When do private actors engage in censorship? Evidence from a correspondence experiment with Russian private media firms

Supplemental Information

May 24, 2021

LIST OF FIGURES

LIST OF TABLES

A. Directories used to compile sample of media firms

We used the following four directories to create a sample of media firms.

ABZY News Links. 2013. https://goo.gl/cd9yNd. Retrieved December 20, 2013.

Kidon. 2013. https://goo.gl/gc9C1G. Retrieved December 20, 2013.

Onlinenewspapers.com. 2013. https://goo.gl/Kre2KR. Retrieved December 20, 2013.

Public.ru Internet Media Library. 2013. https://goo.gl/ChnKBg. Retrieved December 20, 2013.

B. Media bias measure

We constructed a media bias measure in July-August 2014 to partially address the possibility that the media firms in our sample might voluntarily censor our advertisements criticizing the lack of free speech in Russia. To construct the media bias measure, we investigated how websites reported on the 2014 blogging law, the shutdown of *TV Dozhd,* and the blocking of Alexey Navalny's blog. If websites did not cover any of these events we used the most recent article that included the phrase "freedom of speech."

The media bias measure consists of a score for each website, based on the four component scores for reporting on the 2014 blogging law, the shutdown of *TV Dozhd,* the blocking of Alexey Navalny's blog, and the most recent article on freedom of speech. Websites often were supportive of the regime on some issues but critical on others. Rather than average these four scores for each website, many of which were missing either by design or since a given website had not reported on a given topic, we used the score most supportive of free speech and thus most critical of the regime.

- *Strongly anti-free speech:* coverage openly calls for repression of free speech, discusses the necessity of repression, or discusses acts of repression in a positive light.

- *Mainly anti-free speech:* coverage generally shows government repression in a positive light, though some counterarguments might appear. Phrases that signal government support are softer than in texts coded as strongly anti-free speech.

- *Neutral:* coverage includes a neutral account of events. Coverage is neither pro-free speech nor anti-free speech.

- *Mainly pro-free speech:* coverage generally shows government repression in a negative light, though some counterarguments might appear. Phrases that signal free speech support are softer than in texts coded as strongly pro-free speech.

- *Strongly pro-free speech:* coverage openly criticizes the repression of free speech or

discusses acts of repression and their negative effects.

- *Missing:* No information available.

For the empirical analysis we collapsed the strongly and mainly positive categories and the strongly and mainly negative categories. This has no effect on our (null) results but increases statistical power when searching for treatment effect heterogeneity.

In addition to the quantitative results reported in the paper, one might wonder if the experimental treatments prompted any other revealing differences among the treatment groups. One natural place to investigate is firms' stated motivations for declining to publish our ads. We note that reply emails can only provide limited information since responses may not fully reveal respondents' motivations for rejecting the ad request. We also do not know what non-responding firms would have written had they replied. Nonetheless, it is still informative to at least informally map the content of the few rejection emails we received onto our experimental treatments.

When firms rejected a non-political ad, they typically cited the ad's thematic content. One firm wrote, "Unfortunately, we do not place ads on unrelated subjects"[24] while another wrote that "We only place tourism-related advertising on our portal, unfortunately."[25] In marked contrast, firms rejecting the politically sensitive ads often cited the individual ad, or the information it contained, as problematic. For example, one firm wrote, "We cannot place this kind of information on our website."[26] Other firms expressed support for the ad's political message but still rejected it based on content: "Unfortunately, given the specific [content] of the ad (although we support freedom of speech and are trying to protect it in our site), we are forced to refuse placing your ad."[27] Somewhat surprisingly, the diction of rejection emails also varied systematically by treatment condition. Firms that received non-political ads indicated that they 'do not' ('не размещаем') place such ads, suggesting that some firms rejected the non-political ads because they did not fit the site's topic. On the other hand, firms receiving a politically sensitive ad often indicated that they 'cannot' ('сможем') publish the advertisement, suggesting perhaps that they felt constrained in their

---

[24]"К сожалению, мы не размещаем не тематическую рекламу."

[25]"Мы размещаем на нашем портале только туристическую рекламу, к сожалению."

[26]"Подобную информацию мы не сможем разместить на нашем сайте."

[27]Сожалению учитывая специфику рекламы (хотя мы поддерживаем свободу слова и пытаемся на своих страницах ее беречь) мы вынуждены отказать Вам в размещении.

editorial decisions.[28] These observations are merely suggestive, and we do not wish to make too much of them, but in terms of both stated reasons and word choice describing those reasons, the limited qualitative evidence available within the reply emails is consistent with the quantitative results.

---

[28]The verb 'сможем' was not used by firms that rejected the non-political ads.

Figure SI-6: Map
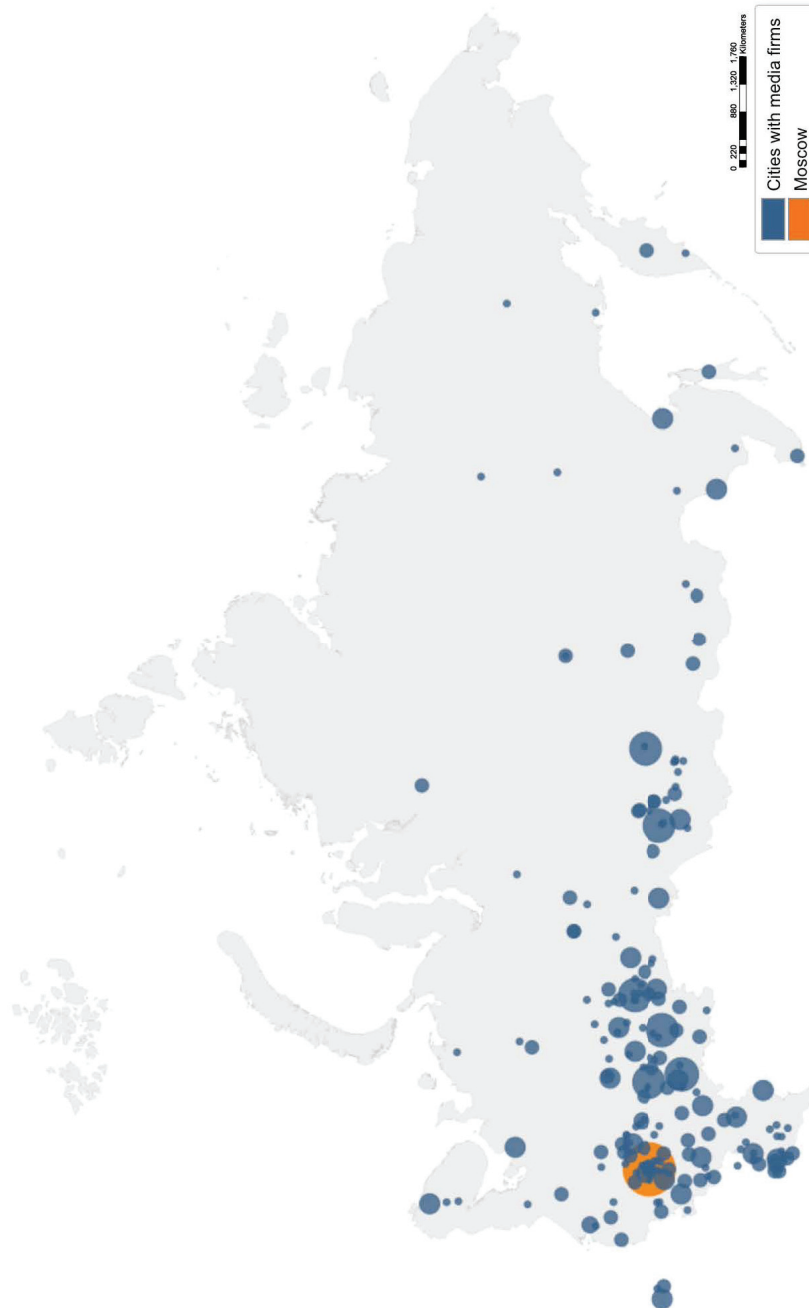
The map shows the location of the media firms in our sample. Larger bubbles denote a larger number of media firms in the same city.

Figure SI-7: Advertisement request email (Russian version)

Subject: Запрос на размещение рекламы

Добрый день,

Я хотел бы поинтересоваться Вашими рекламными предложениями.
Наша компания заинтересована в размещении рекламы на Вашем сайте
(.png версию предварительного макета рекламы Вы можете найти в
приложении). Данный размер рекламы ориентировочный; его можно
изменить в соответствии с Вашей размерной сеткой. Могли бы Вы
прислать мне Ваше ценовое предложение за месяц и за количество показов?
Также, я был бы признателен, если бы Вы посмотрели приложение
и посоветовали мне, какой размер рекламы и место размещения
на сайте будет оптимальным, учитывая предполагаемый дизайн
рекламы и ее содержание.

Мы планируем начать нашу рекламную кампанию в ближайшем будущем и будем
признательны Вам за быстрый ответ на наш запрос.

С благодарностью,

Антон Волков
Рекламный директор
Наш Альянс
volkov@nashalyans.com

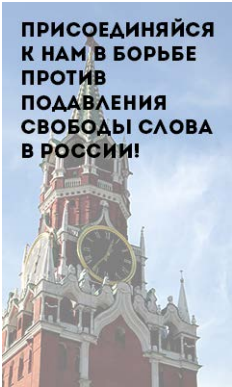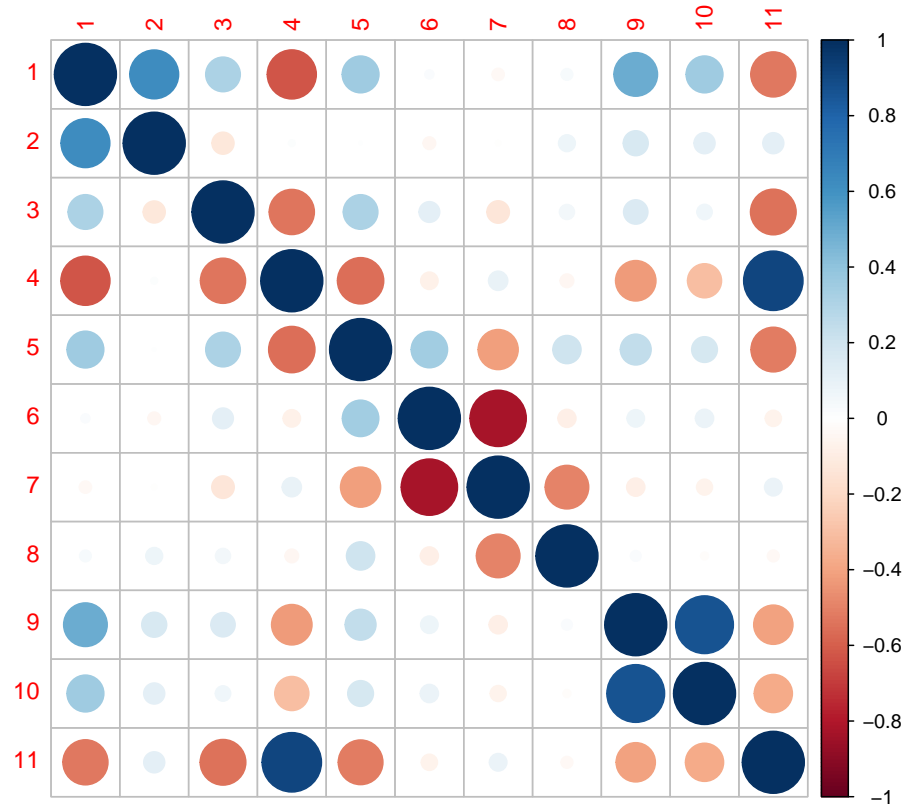Figure SI-8: Advertisements (Russian version)

Figure SI-9: Covariate correlation matrix

Note: The plot shows a correlation matrix for the covariates listed in Table SI-4. For covariate labels please refer to Table SI-4.

Figure SI-10: Predicted probabilities
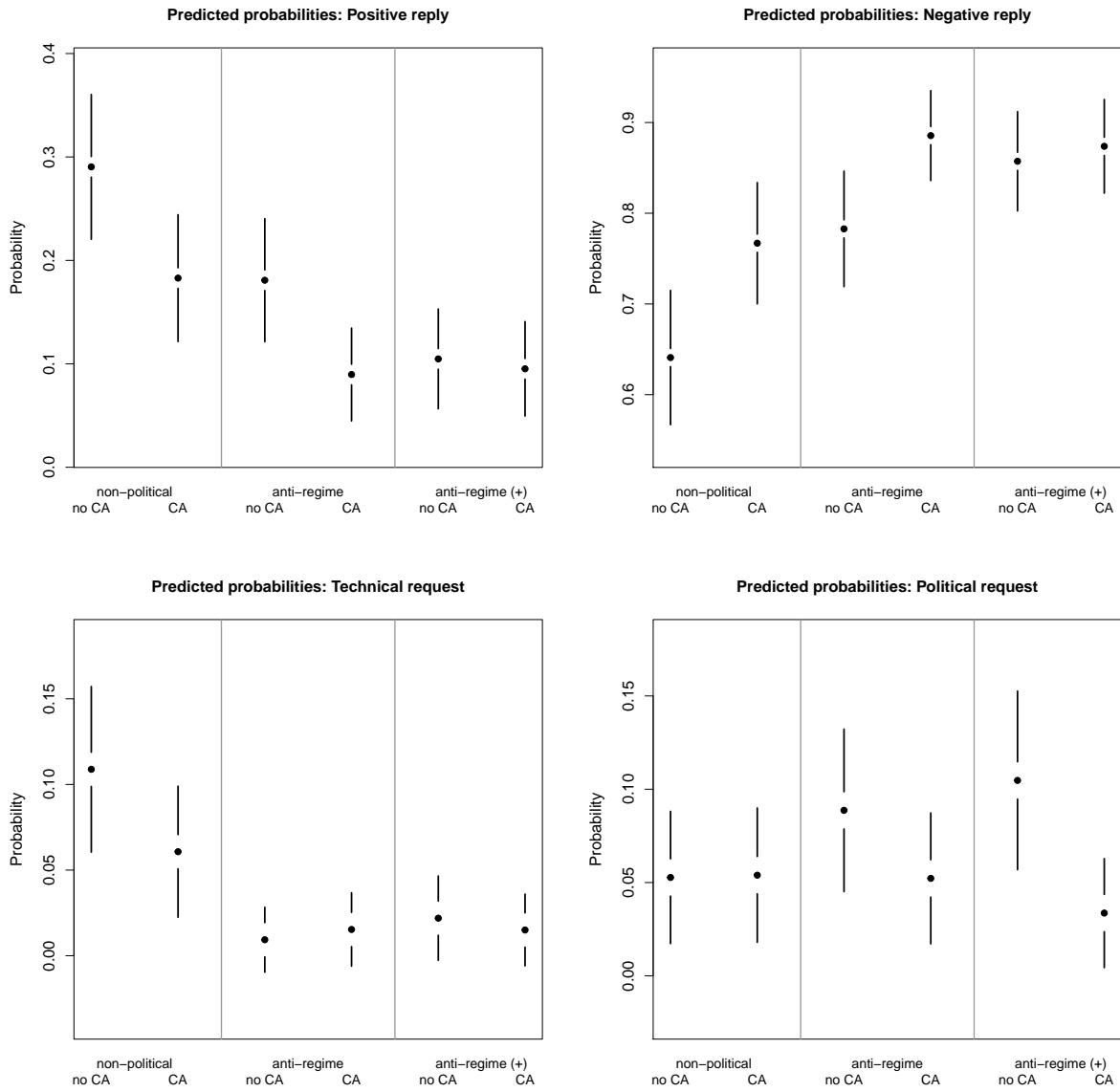
Predicted probabilities: Positive reply

Predicted probabilities: Negative reply

Predicted probabilities: Technical request

Predicted probabilities: Political request

Note: The plots display simulated probability estimates from four probit models. Point estimates and 95% confidence intervals are shown. Note that vertical axes differ across plots.

Figure SI-11: Predicted probabilities
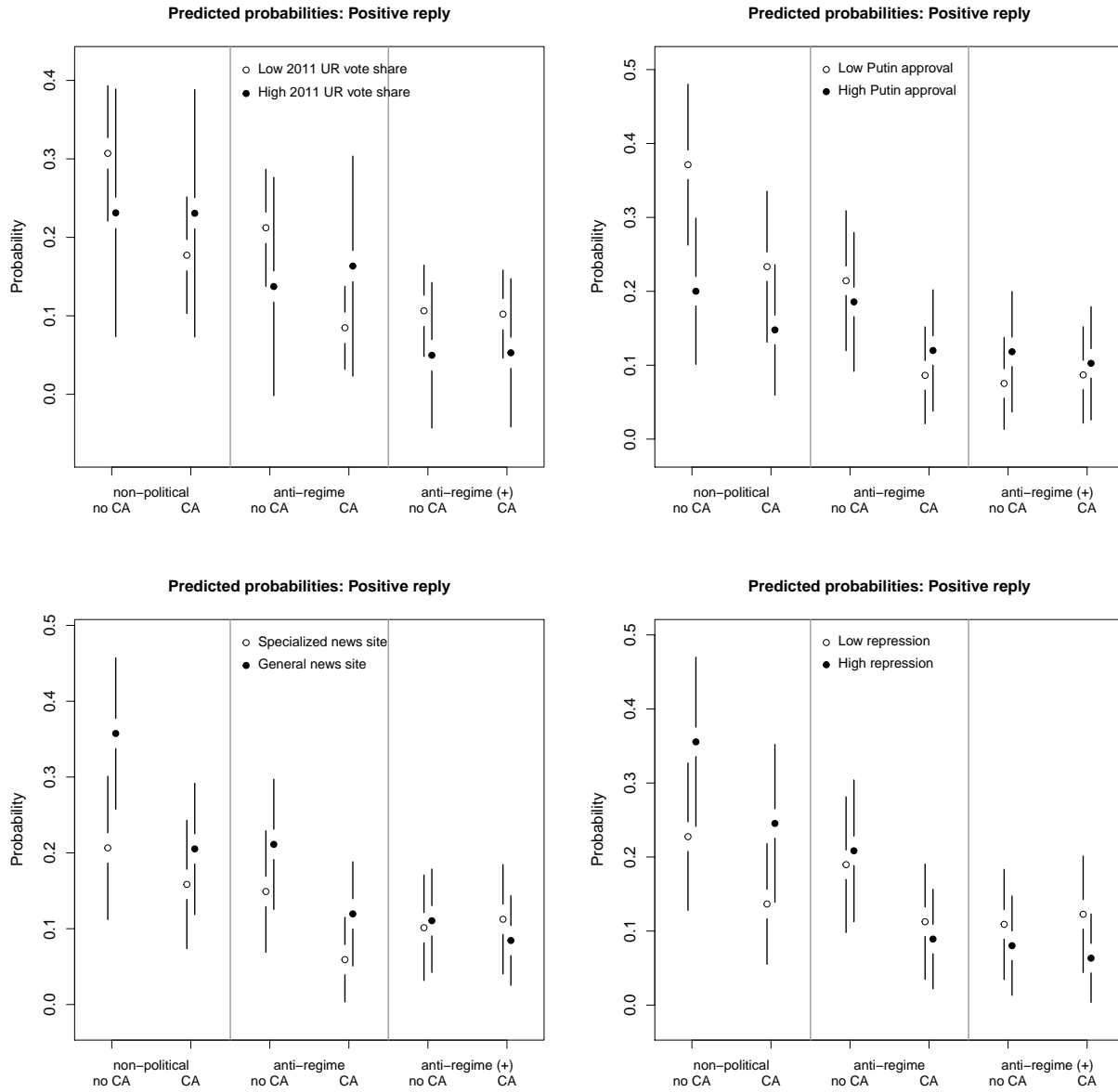
Note: The plots display simulated probability estimates from four probit models. Point estimates and 95% confidence intervals are shown. Note that vertical axes differ across plots.

Figure SI-12: Predicted probabilities

Note: The plots display simulated probability estimates from four probit models. Point estimates and 95% confidence intervals are shown. Note that vertical axes differ across plots.
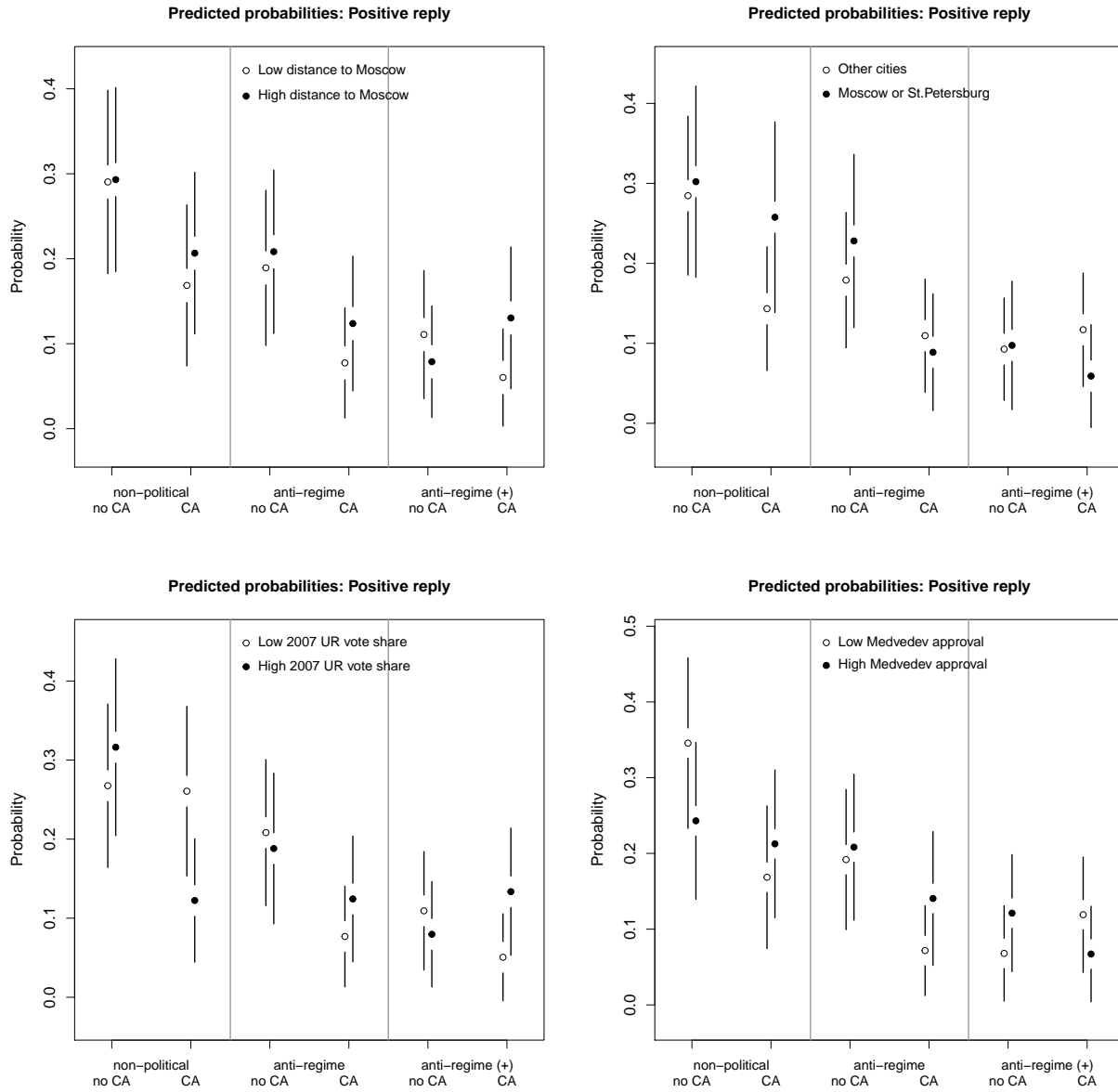
Figure SI-13: Predicted probabilities

Note: The plots display simulated probability estimates from four probit models. Point estimates and 95% confidence intervals are shown. Note that vertical axes differ across plots.

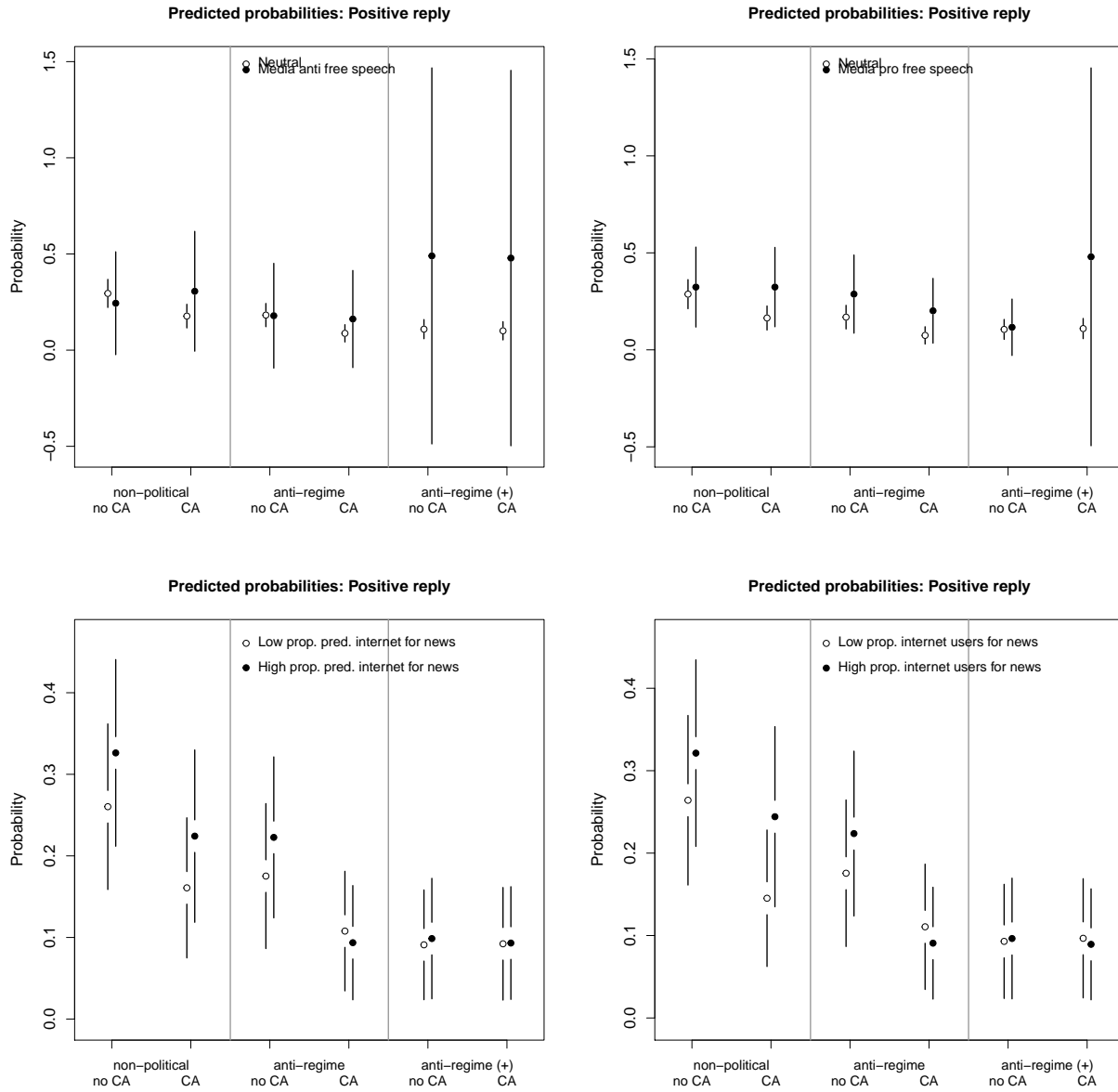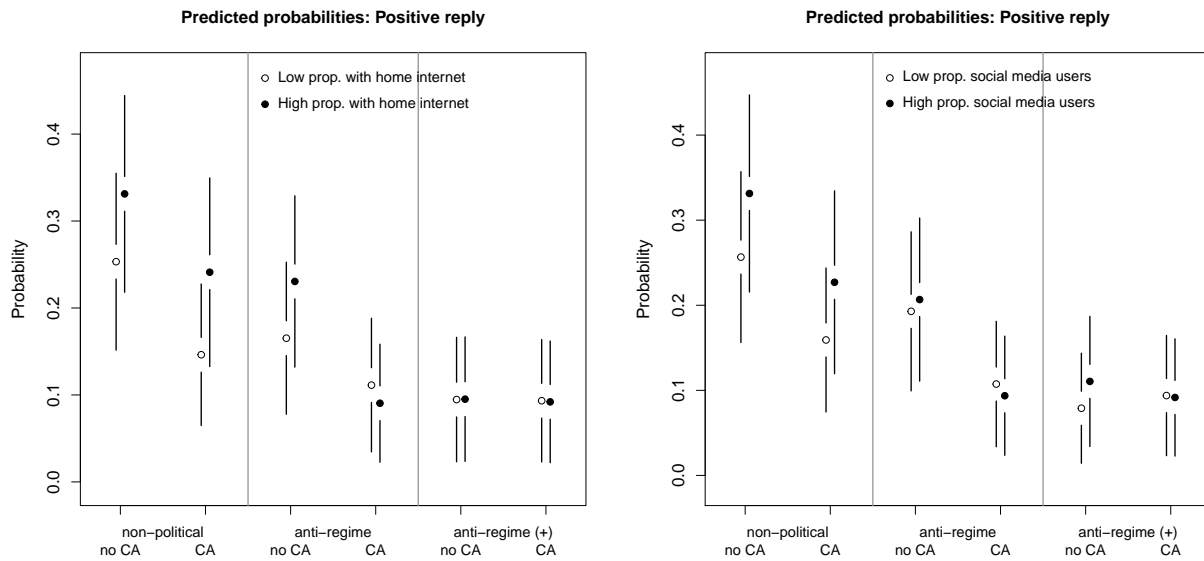Figure SI-14: Predicted probabilities

Note: The plots display simulated probability estimates from four probit models. Point estimates and 95% confidence intervals are shown. Note that vertical axes differ across plots.

Figure SI-15: Treatment effect heterogeneity
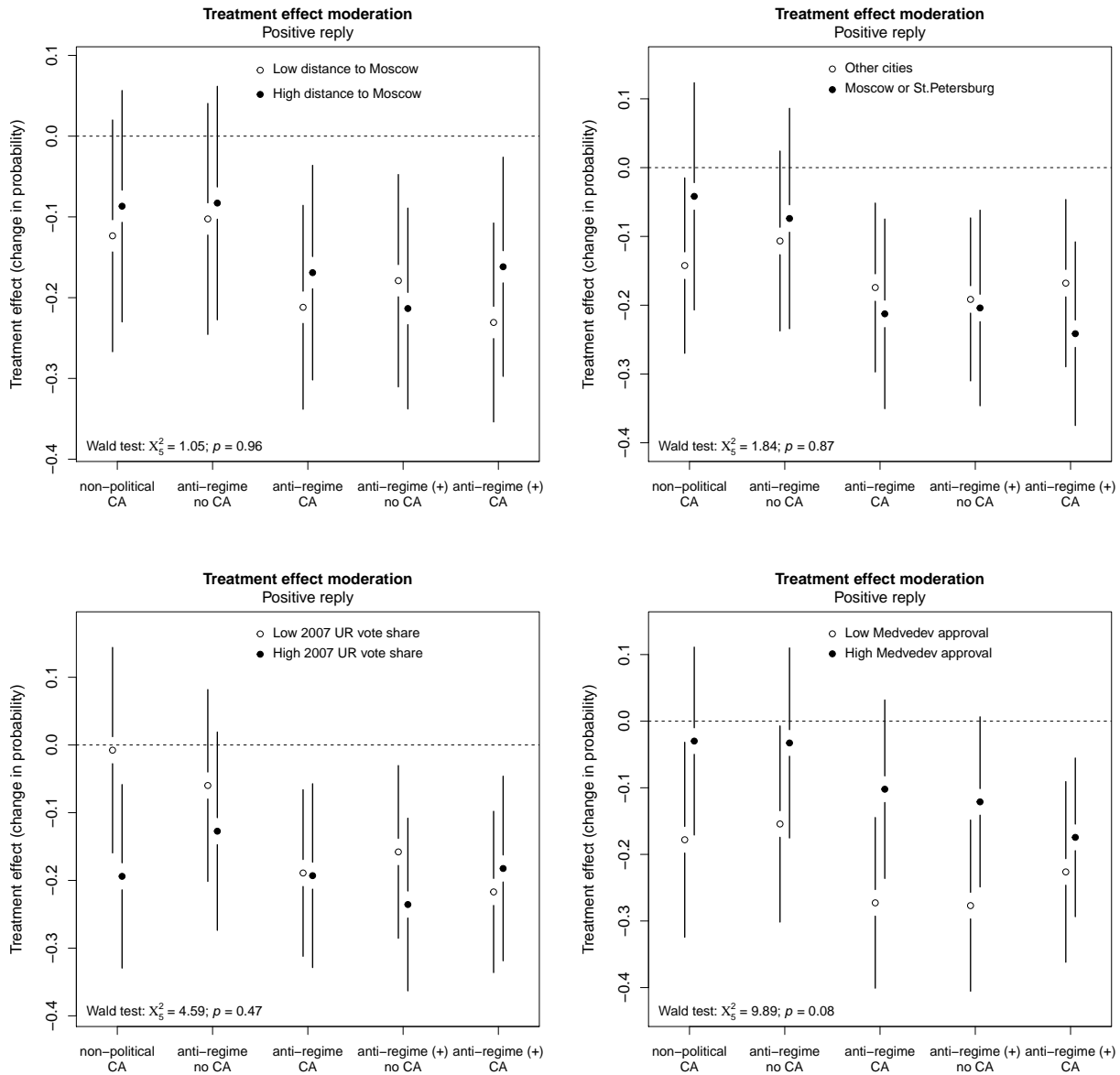
Note: The plots display simulated treatment effect estimates (changes in probabilities) from probit models. Two estimates are shown for each treatment combination: one for firms with covariate values below the sample median (empty circles) and one for firms with covariate values above the sample median (filled circles). Point estimates and 95% confidence intervals are shown. Treatment effects are computed relative to the baseline condition (non-political ad, no collective action framing). The Wald test at the bottom of each plot tests the joint null hypothesis that treatment effects are the same for firms with covariate values below and above the sample median.

Figure SI-16: Treatment effect heterogeneity
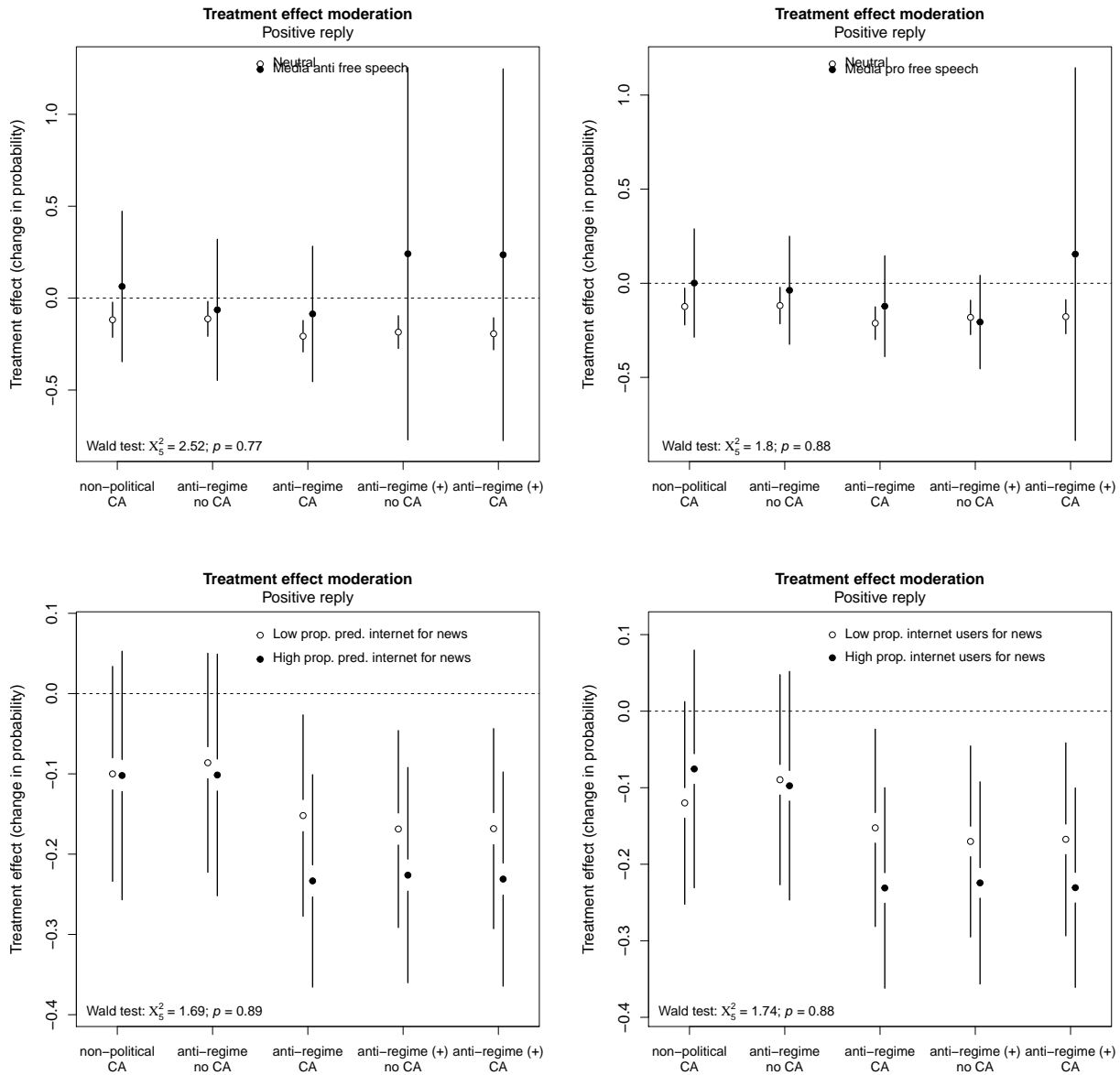
Note: The plots display simulated treatment effect estimates (changes in probabilities) from probit models. Two estimates are shown for each treatment combination: one for firms with covariate values below the sample median (empty circles) and one for firms with covariate values above the sample median (filled circles). Point estimates and 95% confidence intervals are shown. Treatment effects are computed relative to the baseline condition (non-political ad, no collective action framing). The Wald test at the bottom of each plot tests the joint null hypothesis that treatment effects are the same for firms with covariate values below and above the sample median.
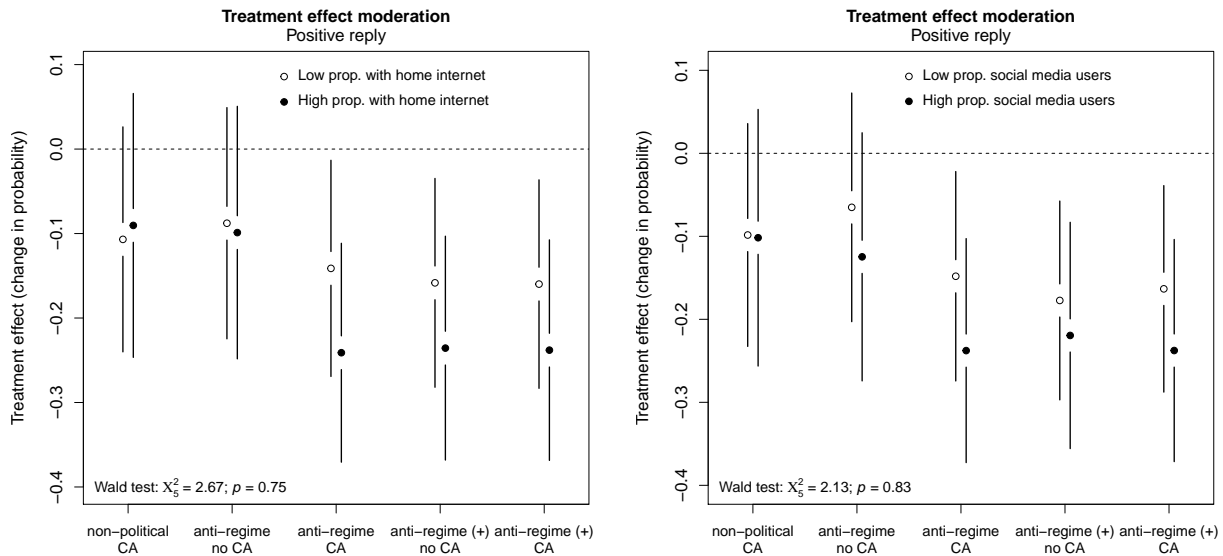
Figure SI-17: Treatment effect heterogeneity

Note: The plots display simulated treatment effect estimates (changes in probabilities) from probit models. Two estimates are shown for each treatment combination: one for firms with covariate values below the sample median (empty circles) and one for firms with covariate values above the sample median (filled circles). Point estimates and 95% confidence intervals are shown. Treatment effects are computed relative to the baseline condition (non-political ad, no collective action framing). The Wald test at the bottom of each plot tests the joint null hypothesis that treatment effects are the same for firms with covariate values below and above the sample median.
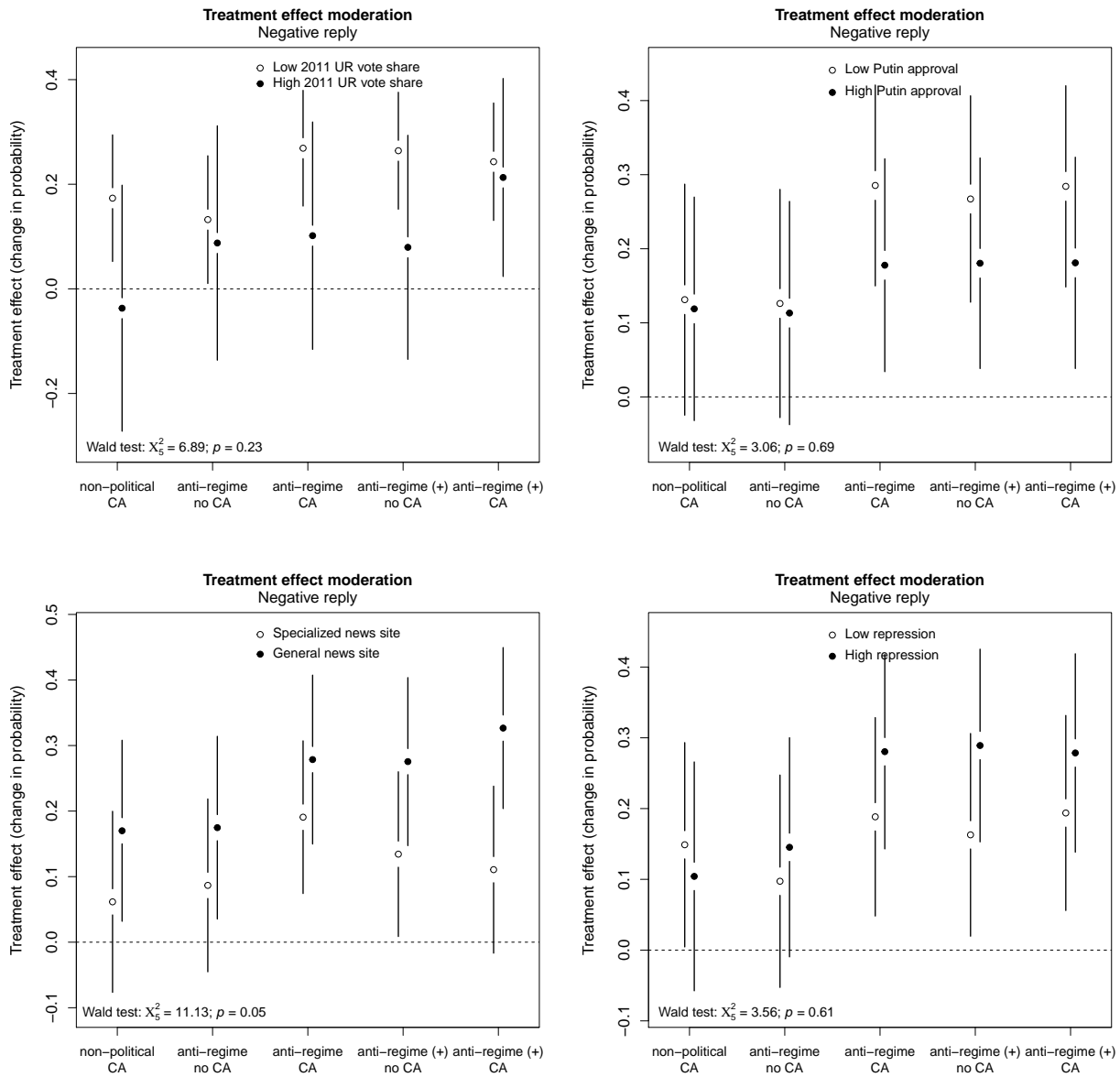
Figure SI-18: Treatment effect heterogeneity

Note: The plots display simulated treatment effect estimates (changes in probabilities) from probit models. Two estimates are shown for each treatment combination: one for firms with covariate values below the sample median (empty circles) and one for firms with covariate values above the sample median (filled circles). Point estimates and 95% confidence intervals are shown. Treatment effects are computed relative to the baseline condition (non-political ad, no collective action framing). The Wald test at the bottom of each plot tests the joint null hypothesis that treatment effects are the same for firms with covariate values below and above the sample median.
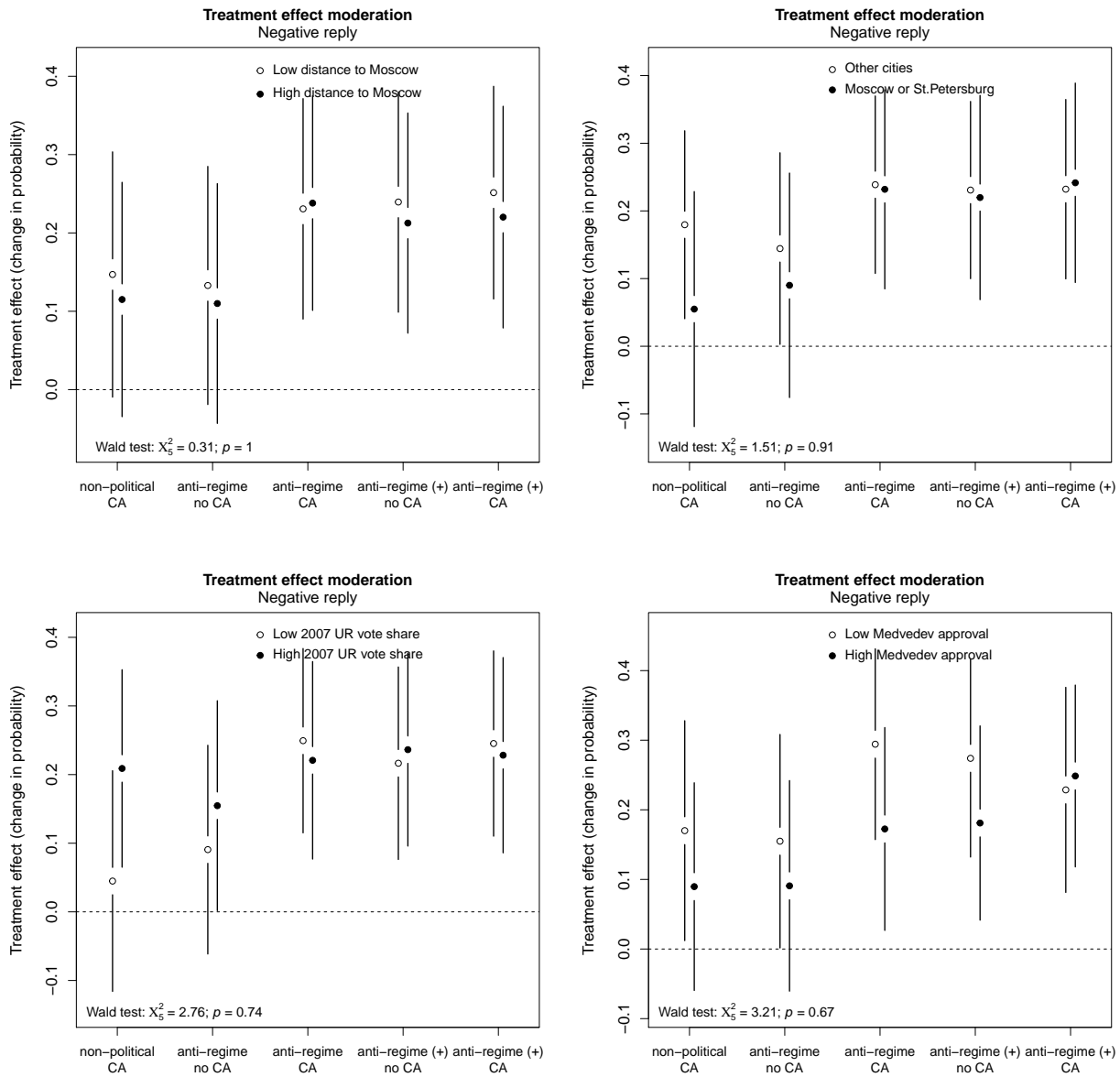
Figure SI-19: Treatment effect heterogeneity

Note: The plots display simulated treatment effect estimates (changes in probabilities) from probit models. Two estimates are shown for each treatment combination: one for firms with covariate values below the sample median (empty circles) and one for firms with covariate values above the sample median (filled circles). Point estimates and 95% confidence intervals are shown. Treatment effects are computed relative to the baseline condition (non-political ad, no collective action framing). The Wald test at the bottom of each plot tests the joint null hypothesis that treatment effects are the same for firms with covariate values below and above the sample median.
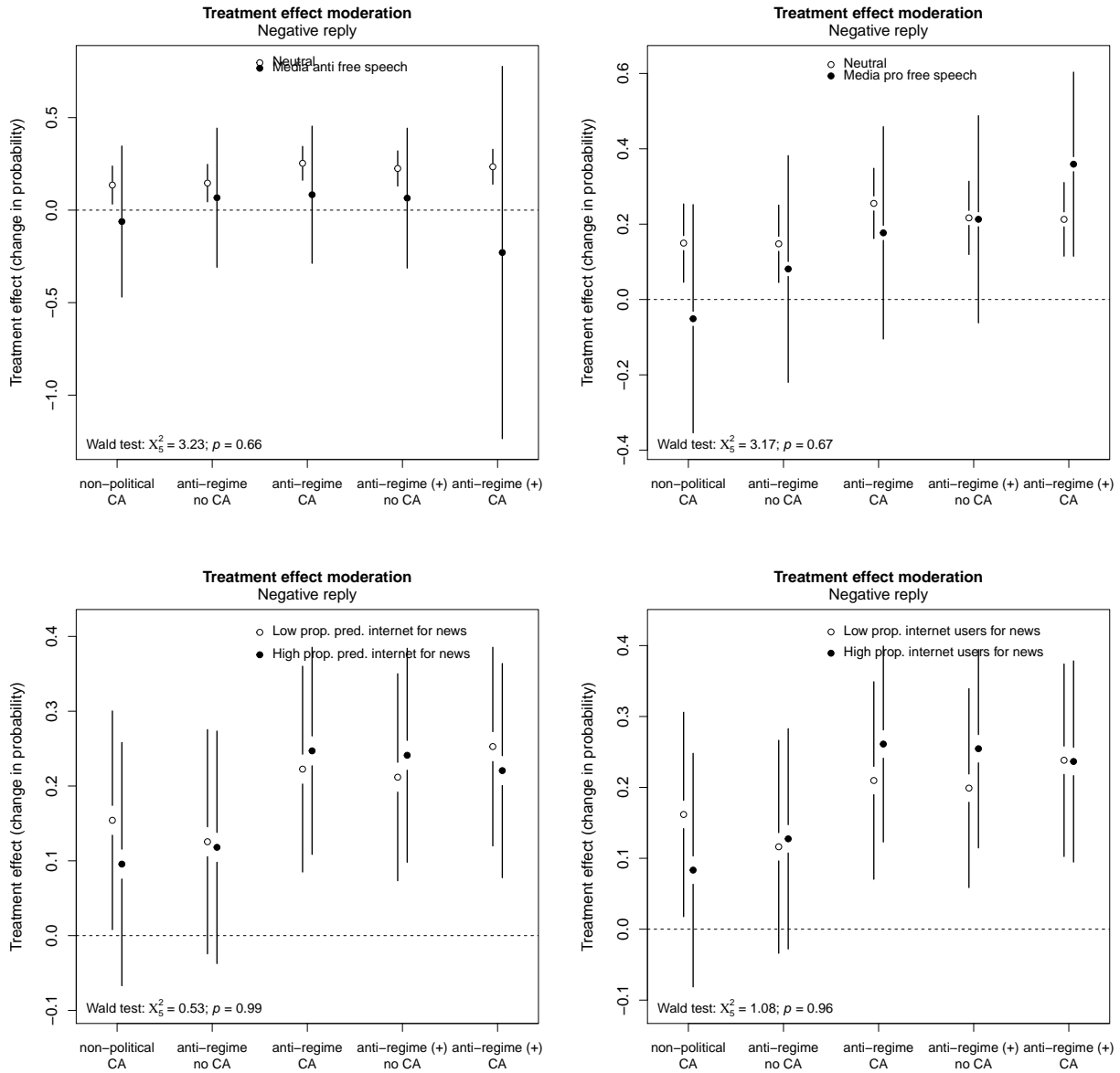
Figure SI-20: Treatment effect heterogeneity

Note: The plots display simulated treatment effect estimates (changes in probabilities) from probit models. Two estimates are shown for each treatment combination: one for firms with covariate values below the sample median (empty circles) and one for firms with covariate values above the sample median (filled circles). Point estimates and 95% confidence intervals are shown. Treatment effects are computed relative to the baseline condition (non-political ad, no collective action framing). The Wald test at the bottom of each plot tests the joint null hypothesis that treatment effects are the same for firms with covariate values below and above the sample median.

Figure SI-21: Treatment effect heterogeneity

Treatment effect moderation
Negative reply

○ Low prop. with home internet
● High prop. with home internet

Treatment effect (change in probability)

Wald test: $X_5^2 = 1.7$; $p = 0.89$

non–political
CA

anti–regime
no CA

anti–regime
CA

anti–regime (+)
no CA

anti–regime (+)
CA

Treatment effect moderation
Negative reply

○ Low prop. social media users
● High prop. social media users

Treatment effect (change in probability)

Wald test: $X_5^2 = 0.72$; $p = 0.98$

non–political
CA

anti–regime
no CA

anti–regime
CA

anti–regime (+)
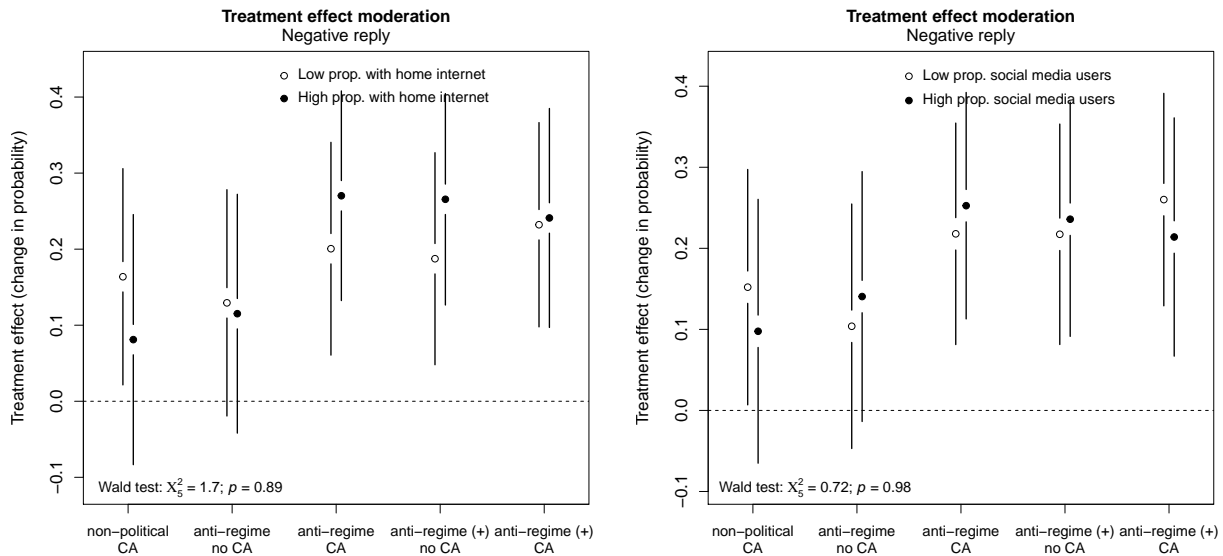no CA

anti–regime (+)
CA

Note: The plots display simulated treatment effect estimates (changes in probabilities) from probit models. Two estimates are shown for each treatment combination: one for firms with covariate values below the sample median (empty circles) and one for firms with covariate values above the sample median (filled circles). Point estimates and 95% confidence intervals are shown. Treatment effects are computed relative to the baseline condition (non-political ad, no collective action framing). The Wald test at the bottom of each plot tests the joint null hypothesis that treatment effects are the same for firms with covariate values below and above the sample median.

Table SI-4: Covariate summary statistics

	covariate	mean	sd	min	max	% missing
1	Regional United Russia vote share 2007	0.59	0.07	0.49	0.83	0.15
2	Regional United Russia vote share 2011	0.45	0.10	0.29	0.78	0.15
3	Distance to Moscow (km)	913.20	1299.71	0.00	6798.83	0.15
4	Moscow or Saint Petersburg	0.42	0.49	0.00	1.00	0.15
5	General news site	0.54	0.50	0.00	1.00	0.00
6	Media bias: pro-free speech	0.12	0.33	0.00	1.00	0.00
7	Media bias: neutral	0.83	0.38	0.00	1.00	0.00
8	Media bias: anti-free speech	0.05	0.21	0.00	1.00	0.00
9	Putin approval	0.75	0.05	0.57	0.87	0.16
10	Medvedev approval	0.70	0.05	0.51	0.81	0.16
11	Incidents of repression against journalists	41.92	48.54	0.00	107.00	0.15

Table SI-5: Experimental estimates: Probit

	Positive reply		Negative reply		Technical request		Political request	
	coef.	se	coef.	se	coef.	se	coef.	se
Intercept	−0.56***	(0.11)	0.36***	(0.10)	−1.24***	(0.13)	−1.64***	(0.17)
Non-political + CA	−0.36**	(0.16)	0.37**	(0.15)	−0.33	(0.21)	0.01	(0.24)
Anti-regime	−0.36**	(0.16)	0.42***	(0.15)	−1.26***	(0.38)	0.28	(0.22)
Anti-regime + CA	−0.80***	(0.18)	0.85***	(0.17)	−1.00***	(0.30)	0.00	(0.24)
Anti-regime (+)	−0.71***	(0.17)	0.71***	(0.16)	−0.83***	(0.27)	0.37*	(0.22)
Anti-regime (+) + CA	−0.77***	(0.17)	0.79***	(0.16)	−1.00***	(0.30)	−0.22	(0.26)

Note: $N = 953$. The table shows probit estimates from four specifications. Standard errors are shown in parentheses. The first treatment combination, non-political advertisement without collective action frame, is the omitted category. * denotes statistical significance at 0.10 level. ** denotes statistical significance at 0.05 level. *** denotes statistical significance at 0.01 level.

Table SI-6: Experimental estimates: OLS

	Positive reply			Negative reply			Technical request			Political request		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Intercept	0.29***	0.21	−0.08	0.64***	0.33	0.61	0.11***	0.16*	0.15	0.05***	0.23	0.19
	(0.04)	(0.28)	(0.56)	(0.04)	(0.32)	(0.60)	(0.02)	(0.08)	(0.27)	(0.02)	(0.22)	(0.36)
Non-political + CA	−0.11**	−0.10*	−0.11**	0.13**	0.13**	0.13**	−0.05	−0.06	−0.05*	0.00	0.02	0.02
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
Anti-regime	−0.11**	−0.09*	−0.10**	0.14***	0.12*	0.13**	−0.10***	−0.10***	−0.11***	0.04	0.06**	0.06**
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Anti-regime + CA	−0.20***	−0.19***	−0.20***	0.25***	0.24***	0.24***	−0.09***	−0.10***	−0.10***	0.00	0.02	0.02
	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
Anti-regime (+)	−0.19***	−0.20***	−0.21***	0.22***	0.22***	0.24***	−0.09***	−0.10***	−0.10***	0.05*	0.07**	0.08**
	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Anti-regime (+) + CA	−0.20***	−0.19***	−0.20***	0.23***	0.23***	0.24***	−0.09***	−0.10***	−0.10***	−0.02	0.00	0.00
	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
Covariates	no	yes	yes	no	yes	yes	no	yes	yes	no	yes	yes
Block FE	no	no	yes	no	no	yes	no	no	yes	no	no	yes
N	953	801	801	953	801	801	953	801	801	953	801	801

Note: The table shows OLS estimates from 12 specifications, three for each of four outcome variables. For each outcome variable, the first specification only contains treatment indicators, the second specification adds the covariates in Table 2, and the third specification also adds block dummies. Coefficient estimates for covariates and block dummies have been omitted. Robust standard errors are shown in parentheses. The first treatment combination, non-political advertisement without collective action frame, is the omitted category, so estimates and statistical tests are relative to this treatment combination.
* denotes statistical significance at 0.10 level. ** denotes statistical significance at 0.05 level. *** denotes statistical significance at 0.01 level.