Electoral Effects of Biased Media: Russian Television in Ukraine

Online Appendix

1. SENTIMENT OF THE COVERAGE OF UKRAINE ON RUSSIAN NEWS

In the paper, we have illustrated how he volume of coverage of Ukraine on Russian television increased substantially following the Maidan protests in late 2013. In this appendix, we report how the tone of this coverage shifted substantially in this period to being considerably more negative. To investigate this claim we undertake sentiment analysis of all news transcripts from 2010 to 2015 and then check whether news reports on Ukraine carried significantly more negative sentiment relative to news on other topics.

For sentiment measurement, we used Russian sentiment lexicon *RuSentiLex 2017*, available at http://www.labinform.ru/pub/rusentilex/index.htm. This lexicon specifies three types of sentiments: fact sentiment (mentions of good/bad events, like acts of terrorism or war), opinion sentiment (captured by words such as 'absurd,' 'nonsense' and so on) and feeling sentiment (captured by words such as 'nervous,' 'pleasant,' etc.). The three types of sentiments are highly correlated in that if a report has many words reflecting negative factual sentiment, then it also has many words representing a negative feeling and opinion sentiment. We use all three types of sentiments by summing negative and positive words across the three sentiment categories. Doing this analysis for each type of sentiment separately yields very similar results as these sentiment types are highly correlated.

We first lemmatize words in the news reports using Python based lemmatizer *pymystem3* developed by Yandex. We then calculate the number of positive-sentiment and negative-sentiment words for each news report in the database in the following fashion: let w_i^+ and w_i^-, denote the number of positive and negative words in a news report i, respectively. We then calculate the overall sentiment of the news report i using the formula

$$s_i = \frac{w_i^+}{w_i^+ + w_i^-}.$$

In cases where $w_i^+ = 0$ and $w_i^- = 0$, we set $s_i = 1/2$, to avoid division by zero (this essentially means that the sentiment is neutral when no positive or negative words are used). The measure s_i is equal to one if the news report is overwhelmingly positive and is equal to zero if the news report is overwhelmingly negative. The report has a neutral sentiment as s_i approaches 0.5.

Analyzing raw values of the sentiment measure s_i makes little sense without the context in which these news are reported – it could be that news on Ukraine are very negative but that all other news reports are negative too. To circumvent this problem we use a measure of *relative sentiment* that is constructed as follows: (1) for each day t, we first calculate the median sentiment of all news that *do not* mention Ukraine, (2) then we calculate the difference between the sentiment of news report s_i (that does mention Ukraine) and the median sentiment $\hat{s}_t$ of all news that do no mention Ukraine on the same day as report i:

$$\triangle_{i,t} = s_{i,t} - \hat{s}_t.$$

Thus, measure $\triangle_{i,t}$ tells us the positivity of the sentiment of news about Ukraine relative to the baseline sentiment on that same day. $\triangle_{i,t}$ ranges from -1 to 1: e.g., when it is equal to -1, the news report on Ukraine is most negative relative to other news.

Note that news reports on topics other than Ukraine include many topics that Russian state-controlled media tends to cover negatively (e.g., foreign/Western governments). Thus, our measure of relative sentiment likely underestimates the degree of negativity and overestimates the positivity of news on Ukraine. Similarly, news reports about Ukraine that are scored as positive often cover events that positively depict the actions of the Russian government or citizens in response to Ukrainian events that are colored in negative tones by implication. Unfortunately, it is not clear how one could measure such implied sentiment. However, it is clear that such potential bias in the measurement would once again lead to underestimation of the negativity of the coverage of Ukraine.

After the relative sentiment scores are calculated for each news report i appearing on day t we fit the following semi-parametric regression model:[1]

$$\triangle_{i,t} = \alpha_0 + \alpha_1 \text{Post-Maidan} + \alpha_2 \text{Ukraine} \times \text{pre-Maidan} \\ + \alpha_3 \text{Ukraine} \times \text{post-Maidan} + s(t) + \epsilon_{it}.$$

The variable *Post-Maidan* is an indicator equal to one if the time period is after Nov 21, 2013 (the first day of Maidan protests) and equal to zero otherwise. Thus, the coefficient α_1 measures the average change in the tone of news reports in the post-Maidan period relative to the pre-Maidan period. The coefficient α_2 measures the difference between the sentiment of news on Ukraine relative to other news in the pre-Maidan period, whereas α_3 is the same measure for the post-Maidan period. Finally, $s(t)$ is a smooth function of time capturing (possibly non-linear) time trends. We approximate this function using regression splines and estimate the model in the GAM framework (Wood, 2006).

Results are reported in Table 1.1. In column 1, we show the estimated parameters of the model where we use an indicator equal to one if a news report mentions the word 'Ukraine' or its derivatives to measure whether the news report is about Ukraine or something else. Prior to the Maidan protests, news reports that mentioned Ukraine were similar in terms of the average sentiment to news reports on other topics as indicated by a tiny and statistically insignificant coefficient. However, following the Maidan protests, news reports about Ukraine had significantly more negative sentiment than news on other topics even after adjusting for time-trends. In substantive terms, reports mentioning Ukraine in the post-Maidan period were more negative by about 0.23 of the standard deviation. Similar message is born out in column 2, where we use a continues measure for the coverage of Ukraine – the number of times that the word 'Ukraine' or its derivatives are

[1] Note that this is a re-parameterized standard interactive regression model of the form

$$\triangle_{i,t} = \beta_0 + \beta_1 \text{Post-Maidan} + \beta_2 \text{Ukraine} + \beta_3 \text{Ukraine} \times \text{Post-Maidan} + s(t) + \epsilon_{it}.$$

Thus, $\alpha_2 = \beta_2$ and $\alpha_3 = \beta_2 + \beta_3$. The specification we use allows us to directly interpret the coefficient estimates as the quantities of interest, α_2 and α_3.

	(1)	(2)
Post-Maidan	0.003	-0.010
	(0.010)	(0.010)
Ukraine mentioned (indicator) pre-Maidan	-0.019	
	(0.017)	
Ukraine mentioned (indicator) post-Maidan	-0.247^{***}	
	(0.011)	
Times Ukraine mentioned (log) pre-Maidan		-0.0001
		(0.013)
Times Ukraine mentioned (log) post-Maidan		-0.140^{***}
		(0.007)
Observations	141,844	141,844
Adjusted R^2	0.005	0.004
Log Likelihood	$-200{,}928.600$	$-200{,}957.500$
UBRE	0.995	0.996
Note:		$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table 1.1: Regression results. The dependent variable is the normalized sentiment score. Both specifications control for time trends using smoothing splines.

mentioned in the report (on the log scale). Again, we see that there is no association between the intensity with which Ukraine was mentioned in news reports and their sentiment in the pre-Maidan period. Following the Maidan protests, however, news reports that mentioned Ukraine more frequently were significantly more likely to be negative in tone.

2. A MODEL OF HETEROGENOUS INFORMATION UPDATING

The theoretical literature on biased media (e.g., Besley and Prat, 2006; Gehlbach and Sonin, 2014) or, more generally, on rational persuasion (e.g., Kamenica and Gentzkow, 2011; Alonso and Camara, 2016) is very rich and it would be difficult to summarize it in this format. Generally, though, in these models the sender provides the receiver with partially biased information, and as long as that bias is not too strong the receiver can be persuaded to change her beliefs. An important component in these models is that the receiver can back out or observe the degree to which the sender's message is biased. Below we explore how a rational receiver updates information by adding another component: suppose that the receiver does not know the degree of to which the message she receives is biased. The receiver will then update on two quantities – the state of the world (as in standard models of persuasion) and the nature of the media source that is transmitting the message (i.e. whether the source is biased or not). As we demonstrate below, if the receiver has to update on both of these quantities simultaneously then the same message can generate divergent effects on two receivers with different priors. In a recent working paper, Fryer, Harms and Jackson (2016) arrive at a similar result using a different modeling approach.

Let $\theta \in R$ represent the unknown state variable (e.g., how incompetent the Ukrainian government is). The receiver does not know the value of this state variable but has a prior probability distribution over the θ given by $\theta \sim \mathcal{N}(\mu, \sigma^2)$. A media source provides the receiver with the information about the state variable θ. The media source can be either truthful or biased, but the receiver does not know before observing the message whether the source is biased or not. Let s be an indicator variable denoting whether the source is truthful or biased, but the consumer assigns probability π that the source is biased and probability $1 - \pi$ that the source is truthful.

The receiver observes a message $y \in \mathbb{R}$ that is generated by the following stochastic process:

$$y = s\theta + (1-s)(-\theta) + \epsilon,$$

where ϵ is a standard normal random variable representing the noise in the news. We see that when media is truthful ($s = 1$) its message y is positively correlated with the true state θ since then $y = \theta + \epsilon$, but when media is biased its message represents a 'lie' since it is negatively correlated with the true state θ as we have $y = -\theta + \epsilon$.

We are interested in how the agent's updated posterior belief $E(\theta|y)$ depends on the observed signal y conditional on his prior expectation $\mu = \mathbb{E}(\theta)$. The proposition below makes two claims: first, the updating is weaker when the receiver has a strong prior (μ very large negative or very large positive), and second, when two receivers observe a sufficiently strong signal (a large value of y in absolute terms) they update in different directions if their priors are sufficiently different.

Proposition 1 (Heterogenous Updating). *For any $\pi > 0$, there exists a cut-off $\hat{y}$ such that if $y > \hat{y}$ then the posterior is increasing in y for μ sufficiently large and decreasing in y for μ sufficiently small.*

Proof. After having observed the message y the receiver updates his prior beliefs about the expected value of the state variable y given by:

$$\mathbb{E}(\theta|y) = \mathbb{E}(\theta|y, s=0)\Pr(s=0|y) + \mathbb{E}(\theta|y, s=1)\Pr(s=1|y).$$

The standard Bayesian updating procedure (Gelman et al., 2003) yields

$$\mathbb{E}(\theta|y, s=0) = \frac{\mu - \sigma^2 y}{\sigma^2 + 1},$$
$$\mathbb{E}(\theta|y, s=1) = \frac{\mu + \sigma^2 y}{\sigma^2 + 1}.$$

The posterior probability that the message is arriving from a truthful media source given that the value of signal is given by

$$\Pr(s=1|y) = \frac{f(y|s=1)(1-\pi)}{f(y|s=1)(1-\pi)+f(y|s=0)(1-\pi)}$$
$$= \frac{1}{1+\frac{\int_\theta \phi(y+\theta)\phi(\frac{\theta-\mu}{\sigma})\,d\theta}{\int_\theta \phi(y-\theta)\phi(\frac{\theta-\mu}{\sigma})\,d\theta}\frac{\pi}{1-\pi}}.$$

Since the convolution of the normal random variables is the normal random variable, the marginal density $f(y|s=1) = \int_\theta \phi(y-\theta)\phi(\frac{\theta-\mu}{\sigma})\,d\theta$ must also be normal. By the law of iterated expectation we have $\mathbb{E}(y|s=1) = \mathbb{E}(E(y|\theta)) = \mu$. By the law of iterated variance we have

$$Var(y|s=1) = \mathbb{E}(Var(y|\theta)) + Var(\mathbb{E}(y|\theta)) = 1 + Var(\theta) = 1+\sigma^2.$$

Similarly, we have $\mathbb{E}(y|s=0) = -\mu$ and $Var(y|s=0) = 1+\sigma^2$. Putting this all together, we can write:

$$\Pr(s=1|y) = \left(1+\frac{\phi\left(\frac{y+\mu}{\sqrt{1+\sigma^2}}\right)}{\phi\left(\frac{y-\mu}{\sqrt{1+\sigma^2}}\right)}\frac{\pi}{1-\pi}\right)^{-1},$$
$$= \left(1+\exp\left\{\frac{-2\mu y}{1+\sigma^2}\right\}\frac{\pi}{1-\pi}\right)^{-1}.$$

By inspection, if $\mu > 0$, then $\Pr(s=1|y)$ is increasing in y for all y, and it is decreasing otherwise. Differentiating the posterior expectation of θ with respect to signal y we get

$$\frac{\partial}{\partial y}\mathbb{E}(\theta|y) \propto \Pr(s=1|y) + y\frac{\partial}{\partial y}\Pr(s=1|y).$$

When $\mu > 0$ the above expression is positive for all $y > 0$. For $\mu < 0$, since the first term is decreasing in y with the zero limit, and the second term is negative and increasing in y, there is a value of y such that the expression is negative. $\square$

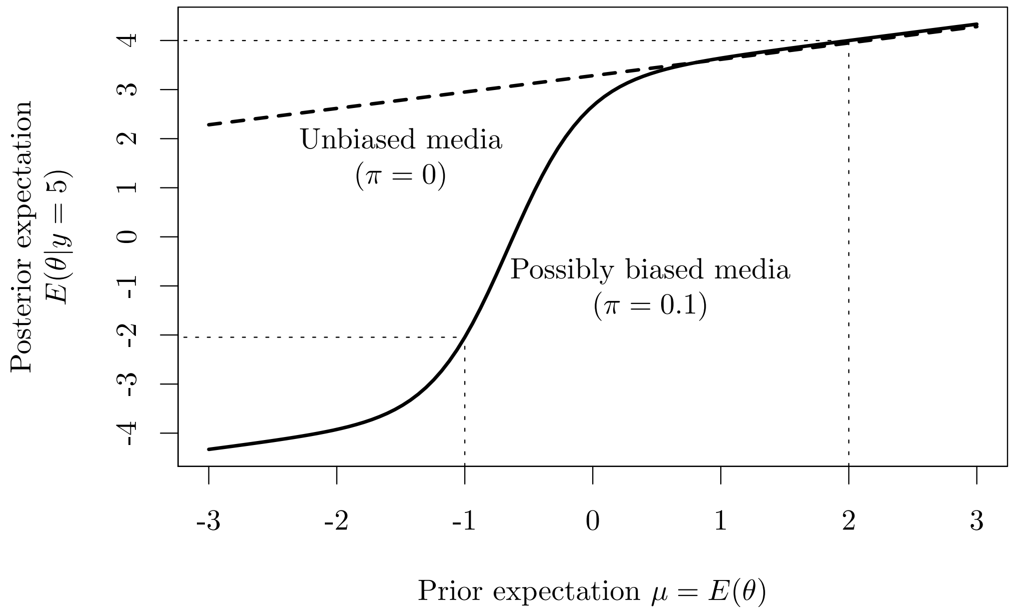

Figure 2.1: Belief-updating as a function of prior expectation of the state variable given a high positive signal y of the state variable θ.

Figure 2.1 provides some additional intuition on the mechanics of updating. The horizontal axis represents the receiver's priors, whereas the vertical axis represents the updated belief about the value of θ after observing a large positive signal ($y = 5$). When receivers know that a media source is unbiased ($\pi = 0$) they update strongly in the direction of the signal (in this case, in the positive direction), and their posteriors do not show a lot of heterogeneity. In fact, any heterogeneity fades away as signal strength increases when $\pi = 0$.

However, when $\pi > 0$ so that receivers expect a small chance that the media source is biased against their interests, posterior beliefs exhibit very strong heterogeneity. If a receiver expected that the value of $\theta = 2$, then his updated expectation is $E(\theta|y = 5) \approx 4$, thus he updates in the direction of the signal. However, if the receiver has a prior that strongly contradicts the content of the message – he beliefs a priori that $\mathbb{E}(\theta) = -1$ – then his updated belief is $E(\theta|y = 5) \approx -2$; thus, he is updating in the direction opposite to the content of the message.

The main intuition behind these results is that whenever a receiver observes a message that strongly contradicts his prior beliefs, he not only revises his prior belief about the state variable θ but also his belief as to whether he is facing a media source that is telling the truth or lying. In a population with highly diverse priors (variable values of μ in this case), observing the same message y can lead to highly polarized posterior, especially if the message is so strong (the value of y is high) that it raises that the concern that it can be biased.

The analysis indicates that the following conditions are necessary to generate heterogeneous updating in the population exposed to the same message: (1) the population has sufficiently divergent priors μ about the true state of the world, (2) the messages generated by the media source should not be subtle but rather tending toward judgmental and extreme (y should take large, in absolute terms, values). Interestingly, there are no strong conditions for how biased the media source is expected to be – as long as that probability is not zero ($\pi > 0$), it is possible to generate a divergent effect of media influence, as long as priors are sufficiently divergent and the message is sufficiently extreme. Given our discussion of ex ante polarization in Ukrainian politics driven by ethnolinguistic cleavages and highly loaded nature of Russian reporting on Ukraine we believe that the two necessary conditions hold in the context of our study.

Party name	*Classification*	*No. registered candidates*	*% national vote*
Party of Regions/ Партія регіонів	Pro-Russian	221	30.00
All-Ukrainian Union "Fatherland"/ Всеукраїнське об'єднання "Батьківщина"	Pro-Western	203	25.55
UDAR of Vitaliy Klychko/ "УДАР Віталія Кличка"	Pro-Western	208	13.96
Communist Party of Ukraine/ Комуністична партія України	Pro-Russian	214	13.18
All-Ukrainian Union Svoboda/ Всеукраїнське об'єднання "Свобода"	Pro-Western	217	10.44
Below 5% party-list threshold for entry into parliament:			
Ukraine Forward! Of Natalia Korolevska/ Партія Наталії Королевської "Україна – Вперед!"	Pro-Western	149	1.58
"Our Ukraine"/"Наша Україна"	Pro-Western	185	1.11
Radical Party of Oleh Liashko/ Радикальна Партія Олега Ляшка	Pro-Western	139	1.08
Pensioners' Party of Ukraine/ Партія Пенсіонерів України	Pro-Russian	29	0.56
Socialist Party of Ukraine/ Соціалістична партія України	Pro-Russian	155	0.45
Party of Greens of Ukraine/ Партія Зелених України	Pro-Western	78	0.34
Ukrainian Party "Green Planet"/ Українська партія "Зелена планета"	Pro-Western	225	0.34
"Russian Bloc"/"Руський блок"	Pro-Russian	34	0.31
Greens/Політична партія "Зелені"	Pro-Russian	56	0.25
Ukraine of the Future/ Політична партія "Україна Майбутнього"	Pro-Western	30	0.18
Political association "Native Fatherland"/ Політичне об'єднання "Рідна Вітчизна"	?	106	0.16
People's Labor Union of Ukraine/ "Народно-трудовий союз України"	Pro-Russian	17	0.11
"New Politics"/"Нова Політика"	?	69	0.10
All-Ukrainian Association "Community"/ Всеукраїнське об'єднання "Громада"	Pro-Western	41	0.08
Ukrainian National Assembly/ Українська Національна Асамблея	Pro-Western	114	0.08
Liberal Party of Ukraine/ Ліберальна партія України	Pro-Western	55	0.07

Figure 3.1: The 2012 parliamentary election.

Table A2: Classification of political parties in the 2014 parliamentary election

Party name	*Classification*	*No. registered candidates*	*% national vote*
People's Front/"Народний Фронт"	Pro-Western	219	22.14
Petro Poroshenko Bloc/ "Блок Петра Порошенка"	Pro-Western	193	21.82
Self-Reliance Union/ "Об'єднання "Самопоміч"	Pro-Western	60	10.97
Opposition Bloc/ "Опозиційний блок"	Pro-Russian	194	9.43
Radical Party of Oleh Liashko/ Радикальна Партія Олега Ляшка	Pro-Western	215	7.44
All-Ukrainian Union "Fatherland"/ Всеукраїнське об'єднання "Батьківщина"	Pro-Western	212	5.68
Below 5% party-list threshold for entry into parliament:			
All-Ukrainian Union Svoboda/ Всеукраїнське об'єднання "Свобода"	Pro-Western	206	4.71
Communist Party of Ukraine/ Комуністична партія України	Pro-Russian	204	3.88
Serhiy Tihipko's "Strong Ukraine"/ Партія Сергія Тігіпка "Сильна Україна"	Pro-Russian	201	3.11
Anatoliy Hrytsenko's "Civic Position"/ Партія "Громадянська позиція (Анатолій Гриценко)"	Pro-Western	146	3.10
All-Ukrainian Agrarian Union "Spade"/ "Всеукраїнське Аграрне Об'єднання "Заступ"	Pro-Western	183	2.65
"Right Sector"/"Правий Сектор"	Pro-Western	32	1.80
Solidarity of the Women of Ukraine/ Партія "Солідарність жінок України"	Pro-Russian	61	0.66
Party "5.10"/Політична Партія "5.10"	?	173	0.42
Internet Party of Ukraine/ "Інтернет партія України"	?	17	0.36
Party of Greens of Ukraine/ Партія Зелених України	Pro-Western	52	0.25
Ukrainian Party "Green Planet"/ Українська партія "Зелена планета"	Pro-Western	98	0.23
Revival Party/Партія "Відродження"	?	89	0.19
"One Country"/ "Єдина Країна"	Pro-Western	27	0.17
All-Ukrainian Union "Ukraine-One Country"/Всеукраїнське Політичне Об'єднання "Україна – Єдина Країна"	Pro-Western	92	0.12
"New Politics"/"Нова Політика"	?	36	0.12
Політична партія "Сила Людей"	Pro-Western	37	0.11
Ukraine of the Future/ Політична партія "Україна Майбутнього"	Pro-Western	51	0.08
"Strength and Honor"/"Сила і Честь"	Pro-Western	72	0.08
Ukrainian Civil Movement/ Громадянський рух України	Pro-Western	34	0.08
Bloc of Left Forces of Ukraine/ "Блок Лівих Сил України"	Pro-Western	109	0.07

Figure 3.2: The 2014 presidential election.

Candidate name:	*Classification*	*% national vote*
Petro Poroshenko/Порошенко Петро Олексійович	Pro-Western	54.70
Yulia Tymoshenko/Тимошенко Юлія Володимирівна	Pro-Western	12.81
Oleh Lyashko/Ляшко Олег Валерійович	Pro-Western	8.32
Anatoliy Hrytsenko/Гриценко Анатолій Степанович	Pro-Western	5.48
Serhiy Tihipko/Тігіпко Сергій Леонідович	Pro-Russian	5.23
Mykhailo Dobkin/Добкін Михайло Маркович	Pro-Russian	3.03
Vadim Rabinovich/Рабінович Вадим Зіновійович	?	2.25
Olha Bohomolets/Богомолець Ольга Вадимівна	Pro-Western	1.91
Petro Symonenko/Симоненко Петро Миколайович	Pro-Russian	1.51
Oleh Tyahnybok/Тягнибок Олег Ярославович	Pro-Western	1.16
Dmytro Yarosh/Ярош Дмитро Анатолійович	Pro-Western	0.70
Andriy Hrynenko/Гриненко Андрій Валерійович	?	0.40
Valeriy Konovalyuk/Коновалюк Валерій Ілліч	Pro-Russian	0.38
Yuriy Boyko/Бойко Юрій Анатолійович	Pro-Russian	0.19
Mykola Malomuzh/Маломуж Микола Григорович	Pro-Russian	0.13
Renat Kuzmin/Кузьмін Ренат Равелійович	Pro-Russian	0.10
Vasyl Kuybida/Куйбіда Василь Степанович	Pro-Western	0.06
Oleksandr Klymenko/Клименко Олександр Іванович	Pro-Western	0.05
Vasyl Tsushko/Цушко Василь Петрович	Pro-Russian	0.05
Volodymyr Saranov/Саранов Володимир Георгійович	?	0.03
Zorian Shkiryak/Шкіряк Зорян Несторович	Pro-Western	0.02

Figure 3.3: The 2014 parliamentary election.

For 2006, 2007, and 2010 elections, which we use for balance tests, we code the Party of Regions and the Communist party as pro-Russian parties, and use their combined vote-shares as our measure of pro-Russian voting. We were not able to credibly code smaller political parties for these earlier elections, because scarce availability of their election manifestos and public statements.

4. MEASURING SIGNAL QUALITY AND TV RECEPTION

Let π_i denote the probability that a respondent i reports receiving Russian television.[2] Let $\boldsymbol{s}_i = \{s_{i,1}, ..., s_{i,T}\}$ denote the strength of signals at location i as predicted by the Irregular Terrain Model from $t = 1, ..., T$ Russian transmitters and relay stations. Let $s_i^{(k)}$ denote field strength of the kth strongest signal at location i, and let K denote the number of highest-quality signals to be averaged over. We then fit the following probit regression model:

$$\Phi^{-1}(\pi_i) = \lambda\left(S_i\right), \text{where } S_i = \frac{1}{K}\sum_{k=1}^{K} s_i^{(k)}, \tag{1}$$

where Φ is the standard normal distribution function and λ is an unknown continuous function, which allows television reception to very non-linearly with signal strength. This probit regression is estimated by approximating the function λ with penalized thin plate regression splines (Wood, 2003) in the generalized additive modeling framework (Wood, 2006). For comparison, Enikolopov, Petrova and Zhuravskaya (2011) use specification $\pi = \Phi(\alpha_0 + \alpha_1 s_i^{(1)})$, which is a special case of our measurement approach that assumes $K = 1$ and λ is a linear function.

The optimal number of signals is then identified by identifying the number K that yields the most optimal classification of actual Russian television availability as reported in the survey. Since variation in Russian television signal quality can only impact those viewers who watch television via analog antennae, to construct these measurements we exclude respondents who watch television through cable and satellite. The optimal K is found through the following steps

1. Fix K and estimate the model in equation 1;

[2] In the survey, we did not ask 'Do you receive Russian television?' To mitigate against social desirability bias we listed various television channels (some of them Russian) and asked whether respondents receive those channels and whether they watch them.

2. Calculate the area under Receiver Operating Characteristic (ROC) curve as a measure of how accurately the model can identify self-reported Russian television reception;

3. Repeat steps 1-2 for $K = 1, ..., T$, and select K^* that yields the largest area under the ROC curve.

Figure 4.1 shows the predictive accuracy of the model for various values of $K = 1, ..., 10$. We see that the optimal value of $K = 4$ in our data because this is the point at which the predictive accuracy of the model is maximized. This means that we can predict signal reception better by averaging over four strongest signals at a given location rather than by using maximum signal value there. Intuitively, even if a single signal is strong it might not be reliable, and thus viewers at a given location might not form a habit of watching the channel. This approach yields a more accurate measure of television reception at no extra data collection cost.

Having identified the optimal value of K (K^*), we then calculate (raw) signal strength at location i ($Signal_i$). To get a sense of how well signal strength predicts actual television reception and viewership, Figure 4.2 shows the *cumulative proportion* of precincts with respondents reporting to receive or watch Russian television as a function of signal strength. We see that starting at about 30 dBmuV's, respondents increasingly report receiving and watching Russian television. This threshold around which respondents begin to report receiving Russian television is roughly the threshold for analog TV availability suggested by the United States Federal Communications Commission, which ranges from 11 to 45 dBmuV's, (FCC, 2002). Note that the overall cumulative proportion of precincts that receive Russian television is about 0.4 (note that this proportion includes all precincts below and at the maximum signal strength). Since about 40 percent of respondents in the sample have analog antennae, the measure picks up actual reception quite accurately.

Using the measure of raw signal strength ($Signal_i$), we calculate the probability of receiving Russian television at location i ($Reception_i$). This probability is simply a fitted

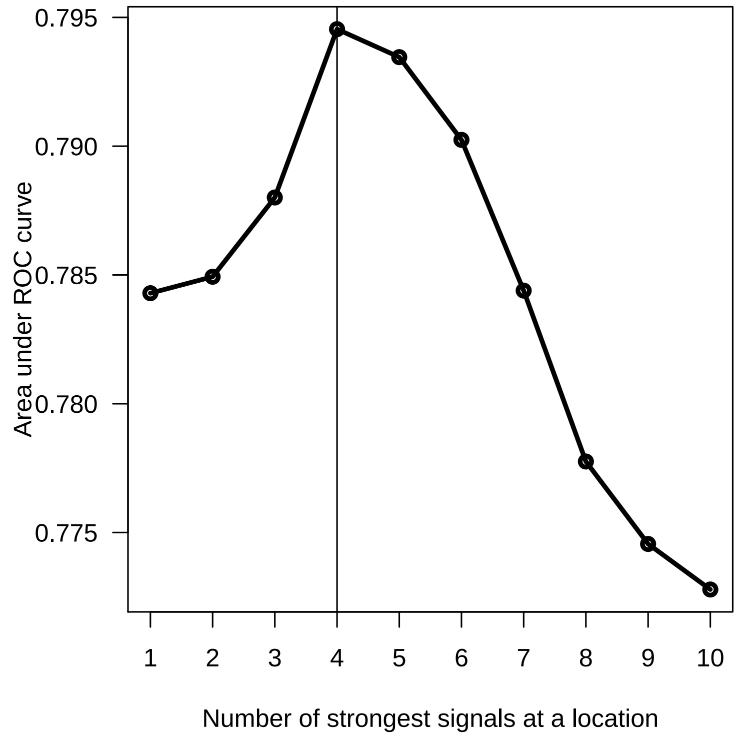

Figure 4.1: This figure shows how accurately we can predict self-reported reception of Russian television as a function of the number of strongest signals that we average over.

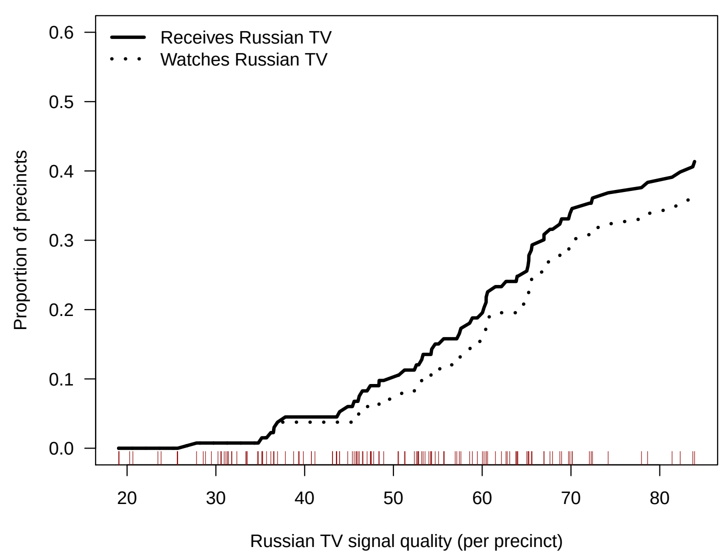

Figure 4.2: This figure shows the *cumulative proportion* of precincts with respondents reporting to receive or watch Russian television as a function of signal strength.

value from the probit regression, $\Phi(\hat{\lambda}(Signal_i))$, where $\hat{\lambda}$ is the estimate of λ. Figure 4.3 shows how raw values of signal strength ($Signal_i$) map onto the probability of receiving Russian television ($Reception_i$). The non-linear nature of this relationship makes sense – signal strength has small impact on reception at low values of the signal and a larger impact at high values of the signal.

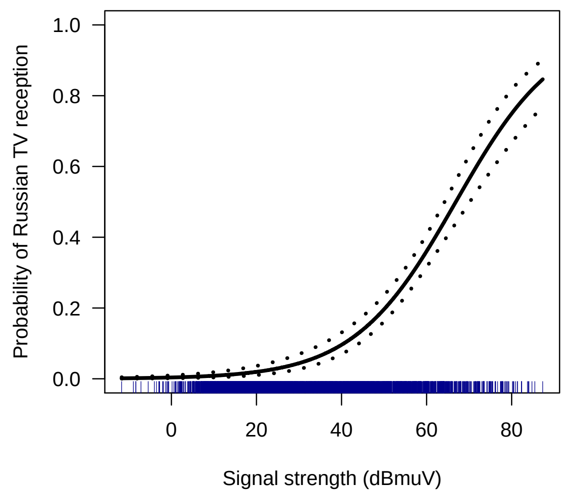

Figure 4.3: Russian television reception (predicted probability of receiving Russian television with 95 percent confidence bounds) as a function of signal strength.

Figure 4.4 shows how well our *Reception* measure predicts reception of each of the four major Russian channels which we asked about in the survey. For the two most popular channels the predicted probability is increasing steeply in step with signal quality. Furthermore, the area under the receiver operator curve (ROC) reported in each figure also shows reception quality estimated by the ITM method, and the model predicts the availability of Russian TV channels quite well. The lower panel of Figure 4.4 shows that reception also predicts well the probability that a respondent *watches* Russian television. We should note that the reception measure predicts best the availability and the propensity to watch Channel 1 and Rossiya 1 – Russia's leading news channels and ones most important for the purposes of this study. The area under the ROC curve for these two channels is 0.81-0.82, indicating that though imperfect, the reception measure is a good predictor of self-reported TV reception and consumption.

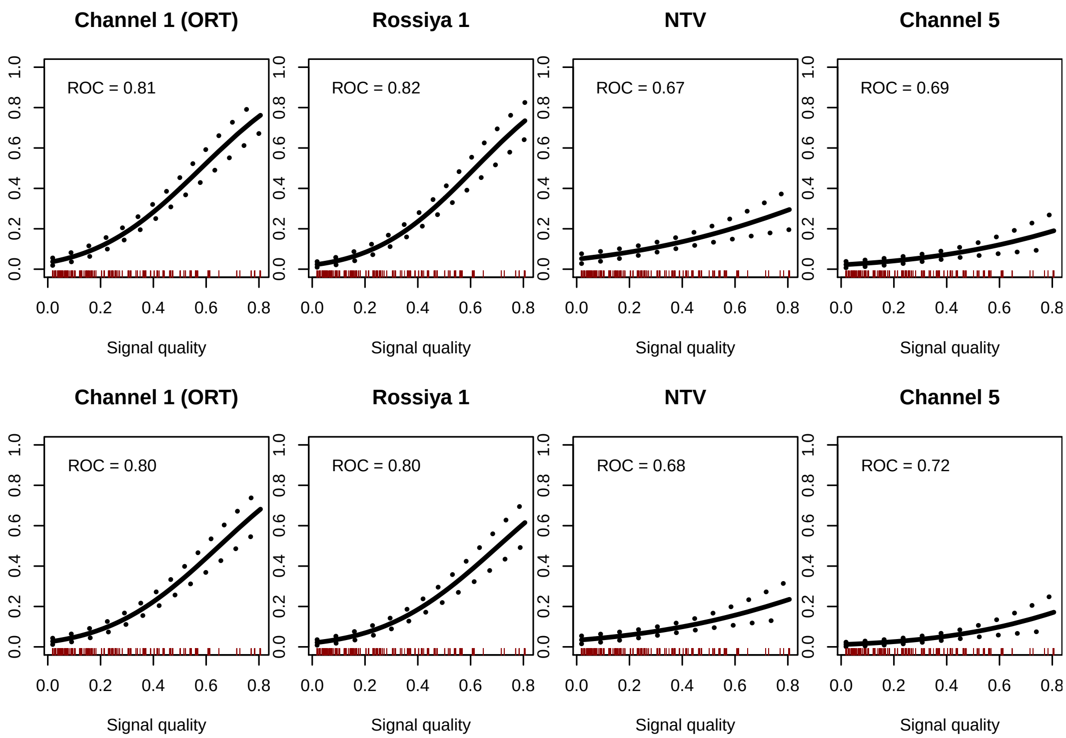

Figure 4.4: Predicted probabilities of receiving (upper panel) and watching (lower panel) Russian TV channels with 95 percent confidence intervals.

5. SURVEY

The survey was fielded in 160 randomly selected precincts located within 50km (31mi) of the Ukrainian-Russian border. We sampled a set of precincts from the overall precinct population and then randomly selected respondents from within each precinct. Since our goal was not to estimate unbiased univariate population parameters but rather to explore the causal relationship between television signal quality and electoral behavior, the sampling scheme was designed to insure high within-sample variation in the quality of Russian television signal. We first sorted all precincts into five equal bands/quintiles corresponding to the distribution of Russian television signal quality and then randomly sampled precincts from within each band. The number of respondents within each band was made proportionate to the number of precincts in it. To make sure that we have enough respondents who actually receive Russian analog television, we oversampled precincts in the band where Russian television reception was very good (above 0.7 on our reception scale). Precincts with fewer than 200 registered voters were excluded because they correspond to very small settlements and are therefore difficult to reach due to poor road conditions. Precincts that made it into the sample are marked in Figure 5.1.

We randomly selected streets within the precinct from which households were sampled wherever Ukraine's Central Election Commission (CEC) provided specific street addresses in its description of the precincts (precincts in cities, towns, and large villages). Otherwise, in cases where a precinct encompassed an entire settlement (i.e. small and medium-sized villages), interviewers were instructed to pick a street at random on their own initiative. Interviewers were instructed to interview five respondents per street (six in villages). The CEC distinguishes between small, medium, and large precincts by number of registered voters. Six respondents were selected at random from small precincts (exclusively villages), 10 from medium-sized precincts, and 15 from large precincts (mostly cities). Interviewers were instructed to select an initial building at random on a given street. In villages where people live mostly in single-family stand-alone homes, interviewers would knock on every fifth door counting from the building that was approached initially. In cities where apartment buildings predominate, interviewers were asked to knock on every fifteenth apartment door. Once contact was made with a specific household, interviewers selected a respondent at random from among all the adults resident at that address following the nearest birthday method (the individual whose birthday is closest to the date of the interview was interviewed). Response rates were very high; over 80% across all settlement types.

The survey was in the field in January-March 2015, and it was implemented by Sotsioinform, a Lviv-based public opinion firm with a national interviewer network. The project was reviewed and approved by the Institutional Review Board at New York University-Abu Dhabi (protocol #123-2014).

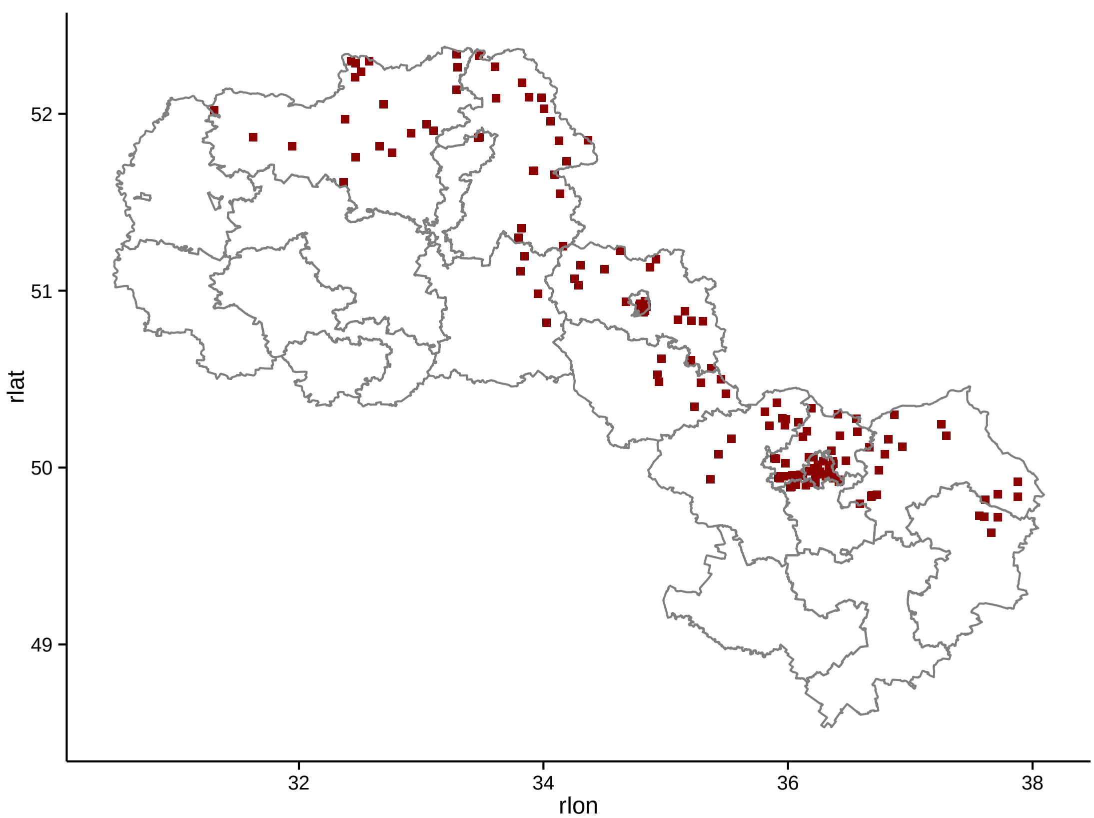

Figure 5.1: Locations of surveyed precincts.

The survey instrument was offered to respondents in Russian and Ukrainian. It contained 136 questions with blocks on television viewing patterns, political attitudes and behavior, attitudes toward the secessionist conflict in eastern Ukraine, and demographics. Below are English-language translations of survey questions (along with answer options) that we used in this paper. The survey instrument in its entirety is available on request.

1. Television viewing:

   (a) Some people have room and house antennae, others have cable TV, satellite, and others yet watch television on the internet. How do you watch television? (a) Room antenna, (b) House antenna, (c) Cable, (d) Satellite, (e) Via the internet. Which one of these methods of television reception do you use most commonly? (a) Room antenna, (b) House antenna, (c) Cable, (d) Satellite, (e) Via the internet.

   (b) I will now read out the names of several television channels. Please tell me whether you receive this channel and describe the quality of television reception. Russian channels: (a) First Channel (ORT), (b) Russia 1, (c) NTV, (d) 5th Channel (Russia).

2. Political behavior:

   (a) Did you vote in the last parliamentary election of 26 October 2014? (1) Yes, (0) No.

   (b) Which political party did you vote for? (a) People's Front, (b) Poroshenko Block, (c) Opposition Block, (d) Radical Party of Oleh Liashko, (e) Fatherland, (f) Svoboda, (g) Strong Ukraine, (h) Communist Party of Ukraine, (i) Against all/spoilt ballot.

   (c) Did you vote in the last presidential election in May 2014? (1) Yes, (0) No.

(d) Which candidate did you vote for? (a) Petro Poroshenko, (b) Yulia Tymoshenko, (c) Oleh Liashko, (d) Anatoliy Hrytsenko, (e) Serhiy Tihipko, (f) Mikhailo Dobkin, (g) Against all/spoilt ballot.

3. Demographic information:

   (a) What language do you speak in daily life/at home? (1) Only Russian, (2) Mostly Russian with a few Ukrainian words interspersed, (3) Equal measure Russian and Ukrainian, (4) Mostly Ukrainian with a few Russian words interspersed, (5) Exclusively Ukrainian.

   (b) How often do you generally travel to Russia? (5) One a week or more frequently, (4) Once a month or more frequently, (3) Once or several times every six months, (2) One or several times every twelve months, (3) Never.

   (c) What is your education level? (1) Incomplete primary, (2) Primary or incomplete secondary, (3) Secondary, (4) Specialized secondary, (5) Professional or technical diploma (polytechnic), (6) Incomplete higher, (7) Higher.

   (d) How would you describe your family's income level: is it low, average, or high? (1) Low, (2) Average, (3) High.

Variable	Mean	Min	Max	Obs.
Precinct- or settlement-level (Sources: Ukrainian Electoral Commission, ITU, 2001 census)				
% Pro-Russian votes (2014 parl.)	26.72	0	78.9	3,589
% Pro-Russian votes (2014 pres.)	22.48	0	75.71	3,589
% Pro-Russian votes (2012 parl.)	51.52	11.69	95.17	3,589
Russian TV signal ($dB\mu V$)	32.68	-11.7	87.31	3,589
Probability of Russian TV reception	0.11	0	0.85	3,589
Voting population	1091.25	40	2516	3,589
Distance to Russia (km)	62.21	0.13	180.26	3,589
Rural precinct	0.56	0	1	3,589
Road density	35.74	0	164.83	3,589
% Ukrainian speakers	88.07	1.92	100	1,717
Individual-level (Source: survey)				
Russian TV available (self-reported)	0.39	0	1	1,676
Watches Russian TV (entire sample)	0.31	0	1	1,676
Watches Russian TV (if available)	0.79	0	1	648
Uses Ukrainian language	1.58	0 (Never)	4 (Always)	1,663
Income category	1.37	1 (Low)	3 (High)	1,662
Education	2.2	1 (Primary)	3 (Higher)	1,674
Travel to Russia	1.17	1 (Never)	5 (Weekly)	1,614

Table 6.1: Summary statistics of the main variables.

7. ADDITIONAL BALANCE TESTS

In the table below, we report eleven additional balance tests. The first four variables are pro-Russian vote and turnout in the 2006 and 2007 parliamentary elections. We also check whether the sample is balanced with respect to several geographic variables: longitude, latitude, the interaction of longitude and latitude (to check whether geographic confounding might occur along the south-east to north-west axis, for example), and distance to the provincial capital. Finally, we also check balance with respect to population density in the precinct (measured by the number of registered voters in 2012), the precinct type as classified by the Ukrainian election commission (small, intermediate, large), and the number of settlements per precincts.

	County fixed effects			District fixed effects			
	Est.	S.E.	p-val.	Est.	S.E.	p-val.	Obs.
1. Pro-Russian vote, 2006	1.35	4.78	0.78	0.99	5.56	0.86	3,747
2. Pro-Russian vote, 2007	3.94	9.85	0.69	0.43	6.03	0.94	3,763
3. Turnout, 2007	−1.05	3.14	0.74	1.08	2.89	0.71	3,747
4. Turnout, 2006	−1.47	4.24	0.73	−0.45	2.83	0.87	3,763
5. Longitude	0.15	0.87	0.86	0.27	0.70	0.70	3,589
6. Latitude	−0.01	0.28	0.97	0.01	0.28	0.96	3,589
7. Longitude×Latitude	0.09	0.44	0.83	0.18	0.35	0.60	3,589
8. km to provincial capital (log)	0.55	0.78	0.48	0.39	0.29	0.18	3,589
9. Number of voters in 2012	−0.03	0.52	0.96	−0.14	0.22	0.51	3,589
10. Precinct type	0.09	0.51	0.86	−0.07	0.19	0.72	3,589
11. Settlements per precinct	0.16	0.46	0.74	0.42	0.26	0.10	3,589

Table 7.1: Balance tests. OLS coefficients for residualized Russian television reception. Standard errors clustered by county.

The results indicate that the balance on these eleven variables is good: the coefficient estimates are generally low, and p-values never even approach conventional levels of statistical significance. Importantly, consistent with the balance test in the paper, we do not find significant association between Russian television reception and pro-Russian vote in the 2006 and 2007 parliamentary elections. The coefficient for the 2007 election in the specification with county fixed effects is larger that in other elections and other specifications. However, the fact that it is estimated so imprecisely and that it drops in magnitude by a factor of about 10 when district fixed effects are used indicates that the size of the coefficient is likely to be a statistical abnormality. We note that only in one of the eight balance tests on previous election results that we conducted does the coefficient exceed 1.5 percent in magnitude, and the smallest p-value across these eight tests is equal to 0.69.

We also note that there is no association between turnout in the 2006 and 2007 elections and Russian television reception, as shown in rows 3 and 4 of the table. These coefficient estimates for turnout are smaller than the ones for 2010 and 2012 elections, though in all of these cases the p-values are too large to suggest that there might be any systematic relationship between Russian television reception and turnout in previous elections. Finally, we also observe very strong balance with respect to all the remaining variables reported in the table.

The table below reports the estimated impact of the availability of Russian television on the level of electoral turnout at precinct level in the two 2014 elections. The estimates are very small across all specifications and statistically not significant. We also do not find any evidence of heterogeneous effects.

	Presidential		Parliamentary	
	Baseline	Full	Baseline	Full
Russian TV reception	-0.89	2.03	-0.22	1.37
	(2.51)	(2.04)	(3.39)	(2.65)
Percent Ukrainian speakers		0.00		-0.03^{*}
		(0.01)		(0.01)
Pro-Russian vote in 2012		-0.17^{***}		-0.10^{***}
		(0.01)		(0.02)
Turnout in 2012		0.69^{***}		0.73^{***}
		(0.03)		(0.03)
Log(Number of Voters)		0.24		-0.05
		(0.38)		(0.39)
Rural precinct		-1.49^{***}		-2.43^{***}
		(0.29)		(0.30)
Road density		0.41^{*}		0.39^{*}
		(0.17)		(0.19)
Adjusted R^2	0.59	0.84	0.45	0.78
Observations	$3,589$	$3,567$	$3,589$	$3,567$

Standard errors (in parentheses) clustered by county; *p<0.05; **p<0.01; ***p<0.001.

Table 8.1: Precinct-level regression results. The dependent variables are turnout-percentages in each election. All specifications control for county fixed effects and smoothing splines for distance to Russia.

DellaVigna and Kaplan (2007), who proposed the idea of persuasion rates, use the formula

$$f = 100\frac{y_1 - y_0}{e_1 - e_0}\frac{1}{1 - y_0},$$

where y_0 and y_1 stand for the share of those who do not and do receive the media message, respectively, and $e_1 - e_0$ is the share of those who are exposed to the message. However, this formula only works for binary treatment and exposure. Enikolopov, Petrova and Zhuravskaya (2011) propose a continuous version of the formula, which is

$$f = 100\frac{1}{1 - v_0 t_0}\left(t\frac{dv}{de} + v\frac{dt}{de}\right),$$

where v_0 and t_0 is the predicted vote-share for pro-Russian parties and turnout, respectively, when reception probability is set to zero, t is turnout rate, dv/de is the rate of change in vote-shares as a function of the change in exposure, and dt/de the rate of change in turnout as a function of the change in exposure. In our estimations, we do not find statistically significant impact of Russian TV reception on turnout and thus we set $dt/de = 0$. Enikolopov, Petrova and Zhuravskaya (2011) suggest to calculate dv/de as a product of the regression coefficient and the inverse of the probability that a given voter watches Russian television when it is available. From our survey data, we estimate this probability to be 0.79. Finally, since turnout is not affected by TV reception in our data, we set $t_0 = t = \hat{t}$ - the average turnout rate in a given election. Thus, the final formula for persuasion rate is

$$f = 100\frac{1}{1 - v_0\hat{t}}\frac{\hat{t}\hat{\gamma}}{0.79},$$

where $\hat{\gamma}$ is the estimated regression coefficient for Russian TV availability. Following the usual practice (DellaVigna and Kaplan, 2007; DellaVigna and Gentzkow, 2010; Enikolopov, Petrova and Zhuravskaya, 2011) we calculate v_0 as the average predicted pro-Russian vote-share at zero probability that Russian TV is available.

10. MATCHING-BASED ESTIMATES

In this appendix, we estimate the effects of Russian television reception on election outcomes using matching methodology. Since the treatment *Russian TV reception* is not dichotomous but continuous we cannot implement the standard matching methodologies like propensity score matching, coarsened matching, or genetic matching. We use a recently developed methodology of covariate balancing generalized propensity score (CBGPS), which can be applied to continuous treatments (Fong, Hazlett and Imai, 2017). The basic idea behind this methodology is to find observation-specific weights that minimize the association between the treatment and the covariates. The estimated weights can then be used in the standard regression setting as variance weights so that the observations that increase disbalance are weighted down in the estimation of the parameters of interest. The method is implemented using the software in Fong et al. (2016).

We consider two different approaches. In Approach 1, the treatment assignment equation is specified as follows:

$$\text{Reception}_i = f(\text{Distance to Russia}_i) + \text{County}_{j[i]} + \boldsymbol{\beta}'\boldsymbol{x}_i + \epsilon_i, \tag{2}$$

This is the same equation as in the main regression models of the section *Biased Media and Mass Electoral Behavior*, except that the outcome variable is Russian TV reception. The vector $\boldsymbol{x}_i$ represents all of the control variables, including prior pro-Russian vote, turnout, urban/rural constituency dummy, number of voters, percent of Ukrainian speakers, and density of roads. This way, the propensity score weights induce balance on all the covariates including distance and county fixed effects.

In Approach 2, we amend the treatment assignment equation as follows:

$$\begin{aligned}\text{Reception}_i = & g(\text{prior pro-Russian vote}_i) + f(\text{Distance to Russia}_i) \\ & + \text{County}_{j[i]} + \boldsymbol{\beta}'\boldsymbol{x}_i + \epsilon_i,\end{aligned} \tag{3}$$

where g is an unknown smooth function, which we approximate by natural cubic splines with three knots (it does not make much difference how many knots we use). The difference here is that prior pro-Russian vote (in the 2012 election) enters the treatment assignment equation in a potentially non-linear fashion. As this pre-treatment covariate is arguably the most important for our analysis, we want to insure that we achieve balance not only with respect to a linear function of this variable (which is done in Approach 1) but also with respect to non-linear functions of this variable.

After calculating CBPS weights for each of the two approaches we fit weighted least squares regressions using the same specification as in the main regression models in Table 2 of the section *Biased Media and Mass Electoral Behavior*. The results are reported in Table 10.1 below. For brevity, the table only reports the quantities of interest – the coefficients and clustered standard errors of the *Russian TV reception* variable. We see that, irrespective of the matching approach used, Russian television reception has a positive and significant effect on pro-Russian voting in both 2014 elections. We note that the magnitude of the coefficients in the dataset weighted by CBPS matching is uniformly larger than the respective coefficients reported in Table 2 of the main text.

	Presidential		Parliamentary	
	Baseline	Full	Baseline	Full
Russian TV reception	13.82***	11.23***	13.73***	8.87***
(using approach 1)	(2.33)	(3.12)	(3.87)	(2.06)
Russian TV reception	12.03***	10.31***	12.04**	8.54***
(using approach 2)	(2.36)	(2.92)	(3.89)	(2.02)

Standard errors (in parentheses) clustered by county; *p<0.05; **p<0.01; ***p<0.001.

Table 10.1: The effects of Russian television reception on voting outcomes in regressions using CBPS weights. Dependent variables are vote-percentages for pro-Russian parties. All specifications control for county-level fixed effects, smoothing splines for distance to Russia; 'full' specifications include all of the covariates used in the main text.

11. IV ANALYSES: SUPPLEMENTARY RESULTS

11.1. *Full IV output*

	Vote pres.	Vote parl.	Maidan illegitimate	Trust Putin
	(1)	(2)	(3)	(4)
Mostly Ukrainian speaker	−0.07	−0.03	−0.10	-0.11^{**}
	(0.07)	(0.09)	(0.06)	(0.04)
Mixed speaker	-0.15^{*}	-0.17^{*}	−0.04	-0.08^{*}
	(0.07)	(0.08)	(0.05)	(0.04)
Mostly Russian speaker	-0.21^{**}	-0.20^{*}	−0.11	−0.07
	(0.07)	(0.08)	(0.06)	(0.05)
Exclusively Russian speaker	−0.15	−0.09	−0.05	−0.08
	(0.10)	(0.10)	(0.08)	(0.05)
Middle income	−0.07	−0.06	0.01	0.01
	(0.05)	(0.06)	(0.03)	(0.03)
High school education	−0.01	−0.09	0.06	0.09^{*}
	(0.06)	(0.08)	(0.06)	(0.05)
Higher education	−0.04	−0.02	−0.03	0.09
	(0.08)	(0.11)	(0.07)	(0.06)
Travels to Russia once a year	0.05	−0.004	−0.01	0.04
	(0.08)	(0.09)	(0.05)	(0.04)
... twice a year	−0.07	−0.01	0.05	0.04
	(0.11)	(0.15)	(0.09)	(0.10)
... every month	-0.35^{*}	-0.54^{*}	−0.17	-0.26^{***}
	(0.16)	(0.21)	(0.14)	(0.08)
... every week	−0.04	−0.09	-0.32^{***}	−0.04
	(0.04)	(0.05)	(0.03)	(0.04)
Watches Russian TV	0.26	0.46^{*}	0.43^{**}	0.30^{**}
	(0.16)	(0.22)	(0.13)	(0.11)
N	346	341	499	566
R^2	0.22	0.24	0.16	0.16
Adjusted R^2	0.12	0.15	0.09	0.10
Residual Std. Error	0.34 (df = 307)	0.40 (df = 302)	0.32 (df = 460)	0.25 (df = 527)

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table 11.1: Full output for IV regressions (specifications include county fixed effects). Standard errors are clustered by precinct.

11.2. *IV Regressions with Matching*

We now re-estimate our main IV regression augmenting it with matching. The idea here is to match the instrument *Russian TV reception* on the covariates to generate a more balanced sample so that the assumption of as-if random assignment of the instrument is better justified. Since the instrument we use is continuos, the standard matching techniques cannot be applied here. As in Appendix 10 we use the covariate balancing generalized propensity score (CBGPS) approach to identify propensity weights (Fong, Hazlett and Imai, 2017).

We first compute CBGPS weights by matching *Russian TV reception* on all of the covariates used in the first stage IV regressions. We do not use county indicators when calculating matching weights and we enter the covariates as linear terms not as factors, because otherwise we begin encountering sparsity problems in CBGPS weight calculation. Once the propensity weights are calculated, we replicate the IV analyses following the same steps as in the main analyses except that now we use weights in both stages of the TSLS estimation (the TSLS includes county fixed effects and conditions on covariates as factors).

Main outcomes	Estimate	S.E.	p-value	First stage F	Obs.
Vote pro-Russian (pres.)	0.21	0.14	0.13	17.31	346
Vote pro-Russian (parl.)	0.42	0.22	0.06	14.29	341
Post-Maidan government illegitimate	0.43	0.13	0.00	24.56	499
Trust Vladimir Putin	0.26	0.10	0.01	29.60	566
'Placebo' outcomes					
Favors state-owned property	0.13	0.08	0.09	36.21	598
Positive towards Lenin	0.07	0.10	0.48	25.29	575
Positive towards Stalin	-0.02	0.11	0.88	26.00	567

Table 11.2: Second stage IV coefficients for watching Russian TV news, using CBPS weights. All specifications include the covariates (levels for language, income, education, frequency of traveling to Russia) and county fixed effects. Standard errors clustered by precinct.

The table above reports second stage coefficients for *Watching Russian TV news*. The coefficients for the main outcomes are very similar in terms of magnitude and statistical significances. The results for the 'placebo outcomes' are also very similar to what is reported in the main body of the paper.

11.3. *Estimating Individual-Level Effects Using OLS*

The table below presents the OLS regressions for individual-level effects of Russian television consumption. The overall patterns are consistent with the IV regressions. We see significant effects of Russian news consumption on both voting and attitudes. As far as placebo outcomes are concerned, the effects are are substantially smaller, and statistically not significant in all but one case (positive attitudes towards Lenin), but the magnitude of the latter coefficient is very similar the IV estimate (which is equal to 0.07).

Main outcomes	Estimate	S.E.	p-value	Obs.
Vote pro-Russian (pres.)	0.11	0.05	0.05	346
Vote pro-Russian (parl.)	0.17	0.07	0.01	341
Post-Maidan government illegitimate	0.17	0.06	0.00	499
Trust Vladimir Putin	0.13	0.03	0.00	566
'Placebo' outcomes				
Favors state-owned property	0.03	0.02	0.27	598
Positive towards Lenin	0.08	0.04	0.03	575
Positive towards Stalin	0.05	0.04	0.14	567

Table 11.3: OLS coefficients for watching Russian news. All specifications include standard covariates and county fixed effects. Standard errors are clustered by precinct.

If people are self-selecting into watching Russian news, then we should expect the OLS estimates (which do not account for such self-selection) to be larger in magnitude than the IV estimates. However, this is the opposite to what we observe here. There are several explanations for this: First, the IV approach estimates the average local treatment effect (Angrist and Pischke, 2008); that is, the IV estimates represent the effect of watching Russian news in the subpopulation of respondents whose television watching habits are affected by the variation in the strength of the analog signal. The effects in this subpopulation could be substantially larger than the effects averaged across all strata of the population. Second, in addition to self-selection, which should push the OLS estimates upwards, there could be omitted variables that could cause the OLS estimates to attenuate. Finally, the magnitude of the IV estimates (relative to the OLS) could be inflated because the instrument is not valid. Although multiple placebo tests indicate that this is unlikely, we certainly cannot rule out this possibility.

11.4. *Controlling for Prior Individual Voting*

The table below presents the second-stage IV regression coefficients after controlling for how individuals voted in the first round of the 2010 presidential election. The vote is coded as 'pro-Russian' if the respondent voted either for Viktor Yanukovich or Serhiy Tihipko. Inclusion of the 2010 control reduces the sample size by quite a bit, which results in a more noisy estimate. However, the coefficients are within the margin of error from those reported in the paper. Moreover, the effects on placebo attitudes are statistically indistinguishable from zero as in the main results.

Main outcomes	Estimate	S.E.	p-value	First stage F	Obs.
Vote pro-Russian (pres.)	0.23	0.12	0.06	13.08	276
Vote pro-Russian (parl.)	0.32	0.18	0.07	11.22	269
Post-Maidan government illegitimate	0.52	0.15	0.00	16.35	307
Trust Vladimir Putin	0.39	0.14	0.01	20.69	340
'Placebo' outcomes					
Favors state-owned property	0.06	0.09	0.51	21.33	352
Positive towards Lenin	0.07	0.11	0.52	16.72	350
Positive towards Stalin	0.08	0.12	0.52	16.96	343

Table 11.4: Second stage IV coefficients for watching Russian TV news, after controlling for individual's pro-Russian vote in 2010 presidential election. All specifications include the covariates (levels for language, income, education, frequency of traveling to Russia) and county fixed effects. Standard errors clustered by precinct.

11.5. *Effects on the Intensive Margin*

We now consider the effect of Russian news consumption on the intensive margin. The treatment variable here is the frequency with which the respondent reports to watch news on the four main national Russian television channels. This measure is an additive index across four Likert scales. That is, for each channel, we asked how often on the scale from one (never) to five (every day) the viewer watches news on Russian channels. We then added these numbers across the four channels. To have results on an interpretable scale, we rescaled the treatment variable to range from zero to one, where zero represents the lowest category and 1 represents that maximum frequency in the sample.

Main outcomes	Estimate	S.E.	p-value	First stage F	Obs.
Vote pro-Russian (pres.)	0.77	0.39	0.05	10.77	346
Vote pro-Russian (parl.)	1.10	0.45	0.02	8.15	341
Post-Maidan government illegitimate	1.17	0.33	0.00	17.43	499
Trust Vladimir Putin	0.86	0.28	0.00	16.42	566
'Placebo' outcomes					
Favors state-owned property	0.26	0.21	0.21	22.34	598
Positive towards Lenin	0.18	0.32	0.57	16.65	575
Positive towards Stalin	-0.05	0.31	0.88	14.95	567

Table 11.5: Second stage IV coefficients for watching Russian TV news – intensive margin. The treatment variable is the frequency of watching Russian television news (across all channels) scaled to a unit interval. All specifications include the covariates (levels for language, income, education, frequency of traveling to Russia) and county fixed effects. Standard errors clustered by precincts.

The results, shown in Table 11.5, are very similar to and in some ways even stronger than the ones reported in the paper. Watching more Russian news makes respondents more likely to vote for pro-Russian parties and hold pro-Russian attitudes, but not for placebo attitudes, where all coefficients are at least three times smaller than for the main outcomes. Thus, the more intensive is consumption of Russian news, the more likely are people to hold pro-Russian attitudes and display pro-Russian behaviors.

While the two assumptions cannot be tested, our data allow us to indirectly asses their validity through the following placebo test. Among survey respondents, 40% watch television *exclusively* via analog, whereas 54% watch television *exclusively* via cable, satellite, or the internet (the six remaining percent have access to analog and non-analog television). The individuals who do not have analog effectively constitute the placebo group because their likelihood of watching Russian news should not be affected by the quality of Russian analog signal. Furthermore, if our identifying assumptions are valid, we should also expect to see no reduced form relationship between the variation in the quality of the Russian television signal and attitudes and behavior for this placebo group.

	With analog (N = 673)			Without analog (“placebo” group, N = 903)		
Dependent variable	Coef.	S.E.	p-value	Coef.	S.E.	p-value
Watches Russian TV	0.85	0.17	0.00	−0.18	0.32	0.56
Vote pro-Russian (pres.)	0.21	0.13	0.11	−0.07	0.15	0.66
Vote pro-Russian (parl.)	0.40	0.16	0.01	−0.03	0.16	0.84
Post-Maidan government illegitimate	0.37	0.13	0.00	0.01	0.18	0.96
Trust Vladimir Putin	0.26	0.08	0.00	0.08	0.15	0.61

Table 11.6: The table shows coefficients for Russian TV reception on respondents with and without analog antennae from reduced-form linear probability regressions with the covariates (as factors) and county fixed effects.

Table 11.6 shows the results of these placebo tests for five dependent variables: the propensity to watch news on Russian television and the four outcome variables (voting preferences in the two elections and the two attitudinal variables). Each of the five dependent variables were regressed using logistic model on signal strength, the covariates, and district-level fixed-effects. For those respondents who have analog antennae, better quality of Russian television signal increases their likelihood of watching Russian news *and* also increases their likelihood of voting for pro-Russian parties and holding pro-Russian attitudes. However, for the placebo group that does not have analog antennae, the propensity of having access to strong Russian analog signal is not associated either with the propensity to watch Russian news or with voting for pro-Russian parties or having pro-Russian attitudes. These placebo tests provide strong support for our identifying assumptions.

12. DOUBLE SOCIAL DESIRABILITY BIAS

There exists a possibility that the respondents, because of social desirability bias, may misrepresent both their attitudes/voting preferences *as well as* their propensity to watch Russian television. The double misreporting bias is likely to be highly asymmetric: those who watch Russian television and those who have pro-Russian attitudes are likely to say otherwise, but not vice versa. Here we investigate the consequences for our findings of the presence of a double social desirability bias.

Let $y_i^* \in \{0, 1\}$ denote the respondent's *true* i attitude/behavior with $y_i^* = 1$ representing pro-Russian attitude/behavior, which is more likely to be stigmatized. Let $x_i^* \in \{0, 1\}$ denote the variable measuring whether the respondent i actually watches Russian television ($x_i^* = 1$) or not ($x_i^* = 0$). Neither y_i^* nor x_i^* are observable directly: the respondent provides survey answers y_i and x_i which may or may not represent the truth. The survey answers are generated by the following measurement model:

$$\begin{aligned} \Pr(y_i = 1 | y_i^*) &= (1 - \epsilon_y) y^* \\ \Pr(x_i = 1 | x_i^*) &= (1 - \epsilon_x) x^* \end{aligned}$$

Hence, whenever the respondent has non-stigmatized attitudes/behavior, he reports the truth. Otherwise, he lies about not watching Russian television with the probability ϵ_x and not having pro-Russian attitudes with the probability ϵ_y. Let the a priori probability $\Pr(x_i^* = 1) = a$, so that a is an unknown fraction of respondents who watch Russian television (but may lie about it in the survey).

The desired estimand is the effect of *actually* watching Russian television on *actually* having pro-Russian attitudes:

$$\delta^* = \mathbb{E}(y^* | x^* = 1) - \mathbb{E}(y^* | x^* = 0),$$

to be estimated from the observed data (x_i, y_i), $i = 1, ..., n$. If we were using the observed data as if it was not subject to measurement error, we would estimate instead

$$\delta = \mathbb{E}(y|x=1) - \mathbb{E}(y|x=0).$$

We now show that measurement errors of the above kind attenuate the estimated causal effect, that is, $|\delta| < |\delta^*|$. Using the law of iterated expectations, we have

$$\begin{aligned}\mathbb{E}(y|x=1) =& \mathbb{E}(y|x=1, y^*=1)\Pr(y^*=1|x=1) \\ =& (1-\epsilon_y)\Pr(y^*=1|x=1, x^*=1)\Pr(x^*=1|x=1) \\ =& (1-\epsilon_y)\mathbb{E}(y^*|x^*=1)\end{aligned}$$

Using the law of iterated expectations and the Bayes rule we also have

$$\begin{aligned}\mathbb{E}(y|x=0) =& (1-\epsilon_y)\Pr(y^*=1|x=0) \\ =& (1-\epsilon_y)\left[\mathbb{E}(y^*|x^*=0)\Pr(x^*=0|x=0) + \mathbb{E}(y^*|x^*=1)\Pr(x^*=1|x=0)\right] \\ =& (1-\epsilon_y)\left[\mathbb{E}(y^*|x^*=0)\frac{1-a}{1-a+\epsilon_x a} + \mathbb{E}(y^*|x^*=1)\frac{\epsilon_x a}{1-a+\epsilon_x a}\right]\end{aligned}$$

Subtracting $E(y|x=0)$ from $E(y|x=1)$, yields, after some algebra,

$$\delta = \delta^* \frac{(1-a)(1-\epsilon_y)}{1-a+\epsilon_x a} \tag{4}$$

The fraction in the of the above expression following δ^* is strictly smaller than 1 for any a and any $\epsilon_x > 0$ and/or $\epsilon_y > 0$. Thus, the average effect estimated from the data contaminated by measurement error, δ, is strictly smaller (in absolute value) than the true average effect δ^*. This attenuation bias *increases* with the measurement errors ϵ_x and ϵ_y. Thus, if we believe that respondents in our survey underreported watching Russian television when it is available ($\epsilon_x > 0$) and/or voting for pro-Russian parties / having pro-Russian attitudes ($\epsilon_y > 0$) then the individual-level estimates reported in the paper constitute the lower bound on the true effects.

13. EFFECT-HETEROGENEITY: ADDITIONAL RESULTS

13.1. *Full Output of the Interactive Models*

The table below presents full output (except estimates of county effects and regression splines for distance to Russia) pertaining to Figure 3 in the paper. Note that only the interaction with pro-Russian vote in 2012 is consistently significant across the two models at conventional confidence levels.

	Presidential	Parliamentary
	(1)	(2)
Reception	-45.14^{*} (18.04)	-53.35^{*} (26.54)
Ukrainian speakers	-0.06^{***} (0.02)	-0.10^{***} (0.01)
Pro-Russian vote in 2012	0.37^{***} (0.05)	0.43^{***} (0.05)
Turnout in 2012	-0.02 (0.02)	-0.07^{**} (0.02)
Rural precinct	-0.62 (0.48)	-2.01^{**} (0.71)
Voting population	0.38 (0.35)	1.06^{*} (0.47)
Road density	-0.19 (0.25)	0.31 (0.26)
Reception x Ukrainian speakers	0.08 (0.05)	0.08 (0.05)
Reception x Pro-Russian vote in 2012	0.59^{***} (0.10)	0.45^{***} (0.09)
Reception x Turnout in 2012	-0.03 (0.14)	-0.04 (0.18)
Reception x Rural precinct	0.83 (3.03)	-0.98 (3.92)
Reception x Voting population	2.70 (1.65)	6.45^{*} (2.97)
Reception x Road density	-2.21 (1.77)	-3.92 (2.46)
N	3,567	3,567
R^2	0.92	0.92
Adjusted R^2	0.92	0.92
Residual Std. Error (df = 3485)	5.03	5.47

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table 13.1: Full output of the fully interactive regressions. Both specifications include splines for distance to Russia and county fixed effects. Standard errors are clustered by county.

The table below presents results of a simpler regression specification where *Reception* is interacted only with pro-Russian vote in 2012. Coefficient estimates remain very similar for the presidential election and somewhat smaller but qualitatively similar for the parliamentary election.

	Presidential (1)	Parliamentary (2)
Reception	-27.94^{***} (7.16)	-14.31^{***} (4.03)
Ukrainian speakers	0.38^{***} (0.05)	0.45^{***} (0.05)
Pro-Russian vote in 2012	-0.04^{***} (0.01)	-0.08^{***} (0.01)
Turnout in 2012	-0.03 (0.02)	-0.08^{**} (0.03)
Rural precinct	-0.55 (0.46)	-2.03^{***} (0.59)
Voting population	0.67^{*} (0.34)	1.75^{***} (0.39)
Road density	-0.37 (0.23)	0.01 (0.19)
Reception x Ukrainian speakers	0.54^{***} (0.11)	0.33^{***} (0.07)
N	3,567	3,567
R^2	0.92	0.92
Adjusted R^2	0.92	0.92
Residual Std. Error (df = 3490)	5.04	5.51

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table 13.2: Full output of the fully interactive regressions. Both specifications include splines for distance to Russia and county fixed effects. Standard errors are clustered by county.

The table below presents results of a more flexible interactive model, where we not only interact *Reception* with other covariates, but also *Pro-Russian vote in 2012* with all other covariates. The results remain very similar to those reported in Table **??**. As noted before, in the fully interactive specification, the coefficient for *Reception* does not have a meaningful interpretation, as it refers to the effect of reception when all other covariates are at zero value, which cannot happen in this sample.

	Presidential	Parliamentary
	(1)	(2)
Reception	-47.38^{*} (20.13)	-41.89 (29.32)
Ukrainian speakers	0.13^{*} (0.06)	0.03 (0.05)
Pro-Russian vote in 2012	0.76^{***} (0.18)	0.36^{*} (0.17)
Turnout in 2012	0.10 (0.05)	0.05 (0.05)
Rural precinct	-4.37^{***} (0.90)	-2.60^{*} (1.16)
Voting population	0.66 (0.71)	-1.04 (0.71)
Road density	-0.17 (0.54)	-0.38 (0.50)
Reception x Ukrainian speakers	0.12^{**} (0.05)	0.08 (0.05)
Reception x Pro-Russian vote in 2012	0.52^{***} (0.09)	0.41^{***} (0.10)
Reception x Turnout in 2012	0.06 (0.15)	0.04 (0.19)
Reception x Rural precinct	-2.81 (3.21)	-1.64 (4.15)
Reception x Voting population	2.67 (1.82)	4.09 (3.02)
Reception x Road density	-1.76 (1.96)	-3.53 (2.83)
Pro-Russian 2012 x Ukrainian speakers	-0.003^{**} (0.001)	-0.002^{*} (0.001)
Pro-Russian 2012 x Turnout in 2012	-0.002 (0.001)	-0.002 (0.001)
Pro-Russian 2012 x Rural precinct	0.07^{***} (0.02)	0.002 (0.03)
Pro-Russian 2012 x Voting population	-0.01 (0.02)	0.05^{**} (0.02)
Pro-Russian 2012 x Road density	0.001 (0.01)	0.02 (0.01)
N	3,567	3,567
R^2	0.92	0.92
Adjusted R^2	0.92	0.92
Residual Std. Error (df = 3480)	4.95	5.35

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table 13.3: Full output of the fully interactive regressions. Both specifications include splines for distance to Russia and county fixed effects. Standard errors are clustered by county.

The table below shows complete second-stage IV regression results (county fixed effects are not reported) relating to Table 5 in the paper.

	Vote pres. (1)	Vote parl. (2)	Maidan illegitimate (3)	Trust Putin (4)
Mostly Russian	0.03	0.02	−0.09	−0.03
	(0.08)	(0.12)	(0.06)	(0.05)
Mixed	0.01	−0.10	−0.03	0.04
	(0.10)	(0.13)	(0.07)	(0.05)
Mostly Ukrainian	−0.001	−0.10	−0.09	0.09
	(0.11)	(0.16)	(0.09)	(0.06)
Only Ukrainian	0.03	−0.01	−0.04	0.06
	(0.13)	(0.14)	(0.09)	(0.06)
Middle income	−0.07	−0.06	0.01	0.01
	(0.05)	(0.06)	(0.03)	(0.03)
Incomplete highschool	0.01	−0.08	0.06	0.10*
	(0.07)	(0.08)	(0.06)	(0.04)
Highschool	0.01	0.003	−0.03	0.10
	(0.08)	(0.10)	(0.07)	(0.06)
Goes to Russia once a year	0.08	0.004	−0.01	0.04
	(0.08)	(0.09)	(0.05)	(0.05)
...twice a year	−0.18	−0.08	0.04	−0.03
	(0.15)	(0.15)	(0.10)	(0.09)
... every month	−0.51***	−0.62**	−0.18	−0.26**
	(0.15)	(0.19)	(0.14)	(0.08)
... every week	−0.01	−0.08	−0.32***	−0.02
	(0.04)	(0.05)	(0.03)	(0.04)
Ukrainian x Watches Russian TV	0.73**	0.69*	0.47**	0.62***
	(0.25)	(0.29)	(0.17)	(0.16)
Watches Russian TV	−0.23*	−0.12	−0.02	−0.18**
	(0.11)	(0.15)	(0.08)	(0.06)
N	346	341	499	566
R^2	0.12	0.21	0.17	0.10
Adjusted R^2	0.01	0.11	0.10	0.03
Residual Std. Error	0.36 (df = 306)	0.41 (df = 301)	0.32 (df = 459)	0.26 (df = 526)

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

Table 13.4: IV regressions for heterogenous effects, including county fixed effects. Standard errors clustered by precinct.

The table below reports a more flexible specification where we interact the indicator for *Watching Russian TV* with language groups treated as factors. While this model is more flexible, we face a problem of sparsity because some cells might become very small and the resulting estimates become very noisy. Nonetheless, we see sufficient evidence consistent with our linear specification in the paper: the coefficient for *Watching* is large and positive for respondents who do not speak Ukrainian frequently. Moreover, in three specifications, the coefficient for *Only Ukrainian* group is negative, though it is estimated very imprecisely.

	Vote pres.	Vote parl.	Maidan illegitimate	Trust Putin
	(1)	(2)	(3)	(4)
Mostly Russian	0.02 (0.14)	0.06 (0.16)	−0.04 (0.08)	−0.02 (0.06)
Mixed	−0.08 (0.12)	−0.17 (0.13)	0.001 (0.07)	0.06 (0.06)
Mostly Ukrainian	−0.11 (0.13)	−0.11 (0.15)	−0.09 (0.08)	0.13 (0.07)
Only Ukrainian	0.11 (0.16)	0.01 (0.15)	0.02 (0.10)	0.003 (0.08)
Middle income	−0.09 (0.06)	−0.05 (0.07)	0.01 (0.04)	0.01 (0.03)
Incomplete highschool	0.03 (0.08)	−0.06 (0.08)	0.06 (0.06)	0.09^{*} (0.04)
Highschool	0.01 (0.10)	0.01 (0.11)	−0.03 (0.07)	0.10 (0.05)
Visits Russia once a year	0.07 (0.08)	−0.01 (0.08)	−0.01 (0.05)	0.04 (0.04)
...twice a year	−0.18 (0.20)	−0.03 (0.20)	0.02 (0.10)	−0.04 (0.08)
... every month	−0.32 (0.49)	−0.43 (0.50)	−0.10 (0.22)	−0.28 (0.16)
... every week	−0.03 (0.42)	−0.05 (0.45)	−0.33 (0.34)	−0.02 (0.29)
Only Russian x Watching	0.55^{*} (0.28)	0.60 (0.33)	0.52^{***} (0.14)	0.67^{***} (0.13)
Mostly Russian x Watching	0.21 (0.35)	0.33 (0.30)	0.30 (0.21)	0.51^{**} (0.17)
Mixed x Watching	0.34 (0.29)	0.71^{*} (0.32)	0.36^{*} (0.17)	0.22 (0.14)
Mostly Ukrainian x Watching	0.17 (0.22)	0.17 (0.26)	0.50^{**} (0.19)	−0.01 (0.15)
Only Ukrainian x Watching	−2.40 (1.45)	−1.95 (2.52)	−0.13 (0.79)	0.89 (0.70)
Constant	−0.01 (0.18)	0.17 (0.20)	0.04 (0.12)	−0.16 (0.10)
N	346	341	499	566
R^2	0.01	0.17	0.20	−0.002
Adjusted R^2	−0.13	0.05	0.13	−0.08
Residual Std. Error	0.39 (df = 303)	0.42 (df = 298)	0.31 (df = 456)	0.27 (df = 523)

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table 13.5: IV regressions with coefficients estimated separately for each language group.

13.6. *Individual-level Heterogeneity: Fully Interactive Specification*

We now fit an IV model where the indicator for *Watches Russian TV* is interacted with each "background" covariate – education, income, frequency of travel to Russia, and usage of Ukrainian. To avoid sparsity we use each covariate as a linear term rather than as a factor. Since our main results do not depend on whether we use each covariate as a linear term or a factor we do not expect this to cause major biases here as well. The results are largely consistent with our estimates reported in the paper: the estimate for the interactive term is large and (in two specifications, as in the paper) significant, and this holds *only* for the *Ukrainian usage* variable. Note that the linear term *Watches Russian TV* does not have a clear interpretation in this fully interactive model, and it should not be compared to our estimates in the paper (its magnitude and significance also cannot be interpreted directly from this output).

	Vote pres.	Vote parl.	Maidan illegitimate	Trust Putin
	(1)	(2)	(3)	(4)
Watches Russian TV	1.21	−0.01	0.12	0.52
	(1.34)	(1.26)	(0.51)	(0.36)
Ukrainian usage	−0.01	−0.01	−0.01	0.02
	(0.03)	(0.03)	(0.02)	(0.02)
Income	−0.07	−0.03	0.01	0.01
	(0.18)	(0.17)	(0.08)	(0.06)
Education	0.05	−0.05	−0.05	0.02
	(0.09)	(0.11)	(0.05)	(0.04)
Travel to Russia	0.02	−0.04	−0.07	−0.005
	(0.08)	(0.09)	(0.06)	(0.04)
Ukrainian usage x Watches Russian TV	−0.20*	−0.14	−0.04	−0.17**
	(0.10)	(0.14)	(0.07)	(0.06)
Income x Watches Russian TV	0.01	−0.10	−0.01	−0.02
	(0.54)	(0.52)	(0.25)	(0.19)
Education x Watches Russian TV	−0.30	0.48	0.14	0.09
	(0.51)	(0.66)	(0.21)	(0.14)
Travel to Russia x Watches Russian TV	−0.05	−0.03	0.09	−0.03
	(0.13)	(0.15)	(0.10)	(0.08)
Constant	−0.02	0.18	0.26	−0.14
	(0.31)	(0.30)	(0.16)	(0.13)
N	346	341	499	566
R^2	0.16	0.13	0.15	0.05
Adjusted R^2	0.06	0.03	0.09	−0.01
Residual Std. Error	0.35 (df = 310)	0.43 (df = 305)	0.32 (df = 463)	0.26 (df = 530)

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

Table 13.6: IV regressions, fully interactive specification.

To estimate the heterogeneous effect of Russian television reception at precinct-level in the main text of the paper we used an interactive model where signal strength is interacted with all of the covariates in the model. Here, we implement a more flexible analysis using the kernel regularized least squares (KRLS) approach (Hainmueller and Hazlett, 2014). The KRLS method fits a highly flexible regression model in which each independent variable is allowed to have a non-linear and interactive effect on the outcome; thus, we are not required to make assumptions about which variables should enter the model in a linear fashion or which ones should be interacted.

The results of KRLS analyses are summarized in Figure 13.1. The distribution of marginal effects of Russian television reception on vote percentages for pro-Russian parties in the two 2014 elections are reported in the figure's upper panel. These marginal effects are estimated individually for each precinct. The histograms indicate that there was a significant degree of variation in how Russian television reception impacted election outcomes. In the majority of precincts (represented in blue), marginal effects are positive and, in some cases, as large as 20% or higher. However, in a significant proportion of precincts (represented in red), marginal effects are negative and, in some cases, as large as -10%. More precisely, in 27% and 16% of precincts in presidential and parliamentary elections respectively, good Russian television reception is associated with *negative* support for pro-Russian parties.

We explore the source of that heterogeneity in the lower panel of Figure 13.1. There we plot the relationship between the percentage of votes cast for pro-Russian parties in 2012 in a given precinct (our measure of pro-Russian priors) and the estimated marginal effect of the availability of Russian television on pro-Russian vote in 2014. These estimates are shown as a smoothed scatterplot, where darker pixels represent higher density points. While the presence of heterogeneity is quite clear from the scatterplots alone, for ease of

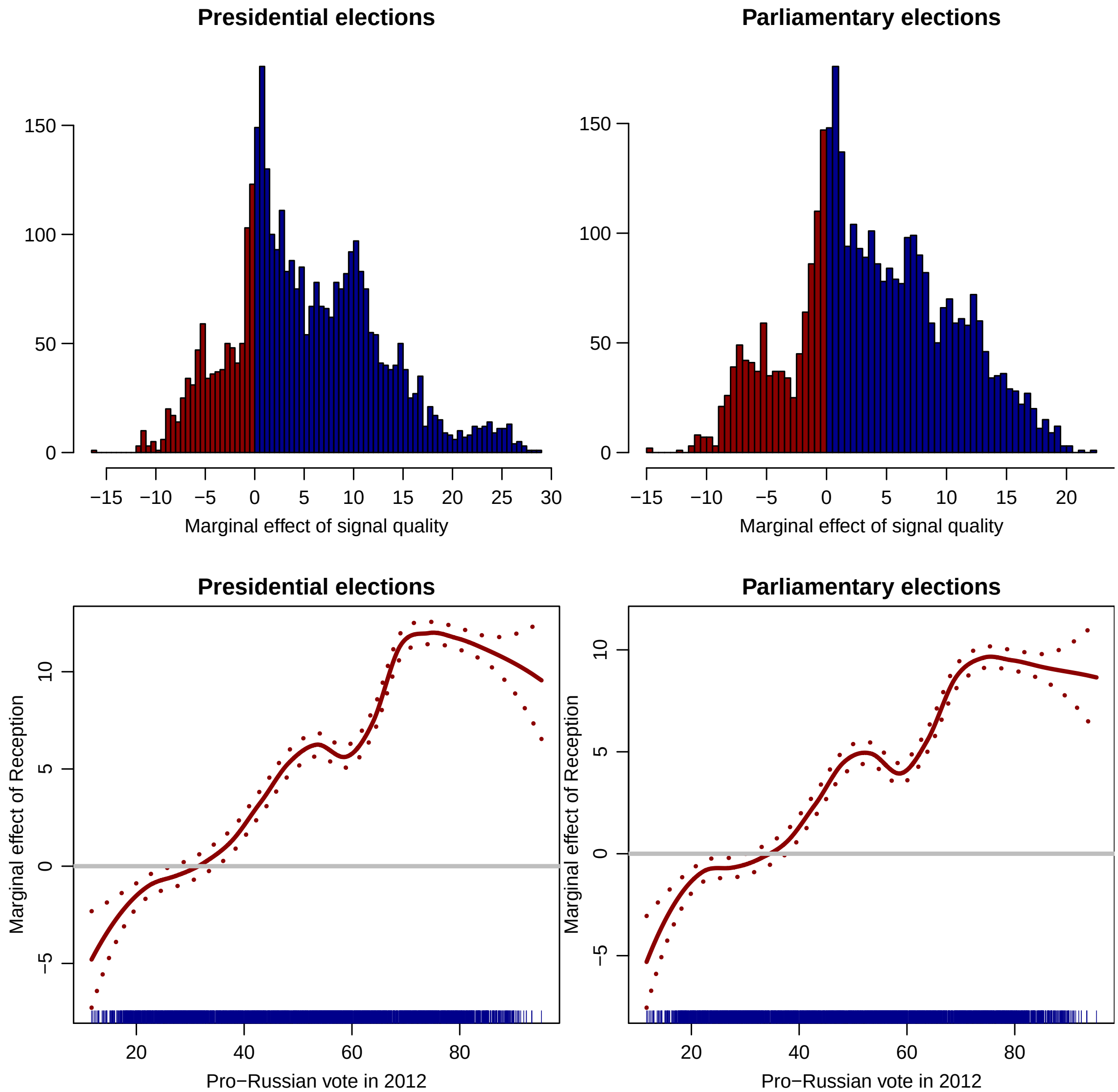

Figure 13.1: Heterogeneous impact of the availability of Russian analog signal on precinct-level electoral results in 2014 (marginal effects from KRLS regressions). Upper panel: distribution of estimated marginal effects (histogram). Lower panel: relationship between pro-Russian vote in 2012 and marginal effects of signal quality (smoothed scatterplot with a non-parametric regression curve and 95% point-wise confidence bounds.)

visualization we also add a non-parametric local regression curve to the plots (Loader, 1999), which indicates how the effect of Russian television signal changes as a function of pro-Russian support in 2012.

In both elections Russian television reception had the largest impact in those precincts that voted overwhelmingly for pro-Russian parties in 2012. In most extreme cases –

precincts where pro-Russian parties received more than 80% of the vote in 2012 – the presence of Russian television signal increased the vote share for pro-Russian parties in 2014 on average by 11% in the presidential contest and by 12% in the parliamentary election. The size of these effects decreases quite steeply as we move to historically less pro-Russian precincts. In precincts where pro-Russian parties received about 40% of the vote in 2012, the effect of Russian television availability in 2014 is statistically indistinguishable from zero. Finally, in historically pro-Western precincts – those where pro-Russian parties received less than 25-30% of the vote in 2012 – the availability of Russian television signal had a negative effect on electoral support for pro-Russian parties in 2014. All in all, the availability of Russian television has substantially different effects on different communities depending on their priors. Russian television is most persuasive in those communities where there are already many voters who are inclined to accept its message. On the other hand, in communities where there are many voters with strongly pro-Western preferences, we observe small but meaningful dissuasive effects of Russian television availability.

In Figure 13.2, we explore how the effect-heterogeneity varies with respect to other covariates. We see some evidence of heterogeneity with respect to the use of the Ukrainian language in the sense that the effect is larger in places where fewer respondents self-identified as Ukrainian speakers in the 2001 population census. However, we do not see the marginal effect changing the sign. The evidence of heterogeneity is much weaker, or even nonexistent, when it comes to economic modernization as measured by road density and population size. The heterogeneity with respect to distance to Russia is highly non-monotonic and quite difficult to interpret. All in all, it seems like the starkest heterogeneity is with respect to voting in the 2012 election.

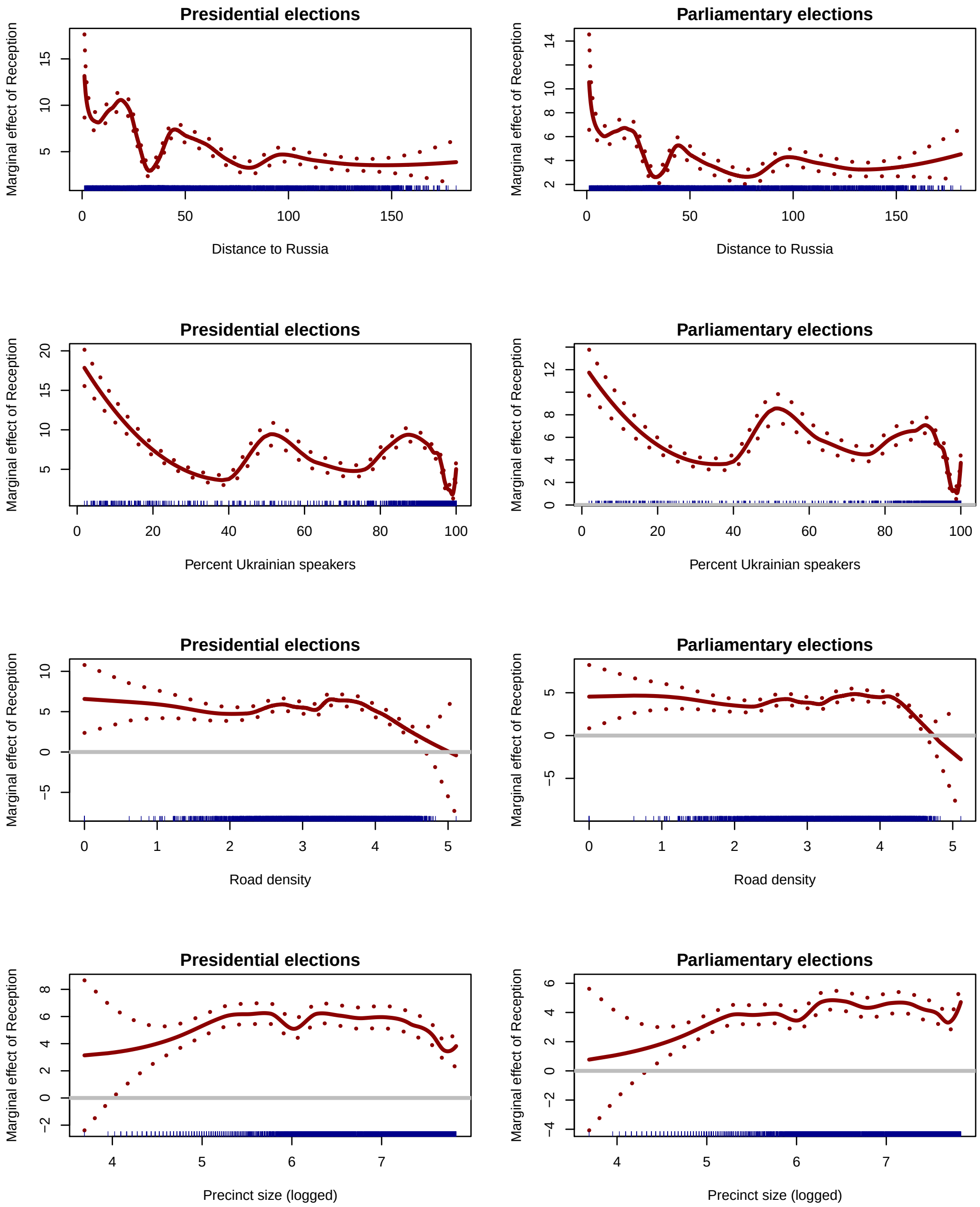

Figure 13.2: Conditional heterogeneity of the Russian television reception effect.

REFERENCES

Alonso, Ricardo and Odilon Camara. 2016. "Bayesian persuasion with heterogeneous priors." *Journal of Economic Theory* 165:672–706.

Angrist, Joshua D and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Besley, Timothy and Andrea Prat. 2006. "Handcuffs for the grabbing hand? Media capture and government accountability." *The American Economic Review* 96(3):720–736.

DellaVigna, Stefano and Ethan Daniel Kaplan. 2007. "The Fox News Effect: Media Bias and Voting." *The Quarterly Journal of Economics* 122(3):1187–1234.

DellaVigna, Stefano and Matthew Gentzkow. 2010. "Persuasion: Empirical Evidence." *Annual Review of Economics* 2(1):643–669.

Enikolopov, Ruben, Maria Petrova and Ekaterina Zhuravskaya. 2011. "Media and political persuasion: Evidence from Russia." *The American Economic Review* 101(7):3253–3285.

FCC. 2002. *The ILLR Computer Prorgam*. Federal Communications Commission.

Fong, Christian, Chad Hazlett and Kosuke Imai. 2017. "Covariate Balancing Propensity Score for a Continuous Treatment: Application to the Ecacy of Political Advertisements." *Unpublished manuscript* .
URL: *https://imai.princeton.edu/research/files/CBGPS.pdf*

Fong, Christian, Marc Ratkovic, Chad Hazlett, Xiaolin Yang and Kosuke Imai. 2016. *CBPS: Covariate Balancing Propensity Score*. R package version 0.13.
URL: *https://CRAN.R-project.org/package=CBPS*

Fryer, Roland, Philipp Harms and Matthew Jackson. 2016. "Updating Beliefs with Ambiguous Evidence: Implications for Polarization." *NBER Working paper* .
URL: *http://www.nber.org/papers/w19114*

Gehlbach, Scott and Konstantin Sonin. 2014. "Government control of the media." *Journal of Public Economics* 118(October):163–171.

Gelman, Andrew, John B. Carlin, Hal B. Stern and Donald B. Rubin. 2003. *Bayesian Data Analysis*. 2 ed. Chapman and Hall.

Hainmueller, Jens and Chad Hazlett. 2014. "Kernel Regularized Least Squares: Reducing Misspecification Bias with a Flexible and Interpretable Machine Learning Approach." *Political Analysis* 22(2):143–168.

Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian persuasion." *American Economic Review* 101(6):2590–2615.

Loader, Clive. 1999. *Local regression and likelihood*. Vol. 47 springer New York.

Wood, Simon. 2006. *Generalized additive models: an introduction with R*. Boca Raton: CRC press.

Wood, Simon N. 2003. "Thin plate regression splines." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65(1):95–114.