<0.05; ***p<0.01

*Estimates are from OLS models. The dependent variable is the change in anti-Shia tweet count in each post-treatment period. Standard Errors are in parentheses. Analysis is limited to subjects (Twitter users) with above the median number of friends in their networks who tweet anti-Shia content.*

# Alternative Model Specifications

Table A16: Effect of Treatment on Anti-Shia Tweet Count
(Negative Binomial Models)

| | *Anti-Shia Tweet Count (Non-Overlapping Treatment Periods)* | | | |
|---|---|---|---|---|
| | Day | Week | Two Weeks | Month |
| Arab ID | 0.034 | −0.005 | −0.089 | −0.261 |
| | (0.312) | (0.229) | (0.204) | (0.210) |
| Religious ID | 0.009 | −0.260 | −0.235 | −0.186 |
| | (0.322) | (0.240) | (0.211) | (0.219) |
| Arab ID (Elite) | 0.042 | −0.319 | −0.347* | −0.255 |
| | (0.316) | (0.237) | (0.210) | (0.219) |
| Religious ID (Elite) | −0.904*** | −1.231*** | −0.732*** | −1.400*** |
| | (0.335) | (0.250) | (0.209) | (0.225) |
| No ID | 0.277 | 0.260 | 0.092 | −0.246 |
| | (0.301) | (0.224) | (0.198) | (0.211) |
| Anti-Shia Pre-Treatment Tweet Count (Day) | 0.400*** | | | |
| | (0.035) | | | |
| Anti-Shia Pre-Treatment Tweet Count (Week) | | 0.166*** | | |
| | | (0.009) | | |
| Anti-Shia Pre-Treatment Tweet Count (Two Weeks) | | | 0.133*** | |
| | | | (0.007) | |
| Anti-Shia Pre-Treatment Tweet Count (Month) | | | | 0.075*** |
| | | | | (0.005) |
| Anti-Shia Pre-Treatment Tweet Count (Two Months) | −1.008*** | 0.076 | 0.573*** | 1.224*** |
| | (0.234) | (0.173) | (0.154) | (0.164) |
| Observations | 952 | 944 | 922 | 795 |
| Log Likelihood | −825.978 | −1,378.359 | −1,746.272 | −1,809.994 |
| $\theta$ | 0.203*** (0.023) | 0.323*** (0.026) | 0.406*** (0.029) | 0.406*** (0.028) |
| Akaike Inf. Crit. | 1,665.956 | 2,770.719 | 3,506.544 | 3,633.988 |

*Note:* *p<0.1; **p<0.05; ***p<0.01

Estimates are from negative binomial models. The dependent variable is the anti-Shia tweet count in each post-treatment period. Model includes control for pre-treatment anti-Shia tweet count. Coefficients are reported as odds ratios. Standard Errors are in parentheses. Analysis is conducted on full sample.

Table A17: Effect of Treatment on Proportion of Anti-Shia Tweets

| | *Change in Prop. Anti-Shia Tweets (Non-Overlapping Treatment Periods)* | | | |
|---|---|---|---|---|
| | Day | Week | Two Weeks | Month |
| Arab ID | 0.0003 | −0.002 | 0.009 | 0.006 |
| | (0.025) | (0.017) | (0.019) | (0.017) |
| Religious ID | −0.018 | −0.032$^{\dagger}$ | −0.013 | −0.001 |
| | (0.026) | (0.018) | (0.020) | (0.018) |
| Arab ID (Elite) | 0.006 | 0.003 | 0.008 | 0.006 |
| | (0.026) | (0.017) | (0.020) | (0.018) |
| Religious ID (Elite) | −0.136$^{***}$ | −0.081$^{***}$ | −0.091$^{***}$ | −0.123$^{***}$ |
| | (0.024) | (0.017) | (0.019) | (0.018) |
| No ID | −0.053$^{*}$ | 0.009 | −0.001 | −0.021 |
| | (0.025) | (0.017) | (0.019) | (0.018) |
| Constant | −0.005 | 0.007 | 0.018 | 0.020 |
| | (0.018) | (0.013) | (0.015) | (0.013) |
| Observations | 711 | 887 | 873 | 769 |
| $R^2$ | 0.071 | 0.050 | 0.049 | 0.108 |
| Adjusted $R^2$ | 0.064 | 0.045 | 0.043 | 0.102 |
| Residual Std. Error | 0.190 (df = 705) | 0.141 (df = 881) | 0.159 (df = 867) | 0.135 (df = 763) |
| F Statistic | 10.755$^{***}$ (df = 5; 705) | 9.289$^{***}$ (df = 5; 881) | 8.928$^{***}$ (df = 5; 867) | 18.397$^{***}$ (df = 5; 763) |

*Note:* $^{\dagger}$p<0.1; $^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Estimates are from OLS models. The dependent variable is the change in the proportion of anti-Shia tweets in each post-treatment period. Standard Errors are in parentheses. Analysis is conducted on full sample.*

# Controlling for Treatment Date

Table A18: Effect of Treatment on Anti-Shia Tweet Count
(Treatment Date Fixed Effect)

| | *Change in Anti-Shia Tweet Count* | | | |
|---|---|---|---|---|
| | Day | Week | 2 Weeks | Month |
| Arab ID | 0.069 | 0.124 | −0.270 | −0.614 |
| | (0.234) | (0.577) | (0.818) | (1.833) |
| Religious ID | −0.099 | −0.443 | −0.304 | 1.117 |
| | (0.244) | (0.601) | (0.849) | (1.928) |
| Arab ID (Elite) | 0.164 | 0.235 | −0.528 | −0.226 |
| | (0.237) | (0.586) | (0.834) | (1.910) |
| Religious ID (Elite) | −0.946*** | −1.561** | −2.762*** | −6.518*** |
| | (0.236) | (0.582) | (0.824) | (1.888) |
| No ID | −0.091 | −0.070 | −0.722 | −1.709 |
| | (0.231) | (0.571) | (0.806) | (1.845) |
| 2/1/18 | −0.795† | 1.568 | 3.875** | 7.278* |
| | (0.409) | (1.003) | (1.394) | (3.194) |
| 2/10/18 | −0.518 | −1.143 | 0.875 | 0.397 |
| | (0.419) | (1.028) | (1.429) | (3.243) |
| 2/11/18 | −0.870* | −0.633 | 0.939 | 0.363 |
| | (0.413) | (1.014) | (1.440) | (3.311) |
| 2/13/18 | −0.612 | 0.066 | 2.202 | 3.061 |
| | (0.410) | (1.006) | (1.398) | (3.110) |
| 2/14/18 | −0.829* | −2.558* | −0.438 | −0.662 |
| | (0.418) | (1.026) | (1.443) | (3.367) |
| 2/15/18 | −0.456 | 0.062 | 2.220 | 1.466 |
| | (0.421) | (1.038) | (1.469) | (3.319) |
| 2/16/18 | −0.154 | −1.733† | 0.373 | −1.591 |
| | (0.408) | (1.001) | (1.405) | (3.366) |
| 2/17/18 | 0.733† | −0.515 | 2.269 | 0.395 |
| | (0.407) | (0.999) | (1.395) | (3.189) |
| 2/18/18 | −0.216 | 0.127 | 2.370† | 4.919 |
| | (0.416) | (1.020) | (1.433) | (3.368) |
| 2/19/18 | −0.558 | −0.441 | 0.769 | 0.638 |
| | (0.415) | (1.019) | (1.440) | (3.335) |
| 2/2/18 | −0.687† | −0.580 | 1.389 | −0.900 |
| | (0.361) | (0.884) | (1.231) | (2.818) |
| 2/3/18 | −0.110 | 0.307 | 2.346 | 2.842 |
| | (0.415) | (1.019) | (1.439) | (3.193) |
| 2/4/18 | −0.696† | −0.828 | 0.350 | −0.223 |
| | (0.413) | (1.024) | (1.415) | (3.195) |
| 2/5/18 | −0.835* | −0.606 | 1.348 | 0.480 |
| | (0.409) | (1.003) | (1.401) | (3.230) |
| 2/6/18 | −0.656 | 0.216 | 2.366 | 2.725 |
| | (0.413) | (1.042) | (1.447) | (3.290) |
| 2/7/18 | −0.908* | −0.558 | 2.061 | 2.220 |
| | (0.416) | (1.020) | (1.417) | (3.274) |
| 2/8/18 | −1.450*** | −1.033 | 0.905 | −0.626 |
| | (0.418) | (1.026) | (1.426) | (3.234) |
| 2/9/18 | −1.029* | −1.326 | −0.258 | −1.246 |
| | (0.423) | (1.038) | (1.452) | (3.299) |
| Constant | 0.600† | 0.753 | −0.090 | 0.813 |
| | (0.332) | (0.816) | (1.140) | (2.648) |
| Observations | 952 | 944 | 922 | 795 |
| $R^2$ | 0.083 | 0.044 | 0.043 | 0.052 |
| Adjusted $R^2$ | 0.060 | 0.020 | 0.019 | 0.024 |
| Residual Std. Error | 2.002 (df = 928) | 4.909 (df = 920) | 6.821 (df = 898) | 14.236 (df = 771) |
| F Statistic | 3.644*** (df = 23; 928) | 1.834** (df = 23; 920) | 1.769* (df = 23; 898) | 1.851** (df = 23; 771) |
| *Note:* | | | | †p<0.1; *p<0.05; **p<0.01; ***p<0.001 |

*Estimates are from OLS models. The dependent variable is the change in anti-Shia tweet count in each post-treatment period. Standard Errors are in parentheses. Analysis is conducted on the full sample.*

Table A19: Effect of Treatment on Anti-Shia Tweet Count
(Leave out Treatment Dates 1)

| | *Change in Anti-Shia Tweet Count* | | | | | |
|---|---|---|---|---|---|---|
| | Each Model Leaves out 1 Treatment Date | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Arab ID | −0.623 | −0.761 | −0.529 | 0.292 | −0.476 | 0.109 |
| | (1.835) | (1.309) | (2.008) | (1.914) | (1.883) | (1.876) |
| Religious ID | 1.543 | −1.082 | 0.937 | 1.901 | 1.365 | 1.926 |
| | (1.921) | (1.375) | (2.078) | (2.000) | (1.978) | (1.966) |
| Arab ID (Elite) | −0.677 | −0.727 | −0.868 | 0.156 | −0.328 | 0.296 |
| | (1.910) | (1.362) | (2.098) | (1.992) | (1.970) | (1.951) |
| Religious ID (Elite) | −6.656*** | −6.715*** | −6.883*** | −5.711** | −6.582*** | −5.675** |
| | (1.891) | (1.345) | (2.044) | (1.972) | (1.942) | (1.937) |
| No ID | −2.180 | −2.055 | −2.043 | −1.198 | −1.357 | −0.957 |
| | (1.842) | (1.313) | (2.005) | (1.921) | (1.891) | (1.886) |
| Constant | 2.063 | 2.082* | 2.359 | 1.191 | 1.857 | 1.286 |
| | (1.429) | (1.015) | (1.549) | (1.495) | (1.459) | (1.453) |
| Observations | 758 | 752 | 712 | 752 | 752 | 754 |
| $R^2$ | 0.033 | 0.048 | 0.029 | 0.027 | 0.030 | 0.026 |
| Adjusted $R^2$ | 0.026 | 0.042 | 0.022 | 0.020 | 0.023 | 0.020 |
| Residual Std. Error | 14.002 (df = 752) | 9.992 (df = 746) | 14.853 (df = 706) | 14.491 (df = 746) | 14.441 (df = 746) | 14.384 (df = 748) |
| F Statistic | 5.083*** (df = 5; 752) | 7.515*** (df = 5; 746) | 4.236*** (df = 5; 706) | 4.095** (df = 5; 746) | 4.557*** (df = 5; 746) | 4.064** (df = 5; 748) |
| *Note:* | | | | | | †p<0.1; *p<0.05; **p<0.01; ***p<0.001 |

*Estimates are from OLS models. The dependent variable is the change in anti-Shia tweet count one month post-treatment. Standard Errors are in parentheses. Each model is conducted on the full sample leaving out subjects treated on one treatment date at a time.*

Table A20: Effect of Treatment on Anti-Shia Tweet Count
(Leave out Treatment Dates 2)

| | *Change in Anti-Shia Tweet Count* | | | | | |
|---|---|---|---|---|---|---|
| | Each Model Leaves out 1 Treatment Date | | | | | |
| | (7) | (8) | (9) | (10) | (11) | (12) |
| Arab ID | 0.459 | −0.426 | −0.642 | −0.360 | −1.281 | −0.416 |
| | (1.890) | (1.879) | (1.897) | (1.873) | (1.926) | (1.901) |
| Religious ID | 2.160 | 1.488 | 1.404 | 1.388 | 0.708 | 1.319 |
| | (1.978) | (1.971) | (1.984) | (1.962) | (1.993) | (1.993) |
| Arab ID (Elite) | 0.090 | −1.037 | −0.558 | −0.460 | −1.096 | −0.448 |
| | (1.956) | (1.979) | (1.980) | (1.954) | (2.025) | (1.982) |
| Religious ID (Elite) | −5.791** | −6.606*** | −6.620*** | −6.643*** | −7.057*** | −6.469** |
| | (1.943) | (1.940) | (1.956) | (1.934) | (1.980) | (1.968) |
| No ID | −0.874 | −1.571 | −1.534 | −1.153 | −2.401 | −1.631 |
| | (1.896) | (1.892) | (1.922) | (1.903) | (1.934) | (1.911) |
| Constant | 1.103 | 1.848 | 1.980 | 1.870 | 2.484 | 1.866 |
| | (1.465) | (1.456) | (1.472) | (1.449) | (1.516) | (1.479) |
| Observations | 757 | 756 | 754 | 757 | 754 | 758 |
| $R^2$ | 0.029 | 0.030 | 0.029 | 0.030 | 0.030 | 0.028 |
| Adjusted $R^2$ | 0.022 | 0.023 | 0.023 | 0.024 | 0.023 | 0.021 |
| Residual Std. Error | 14.432 (df = 751) | 14.491 (df = 750) | 14.569 (df = 748) | 14.489 (df = 751) | 14.460 (df = 748) | 14.570 (df = 752) |
| F Statistic | 4.424*** (df = 5; 751) | 4.587*** (df = 5; 750) | 4.520*** (df = 5; 748) | 4.678*** (df = 5; 751) | 4.566*** (df = 5; 748) | 4.303*** (df = 5; 752) |

*Note:* †p<0.1; *p<0.05; **p<0.01; ***p<0.001

*Estimates are from OLS models. The dependent variable is the change in anti-Shia tweet count one month post-treatment. Standard Errors are in parentheses. Each model is conducted on the full sample leaving out subjects treated on one treatment date at a time.*

Table A21: Effect of Treatment on Anti-Shia Tweet Count
(Leave out Treatment Dates 3)

| | *Change in Anti-Shia Tweet Count:* | | | | | | |
|---|---|---|---|---|---|---|---|
| | Each Model Leaves out 1 Treatment Date | | | | | | |
| | (13) | (14) | (15) | (16) | (17) | (18) | (19) |
| Arab ID | −0.214 | −0.446 | −0.257 | −0.510 | −0.635 | −0.621 | −0.334 |
| | (1.905) | (1.835) | (1.888) | (1.882) | (1.860) | (1.815) | (1.897) |
| Religious ID | 1.726 | 1.488 | 1.470 | 1.398 | 2.273 | 1.136 | 1.558 |
| | (1.997) | (1.911) | (1.970) | (1.977) | (1.979) | (1.893) | (1.996) |
| Arab ID (Elite) | −0.151 | −0.770 | −0.548 | −0.461 | −0.639 | −0.154 | −0.443 |
| | (1.960) | (1.915) | (1.949) | (1.985) | (1.959) | (1.893) | (1.974) |
| Religious ID (Elite) | −6.055** | −6.343*** | −6.192** | −6.086** | −6.524*** | −6.595*** | −6.531*** |
| | (1.960) | (1.884) | (1.952) | (1.960) | (1.913) | (1.876) | (1.964) |
| No ID | −1.909 | −1.077 | −1.678 | −1.741 | −1.940 | −2.565 | −1.661 |
| | (1.916) | (1.830) | (1.888) | (1.894) | (1.875) | (1.826) | (1.915) |
| Constant | 1.535 | 1.812 | 1.740 | 1.918 | 1.870 | 1.873 | 1.814 |
| | (1.464) | (1.409) | (1.452) | (1.467) | (1.438) | (1.395) | (1.479) |
| Observations | 746 | 760 | 758 | 760 | 751 | 760 | 759 |
| $R^2$ | 0.027 | 0.029 | 0.026 | 0.025 | 0.033 | 0.031 | 0.029 |
| Adjusted $R^2$ | 0.021 | 0.023 | 0.020 | 0.019 | 0.027 | 0.025 | 0.023 |
| Residual Std. Error | 14.568 (df = 740) | 14.156 (df = 754) | 14.525 (df = 752) | 14.523 (df = 754) | 14.385 (df = 745) | 14.087 (df = 754) | 14.562 (df = 753) |
| F Statistic | 4.165*** (df = 5; 740) | 4.495*** (df = 5; 754) | 4.073** (df = 5; 752) | 3.896** (df = 5; 754) | 5.138*** (df = 5; 745) | 4.873*** (df = 5; 754) | 4.577*** (df = 5; 753) |

*Note:* †p<0.1; *p<0.05; **p<0.01; ***p<0.001

*Estimates are from OLS models. The dependent variable is the change in anti-Shia tweet count one month post-treatment. Standard Errors are in parentheses. Each model is conducted on the full sample leaving out subjects treated on one treatment date at a time.*

# Survey Experiment

## Measurement of Other Covariates

- **Motivation to Control Prejudice:** To assess one's efforts at controlling negative judgments about outgroup members, we used an 11-item subscale of the original Motivation to Control Prejudice scale (Dunton and Fazio 1997). These items were rated on a five point Likert-type scale ranging from 1 (strongly agree) to 5 (strongly disagree). Because this scale is usually used to study race relations, it was modified to substitute "religious sect" for "Black/White" in the Lebanese context. A sample item is "It's important to me that other people not think I'm prejudiced" El Jarrah (2007).

- **Sectarianism**: We used a five item scale adopted from Harb (2010) to measure sectarianism. Sample items included: "My sect can serve Lebanon better than any other sect" and "I am proud of belonging to my sect." The respondents answered the items on this scale on a scale from 1 (strongly disagree) to 5 (strongly agree).

- **Religiosity**: The current study used a modified version of the Rebeiz and Harb (2010) 8-item religiosity scale (we retained 7 out of the 8 items), which taps into intrinsic religiosity, used and validated in the context of Iraq (Fischer et al. 2008) and Lebanon. Sample items include "I consider myself a religious person," "My religion influences the way I choose to act in my routine life." We added four reverse-coded items to the scale to enhance its psychometric features, including "Sometimes I wonder whether God really exists" and "I do not consider religion to be a priority in my life." Items are rated on a 5-point Likert-type scale ranging from 1 (strongly agree) to 5 (strongly disagree).

- **Sectarian and national group identification**: Group identification was assessed using items from Phinney and Ong (2007) and Fischer et al. (2008). Group identification was measured using 10 items and rated on a five point Likert-type scale. For

each item, "religious sect" replaced "ethnic group" for sectarian identification, and "Lebanon" for national identification, such as "My identity is mostly defined by my belonging to Lebanon (or my religious sect)."

# Sample Details

## I. Sample Description

Distribution of Questionnaires per Mohafaza

| Area | # of Questionnaires | % of Questionnaires |
|---|---|---|
| Beirut | 69 | 13.8 |
| Mount Lebanon | 123 | 24.6 |
| South | 54 | 10.8 |
| North | 116 | 23.2 |
| Nabatieh | 57 | 11.4 |
| Bekaa | 81 | 16.2 |
| ***Total*** | ***500*** | ***100.0*** |

Distribution of Questionnaires per District

| Area | # of Questionnaires | % of Questionnaires |
|---|---|---|
| Beirut I | 14 | 2.8 |
| Beirut II | 16 | 3.2 |
| Beirut III | 39 | 7.8 |
| Baabda | 23 | 4.6 |
| Aley | 17 | 3.4 |
| Jbeil | 12 | 2.4 |
| Metn | 27 | 5.4 |
| Chouf | 28 | 5.6 |
| Keserwan | 14 | 2.8 |
| Sour | 24 | 4.8 |
| Jezzine | 8 | 1.6 |
| Zahrani | 14 | 2.8 |
| Saida | 8 | 1.6 |
| Akkar | 36 | 7.2 |
| Batroun | 9 | 1.8 |
| Koura | 9 | 1.8 |
| Tripoli | 30 | 6 |
| Zgharta | 11 | 2.2 |
| Bcharre | 7 | 1.4 |
| Miniyeh-Dinniyeh | 14 | 2.8 |
| Nabatiyeh | 19 | 3.8 |
| Marjaayoun/ Hasbaya | 21 | 4.2 |
| Bint Jbeil | 19 | 3.8 |
| Zahle | 22 | 4.4 |
| Baalbek | 40 | 8 |
| West Beqaa-Rachaya | 19 | 3.8 |
| ***Total*** | ***500*** | ***100.0*** |

Distribution of the Respondents by Confession

| Confession | Frequency | % |
|---|---|---|
| Maronite | 108 | 21.6 |
| Greek Orthodox | 41 | 8.2 |
| Catholic | 25 | 5 |
| Armenian Orthodox | 17 | 3.4 |
| Armenian Catholic | 4 | 0.8 |
| Sunni | 134 | 26.8 |
| Shiaa | 135 | 27 |
| Druze | 27 | 5.4 |
| Alawite | 4 | 0.8 |
| Christian minorities | 5 | 1 |
| ***Total*** | ***500*** | ***100.0*** |

The sample consisted of 58 % males and 42% females.

## II. Methodology

The questionnaires were distributed proportionally to the number of residents in each governorate (Mohafaza) of Lebanon (North, South, Nabatieh, Mount Lebanon, Beirut and Beqaa).

Each governorate was stratified into districts (26 districts in total) and the capital city of each district was selected for the purpose of the survey.

Information International adopted a multi-stage probability sampling to ensure a random, representative sample for identifying households and main respondents.

The first stage consisted of selecting neighborhoods inside each selected area in a way to represent the make-up of the area, the second stage consisted of selecting households based on a systematic random sample in each selected neighborhood according to the estimated number of buildings in the neighborhood, and finally the third stage consisted of sampling a primary respondent within each household based on the most recent birthday.

The interviewer asked about the total number of adults aged 18 years and above living in the household, and chose the one with the most recent birthday (at the date of the interview) to be the main respondent. If the selected person was not at home, a follow-up up to one additional time was conducted before declaring a non-response. This method ensured that everyone has an equal chance of inclusion, with no one allowed to self-select into the sample.

If the selected respondent accepted to participate in the survey, the respondent was read the oral consent form and explained the objectives of the survey. The interviewers re-assured the respondent that the questionnaire is voluntary, anonymous and confidential.

The field work started on November 9, 2017 and ended on November 23, 2017.

Twenty Two (22) field workers were employed for the purpose of this survey.

## III. Issues in the Field

The following incidents were faced in the field, detailed as below:

1. Convincing the household head that the person to be interviewed is the person aged 18 years and above whose birthday come next at the time of interview constituted a challenge. It took the field workers some effort to explain this methodology and the need for a representative sample of both genders and various age groups.

2. 59 follow up cases were conducted with respondents (based on appointments after the first visit). Details of the follow up cases are presented in the table below.

| *Governorate* | *District* | *1 time follow up (2 visits)* |
|---|---|---|
| **Beirut** | Beirut 1 | 1 |
| | Beirut 2 | 2 |
| | Beirut 3 | 5 |
| **Beirut Total** | | **8** |
| **Mount Lebanon** | Baabda | 2 |
| | Aley | 1 |
| | Jbeil | 2 |
| | Metn | 3 |
| | Chouf | 1 |
| | Keserwan | 2 |
| **Mount Lebanon Total** | | **11** |
| **South Lebanon** | Tyre | 3 |
| | Jezzine | 1 |
| | Saida | 2 |
| | Zahrani | 2 |
| **South Lebanon Total** | | **8** |
| **North Lebanon** | Akkar | 4 |
| | Batroun | 2 |
| | Koura | 2 |
| | Tripoli | 3 |
| | Zgharta | 2 |
| | Bcharre | |
| | Minieh-Dinnieh | 3 |
| **North Lebanon Total** | | **16** |
| **Nabatieh** | Nabatieh | 3 |
| | Marjeyoun/Hasbaya | 2 |
| | Bint Jbeil | 2 |
| **Nabatieh Total** | | **7** |
| **Beqaa** | Zahle | 2 |
| | Baalbek | 4 |
| | West Beqaa/Rachaya | 3 |
| **Bekaa Total** | | **9** |
| ***Grand Total*** | | ***59*** |

3. 72 respondents refused to participate in the survey. The refusal cases are detailed in the following table:

| *Governorate* | *District* | *Refused* |
|---|---|---|
| **Beirut** | Beirut 1 | 2 |
| | Beirut 2 | 2 |
| | Beirut 3 | 6 |
| **Beirut Total** | | **10** |
| **Mount Lebanon** | Baabda | 3 |
| | Aley | 2 |
| | Jbeil | 2 |
| | Metn | 3 |
| | Chouf | 4 |
| | Keserwan | 2 |
| **Mount Lebanon Total** | | **16** |
| **South Lebanon** | Tyre | 4 |
| | Jezzine | 1 |
| | Saida | 1 |
| | Zahrani | 2 |
| **South Lebanon Total** | | **8** |
| **North Lebanon** | Akkar | 6 |
| | Batroun | 2 |
| | Koura | 1 |
| | Tripoli | 5 |
| | Zgharta | 2 |
| | Bcharre | 1 |
| | Minieh-Dinnieh | 3 |
| **North Lebanon Total** | | **20** |
| **Nabatieh** | Nabatieh | 2 |
| | Marjeyoun/Hasbaya | 3 |
| | Bint Jbeil | 2 |
| **Nabatieh Total** | | **7** |
| **Beqaa** | Zahle | 3 |
| | Baalbek | 5 |
| | West Beqaa/Rachaya | 3 |
| **Beqaa Total** | | **11** |
| ***Grand Total*** | | ***72*** |

4. 35 potential households were not eligible for the survey as they were not Lebanese (Syrian refugees or Palestinians), in addition to another 40 households who did not have residents aged between 18-65 years at their premises (only old aged couples living alone).

The non-eligibility cases are detailed in the following table:

| *Governorate* | *District* | *Not Lebanese* | *No one is available at home between 18 & 65 years old* |
|---|---|---|---|
| **Beirut** | Beirut 1 | 0 | 0 |
| | Beirut 2 | 1 | 1 |
| | Beirut 3 | 2 | 3 |
| **Beirut Total** | | **3** | **4** |
| **Mount Lebanon** | Baabda | 3 | 2 |
| | Aley | 1 | 1 |
| | Jbeil | 0 | 2 |
| | Metn | 0 | 3 |
| | Chouf | 0 | 1 |
| | Keserwan | 2 | 2 |
| **Mount Lebanon Total** | | **6** | **11** |
| **South Lebanon** | Tyre | 2 | 2 |
| | Jezzine | 0 | 1 |
| | Saida | 1 | 1 |
| | Zahrani | 2 | 1 |
| **South Lebanon Total** | | **5** | **5** |
| **North Lebanon** | Akkar | 4 | 3 |
| | Batroun | 1 | 1 |
| | Koura | 1 | 1 |
| | Tripoli | 3 | 2 |
| | Zgharta | 0 | 1 |
| | Bcharre | 0 | 0 |
| | Minieh-Dinnieh | 2 | 1 |
| **North Lebanon Total** | | **11** | **9** |
| **Nabatieh** | Nabatieh | 1 | 2 |
| | Marjeyoun/Hasbaya | 2 | 2 |
| | Bint Jbeil | 2 | 1 |
| **Nabatieh Total** | | **5** | **5** |
| **Beqaa** | Zahle | 1 | 2 |
| | Baalbek | 1 | 3 |
| | West Beqaa/Rachaya | 3 | 1 |
| **Beqaa Total** | | **5** | **6** |
| ***Grand Total*** | | ***35*** | ***40*** |

# Regression Tables

Table A22: Effect of Primes on All Tweet Ratings

| | Tweet Rating | Tweet Rating | User Rating | User Rating | Likely to Share | Likely to Share |
|---|---|---|---|---|---|---|
| Religious ID | 0.540 | 0.365 | 0.540 | 0.461 | −0.560 | −0.577 |
| | (0.521) | (0.553) | (0.521) | (0.557) | (0.489) | (0.608) |
| National ID | 0.785 | 0.808 | 0.785 | 0.695 | −0.088 | −0.372 |
| | (0.531) | (0.561) | (0.531) | (0.564) | (0.503) | (0.607) |
| Religious ID (Elite) | −2.590*** | −3.001*** | −2.590*** | −2.919*** | −1.182* | −1.163† |
| | (0.531) | (0.565) | (0.531) | (0.569) | (0.504) | (0.615) |
| National ID (Elite) | −0.037 | −0.138 | −0.037 | −0.063 | −0.612 | −0.602 |
| | (0.531) | (0.553) | (0.531) | (0.556) | (0.490) | (0.600) |
| Sectarianism Index | | 0.687* | | 0.559* | | 0.520† |
| | | (0.267) | | (0.269) | | (0.293) |
| Social Media Use | | 0.199 | | 0.173 | | 0.195 |
| | | (0.169) | | (0.170) | | (0.183) |
| System Justification | | 0.142 | | 0.071 | | 0.498 |
| | | (0.323) | | (0.325) | | (0.355) |
| MCP | | −1.774*** | | −1.659*** | | −0.803* |
| | | (0.356) | | (0.358) | | (0.389) |
| Gender | | 0.244 | | 0.029 | | −0.314 |
| | | (0.360) | | (0.362) | | (0.388) |
| Education | | 0.180 | | 0.280† | | 0.202 |
| | | (0.144) | | (0.145) | | (0.156) |
| Religiosity | | −0.716 | | −0.528 | | −0.864 |
| | | (0.670) | | (0.674) | | (0.730) |
| Internet Use | | −0.163 | | −0.085 | | 0.190 |
| | | (0.107) | | (0.107) | | (0.115) |
| Political Interest | | 0.093 | | 0.098 | | 0.124 |
| | | (0.140) | | (0.141) | | (0.153) |
| Constant | −4.099*** | 0.868 | −4.099*** | 0.502 | −1.888*** | −1.007 |
| | (0.388) | (2.819) | (0.388) | (2.836) | (0.360) | (3.095) |
| Observations | 328 | 274 | 328 | 274 | 377 | 290 |
| $R^2$ | 0.147 | 0.267 | 0.147 | 0.241 | 0.019 | 0.073 |
| Adjusted $R^2$ | 0.136 | 0.230 | 0.136 | 0.204 | 0.008 | 0.029 |

*Note:* ±p<0.01; *p<0.05; **p<0.01; ***p<0.001

Estimates from OLS models with and without covariates. Control group is reference category. Standard Errors in parentheses.

Table A23: Effect of Primes on Sectarian Tweet Ratings

| | Tweet Rating | Tweet Rating | User Rating | User Rating | Likely to Share | Likely to Share |
|---|---|---|---|---|---|---|
| Religious ID | 0.581 | 0.493 | 0.581 | 0.537 | 0.424 | 0.358 |
| | (0.363) | (0.417) | (0.363) | (0.400) | (0.318) | (0.394) |
| National ID | 0.783* | 1.115** | 0.783* | 0.962* | 0.138 | 0.293 |
| | (0.374) | (0.422) | (0.374) | (0.404) | (0.326) | (0.394) |
| Religious ID (Elite) | −1.121** | −1.433*** | −1.121** | −1.401*** | −0.390 | −0.749$^{\dagger}$ |
| | (0.373) | (0.426) | (0.373) | (0.408) | (0.326) | (0.400) |
| National ID (Elite) | 0.323 | 0.232 | 0.323 | 0.233 | 0.094 | 0.157 |
| | (0.377) | (0.419) | (0.377) | (0.401) | (0.325) | (0.393) |
| Sectarianism Index | | 0.521* | | 0.514** | | 0.279 |
| | | (0.202) | | (0.194) | | (0.192) |
| Social Media Use | | 0.207 | | 0.224$^{\dagger}$ | | −0.109 |
| | | (0.127) | | (0.122) | | (0.119) |
| Sectarian System Justification Index | | 0.062 | | 0.038 | | 0.182 |
| | | (0.246) | | (0.236) | | (0.233) |
| MCP | | −0.412 | | −0.395 | | −0.566* |
| | | (0.272) | | (0.260) | | (0.256) |
| Gender | | −0.089 | | −0.194 | | −0.013 |
| | | (0.272) | | (0.260) | | (0.253) |
| Education | | 0.199$^{\dagger}$ | | 0.203$^{\dagger}$ | | 0.012 |
| | | (0.108) | | (0.104) | | (0.101) |
| Religiosity | | −0.154 | | −0.133 | | 0.479 |
| | | (0.509) | | (0.488) | | (0.479) |
| Internet Use | | −0.118 | | −0.082 | | −0.032 |
| | | (0.080) | | (0.077) | | (0.075) |
| Political Interest | | −0.036 | | −0.042 | | −0.045 |
| | | (0.106) | | (0.101) | | (0.100) |
| Constant | 4.395*** | 3.817$^{\dagger}$ | 4.395*** | 3.758$^{\dagger}$ | 2.507*** | 2.191 |
| | (0.272) | (2.138) | (0.272) | (2.049) | (0.237) | (2.024) |
| Observations | 362 | 283 | 362 | 283 | 406 | 297 |
| $R^2$ | 0.089 | 0.173 | 0.089 | 0.173 | 0.018 | 0.069 |
| Adjusted $R^2$ | 0.079 | 0.133 | 0.079 | 0.133 | 0.008 | 0.026 |

*Note:* $^{\pm}$p<0.01; *p<0.05; **p<0.01; ***p<0.001

Estimates from OLS models with and without covariates. Control group is reference category. Standard Errors in parentheses.

Table A24: Effect of Primes on Counter-Sectarian Tweet Ratings

| | Tweet Rating | Tweet Rating | User Rating | User Rating | Likely to Share | Likely to Share |
|---|---|---|---|---|---|---|
| Religious ID | −0.047 | −0.003 | −0.032 | −0.083 | 0.860 | 0.725 |
| | (0.308) | (0.329) | (0.316) | (0.347) | (0.533) | (0.627) |
| National ID | −0.177 | 0.010 | −0.134 | −0.023 | 0.080 | 0.447 |
| | (0.305) | (0.322) | (0.312) | (0.339) | (0.530) | (0.609) |
| Religious ID (Elite) | 1.263*** | 1.479*** | 1.340*** | 1.430*** | 0.721 | 0.406 |
| | (0.310) | (0.331) | (0.317) | (0.349) | (0.537) | (0.625) |
| National ID (Elite) | 0.248 | 0.364 | 0.269 | 0.271 | 0.494 | 0.673 |
| | (0.300) | (0.319) | (0.307) | (0.336) | (0.522) | (0.609) |
| Sectarianism Index | | −0.181 | | −0.076 | | −0.520$^{\dagger}$ |
| | | (0.156) | | (0.165) | | (0.296) |
| Social Media Use | | 0.008 | | 0.043 | | −0.361$^{\dagger}$ |
| | | (0.099) | | (0.104) | | (0.186) |
| Sectarian System Justification Index | | −0.073 | | −0.035 | | −0.184 |
| | | (0.191) | | (0.201) | | (0.363) |
| MCP | | 1.270*** | | 1.189*** | | 0.132 |
| | | (0.209) | | (0.220) | | (0.397) |
| Gender | | −0.281 | | −0.198 | | 0.071 |
| | | (0.209) | | (0.220) | | (0.394) |
| Education | | 0.026 | | −0.061 | | −0.245 |
| | | (0.082) | | (0.086) | | (0.155) |
| Religiosity | | 0.239 | | 0.079 | | 1.396$^{\dagger}$ |
| | | (0.395) | | (0.416) | | (0.746) |
| Internet Use | | 0.041 | | −0.013 | | −0.165 |
| | | (0.060) | | (0.063) | | (0.115) |
| Political Interest | | −0.179* | | −0.164$^{\dagger}$ | | −0.133 |
| | | (0.081) | | (0.085) | | (0.154) |
| Constant | 8.666*** | 4.557** | 8.419*** | 4.882** | 4.577*** | 4.697 |
| | (0.224) | (1.673) | (0.229) | (1.762) | (0.387) | (3.161) |
| Observations | 414 | 330 | 414 | 330 | 433 | 333 |
| $R^2$ | 0.067 | 0.186 | 0.070 | 0.154 | 0.010 | 0.063 |
| Adjusted $R^2$ | 0.058 | 0.153 | 0.060 | 0.119 | 0.0004 | 0.025 |

*Note:* $^{\pm}$p<0.01; *p<0.05; **p<0.01; ***p<0.001

Estimates from OLS models with and without covariates. Control group is reference category. Standard Errors in parentheses.

# References


Abdo, Geneive. 2015. "Salafists and Sectarianism: Twitter and Communal Conflict in the Middle East." *Brookings Institution* .

Dunton, Bridget C and Russell H Fazio. 1997. "An individual difference measure of motivation to control prejudiced reactions." *Personality and Social Psychology Bulletin* 23(3):316–326.

El Jarrah, Rand Tahseen. 2007. The Effect of Confession and Gender on Evaluation Bias and Likeability Among AUB Students: The Role of Motivation to Control Prejudice, Confessionalism, and Religiosity PhD thesis American University of Beirut, Department of Social and Behavioral Sciences.

Fischer, Ronald, Charles Harb, Sarah Al-Sarraf and Omar Nashabe. 2008. "Support for resistance among Iraqi students: An exploratory study." *Basic and Applied Social Psychology* 30(2):167–175.

Harb, Charles. 2010. "Describing the Lebanese Youth: A National and Psycho-social Survey." *Issam Fares Institute for Public Policy and International Affairs, Working Paper Series (1). Beirut: American University of Beirut* .

Phinney, Jean S and Anthony D Ong. 2007. "Conceptualization and measurement of ethnic identity: Current status and future directions." *Journal of counseling Psychology* 54(3):271.

Rebeiz, Maria J and Charles Harb. 2010. "Perceptions of rape and attitudes toward women in a sample of Lebanese students." *Journal of Interpersonal Violence* 25(4):735–752.

Zelin, Aaron Y. and Phillip Smyth. 2014. "The Vocabulary of Sectarianism." *Foreign Policy* . Available at: http://foreignpolicy.com/2014/01/29/the-vocabulary-of-sectarianism/.