INSTITUTE for REPLICATION

No. 86
I4R DISCUSSION PAPER SERIES

A Comment on Manekin & Mitts 2022: Effective for Whom? Ethnis Identity and Nonviolent Resistance

Jan Fabian Dollbaum

Endre Borbáth

Jan Matti Dollbaum

November 2023

INSTITUTE for REPLICATION

I4R DISCUSSION PAPER SERIES

I4R DP No. 86

A Comment on Manekin & Mitts 2022: Effective for Whom? Ethnic Identity and Nonviolent Resistance

Jan Fabian Dollbaum[1], Endre Borbáth[2], Jan Matti Dollbaum[3]

[1]University College Dublin/Ireland and European University Institute Florence/Italy
[2]Freie Universität Berlin/Germany and WZB Berlin/Germany
[3]Ludwig-Maximilians-Universität München/Germany

NOVEMBER 2023

E-Mail: joerg.peters@rwi-essen.de
RWI – Leibniz Institute for Economic Research
Hohenzollernstraße 1-3
45128 Essen/Germany
www.i4replication.org

ISSN: 2752-1931

A comment on Manekin & Mitts 2022: Effective for Whom? Ethnic Identity and Nonviolent Resistance*

Dollbaum, Jan Fabian Borbáth, Endre
Dollbaum, Jan Matti

October 8, 2023

Abstract

Manekin and Mitts (2022) investigate the success chances of minority ethnic groups when engaging in non-violent protests demanding political change. First, using observational data, the authors find that the success rate for non-violent campaign tactics is lower for excluded/minority ethnic groups than for non-excluded/majority ethnic groups. Second, the authors use two original survey experiments to show that non-violent protest by ethnic minorities is perceived as more violent and requiring more policing than identical protest by majorities. This report reproduces the paper computationally and conducts several sensitivity analyses for both the observational and the experimental parts of the paper. We can confirm the general direction of the postulated effects, but evidence becomes less consistent (effect magnitudes and significance levels are not robust to some of the changes).

KEYWORDS: non-violence, group status, protest, social movement success, public opinion, discrimination, racism, minority groups

*Code available at: https://zenodo.org/records/10193470 (DOI: 10.5281/zenodo.10193470) -
Authors: Dollbaum, J.F.: University College Dublin and European University Institute (Florence). Corresponding author. E-mail: jan.dollbaum@ucd.ie. Borbáth, E.: Freie Universität Berlin and WZB Berlin Social Science Center. E-mail: endre.borbath@wzb.eu. Dollbaum, J.M.: Ludwig-Maximilians-Universität München. E-mail: jan.dollbaum@lmu.de.

1 General introduction

Manekin and Mitts (2022), hereafter MM, investigate the success chances of minority ethnic groups when engaging in non-violent protests demanding political change. They first analyze observational data on violent and non-violent campaigns with "maximalist" goals (e.g., regime change), using the datasets Ethnic Barriers to Civil Resistance and NAVCO 2.0. They find that the success rate for non-violent campaign tactics is lower for excluded/minority ethnic groups than for non-excluded/majority ethnic groups, reporting a point estimate for the interaction of non-violence with an ethnic group's status (ordinal measure) of 0.05 units (SE: 0.01).

In a second step, they intend to use two survey experiments to shed light on the causal mechanism underlying this finding. Both experiments (also referred to as study 1 and 2 hereafter) are vignette-based, both were conducted in Israel *and* the US, and both feature relatively large sample sizes (all have N>3000, except the US survey for experiment 1, which had N>2200). The experiments test the effect of protester ethnicity/race on both perceived protest violence and perceived need for police intervention for the populations of Israel and the US. The main result of study 1 is that "nonviolent resistance by ethnic minorities is perceived as more violent and requiring more policing than identical resistance by majorities" (from the article's abstract). In terms of survey respondents' perception of violence (on an 11-point scale), the main results show an effect of 0.24 units (SE: 0.12) for vignettes featuring Black protesters in the US, an effect of 0.43 units (SE: 0.11) for vignettes featuring protesters with Ethiopian migration background in Israel, and an effect of 0.62 units (SE: 0.11) for vignettes featuring Arab protesters in Israel. The effects for perceived requirement of police action (also on an 11-point scale) are: 0.27 (SE: 0.13), 0.18 (SE: 0.12), and 0.96 (SE: 0.12) respectively.

Study 2 tries to disentangle the effect of protester identity (race/ethnicity, again pitting minority vs majority) and protest goal (economic protests vs anti-racism

protests). The authors find that for anti-racism protests "white participants are perceived as less violent than Black participants" (from the abstract). The coefficients for this claim as reported in the main table (table 2 of the original paper) is 0.35 (SE: 0.12). The coefficient for required police action is 0.43 (SE: 0.11), both again measured on 11-point scales.

In the present replication report, we investigate whether MM's analytical results are reproducible and replicable, subjecting them to a variety of robustness checks.

For the observational part of MM's paper, we can confirm computational reproducibility with one exception (larger standard errors in Figure 1, which result in overlapping confidence intervals). Turning to sensitivity analysis, we re-code key IVs (group size, group status), introduce an alternative IV, and exclude some cases from the sample. Some of these changes reduce the theoretically most important coefficient, namely the effect size of the interaction between a group's ethnic status and their chosen tactic (violent vs non-violent); but statistical significance is usually not compromised (exception: re-coding the cutoff for small vs. large ethnic groups). However, we also aggregate the dataset to the level of ethnic groups and, alternatively, to the level of campaigns because after looking at the replication code, we became aware that MM do not account for the multilevel nature of their data (using OLS on an apparent panel dataset). Results are not robust to this change: Point estimates for the interaction of non-violent tactics with an ethnic group's status (ordinal measure) (previously: 0.05) decrease to 0.02 (SE: 0.04) and -0.01 (SE: 0.04), respectively, and statistical significance is lost.

For the experimental part, we introduce weights for both studies. With respect to the paper's main claims (effect of ethnic identity on the perception of violence and required policing) the number of coefficients significant at the 0.05 level drops from 5 to 3 (out of a total of 6) for the US (study 1 & 2) and from 5 to 4 for Israel (study 1, again out of 6). For study 2 for Israel (which does not include Arabs in the sample) all 6 coefficients stay significant.

For the open questions of the surveys / the text-as-data part, we replicate the

analysis with varying number of topics. We identify the number of topics based on various measures of fit. Substantively the results are similar to the ones reported in the paper. However, we find a different set of topics and the statistically significant differences between protesters with varying identity often do not hold up to changing the number of topics.

We are grateful to the authors for providing comprehensive data and clearly written code, and, beyond that, for also responding generously to our request for additional code and data (specifically regarding the open-ended questions of the surveys).

2 Observational Part

2.1 Introduction

2.1.1 Overview of main variables, estimations, and results The response variable is an ethnic group's success in achieving its campaign's stated goals.[1] The main predictors of the study are an ethnic group's chosen tactic/method (violent vs non-violent) and an ethnic group's status.

The paper's main figure (of the first, observational part; named Figure 1) is based on two separate OLS regressions: The tactic/method variable serves to split the sample. Thus, one estimation is based on all observations coded as non-violent, the other based on those coded as violent. The same results are given in tabular form in Table A5. The main analysis run on the *full* dataset features an interaction between tactic/method and group status, and the results are presented in Table A3. This interaction is the key contribution of the observational part to the study's overall argument, and will thus be the first focus of our replication efforts. Hence, for all our replications, we produce table A3 and table A5 (with Figure 1).

MM operationalize ethnic group status in two ways: first as an ethnic group's size

[1]Note that in contrast to the protests described in the experiments' vignettes, which call for policy changes, the Navco 2.0 dataset contains political campaigns with "maximalist" goals entailing some sort of regime change. The latter could be seen as a special case of the former, and the same causal mechanisms could be at work, but this is not immediately obvious.

(expressed as the share of a country's entire population, thus between 0 and 1) and second as an ethnic group's status in the narrower sense (power through inclusion in central government). The latter is an ordinal variable (treated as continuous); in an alternative operationalization, the measure is coded as a dummy. Note that for the ordinal variable the highest score (7) encodes a dominant ethnic group, while for the dummy "1" means that an ethnic group is excluded, thus powerless. Hence, the sign of the main effect of group status (and the interaction with tactic) should be negative for the dummy and positive for the ordinal measure. We replicate results for both size and status (for Table A5 / Figure 1), but focus our robustness checks on the ethnic group's status (controlling for group size) in line with the authors' preferred specification in Table A3. We run all the replications for both the dummy and the ordinal measure of group status.

The structure of the dataset was not clear to us. It seems to be panel data – and MM state on page 2 in the SI that they "focus on the group-year as the unit of analysis". Indeed, there are multiple observations per ethnic group (nested in campaigns). However, there is no time variable (the YEAR variable does not vary between observations of the same campaign so that almost 80% of observations are complete duplicates). In connection with this, MM's estimation does not account for the multilevel structure of the dataset so that those campaigns with many observations in the dataset receive inordinate statistical weight.

We also noticed a high number of missing values for the main independent variables tactic/method (`INIT_NV_ONSET and INIT_V_ONSET`), and an inconsistency between the two; both problems do not seem to derive from the original dataset (Navco 2.0), which has a consistently coded variable for violent vs non-violent campaign tactic/method (called `prim_method`).

2.2 Replication without changes (computational reproducibility)

Point estimates of table A5 do not exactly replicate but are close (see Tables 2 and 3). However, confidence intervals are larger and overlapping for the most important

estimates, that is, the comparison of success probabilities between non-violent ethnic groups of excluded/minority versus non-excluded/majority group status/size. Hence figure 1 of the original paper, which is based on table A5, looks similar with the exception of the confidence intervals (see Figure 1). The replication code for figure 1 as provided by the authors also produces a figure with larger confidence intervals than the original figure, but they do not overlap. However, compared to the corresponding table A5, the estimation is missing two control variables (number of past non-violent and violent conflicts). When inserting these two control variables, the confidence intervals do overlap – as in table A5. The absence of these two control variables seems to be an oversight as they are present not only for the estimation of table A5 but also for those of the main table A3.

The estimates in table A3 replicate perfectly (see Table A1).

2.3 Robustness replicability

2.3.1 IV group size For the left-hand side of Figure 1 (of the original paper) the authors used the size of the ethnic group as a control (while the group's ethnic status is the DV). But for the right-hand side of the figure (where the DV is group size), the regression equations do not include the ethnic group's status. When including it, point estimates change substantially. In particular, the probability of success for non-violent non-excluded ethnic groups is more than halved, so that the confidence intervals overlap widely with those of non-violent excluded ethnic groups (see Table 4 and Figure 2). This change is less pronounced when using the *ordinal* group status variable as control.

Next, we re-code the group size variable, which measures an ethnic group's share of the total population of a country (thus ranging from 0 to 1). The distribution is heavily skewed with the mean at 0.24 and the median at 0.09. We first use the median (instead of the mean, as in the original study) as a cutoff point for small vs large ethnic groups. This changes point estimates substantially: Again the probability of success for non-violent non-excluded ethnic groups is more than

halved, so that the confidence intervals overlap widely with those of non-violent excluded ethnic groups (see Table 5 and Figure 3). Second, we set the group size cutoff at 50% of the population because MM frequently use the distinction minority vs. majority in their arguments. This produces results *similar* to the previous robustness check (not to the original results) (see Table 6 and Figure 4). This is surprising given that the cutoffs for the original results falls *in between* the cutoff for the two robustness checks.

These two robustness checks regarding the influence of group size on success do not necessarily run counter to MM's general argument (that non-violent minority groups do worse than non-violent majority groups). But they do question a linear relationship between a non-violent group's size and its success chances. In particular, size seems to matter much less (or not at all) if accounting for a group's power status in the regime.

2.3.2 Alternative predictor for campaign method/tactic The authors use a variable originally measuring whether an ethnic group initiated the onset of a non-violent campaign (`INIT_ONSET_NV`), and have somehow re-coded it. We cannot assess this further because the code for the merging of the original datasets, as is often the case, is not included in the replication files. However, we do know that there is a variable called `prim_method` provided by Navco 2.0 which specifically measures whether a campaign was primarily violent or non-violent in a given year (the quantity of interest). Including the latter variable instead (it is part of MM's dataset), yields a fifth more observations and the effect size of the interaction is about halved. Other coefficients are also substantially affected (main effect of non-violence, group size) (see Tables 7 and 8). Again, we cannot assess a) the authors' reasons behind the decision to not use this variable, and b) the programming procedure which let to the variable they used instead.

2.3.3 Sub-sampling We exclude from the sample those ethnic groups with the ethnic status "Self-Excluded" because these groups' status is exceptional: The code-

book of Navco 2.0 defines that "they control a particular territory of the state which they have declared independent"). The interactions between non-violence and ethnic group status stay highly significant yet their size is reduced by about a quarter (see Table A3).

2.3.4 Logistic regression We run *logistic* regressions (instead of OLS) because of the dichotomous dependent variable (success). The interactions between non-violence and ethnic group status stay highly significant (see Table A4).

2.3.5 Change in unit of analysis The original unit of analysis is unclear (see above); it seems like it is supposed to be the *ethniccampaign-year* (a specific year of a specific ethnic group engaged in a specific campaign). When aggregating, we take the *mean* value for most variables, but the maximum of the ethnic group status because this should better reflect a group's aggregated chances of success. First changing the unit of analysis to campaign (because this is the level where variance on IV and DV is substantial), the interaction between nonviolence and ordinal group status disappears (even turns negative) (see Table 7). When using a dummy variable for group status, the change is less dramatic, but still considerable: the interaction's effect size is halved and loses significance (see Table 8). Then changing the unit of analysis to *ethniccampaign* (thus the ethnic groups nested within campaigns), the interaction between nonviolence and ordinal group status almost disappears (far from statistical significance). When using the dummy variable for group status, the interaction's effect is almost halved and loses significance. In sum, aggregation leads to the disappearance of the interaction effect for the ordinal group status variable, and the weakening below statistical significance when using the dichotomous group status variable. For the latter case, the effect is substantially still similar to the one proposed by MM, as can be seen in Figures A1 and A2.

3 Experiment 1

3.1 Introduction

For both experiments, we first report the results for computational reproducibility checks. We then limit our changes to the analysis to introducing weights for the (hitherto unweighted) samples of experiment 2 and new, consistent weights for the samples of experiment 1.

3.2 Replication without changes (computational reproducibility)

Tables 1 and Tables A13 through A15 as well as Figure 3 and Figures A6 through A7 (i.e., those that represent the results of the regressions) reproduce exactly like in the paper and the appendix (only exception: in the latter two figures, the models in the dotwhisker plots appear to be arranged in reverse order and are thus assigned different colours compared to the original figures).

3.3 Weighting

Study 1 is weighted, but study 2 is not. Hence, in a first step, we tested whether the results for study 1 replicate for the unweighted samples. In the case of the US the effects of protester identity are about halved in size and lose all statistical significance (Table 9), while for Israel both effect sizes and significance levels hold (Table 10).

Next, we create consistent weights for both studies according to age, gender, ethnicity/race and education, using the raking procedure from the R package "survey". We use official (census) data. For ethnicity in Israel, we supplement this with data from Lewin-Epstein and Cohen (2019). We trim the weights to lie between 0.1 and 5 because some groups are so underrepresented (e.g., Ethiopian Jews or people without high-school diploma) that the maximum weight after initial raking for study 1 lay above 25. However, results for untrimmed weighted samples are mostly very similar to those for the trimmed weighted samples.

We do this for both studies (not just the unweighted study 2) because the application and the composition of the weights for study 1 is not described in the paper, and looking at the weights we found that MM exclusively account for ethnicity (not age, gender, etc.). Also, MM only account for Arabs vs Jews, not the various Jewish subgroups (Mizrachi, Ashkenazi, Ethiopian, etc.), despite the fact that this data is gathered in the survey and is prominent in the study's argument. Furthermore, Arabs are weighted down (0.69) while it seems to us that they should rather be weighted *up*: The study's sample is made up of 17% Arabs, but their share of the population is given as over 20% (Central Bureau of Statistics (2020)). We applied weights for ethnicity (weighting also the Jewish subgroups), age, gender and education.

3.3.1 Results for US survey, new weights The effect of one DV (Recall violence) holds, the effects of the other two become a bit weaker and their significance levels drop from 0.5 to 0.1 (Table 11).

3.3.2 Results for Israel survey, new weights With respect to Ethiopian Jewish protesters, the effects of all three DVs are substantially reduced. This means that, for example, a vignette featuring Ethiopian Jewish protesters causes respondents on average to perceive 0.10 points more need for police intervention on a 0-10 scale (before it was 0.18). In terms of significance, one DV stays insignificant (sic), one stays highly significant, and one switches from $p < 0.01$ to $p < 0.05$ (Table 12).

With respect to Arab protesters, the effects of two DVs hold and stay highly significant while one (recall of violence) is greatly reduced (0.47 to 0.15) and loses all significance.

4 Experiment 2

4.1 Replication without changes (computational reproducibility)

The results in Table 2 replicate perfectly, but in the replication material, the script is incorrectly named "replicate_table_4". Tables A16 through A18 replicate without problems. As above, all tables except A18 need some tweaking for perfect visual reproducibility.

In Figures A10 and A11, the colours are reversed.

4.2 Weighting

The observations of study 2, in contrast to study 1, are not weighted although whites are over-represented in the sample. This could make the detection of an effect easier and thus we generated weights according to the US census (age, gender, education, ethnicity). For the Israel survey of study 2, the sample does not include Arab respondents (in contrast to study 1, where separate, identical surveys were fielded for Jews and Arabs, and the results merged for the analysis). Hence, the reference population for study 2 is only the Jewish part of Israel's population, and therefore differs from the reference population of study 1 (the entire population of Israel). So we can only weight to better represent the Jewish subgroups, but not the Arab population. Overall, the results of study 2 after weighting become patchy for the US but hold for Jewish Israelis.

4.2.1 Results for US survey With respect to ethnic identity of the protesters, the effect of one DV holds, while the effect of the second is reduced (and significance drops to $p < 0.05$) (Table 13). The third DV (recall of violence) stays insignificant but reverses its sign: Survey respondents, when weighted to represent the US population, on average recalled *less* violence for black protesters than for whites.

With respect to the explicit goal of protesting for minority rights, the effects hold. But with respect to protesters' commitment to nonviolence, effect sizes drop

substantially, lose all statistical significance, and one DV's sign is reversed (still Table 13).

4.2.2 Results for Israel survey Effects hold (size and significance level), except for the effect of explicitly committing to nonviolence: All three DVs change substantially; while two lose all significance, one becomes significant ($p < 0.1$) in the unexpected direction. (Table 14).

5 Text-as-data / Open question

In this part we replicate the results of the structural topic model (STM) analysis. Using the same document feature matrix (dfm) as the original authors, our replication focuses on re-estimating the STMs, while varying the number of topics. Using the same design as in the original study, we add the protesters' ethnicity as a document-level covariate in the estimation.

5.1 US survey

Since the number of topics is specified in the STM analysis prior to the estimation, we start by examining the dfm to identify the 'optimal' number of topics. To do so, we calculate standard parameters of performance, namely exclusivity, held-out likelihood, lower bound, semantic coherence, and the STM residuals. Figure 5 shows the results for the US corpus.

As the figure shows, next to the originally specified K=10, there are a number of points where the 'exclusivity-semantic and the coherence trade-off' is optimized. Based on these results, we discuss the results for STM models with K values 6, 8, 12, and 14. For these specifications we replicate the estimation presented in the paper (Figure 8 in the original study). The topic loadings are included in the appendix.

First we include our results for replicating the model with K=10, the specification the authors of the study used. Figure 6 shows the results.

As the figure shows, the results are very similar to the ones presented in the

paper, even if with our estimation the confidence intervals are somewhat larger, and the identifying words also differ. Nevertheless, many of the topics seem to capture substantively similar words, as the ones discussed in the original study. This indicates that the themes respondents bring up are the same as discussed in the original paper.

In the most parsimonious specification, with 6 (Figure 7) and with 8 (Figure 8) topics, white and Black protesters are only significantly different on topics that capture whether the demonstration is violent or peaceful. At the same time, the two groups do not differ on all of the topics that deal with violence. We observe a similar pattern when we increase the number of topics to 12 (Figure 9) or 14 (Figure 10), with the difference that in these specification topics related to violence are more clearly isolated.

5.2 Israeli survey - Arab minority

We now turn to the corpus of the Arab minority in Israel. Similarly to the US survey, we start by examining the 'optimal' number of topics by calculating the same metrics as before for STMs with varying number of K. Figure 11 shows the results.

Based on these results, next to the originally specified K=10, we observe optimal points at K values 7, 12, and 17. For these specifications we replicate the estimation presented in the paper (Figure 9 in the original study). The topic loadings are included in the appendix. First we include our results for replicating the model with K=10, the specification the authors of the study used. We use Google Translate to translate the terms identifying the topics. Figure 12 shows the results.

Similarly to the US case, the results we gain are not exactly the same as in the original study, even if they are substantively rather similar. In our case, the topics most strongly distinguishing the two groups of protesters are related to violence, democratic rights, and the economic situation. Even if the terms differ from the ones found in the original analysis, the topics carry a substantively similar meaning.

By changing the number of topics, we observe the strongest differences between the two groups with K=7 (Figure 13) or 17 (Figure 15), as opposed to less significant differences with K=12 (Figure 14). However, we find that in line with the original study, the topics that differentiate the two groups substantively cluster around issues related to violence, legitimacy, rights, and democracy.

5.3 Israeli survey - Ethiopian minority

Next we turn to the third corpus on the Ethiopian minority in Israel. We proceed the same way as before, and estimate various metrics of fit to gain insights into identifying the optimal number of topics. Figure 16 shows the results.

Based on these results, next to the originally specified K=10, we observe optimal points at K values 7, 11, and 16. For these specifications replicate the estimation presented in the paper (Figure 9 in the original study). The topic loadings are included in the appendix. First we include our results for replicating the model with K=10, the specification the authors of the study used. We use Google Translate to translate the terms identifying the topics. Figure 17 shows the results.

We once again find different topics than the original authors, although the substantive findings are very similar: the two communities are the most strongly distinguished in terms of violence, and support for their cause. Unlike in the original paper where the differences are statistically significant on 8 of the 10 topics, in our analysis the two minorities are statistically significantly different in only three topics of the ten. These results are substantively robust to changing the number of K to 7 (Figure 18), 11 (Figure 19), 16 (Figure 20). In addition to the topic of violence, in these alternative specifications, topics characterized by keywords like racism, legitimacy, justice also appear to be influential in distinguishing the two groups.

6 Conclusion

Throughout this replication exercise, we were impressed by the good organization of the code which made it easy for others to understand what's going on. Furthermore, we would like to stress that there can be perfectly good reasons for the authors' choices that we failed to see, and that with this replication report we aim to start a dialogue rather than pointing to things we see as "problems". In addition, despite careful cross-checking of our code, we cannot rule out mistakes on our part.

With these thoughts in mind, what did we find overall? Taken together, our robustness checks suggest that the main claim of the paper holds ("nonviolent resistance by ethnic minorities is perceived as more violent and requiring more policing than identical resistance by majorities", from the abstract), but that the pattern of results is more patchy than in the original paper.

This applies in particular to the observational data and the replication of the experiments with a focus on the weights. On the former, except for the logistic regression, our changes lead to smaller effect sizes and reduced significance levels, sometimes not substantively, sometimes more so. The fact that the interaction effect disappears when aggregating the dataset casts some doubt on the authors' claim that "nonviolence [is] increasing success only for dominant groups" (from the abstract).

Regarding the weights, we sometimes see reduced effect sizes and levels of significance, while in other cases the effects hold (there seems to be no pattern as to which DVs are more robust than others): In the US data (looking at both studies combined), originally 5 out of 6 coefficients (see Tables 1 and 2 in the paper) are significant at the 5% level. With our new weights this number is reduced to 3 (Tables 11 and 13). For study 1 from Israel, originally 5 out of 6 were significant at 5% as well, which after our weighting is reduced to 4 (Table 12). Finally, for study 2 in Israel (the survey with no Arabs in the sample), 6 out of 6 were significant at the 5% level, and all 6 remain so with the new weights (Table 14).

With respect to the text-as-data part, finally, we largely replicate the results in substance. To us, this all shows that the pattern of results is less robust than the original paper suggests – however, our (sometimes substantial) changes do not change the main message of the paper, which constitutes its central contribution.

References

Central Bureau of Statistics: 2020, Population of Israel on the eve of 2021.
URL: *https://www.cbs.gov.il/en/mediarelease/Pages/2020/Population-of-Israel-on-the-Eve-of-2021.aspx*

Lewin-Epstein, N. and Cohen, Y.: 2019, Ethnic origin and identity in the Jewish population of Israel, *Journal of Ethnic and Migration Studies* **45**(11), 2118–2137.
URL: *https://doi.org/10.1080/1369183X.2018.1492370*

Manekin, D. and Mitts, T.: 2022, Effective for whom? Ethnic identity and nonviolent resistance, *American Political Science Review* **116**(1), 161–180.
URL: *https://www.cambridge.org/core/journals/american-political-science-review/article/effective-for-whom-ethnic-identity-and-nonviolent-resistance/D78EE1F9EE3B41D6F1500311F17B8EA6*

7 Figures

7.1 Observational part

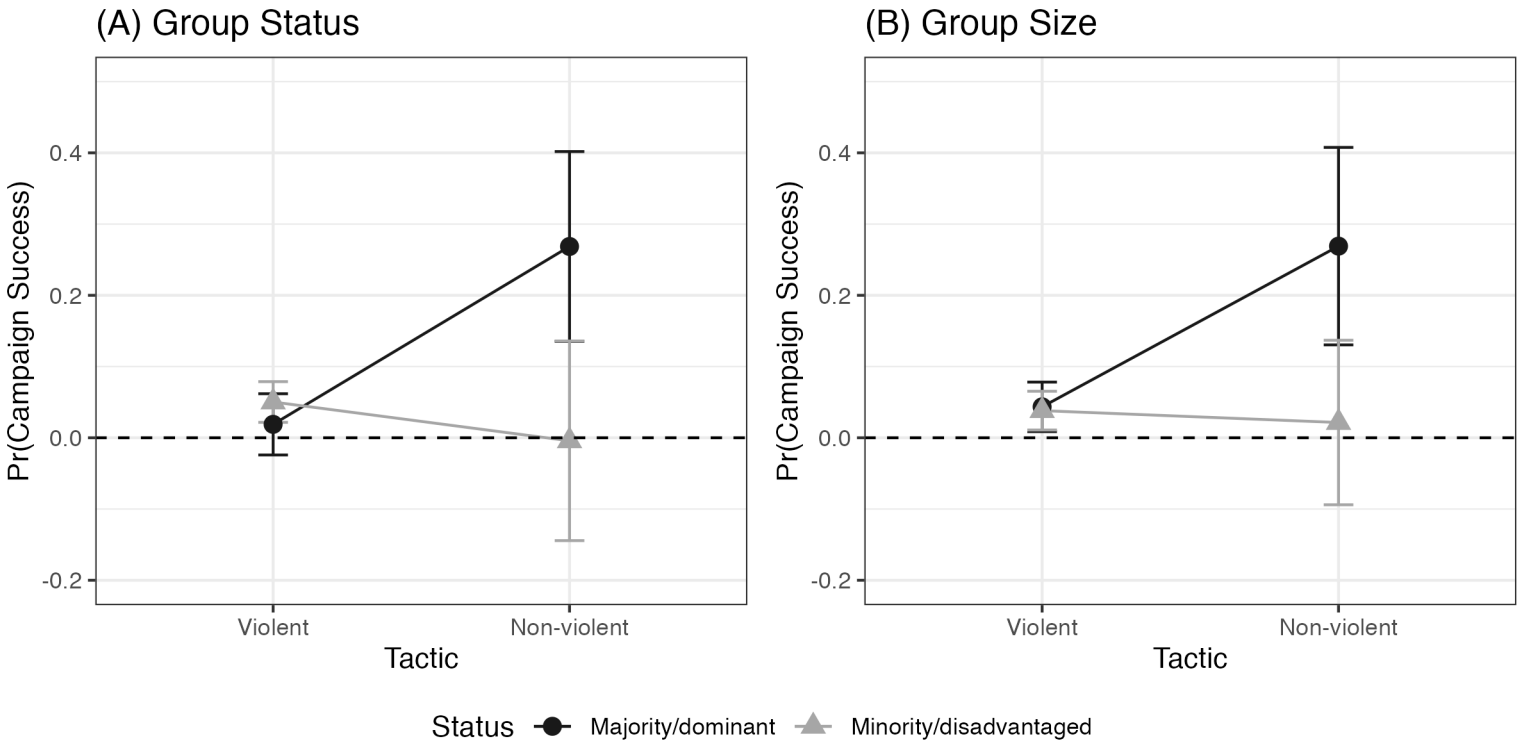

Figure 1: Replication of Figure 1. Same data and method.

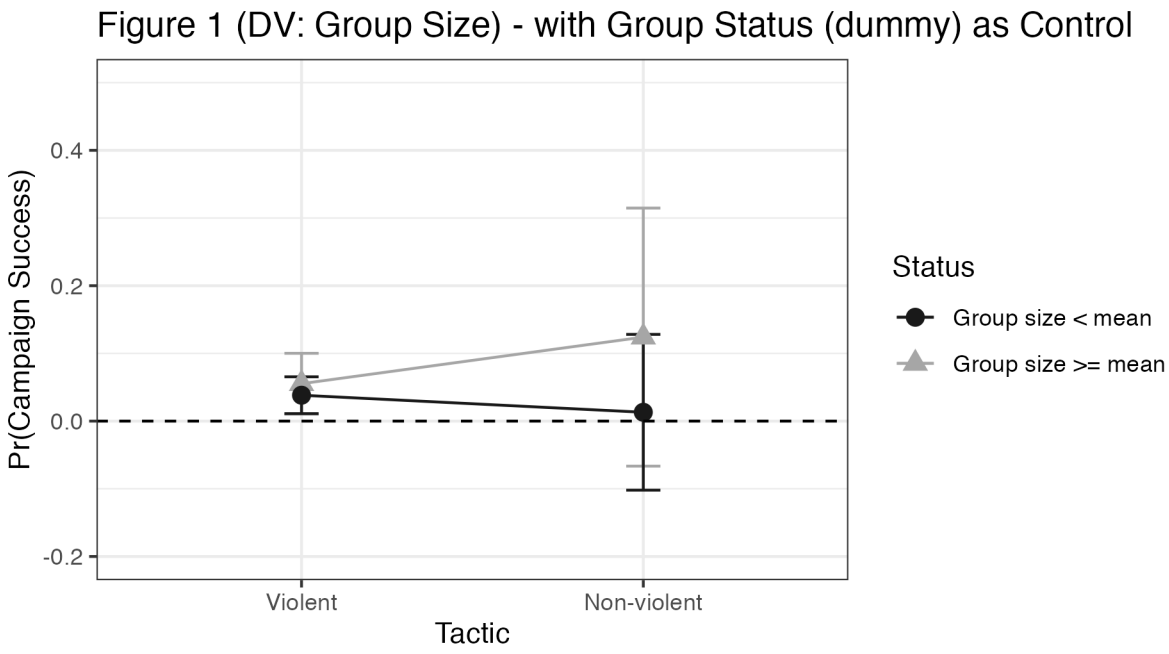

Figure 2: Replication of Figure 1 (right side: Group Size). Same data, but Group Status (dummy) as Control.

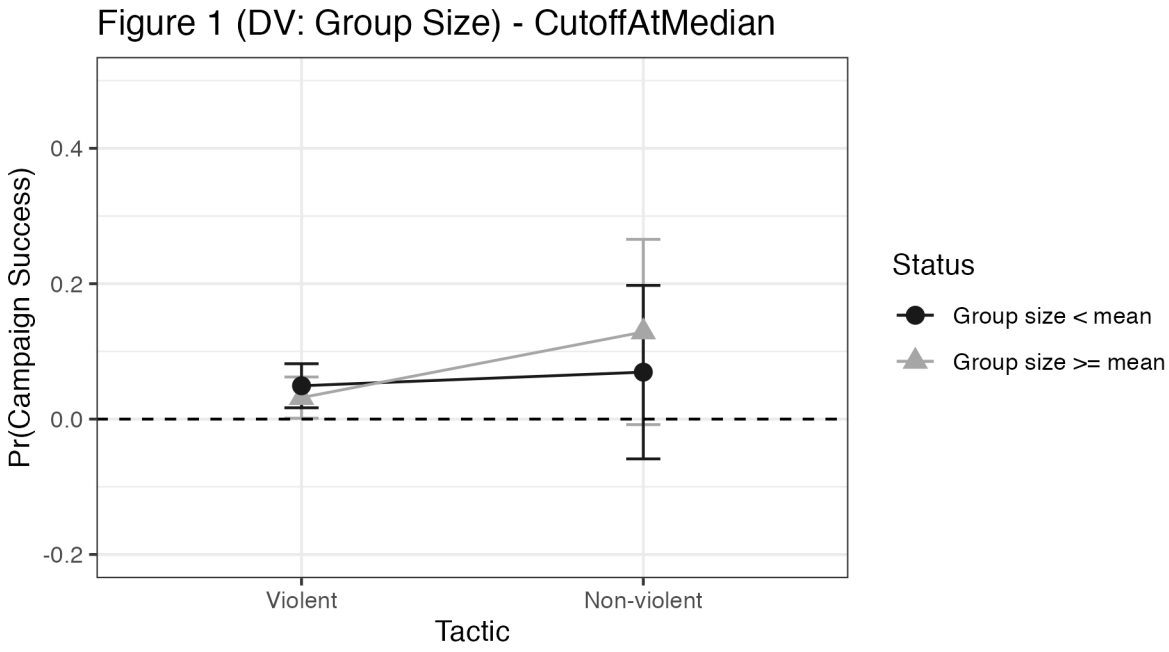

Figure 3: Replication of Figure 1 (right side: Group Size). Same data, but cutoff for small/large at median, not mean.

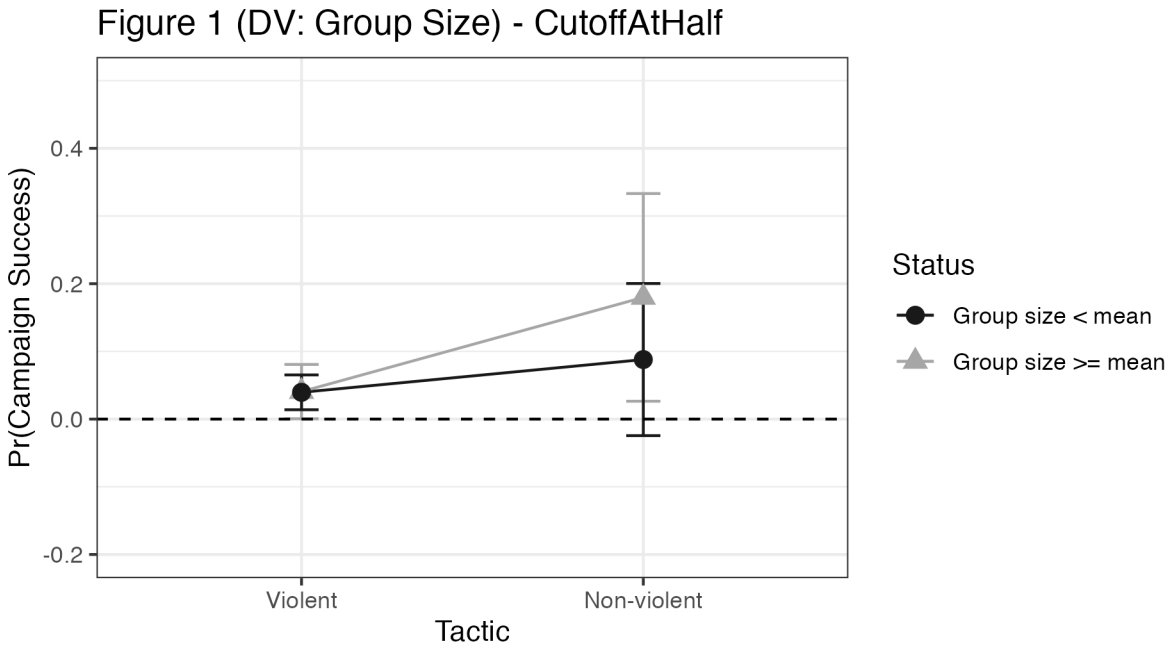

Figure 4: Replication of Figure 1 (right side: Group Size). Same data, but cutoff for small/large at 0.5, that is, 50% of country's population.

7.2 Text-as-data / Open question

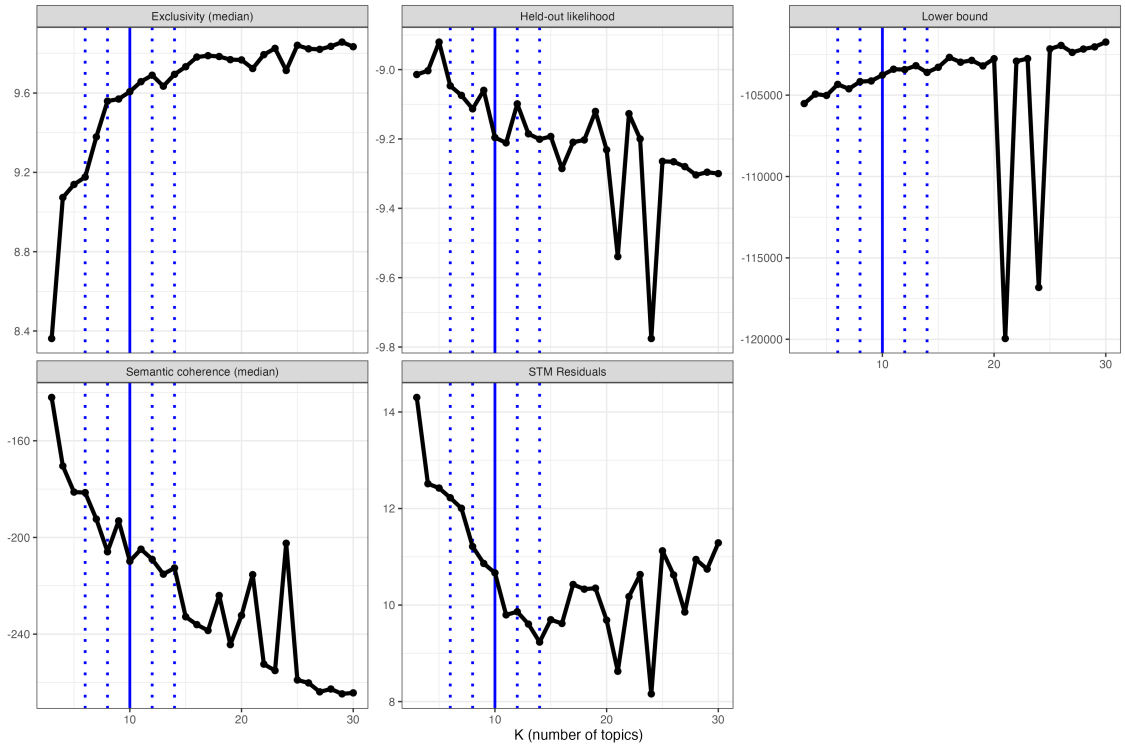

Figure 5: Structural topic model parameters with varying number of topics

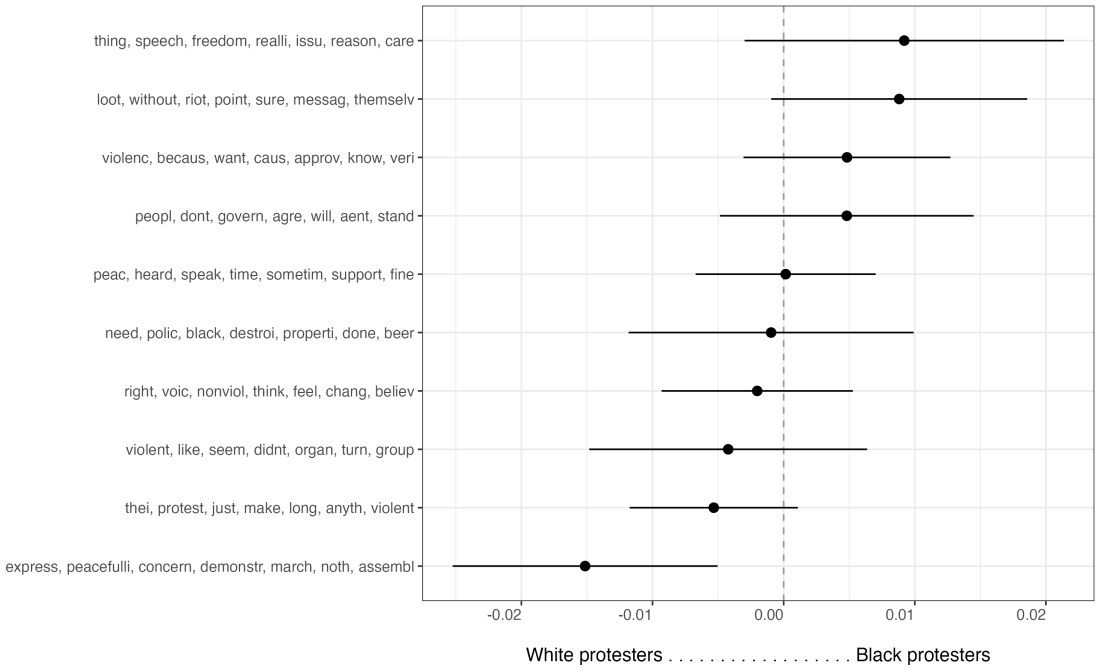

Figure 6: Prevalence estimates by ethnic status based on STM with 10 topics

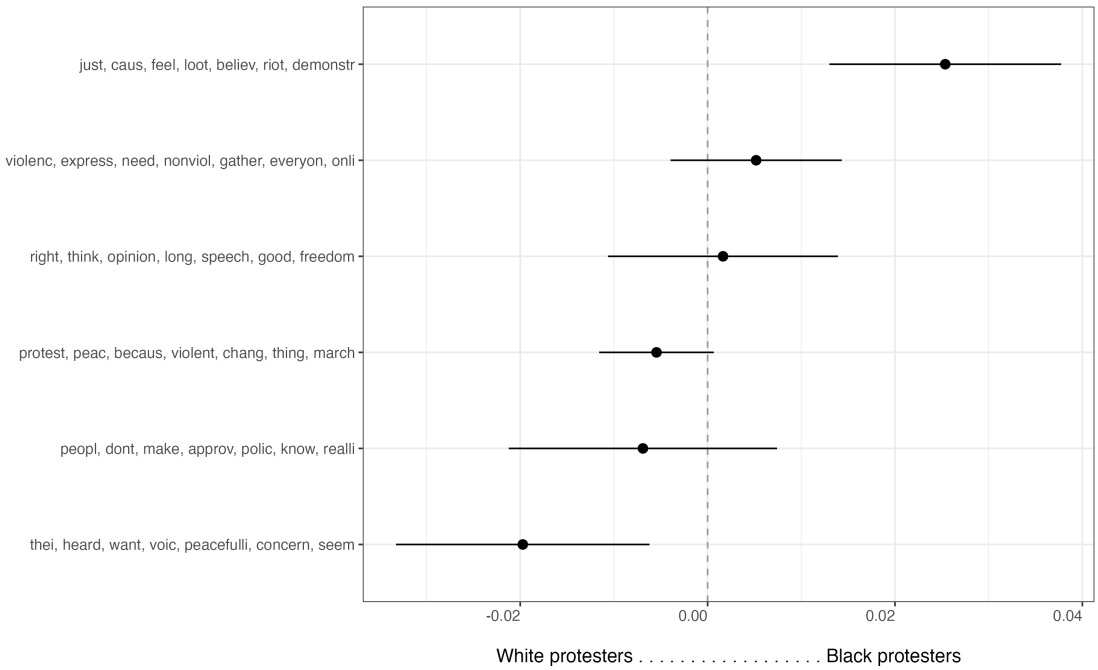

Figure 7: Prevalence estimates by ethnic status based on STM with 6 topics

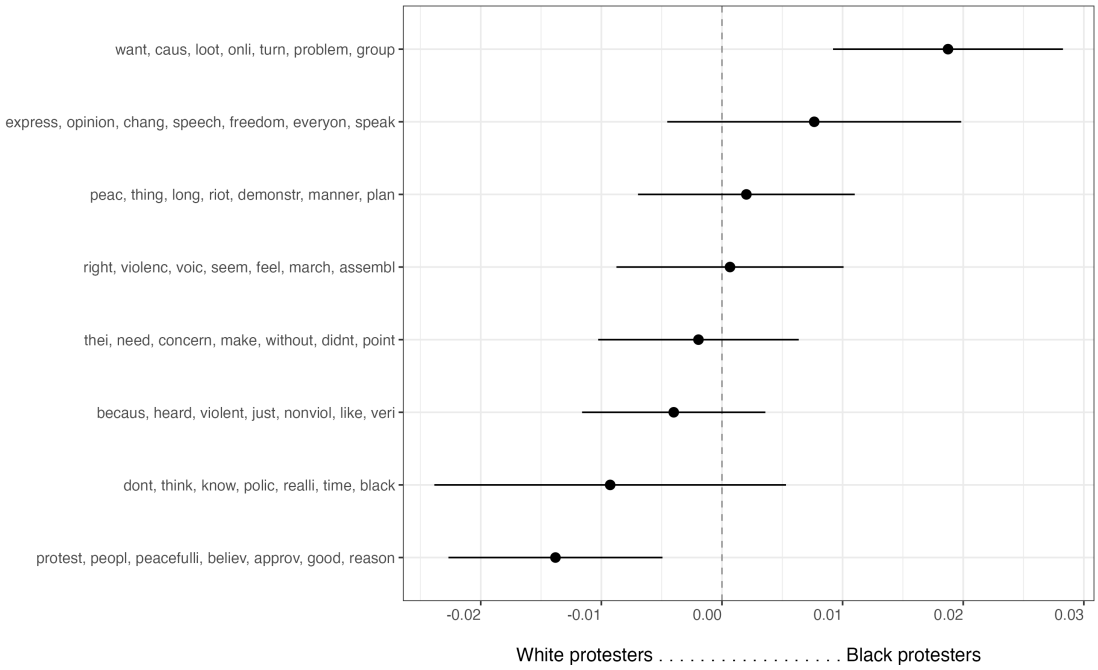

Figure 8: Prevalence estimates by ethnic status based on STM with 8 topics

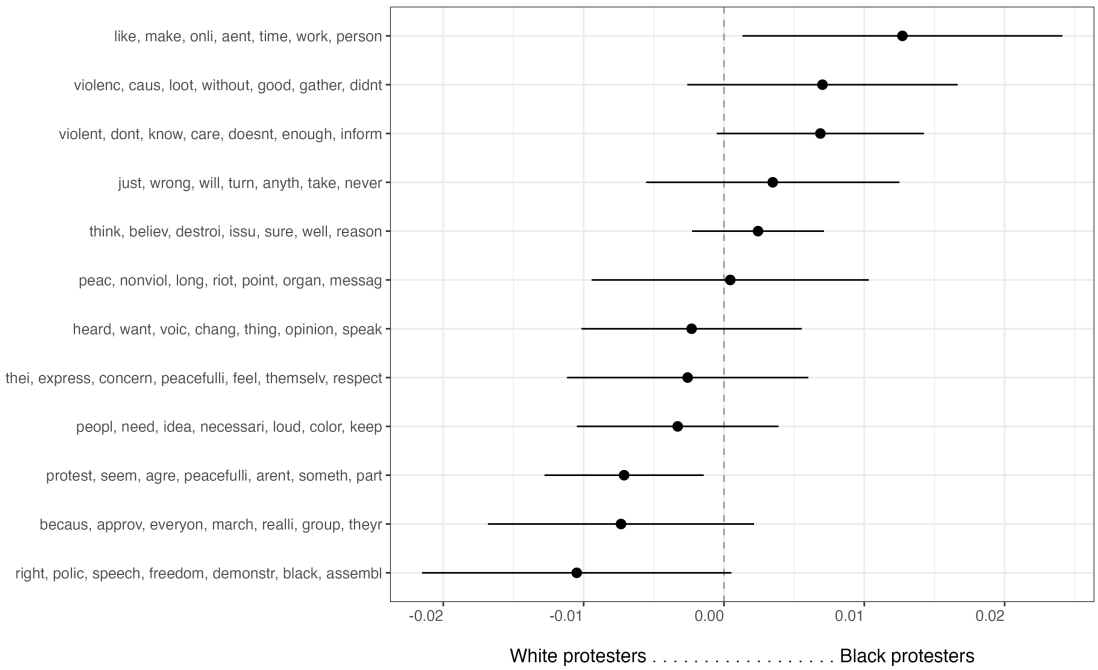

Figure 9: Prevalence estimates by ethnic status based on STM with 12 topics

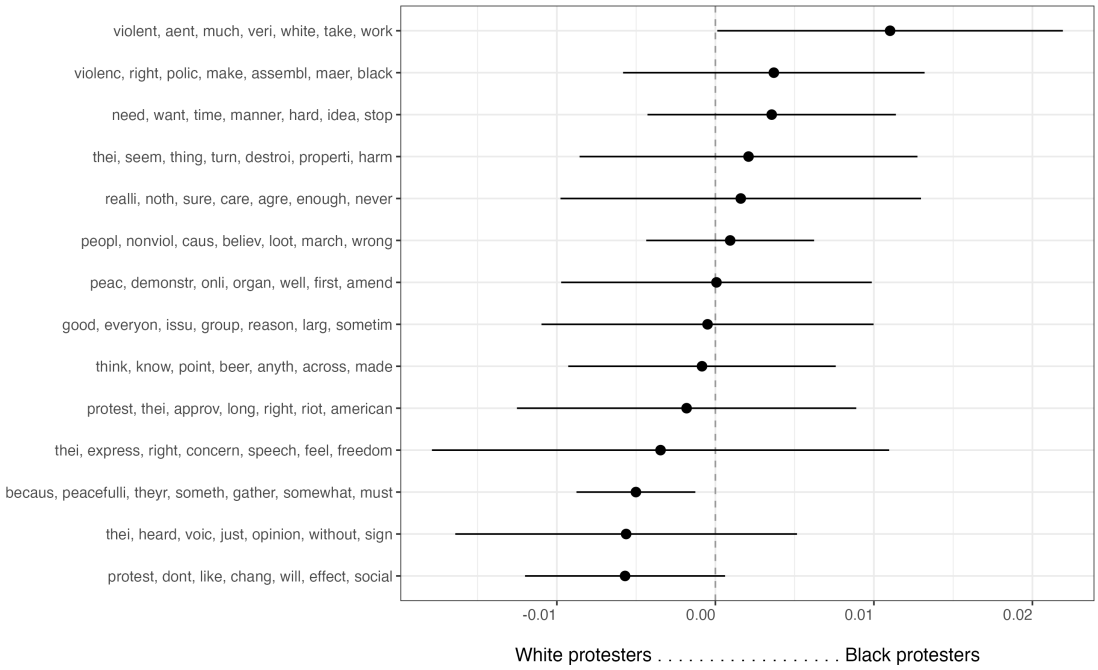

Figure 10: Prevalence estimates by ethnic status based on STM with 14 topics

7.2.1 US survey

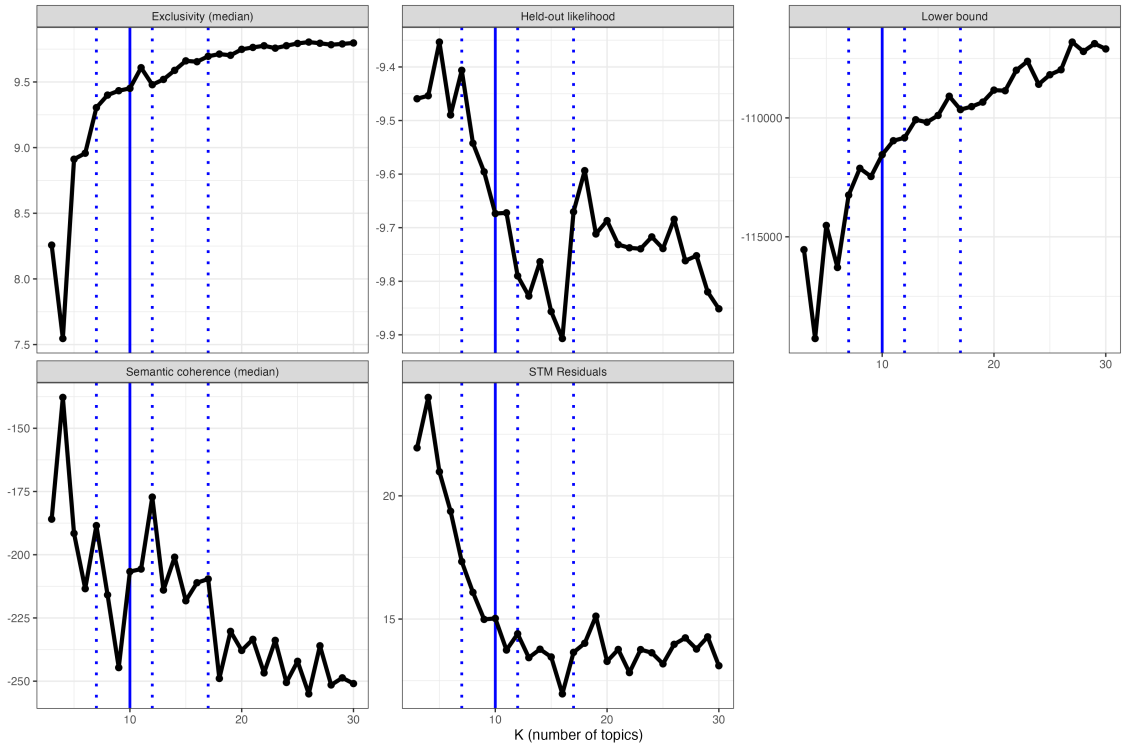

Figure 11: Structural topic model parameters with varying number of topics

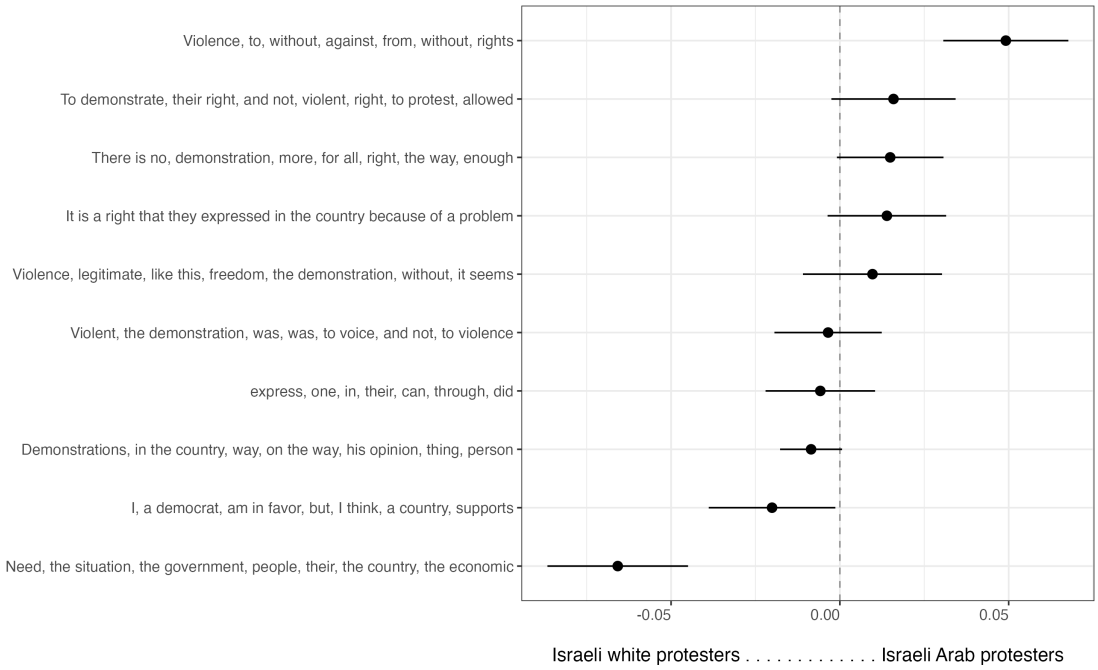

Figure 12: Prevalence estimates by ethnic status based on STM with 10 topics

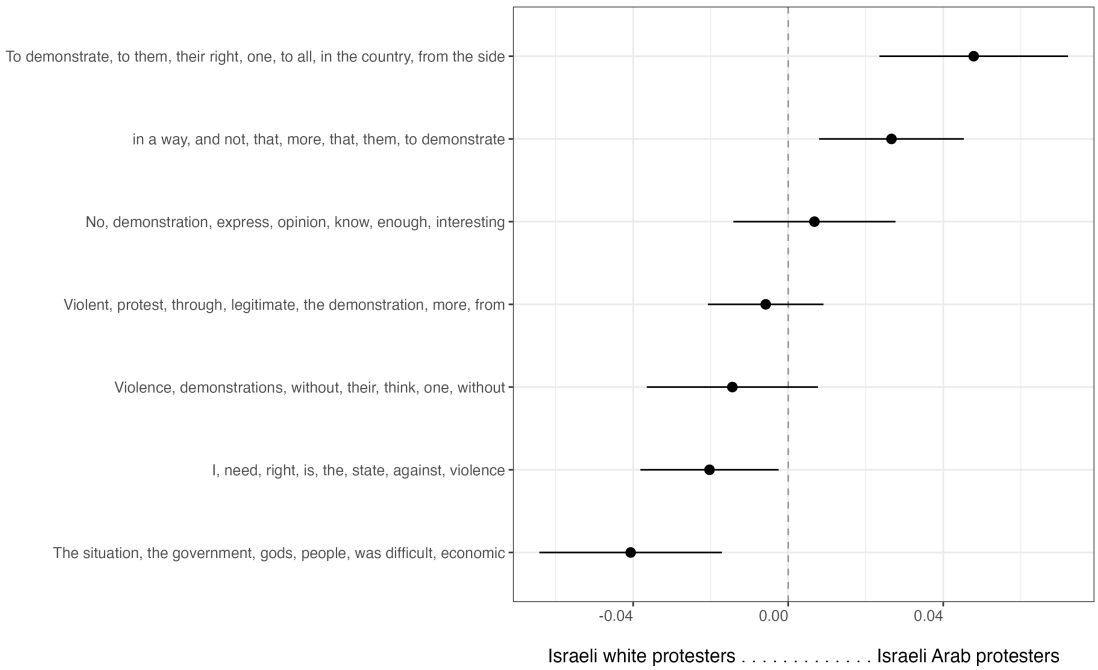

Figure 13: Prevalence estimates by ethnic status based on STM with 7 topics

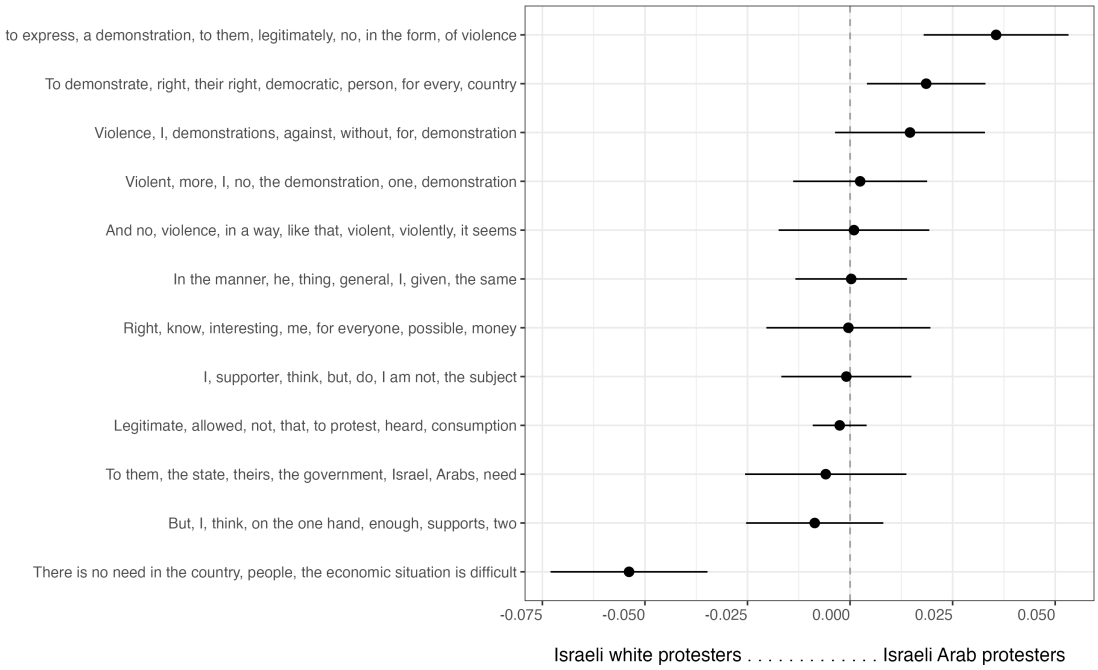

Figure 14: Prevalence estimates by ethnic status based on STM with 12 topics

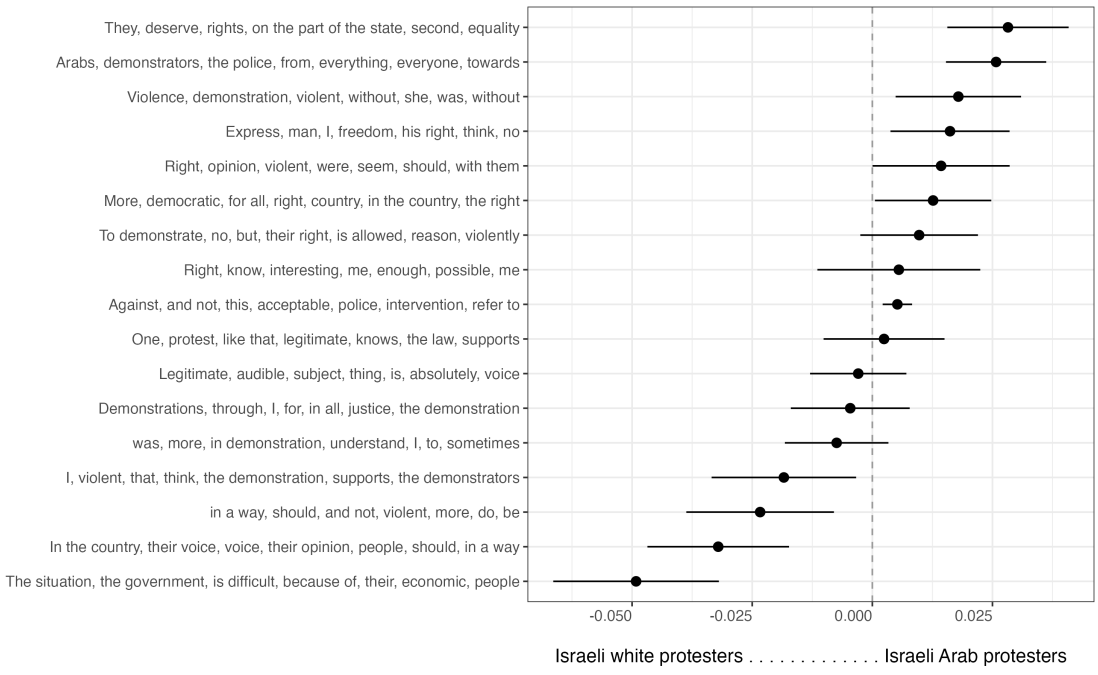

Figure 15: Prevalence estimates by ethnic status based on STM with 17 topics

7.2.2 Israeli Arab survey

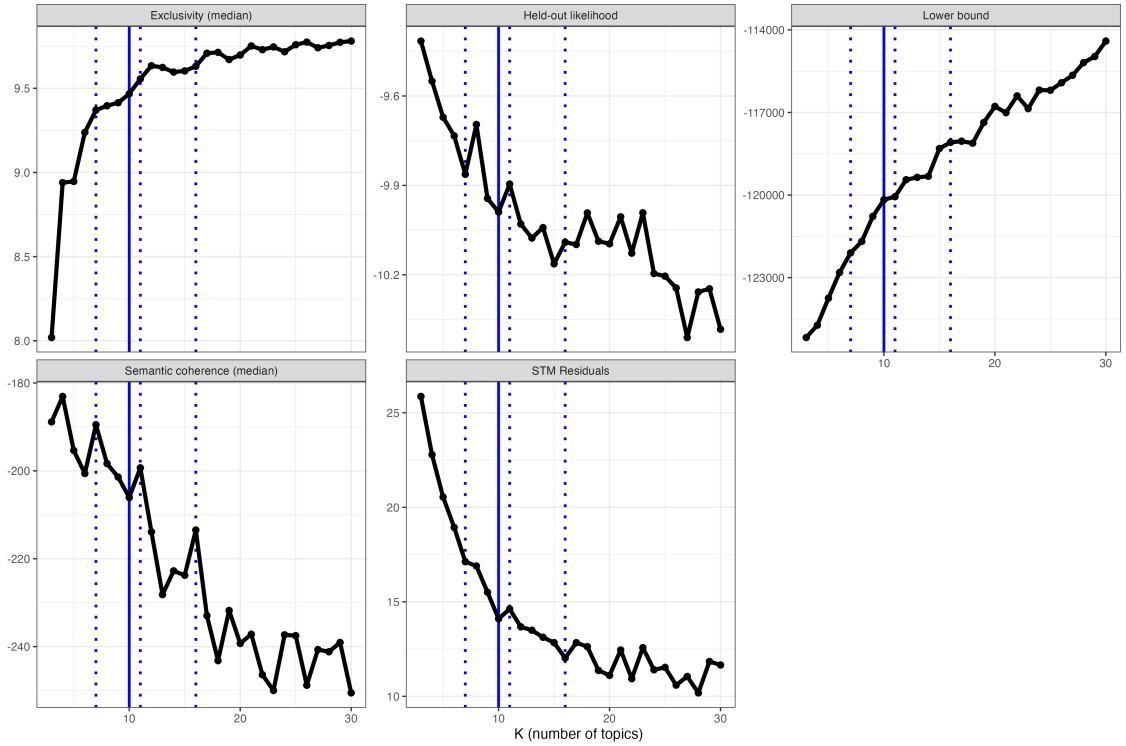

Figure 16: Structural topic model parameters with varying number of topics

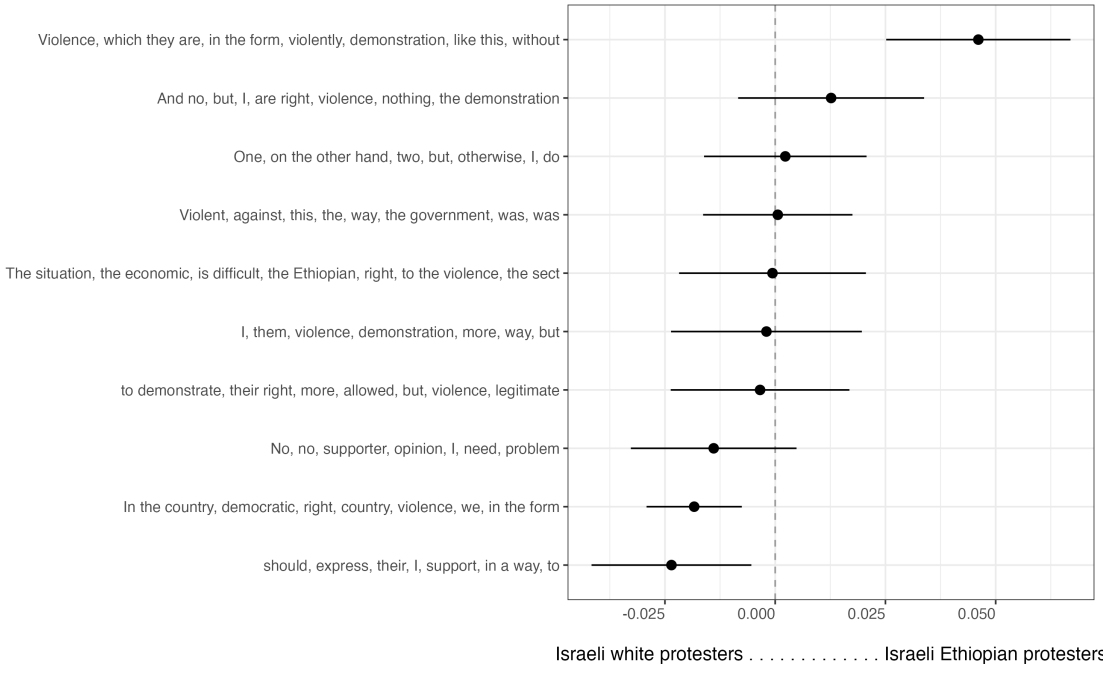

Figure 17: Prevalence estimates by ethnic status based on STM with 10 topics

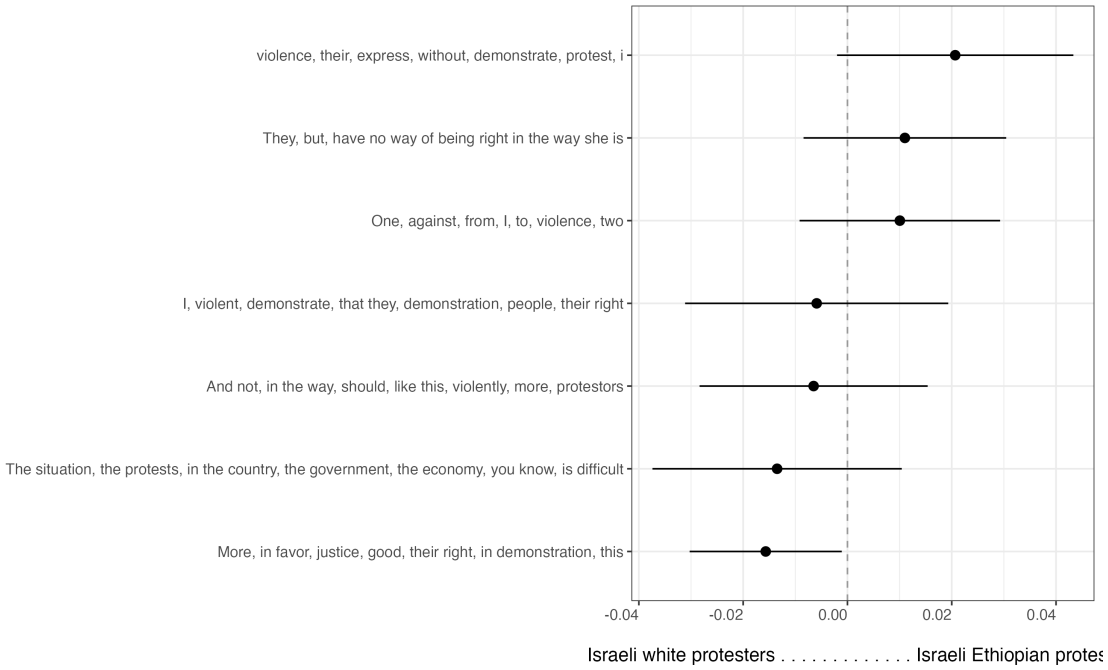

Figure 18: Prevalence estimates by ethnic status based on STM with 7 topics

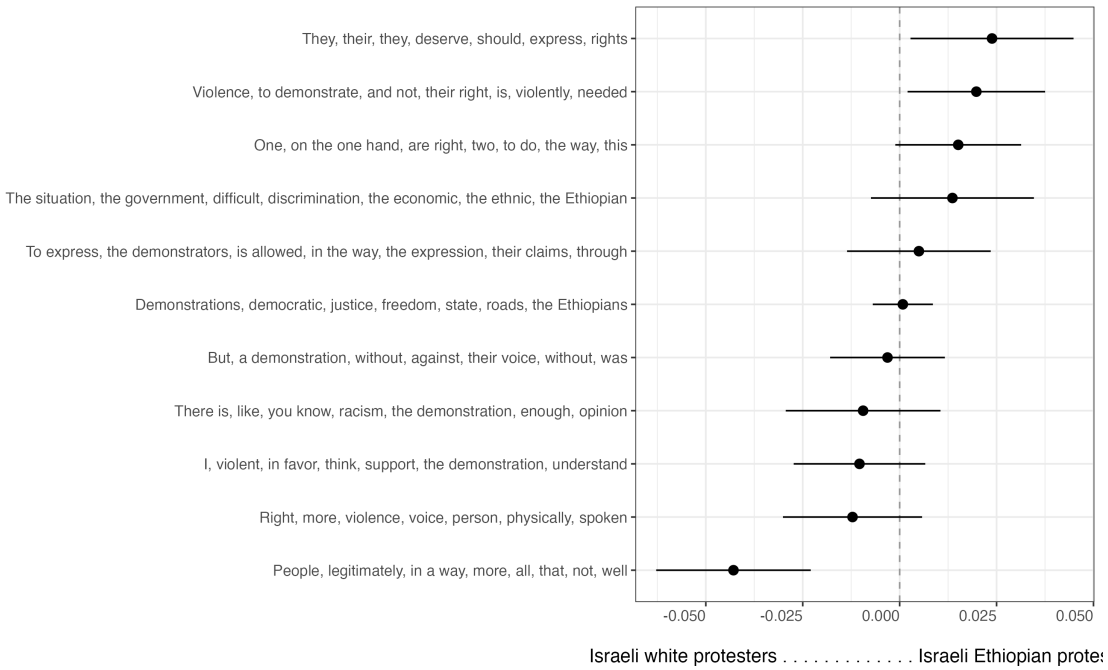

Figure 19: Prevalence estimates by ethnic status based on STM with 11 topics

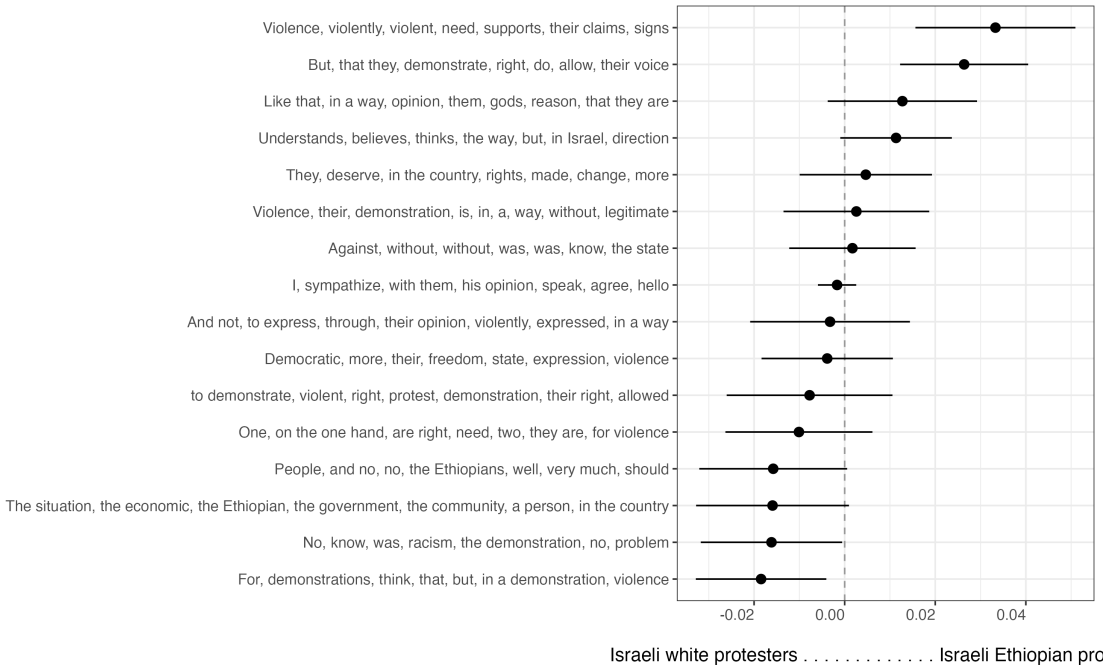

Figure 20: Prevalence estimates by ethnic status based on STM with 16 topics

7.2.3 Israeli Ethiopian survey

8 Tables

8.1 Observational part

Table 1

Statistic	N	Mean	St. Dev.	Min	Max
success	2,465	0.071	0.258	0	1
prim_method	2,493	0.183	0.387	0	1
EPR_STATUS_ORD	2,186	3.162	1.992	1	7
EPR_STATUS_EXCL	2,493	0.530	0.499	0	1
EPR_GROUPSIZE	2,493	0.256	0.312	0.000	1.000
POP_LOG_LAG_EXT	2,493	10.151	1.444	6.006	13.947
GDPPC_LOG_LAG_EXT	2,493	7.610	0.834	5.728	10.613
VDEM_POLYARCHY_LAG	2,486	0.299	0.202	0.025	0.896
VDEM_PHYSINT_LAG	2,487	0.346	0.251	0.022	0.961
EPR_TEK_EGIP	2,493	0.319	0.466	0	1
EPR_DOWNGRADED5	2,186	0.055	0.229	0	1
HORIZ_INEQ	2,010	0.118	0.319	0.000	3.238
NVYEARS	2,493	2.064	1.626	0.000	6.000
VYEARS	2,493	0.922	1.496	0.000	6.000
progress	2,465	0.144	0.351	0	1

Table 2: Replication of Table A5 - Ethnic Group Status - NoChanges

	OrigVsReplic.	Pr(success)	Std. Err	Min95	Max95	Tactic	Status
1	Original	0.25	0.05	0.16	0.35	Non-violent	Majority/dominant
2	Replication	0.27	0.07	0.14	0.40	Non-violent	Majority/dominant
3	Original	0.02	0.05	-0.08	0.11	Non-violent	Minority/disadvantaged
4	Replication	-0.004	0.07	-0.14	0.14	Non-violent	Minority/disadvantaged
5	Original	0.04	0.02	0	0.08	Violent	Majority/dominant
6	Replication	0.02	0.02	-0.02	0.06	Violent	Majority/dominant
7	Original	0.05	0.01	0.02	0.07	Violent	Minority/disadvantaged
8	Replication	0.05	0.01	0.02	0.08	Violent	Minority/disadvantaged

Table 3: Replication of Table A5 - Ethnic Group Size - NoChanges

	OrigVsReplic.	Pr(success)	Std. Err	Min95	Max95	Tactic	Status
5	Original	0.27	0.04	0.18	0.35	Non-violent	Group size >= mean
51	Replication	0.27	0.07	0.13	0.41	Non-violent	Group size >= mean
6	Original	0.03	0.04	-0.05	0.11	Non-violent	Group size <mean
61	Replication	0.02	0.06	-0.09	0.14	Non-violent	Group size <mean
7	Original	0.04	0.01	0.02	0.06	Violent	Group size >= mean
71	Replication	0.04	0.02	0.01	0.08	Violent	Group size >= mean
8	Original	0.03	0.01	0.02	0.05	Violent	Group size <mean
81	Replication	0.04	0.01	0.01	0.07	Violent	Group size <mean

Table 4: Replication of Table A5 - Ethnic Group Size - DataOriginal - wStatusAsControl

	OrigVsReplic.	Pr(success)	Std. Err	Min95	Max95	Tactic	Size
1	Original	0.27	0.04	0.18	0.35	Non-violent	Group size >= mean
2	Replication	0.12	0.10	-0.07	0.31	Non-violent	Group size >= mean
3	Original	0.03	0.04	-0.05	0.11	Non-violent	Group size <mean
4	Replication	0.01	0.06	-0.10	0.13	Non-violent	Group size <mean
5	Original	0.04	0.01	0.02	0.06	Violent	Group size >= mean
6	Replication	0.06	0.02	0.01	0.10	Violent	Group size >= mean
7	Original	0.03	0.01	0.02	0.05	Violent	Group size <mean
8	Replication	0.04	0.01	0.01	0.07	Violent	Group size <mean

Table 5: Replication of Table A5 - Ethnic Group Size - DataOriginal - CutoffAtMedian

	OrigVsReplic.	Pr(success)	Std. Err	Min95	Max95	Tactic	Size
1	Original	0.27	0.04	0.18	0.35	Non-violent	Group size >= mean
2	Replication	0.13	0.07	-0.01	0.27	Non-violent	Group size >= mean
3	Original	0.03	0.04	-0.05	0.11	Non-violent	Group size <mean
4	Replication	0.07	0.07	-0.06	0.20	Non-violent	Group size <mean
5	Original	0.04	0.01	0.02	0.06	Violent	Group size >= mean
6	Replication	0.03	0.02	0.001	0.06	Violent	Group size >= mean
7	Original	0.03	0.01	0.02	0.05	Violent	Group size <mean
8	Replication	0.05	0.02	0.02	0.08	Violent	Group size <mean

Table 6: Replication of Table A5 - Ethnic Group Size - DataOriginal - CutoffAtHalf

	OrigVsReplic.	Pr(success)	Std. Err	Min95	Max95	Tactic	Size
1	Original	0.27	0.04	0.18	0.35	Non-violent	Group size >= mean
2	Replication	0.18	0.08	0.03	0.33	Non-violent	Group size >= mean
3	Original	0.03	0.04	-0.05	0.11	Non-violent	Group size <mean
4	Replication	0.09	0.06	-0.02	0.20	Non-violent	Group size <mean
5	Original	0.04	0.01	0.02	0.06	Violent	Group size >= mean
6	Replication	0.04	0.02	0.001	0.08	Violent	Group size >= mean
7	Original	0.03	0.01	0.02	0.05	Violent	Group size <mean
8	Replication	0.04	0.01	0.01	0.07	Violent	Group size <mean

Table 7: DataOriginal SUMMARY-DV-ordinal

	Dependent variable: Success (0/1)			
	Original	UoA-campaign	UoA-Ethnicgroups	AltIV-Tactic
	(1)	(2)	(3)	(4)
EPR_STATUS_ORD	0.002	0.025	0.019	0.008
	(0.005)	(0.027)	(0.027)	(0.005)
INIT_NV_ONSET	−0.026	0.372*	0.349**	
	(0.029)	(0.212)	(0.175)	
PRIM_METHOD				0.097***
				(0.024)
EPR_GROUPSIZE	−0.027	−0.010	−0.137	−0.107***
	(0.036)	(0.011)	(0.169)	(0.038)
EPR_STATUS_ORD:INIT_NV_ONSET	0.049***	−0.010	0.019	
	(0.007)	(0.041)	(0.035)	
EPR_STATUS_ORD:PRIM_METHOD				0.029***
				(0.006)
Controls	Yes	Yes	Yes	Yes
Cubic polynomials for time	Yes	Yes	Yes	Yes
Observations	1,569	144	189	1,902
R^2	0.125	0.317	0.335	0.147
Residual Std. Error	0.220 (df = 1549)	0.445 (df = 124)	0.429 (df = 169)	0.252 (df = 1882)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 8: DataOriginal SUMMARY-DV-dummy

	Dependent variable: Success (0/1)			
	Original	UoA-campaign	UoA-Ethnicgroups	AltIV-Tactic
	(1)	(2)	(3)	(4)
EPR_STATUS_EXCL	−0.003	0.002	−0.049	−0.020
	(0.021)	(0.111)	(0.106)	(0.022)
INIT_NV_ONSET	0.272***	0.393***	0.482***	
	(0.023)	(0.095)	(0.096)	
PRIM_METHOD				0.272***
				(0.023)
EPR_GROUPSIZE	−0.016	−0.008	−0.105	−0.081**
	(0.033)	(0.007)	(0.149)	(0.034)
EPR_STATUS_EXCL:INIT_NV_ONSET	−0.233***	−0.122	−0.156	
	(0.031)	(0.174)	(0.153)	
EPR_STATUS_EXCL:PRIM_METHOD				−0.135***
				(0.027)
Controls	Yes	Yes	Yes	Yes
Cubic polynomials for time	Yes	Yes	Yes	Yes
Observations	1,569	154	189	1,902
R^2	0.128	0.344	0.337	0.149
Residual Std. Error	0.220 (df = 1549)	0.434 (df = 134)	0.428 (df = 169)	0.251 (df = 1882)

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

8.2 Survey experiments 1 & 2

Table 9: Table1-Study 1 (US) - WeightsRemoved

	Dependent variable:		
	degree_violence	recall_violence2	police_action_required
	(1)	(2)	(3)
identity_protesters1	0.137	0.029	0.171
	(0.118)	(0.018)	(0.127)
tactic1	0.123	0.005	0.258*
	(0.144)	(0.022)	(0.156)
tactic2	3.277***	0.632***	2.849***
	(0.144)	(0.022)	(0.156)
Constant	2.858***	0.065***	3.956***
	(0.118)	(0.018)	(0.128)
Weights	Removed	Removed	Removed
Observations	2,269	2,269	2,269
R^2	0.227	0.322	0.153
Adjusted R^2	0.226	0.321	0.152
Residual Std. Error (df = 2265)	2.807	0.432	3.034
F Statistic (df = 3; 2265)	221.642***	357.798***	136.796***
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 10: Table1-Study 1 (Israel) - WeightsRemoved

	Dependent variable:		
	degree_violence	recall_violence2	police_action_required
	(1)	(2)	(3)
identity_protesters1	0.426***	0.082***	0.165
	(0.108)	(0.023)	(0.119)
identity_protesters2	0.584***	0.045**	0.890***
	(0.108)	(0.023)	(0.120)
tactic1	0.242**	−0.028	0.362***
	(0.109)	(0.023)	(0.121)
tactic2	2.504***	0.404***	2.024***
	(0.107)	(0.023)	(0.119)
Constant	3.891***	0.088***	4.809***
	(0.100)	(0.021)	(0.111)
Weights	Removed	Removed	Removed
Observations	3,063	3,063	3,063
R^2	0.186	0.132	0.114
Adjusted R^2	0.185	0.131	0.113
Residual Std. Error (df = 3058)	2.436	0.516	2.707
F Statistic (df = 4; 3058)	174.236***	116.534***	98.609***
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 11: Table1-Study 1 (US) - WeightsNew

	Dependent variable:		
	degree_violence	recall_violence2	police_action_required
	(1)	(2)	(3)
identity_protesters1	0.199*	0.038**	0.212*
	(0.118)	(0.018)	(0.127)
tactic1	0.332**	0.021	0.416***
	(0.144)	(0.022)	(0.156)
tactic2	3.325***	0.640***	2.813***
	(0.144)	(0.022)	(0.155)
Constant	2.773***	0.050***	3.891***
	(0.118)	(0.018)	(0.128)
Weights	New	New	New
Observations	2,269	2,269	2,269
R^2	0.224	0.320	0.146
Adjusted R^2	0.223	0.319	0.145
Residual Std. Error (df = 2265)	2.801	0.435	3.018
F Statistic (df = 3; 2265)	217.515***	355.225***	129.059***
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 12: Table1-Study 1 (Israel) - WeightsNew

	Dependent variable:		
	degree_violence	recall_violence2	police_action_required
	(1)	(2)	(3)
identity_protesters1	0.332***	0.056**	0.095
	(0.109)	(0.023)	(0.121)
identity_protesters2	0.496***	0.015	0.887***
	(0.110)	(0.023)	(0.121)
tactic1	0.183*	−0.025	0.330***
	(0.110)	(0.023)	(0.122)
tactic2	2.401***	0.390***	1.929***
	(0.109)	(0.023)	(0.120)
Constant	3.911***	0.099***	4.816***
	(0.100)	(0.021)	(0.111)
Weights	New	New	New
Observations	3,063	3,063	3,063
R^2	0.171	0.124	0.107
Adjusted R^2	0.170	0.123	0.106
Residual Std. Error (df = 3058)	2.466	0.512	2.732
F Statistic (df = 4; 3058)	157.979***	108.059***	91.752***
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 13: Table2-Study 2 (US) - WeightsNew

	Dependent variable:		
	degree_violence	recall_violence2	police_action_required
	(1)	(2)	(3)
black	0.247**	−0.014	0.440***
	(0.116)	(0.028)	(0.116)
group_goal	0.610***	0.203***	0.534***
	(0.116)	(0.028)	(0.116)
commitment	−0.162	0.010	−0.073
	(0.116)	(0.028)	(0.115)
Constant	3.414***	0.343***	4.139***
	(0.115)	(0.028)	(0.115)
Weights	yes (new)	yes (new)	yes (new)
Observations	3,013	3,013	3,008
R^2	0.011	0.017	0.012
Adjusted R^2	0.010	0.016	0.011
Residual Std. Error	3.181 (df = 3009)	0.778 (df = 3009)	3.163 (df = 3004)
F Statistic	11.496*** (df = 3; 3009)	17.085*** (df = 3; 3009)	12.215*** (df = 3; 3004)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 14: Table2-Study 2 (Israel) - WeightsNew

	Dependent variable:		
	degree_violence	recall_violence2	police_action_required
	(1)	(2)	(3)
identity_protesters1	0.974***	0.285***	0.782***
	(0.141)	(0.029)	(0.145)
identity_protesters2	0.512***	0.060**	0.785***
	(0.142)	(0.029)	(0.146)
group_goal	0.740***	0.145***	0.589***
	(0.111)	(0.023)	(0.114)
commitment	−0.158	0.033*	−0.042
	(0.098)	(0.020)	(0.101)
Constant	3.385***	0.038	3.879***
	(0.115)	(0.023)	(0.118)
Weights	yes (new)	yes (new)	yes (new)
Observations	3,465	3,465	3,465
R^2	0.044	0.071	0.030
Adjusted R^2	0.043	0.070	0.029
Residual Std. Error (df = 3460)	2.876	0.583	2.954
F Statistic (df = 4; 3460)	39.888***	66.133***	26.723***
Note:			*p<0.1; **p<0.05; ***p<0.01

9 APPENDIX

10 Appendix Figures

10.1 Observational part

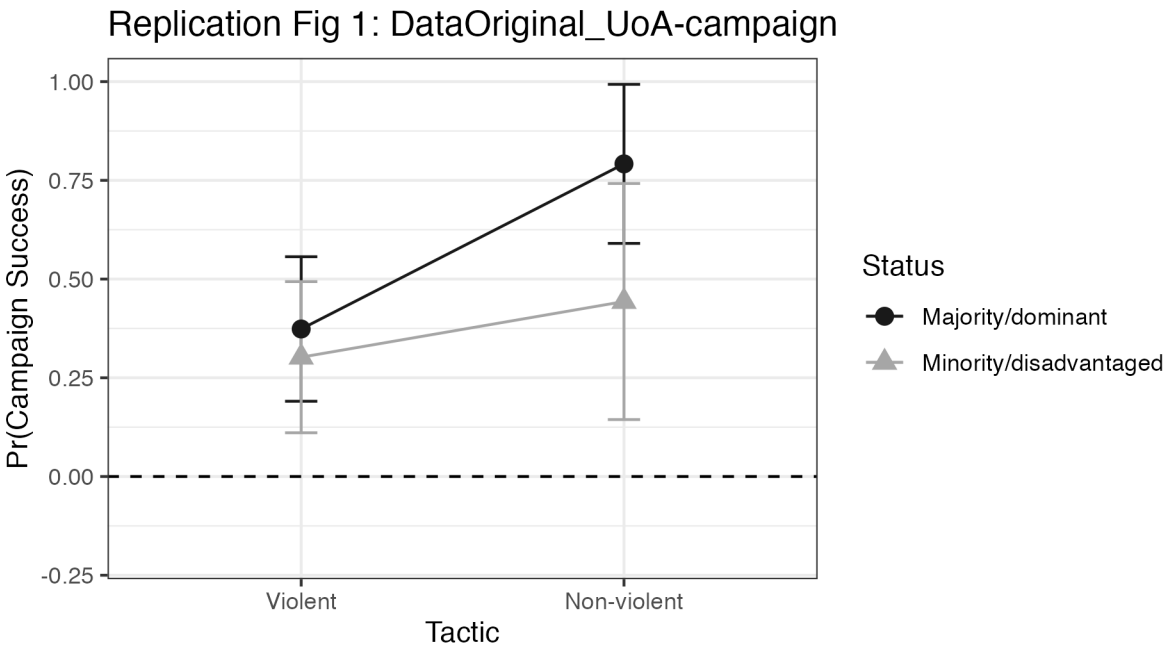

Figure A1: Figure 1 when aggregating data to campaigns.

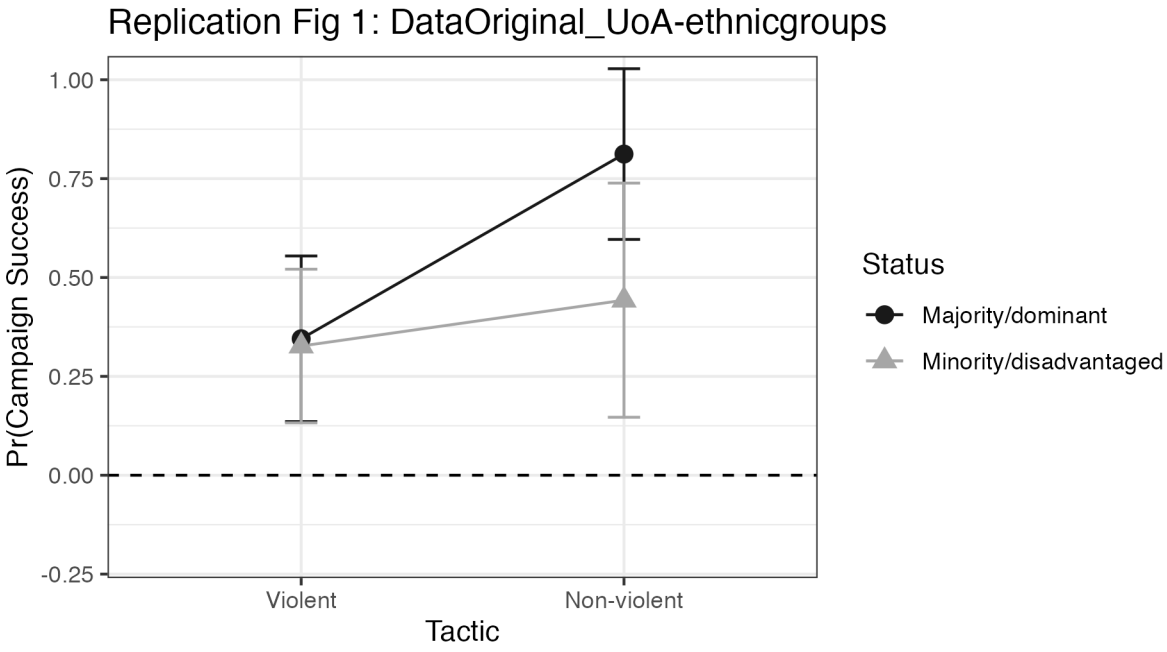

Figure A2: Figure 1 when aggregating data to ethnic groups.

10.2 Text-as-data / Open question

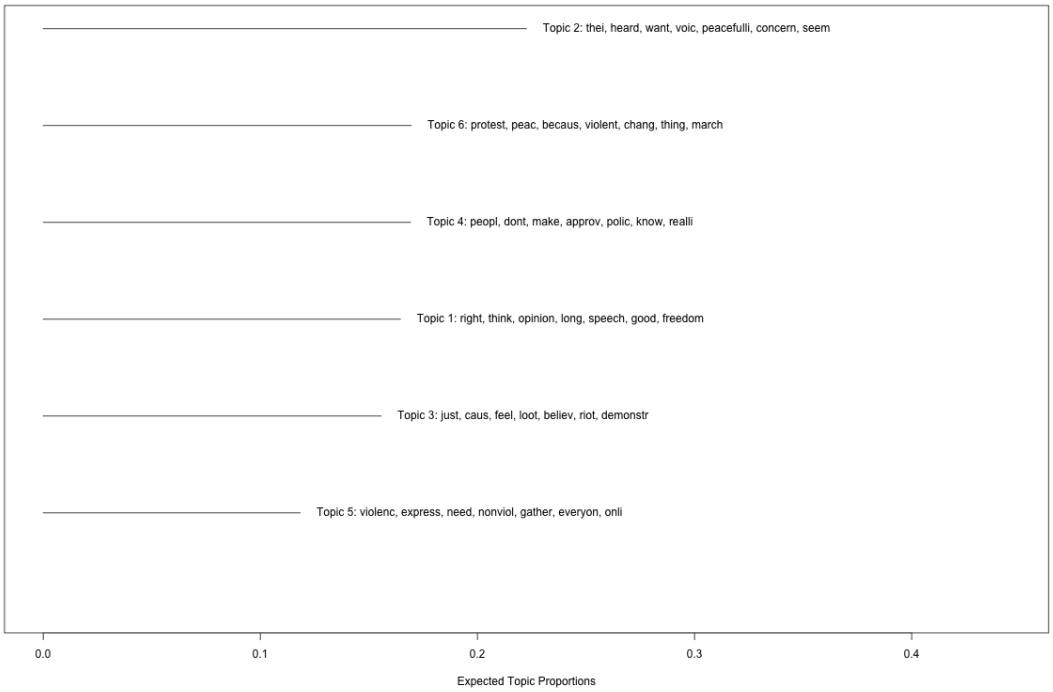

Figure A3: Structural topic model with 6 topics (US survey)

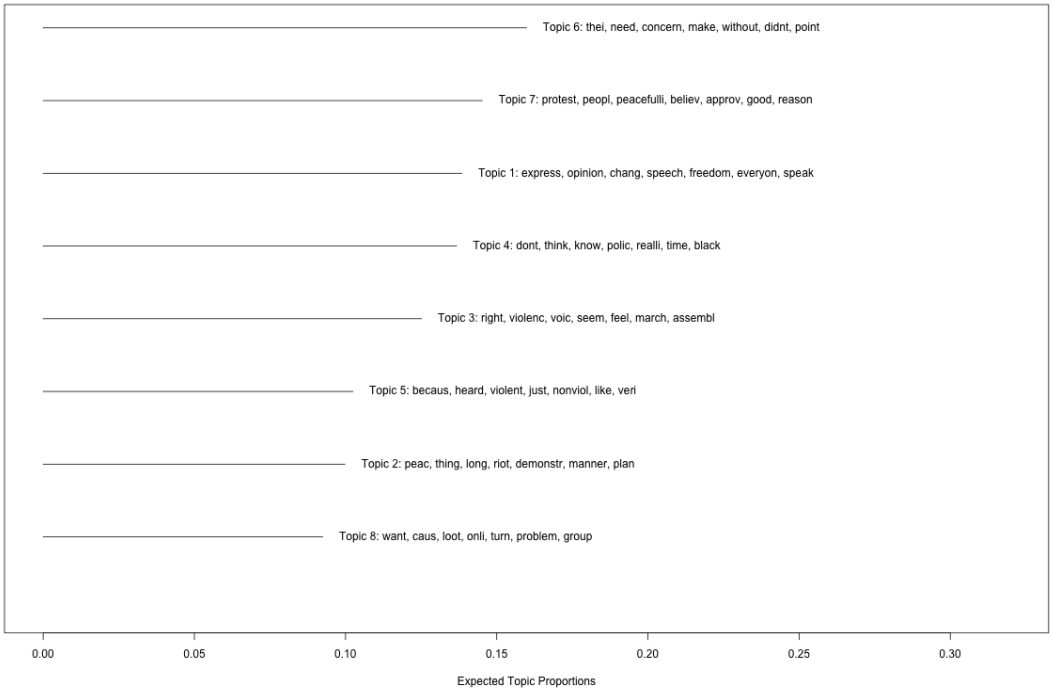

Figure A4: Structural topic model with 8 topics (US survey)

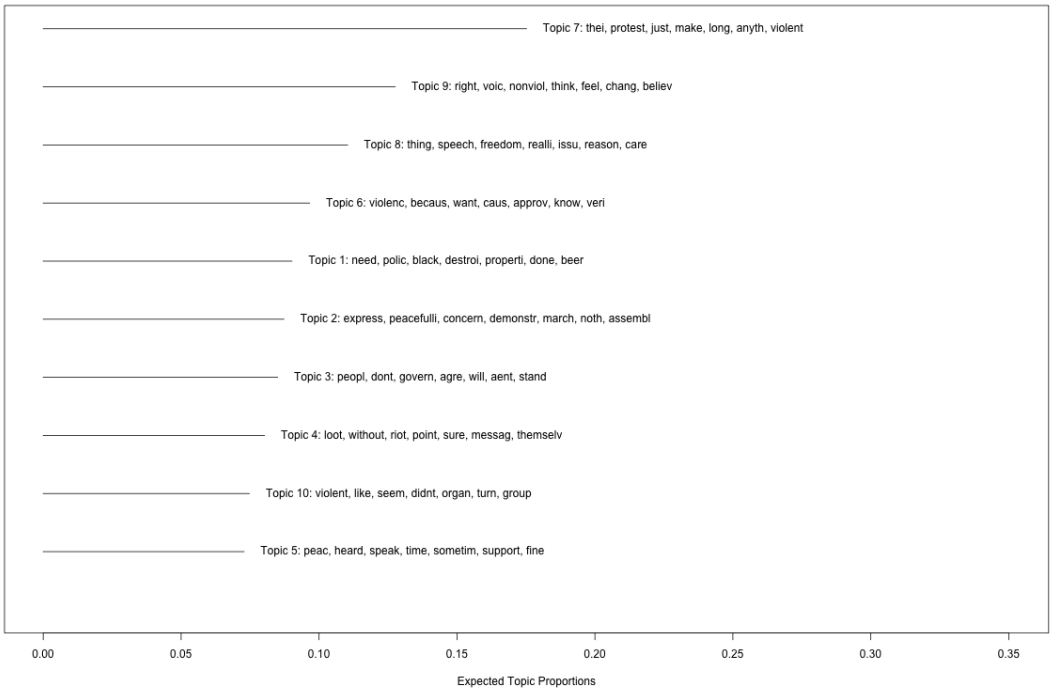

Figure A5: Structural topic model with 10 topics (US survey)

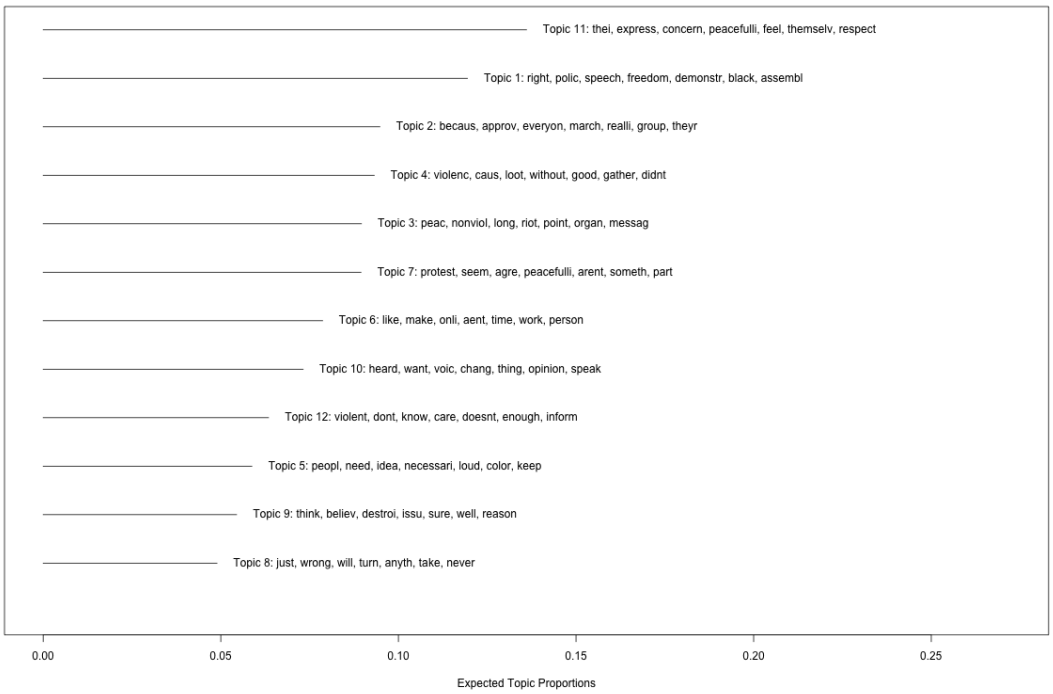

Figure A6: Structural topic model with 12 topics (US survey)

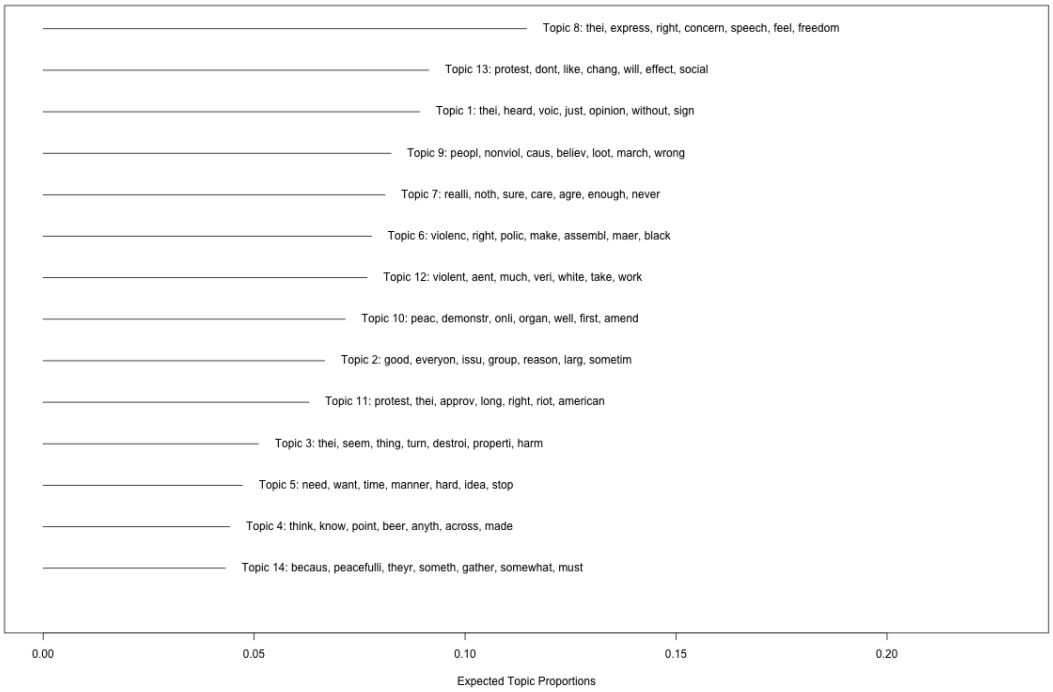

Figure A7: Structural topic model with 14 topics (US survey)

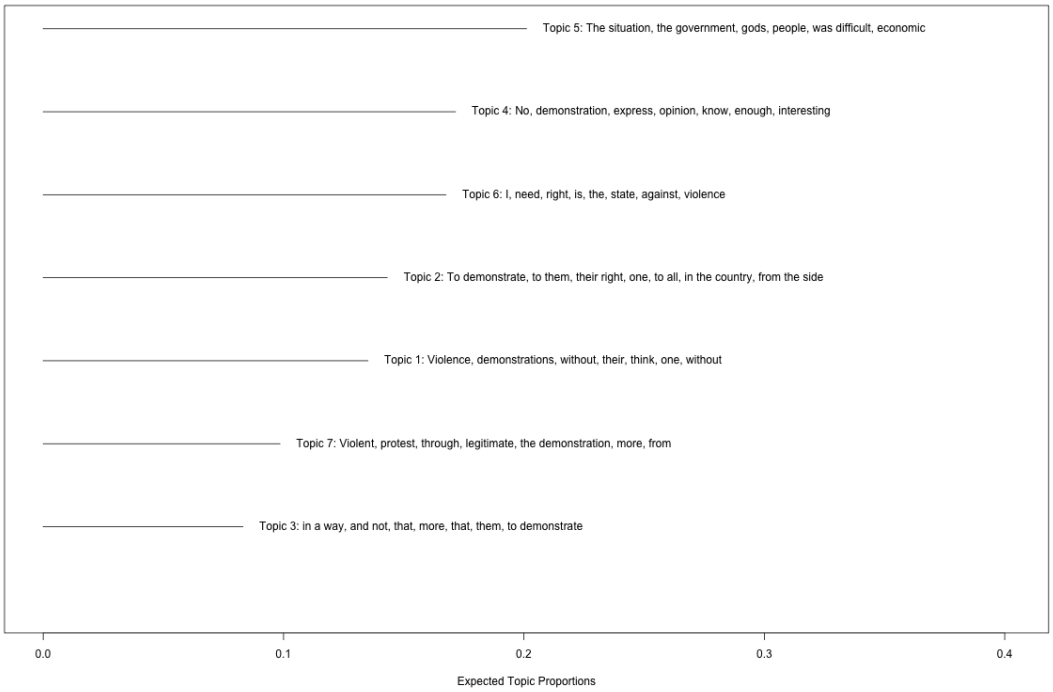

Figure A8: Structural topic model with 7 topics (Israeli Arab survey)

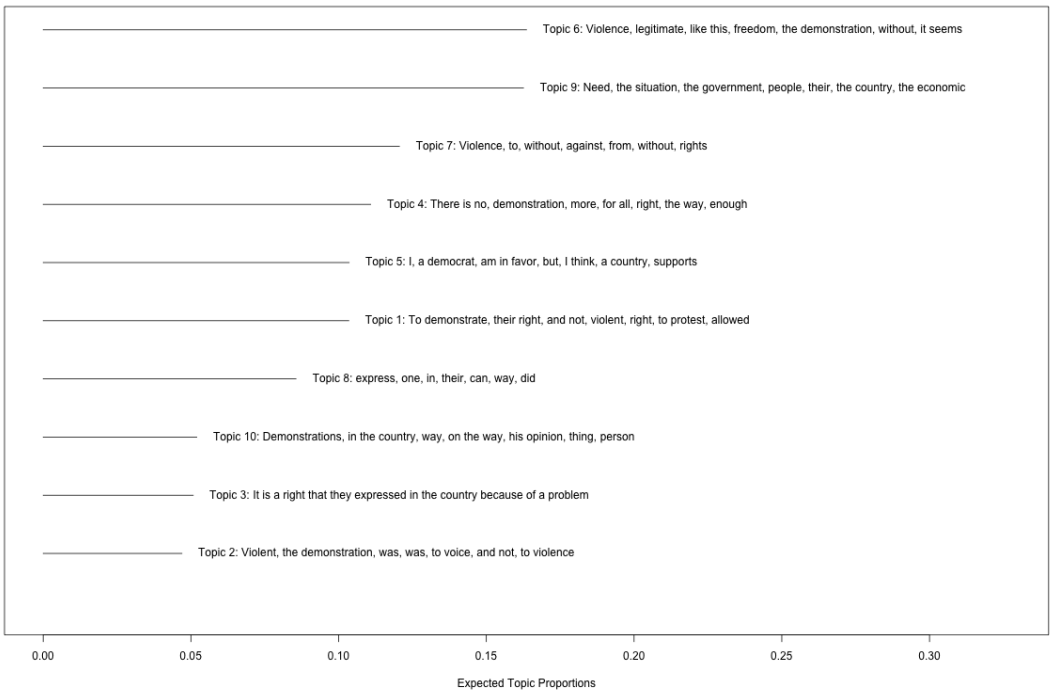

Figure A9: Structural topic model with 10 topics (Israeli Arab survey)

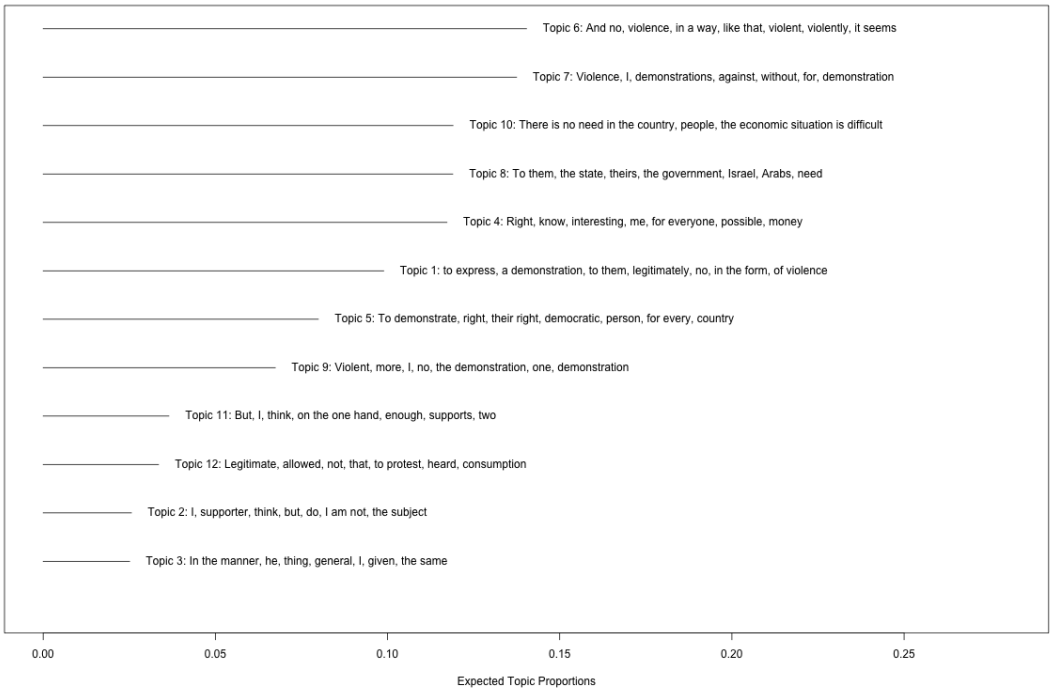

Figure A10: Structural topic model with 12 topics (Israeli Arab survey)

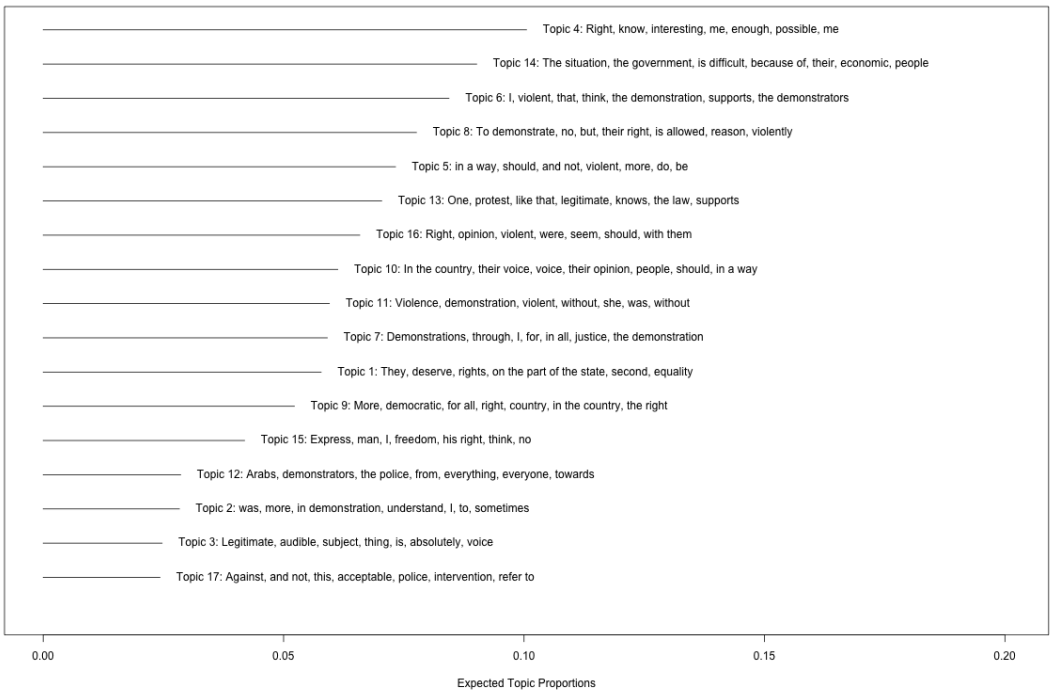

Figure A11: Structural topic model with 17 topics (Israeli Arab survey)

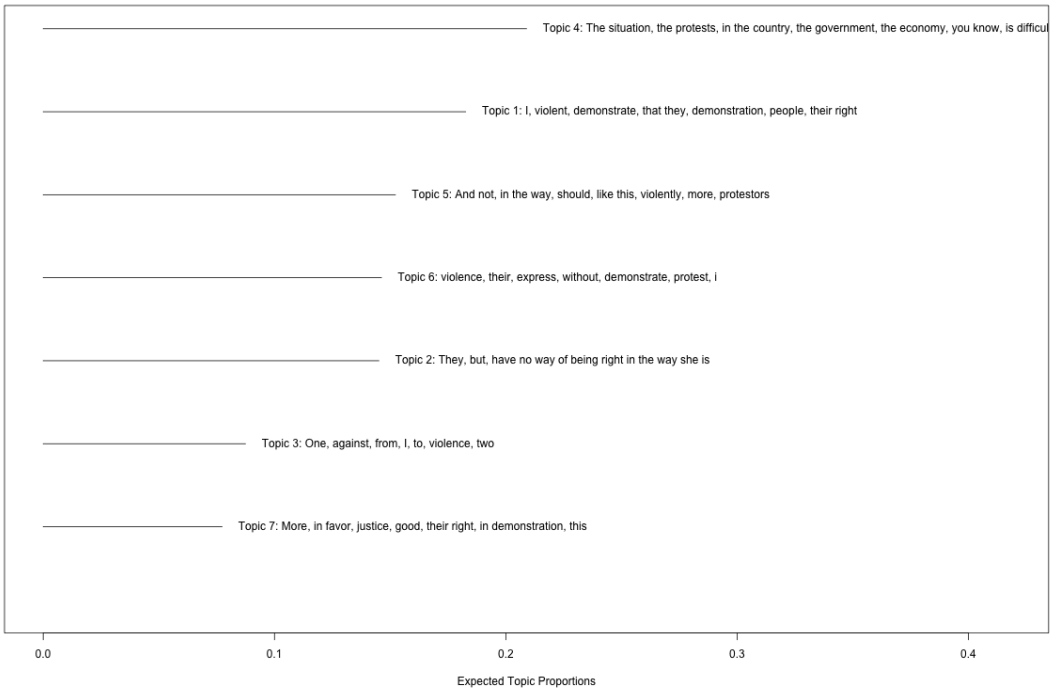

Figure A12: Structural topic model with 7 topics (Israeli Ethiopian survey)

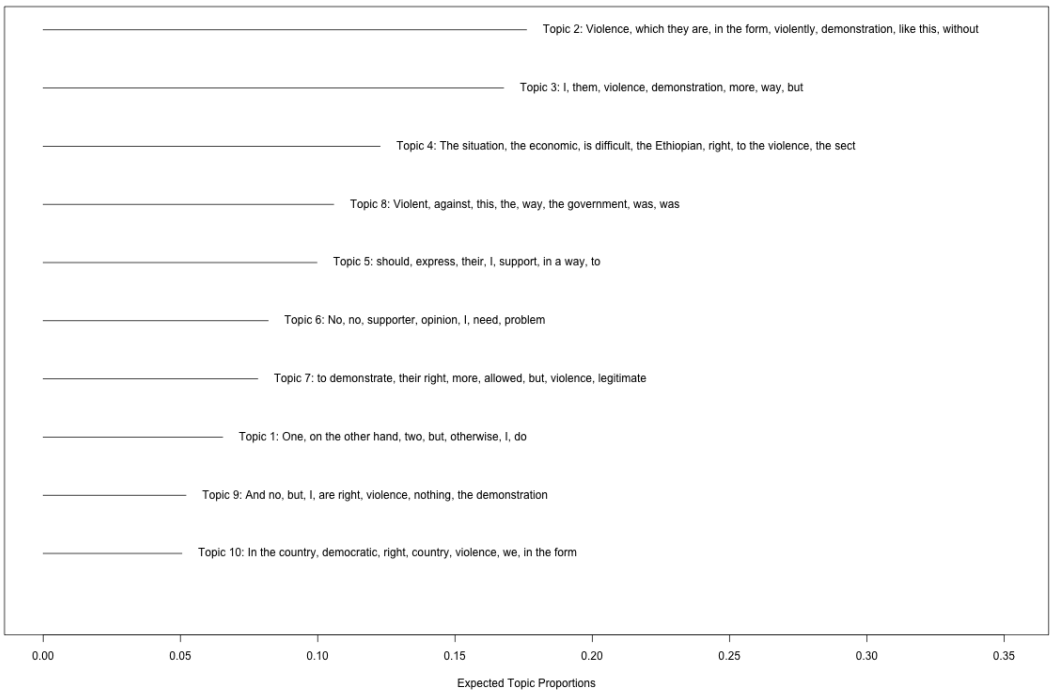

Figure A13: Structural topic model with 10 topics (Israeli Ethiopian survey)

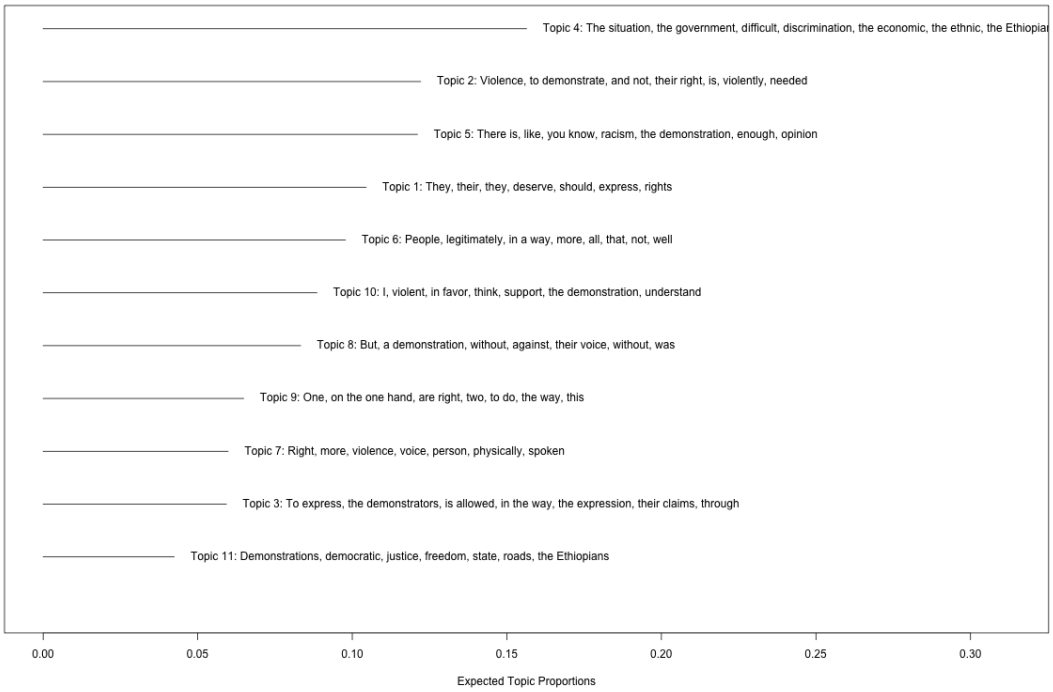

Figure A14: Structural topic model with 11 topics (Israeli Ethiopian survey)

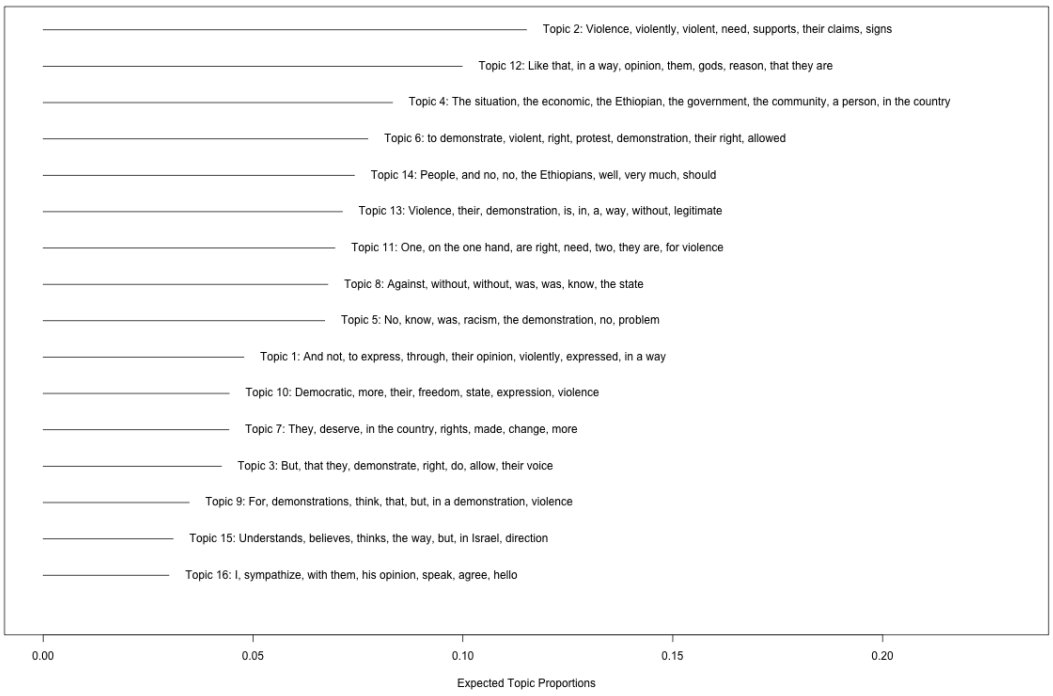

Figure A15: Structural topic model with 16 topics (Israeli Ethiopian survey)

11 Appendix Tables

Table A1: Relication of Table A3 - NoChanges

	Dependent variable:					
	success					
	(1)	(2)	(3)	(4)	(5)	(6)
INIT_NV_ONSET	0.001	−0.001	−0.026	0.226***	0.226***	0.272***
	(0.024)	(0.028)	(0.029)	(0.017)	(0.021)	(0.023)
EPR_STATUS_ORD	0.003	0.004	0.002			
	(0.003)	(0.005)	(0.005)			
EPR_STATUS_ORD:INIT_NV_ONSET	0.031***	0.033***	0.049***			
	(0.006)	(0.006)	(0.007)			
EPR_STATUS_EXCL				−0.008	−0.001	−0.003
				(0.011)	(0.016)	(0.021)
EPR_STATUS_EXCL:INIT_NV_ONSET				−0.202***	−0.188***	−0.233***
				(0.025)	(0.028)	(0.031)
EPR_GROUPSIZE		−0.040	−0.027		−0.008	−0.016
		(0.031)	(0.036)		(0.025)	(0.033)
Observations	1,944	1,746	1,569	2,099	1,901	1,569
R^2	0.072	0.088	0.125	0.096	0.105	0.128
Adjusted R^2	0.071	0.080	0.114	0.095	0.098	0.117
Note:						*p<0.1; **p<0.05; ***p<0.01

Table A2: DataOriginal RecodeIV-PowerlessNotExcluded

	Dependent variable: Success (0/1)			
	success			
	Original	Replication	Original	Replication
	(1)	(2)	(3)	(4)
EPR_STATUS_ORD	0.002	0.002		
	(0.005)	(0.005)		
EPR_STATUS_EXCL			−0.003	−0.001
			(0.021)	(0.020)
INIT_NV_ONSET	−0.026	−0.026	0.272***	0.272***
	(0.029)	(0.029)	(0.023)	(0.023)
EPR_GROUPSIZE	−0.027	−0.027	−0.016	−0.014
	(0.036)	(0.036)	(0.033)	(0.032)
EPR_STATUS_ORD:INIT_NV_ONSET	0.049***	0.049***		
	(0.007)	(0.007)		
EPR_STATUS_EXCL:INIT_NV_ONSET			−0.233***	−0.234***
			(0.031)	(0.031)
Controls	Yes	Yes	Yes	Yes
Cubic polynomials for time	Yes	Yes	Yes	Yes
Observations	1,569	1,569	1,569	1,569
R^2	0.125	0.125	0.128	0.128
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A3: DataOriginal Subsample-WithoutSelfExcludedGroups

	Dependent variable: Success (0/1)			
	success			
	Original	Replication	Original	Replication
	(1)	(2)	(3)	(4)
EPR_STATUS_ORD	0.002	0.005		
	(0.005)	(0.007)		
EPR_STATUS_EXCL			−0.003	0.009
			(0.021)	(0.025)
INIT_NV_ONSET	−0.026	0.065	0.272***	0.293***
	(0.029)	(0.047)	(0.023)	(0.026)
EPR_GROUPSIZE	−0.027	−0.078	−0.016	−0.051
	(0.036)	(0.049)	(0.033)	(0.041)
EPR_STATUS_ORD:INIT_NV_ONSET	0.049***	0.036***		
	(0.007)	(0.009)		
EPR_STATUS_EXCL:INIT_NV_ONSET			−0.233***	−0.185***
			(0.031)	(0.039)
Controls	Yes	Yes	Yes	Yes
Cubic polynomials for time	Yes	Yes	Yes	Yes
Observations	1,569	1,161	1,569	1,161
R^2	0.125	0.147	0.128	0.152
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A4: DataOriginal EstimationBinomial

	Dependent variable: Success (0/1)			
	success			
	OLS Original	*logistic* Replication	*OLS* Original	*logistic* Replication
	(1)	(2)	(3)	(4)
EPR_STATUS_ORD	0.002	0.041		
	(0.005)	(0.115)		
EPR_STATUS_EXCL			−0.003	−0.043
			(0.021)	(0.459)
INIT_NV_ONSET	−0.026	0.815	0.272***	2.915***
	(0.029)	(0.618)	(0.023)	(0.407)
EPR_GROUPSIZE	−0.027	−0.376	−0.016	−0.185
	(0.036)	(0.727)	(0.033)	(0.655)
EPR_STATUS_ORD:INIT_NV_ONSET	0.049***	0.343***		
	(0.007)	(0.129)		
EPR_STATUS_EXCL:INIT_NV_ONSET			−0.233***	−1.768***
			(0.031)	(0.595)
Controls	Yes	Yes	Yes	Yes
Cubic polynomials for time	Yes	Yes	Yes	Yes
Observations	1,569	1,569	1,569	1,569
R^2	0.125		0.128	
Log Likelihood		−278.439		−277.556
Akaike Inf. Crit.		596.879		595.111
Note:				*p<0.1; **p<0.05; ***p<0.01