

THE UNIVERSITY OF CHICAGO

DEMAND FOR ONLINE NEWS UNDER GOVERNMENT CONTROL

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE UNIVERSITY OF CHICAGO
BOOTH SCHOOL OF BUSINESS
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY
ANDREY SIMONOV

CHICAGO, ILLINOIS

JUNE 2017

Copyright © 2017 by Andrey Simonov
All Rights Reserved

To my parents, Irina and Dmitry, who have encouraged me for this long journey, I would not be here without you. Love you with all my heart.

“Maybe there is a beast... maybe it’s only us.” – William Golding, *Lord of the Flies*

TABLE OF CONTENTS

LIST OF FIGURES	vii
LIST OF TABLES	ix
ACKNOWLEDGMENTS	x
ABSTRACT	xii
1 INTRODUCTION	1
2 DATA	7
2.1 Online News Market in Russia	7
2.2 Supply Data	9
2.3 Demand Data	10
2.3.1 IE Toolbar Representativeness	10
3 GOVERNMENT CONTROL AND SENSITIVE NEWS	13
3.1 Types of Government Control	13
3.2 Internal-Sensitive News	14
3.3 Government-Sensitive News about the Ukraine Crisis	19
4 DESCRIPTIVE EVIDENCE	25
5 MODEL	31
5.1 Nature	31
5.2 News Outlets	31
5.3 Demand	33
5.4 Government Control	36
6 ESTIMATION AND RESULTS	37
6.1 Demand Estimation	37
6.2 Estimation Results	38
7 COUNTERFACTUALS	49
7.1 Effectiveness of Government Control	53
8 NEWS CONSUMPTION FOR INFORMATION	55
9 CONCLUSION	61
REFERENCES	65
APPENDIX A NEWS OUTLETS CLASSIFICATION	69

APPENDIX B	COMPARISON OF IE TOOLBAR AND LIVEINTERNET.COM DATA	73
APPENDIX C	ATTRITION OF USERS IN IE TOOLBAR	77
APPENDIX D	CENSORED PROPER NOUNS	78
APPENDIX E	SLANT IN THE INTERNAL-SENSITIVE NEWS	80
APPENDIX F	CENSORSHIP IN UKRAINE-CRISIS NEWS	82
APPENDIX G	SLANT IN UKRAINE-CRISIS NEWS	83
APPENDIX H	SUMMARY OF BROWSING BEHAVIOR	84
APPENDIX I	DESCRIPTIVE EVIDENCE FIGURES	85
APPENDIX J	BASIC SUPPLY MODEL	87
APPENDIX K	ESTIMATION RESULTS: DISTRIBUTIONS OF MATCHING VAL- UES α_{IJ}	89
APPENDIX L	SIMULATED MARKET SHARE, LARGE VOLUME OF SENSITIVE NEWS	90
APPENDIX M	PREFERENCES OF NEWS JUNKIES	92

LIST OF FIGURES

2.1	Normalized number of weekly visitors to rbc.ru, IE Toolbar and LI data	11
3.1	Reporting about internal-sensitive news by the Russian news outlets	18
3.2	Levels of reporting and slant in internal-sensitive news, by news outlets	19
3.3	Share of articles containing the word “Ukraine” in the weekly coverage of news outlets, by types	20
3.4	Amount of the pro-Russia slant in the reporting about the Ukraine Crisis	22
3.5	Amount of the pro-Ukraine slant in the reporting about the Ukraine crisis	23
3.6	Level of the pro-Russia and pro-Ukraine slant in the Ukraine-crisis news coverage by each news outlet	24
4.1	Predicted changes in the news outlets’ market shares with the change in the amount of sensitive news, by news outlet	27
4.2	Histograms of number of consumers and number of consumptions by types of consumers	28
4.3	Predicted changes in the news outlets’ market shares from the infrequent news consumers with the change in the amount of sensitive news, by news outlet	29
6.1	Posterior distributions of preferences for internal-sensitive news and corresponding slant	39
6.2	Posterior distribution of preference for the Ukraine-crisis news and corresponding slant	42
6.3	Joint distribution of posterior means of preferences for pro-Russia and pro-Ukraine slant	43
6.4	Posterior distributions of preferences for internal-sensitive news and corresponding slant, by frequency of consumption: News junkies and occasionals	46
6.5	Posterior distributions of preferences for Ukraine-crisis news and corresponding slant, by frequency of consumption: News junkies and occasionals	47
8.1	Average number of websites consumers visit on a day when they consume news	55
8.2	Posterior distribution of consumers’ preference for ideologically diverse news (ρ_i^x)	59
8.3	Individual-level posterior means of preferences for ideological diversity in government-sensitive news (ρ_i^{Int} and ρ_i^{Ukr})	60
B.1	Normalized number of weekly visitors to top second and third news websites, IE Toolbar and Liveinternet.ru	74
B.2	Normalized number of weekly visitors to top fourth and fifth news websites, IE Toolbar and Liveinternet.ru	75
B.3	Normalized number of weekly visitors to top sixth and seventh news websites, IE Toolbar and Liveinternet.ru	76
C.1	Normalized number of weekly visitors of IE Toolbar data	77

D.1	Distribution of the differences in usage rankings of proper names between government-controlled and independent outlets	78
D.2	Top 100 proper names that are under-used by the government-controlled media, and selected subset used to define the internal sensitive news topics	79
E.1	Slant related to sensitive news	80
F.1	Top 50 proper names that are underused by the government in articles about the Ukraine crisis	82
G.1	List of slanted words used related to the Ukrainian conflict	83
G.2	Pro-Russia and pro-Ukraine slants in the Ukraine conflict	83
I.1	Changes in the market shares with the increase in the amount of the internal-sensitive news	85
I.2	Changes in the market shares with the increase in the amount of the Ukraine-crisis news	86
K.1	Posterior distributions of demand parameters: matching values	89
M.1	Posterior distribution of preference for Internal-sensitive news and corresponding slant: Model with and without the interdependence of choices within a day . . .	93
M.2	Posterior distribution of preference for the Ukraine Crisis news	94
M.3	Posterior distribution of preference for slant in the Ukraine Crisis	95

LIST OF TABLES

2.1	Russian-language online news media by the type of influence in December 2014 .	8
2.2	Number of articles by type of news outlet	9
6.1	Posterior mean point estimates of β and γ parameters: unweighted and weighted by number of news-consumption occasions	40
7.1	Simulated market shares for different levels of government control.	51
B.1	Comparison of website rankings in IE Toolbar and LI.ru	73
H.1	Summary of browsing behavior	84
L.1	Simulated shares, day with a large internal-sensitive news event.	90
L.2	Simulated shares, day with a large Ukraine-crisis news event.	91

ACKNOWLEDGMENTS

This dissertation would not be possible without a tremendous support of people who have surrounded me during the years of my graduate studies. I am grateful to my dissertation committee, Pradeep Chintagunta, Jean-Pierre Dubé, Günter Hitsch and Matthew Gentzkow, for their trust, guidance, and friendship. I am deeply thankful to Pradeep and JP for the time and effort invested in me and for the willingness to accommodate and comment on all my endless requests and ideas. It was fun, and I have learned a lot from our interactions. I am also thankful to Günter and Matthew for their thoughtful and thorough feedback at every meeting, and for the ability to highlight the most important pieces of the puzzle.

I am grateful to all the other people of Booth and the University of Chicago from whom I was fortunate to learn over these years, especially Anita Rao, Brad Shapiro, Oeystein Daljord and Sanjog Misra. I extend a special thank you to Chris Nosko for his support and advice both during his time at Chicago and later on. I was extremely fortunate to have amazing colleagues throughout the program, who have made this experience very memorable and enjoyable: thanks to the marketing folks, especially to Robbie and Xiliang, for always being around ready to provide feedback and think constructively about research, and to other students at Booth and the department of economics for all the amusing conversations and insights. In particular, thanks to Adam, Arvid, Ging Cee, Jack and Russ for making the first year an enjoyable experience, and to Arvid and Winnie for being such great roommates and companions.

I am also grateful to the people of Microsoft Research: NYC, where I have spent time working on this dissertation and other projects. I am especially grateful to Justin Rao for his advice, support and friendship during my time at MSR and beyond.

Another thank you goes to my former advisors at Tilburg University, Bart Bronnenberg and Tobias Klein, for putting me on the right track and encouraging me for this long journey.

I deeply thank my friends and family for their tremendous support. I am proud to have

such great parents, Irina and Dmitry, who have always encouraged me to move forward. I am also extremely lucky to have met my beloved fiancée Anna during my time at Chicago, who has supported me through the darkest times of the Ph.D. program and with whom we have enjoyed the best moments of this time together. Thank you for being here with me and for your love and inspiration.

I thank Philipp Kats for help with some of the data collection, public.ru for providing part of the publication records data, and anonymous media experts for providing the institutional knowledge about the Russian news industry. Research support from Joseph A. and Susan E. Pichler Fellowship, Katherine Dusak Miller Fellowship, Bradley Fellowship from Stigler Center and Kilts Fellowship is gratefully acknowledged, as well as ISMS Doctoral Dissertation Award and ASA Annual Doctoral Research Award. Research support from the Sanford J. Grossman Fellowship in Honor of Arnold Zellner is also gratefully acknowledged; any opinions expressed herein are the author's and not necessarily those of Sanford J. Grossman or Arnold Zellner.

ABSTRACT

Anecdotal evidence suggests consumers navigate to news outlets that are government controlled (GC) and pro-government biased even in the presence of independent news outlets. Does this observation imply a demand for pro-government bias in the news? Or do consumers enjoy other aspects of GC news outlets such as quality and coverage of unbiased stories? To answer these questions, we examine consumers' demand in the Russian online news market. We use publication records of the top 48 online news outlets to characterize the methods of government control in the news and identify the news that is sensitive for the government. We then use temporal variation in the amount of sensitive news and click-level browsing panel data to estimate the demand for news. We find the average consumer prefers news with less pro-government bias, but substantial heterogeneity exists across consumers. In particular, preferences of frequent news consumers differ from the rest of consumers. In a counterfactual simulation, we show that GC news outlets would have higher market shares in the absence of control.

CHAPTER 1

INTRODUCTION

On August 23, 2016, BBC.com published a news story covering the ban on Russian athletes from the 2016 Paralympic games due to a doping scandal. The title of the story was "Rio Paralympics 2016: Russia banned after losing appeal", and discussed that an arbitration court upheld an earlier decision to ban Russian Paralympic team.¹ On the same day, another online news agency, RT.com, published a story on the same court ruling titled "Removing a strong rival? Russia shocked by 'cynical and political' CAS ruling on Paralympic team ban".² The article emphasized that the decision of the arbitration court was political and that there was no hard evidence of Russian athletes using doping.

Such difference in media coverage is defined by economics literature as media slant (e.g. Gentzkow and Shapiro 2010). Media slant can come from the supply side, reflecting preferences of journalists, advertisers, owners of the news outlet, or governments (Baron 2006, Besley and Prat 2006). It can also come from the demand side, reflecting preferences of readers for like-minded news or diverse information sources (Mullainathan and Shleifer 2005, Xiang and Sarvary 2007, Gentzkow and Shapiro 2010, Zhu and Dukes 2014).

In the case of coverage of the Russian Paralympic team ban, there is strong evidence suggesting that media slant comes from the supply side: website RT.com is owned by the Russian government, and "RT" is short for "Russia Today". However, this government-induced bias is different from government capture in the model of Besley and Prat (2006), where a government needs to capture all news sources to suspend information dissemination. If news outlets cover the topic truthfully, and consumers have access to all news outlets, why would the Russian government invest money in RT? Does it attract (and potentially

. A version of this paper is work in progress with a co-author, Justin M. Rao.

1. <http://www.bbc.com/sport/disability-sport/37165427>

2. <https://www.rt.com/sport/356863-paralympic-russia-reaction-rio/>

persuade) customers with different ideological views, or is it read only by customers whose political beliefs are similar to the political position of the Russian government?

In this paper, we examine the demand for news coverage under government control in a case study of the Russian online news market in years 2013-2015. In this period, the Russian online news market had both government-controlled (GC) and independent news outlets, with all news outlets being easily accessible to any user of the internet. A stylized fact in this market is that GC news outlets enjoy high and stable market shares in this period of time: around 25% of news outlets are GC (owned by the government), and their overall market share is around 35-40%.³

The key question of this paper is what drives demand for the GC news outlets in Russia. There are two potential families of explanations. On the one hand, consumers might have a preference for a pro-government news coverage, either because their beliefs align with the government's ideological position, or because they value knowing the ideological position of the government. On the other hand, consumers might have a distaste for the pro-government slant but prefer some other characteristics of the GC news outlets, such as overall quality of the news coverage or brand capital of news outlets.

We start this project with collecting two novel datasets. One dataset contains all publication records for the top 48 online news outlets that write in Russian during the period of April 2013 - April 2015. The resulting panel contains 3.9 million news articles, and for each article we know its URL, text, title and publication date. Another dataset comes from Internet Explorer Toolbar and contains complete browsing records for 284,574 consumers of Russian online news websites in the period of November 2013 - April 2015.

We use the publication records data to find and characterize pro-government bias in the news. For this, we compare publications of GC and independent news outlets. To find sensitive news, we use two potential methods of government control: censorship and

3. Based on the statistics from <http://www.liveinternet.ru>. Historical data is scrapped using the Wayback Machine: <http://web.archive.org/>

propaganda. First, we use the idea of censorship and find topics (identified as proper nouns) that are systematically underused by the GC news outlets compared to the independent news outlets. There is a significant difference in coverage, indicating that the GC news outlets systematically omit news about political opposition and corruption, such as news about opposition leaders (e.g., “Khodorkovsky,” “Navalny”), president Putin’s affiliates related to corruption (e.g., “Rotenberg,” “Timchenko”), and political protests (e.g. “Bolotnaya”). We label censored news as “internal-sensitive news”. Second, we examine news coverage of the Ukraine crisis of 2013-2015, a sensitive news topic widely reported to have a pro-Russia propaganda. We show that the GC news outlets systematically report more news about the Ukraine crisis compared to the independent news outlets, but these news exhibit the pro-Russia (or anti-Ukraine) slant: for example, GC news outlets tend to say that Crimea has “reunited” with Russia, the new Ukrainian government is fascist and “anti-Russian”, and the Ukrainian government is conducting a “punitive” operation against rebels in eastern Ukraine. We find this slant comparing publications of the GC and Ukrainian news outlets that have coverage in Russian. We also find the pro-Ukraine (or anti-Russia) slant, defined as language that is systematically overused by the Ukrainian news outlets: for example, that Russia has “annexed” Crimea, Russia is an “aggressor” country, and the Ukrainian government is conducting an “anti-terrorist” operation against “terrorists” in eastern Ukraine.

We use the identities of sensitive news topics and publication records data for two purposes. First, we characterize the coverage of news outlets on the sensitive news: how much each news outlet reports about a sensitive topic, and what language (slant) it uses when it reports about this sensitive topic. This gives us ideological positions of the news outlets. Second, an overall daily share of publications about the sensitive news topics gives us a measure of the relative importance of the sensitive news for each day, or how much sensitive news happened on this day.

Having ideological positions of news outlets and the relative importance of sensitive news

over time, we build and estimate the demand model for news. Consumer preferences are defined over news outlets, topics that these news outlets cover on a particular day, and the ideological slant in the sensitive news coverage. Demand identification comes from changes in the relative importance of sensitive news over time. On days with no sensitive news, censorship and propaganda do not affect the GC news outlets’ coverage, so their coverage is similar to the coverage of the independent news outlets. On these days consumption is driven by persistent preferences for news outlets, which can reflect the quality of news outlets or their brand capital. On the contrary, on days with sensitive news there are ideological differences between the news outlets, so both the preferences for particular news outlets and the ideological preferences of consumers play a role. If consumers switch to reading the GC news outlets on days with sensitive news, they prefer the pro-government bias in the news. If they switch away from the GC news outlets on these days, they have a distaste for pro-government bias. Finally, if consumers are more likely to visit both the GC and independent news outlets on days with sensitive news, it suggests that consumers value knowing the government’s position about sensitive issues and are “conscientious” (Mullainathan and Shleifer 2005, Xiang and Sarvary 2007).

We estimate the structural model of demand for news using Internet Explorer browsing data. To get a reliable measure of demand heterogeneity, we use a Bayesian hierarchical model with consumer heterogeneity approximated with a mixture of normal distributions (Rossi 2014).

Estimates reveal that the majority of consumers in the market prefer the coverage of sensitive news by the independent news outlets. This is true even for consumers of the GC news outlets, indicating that they visit the GC news outlets due to persistent outlet preferences. To measure the importance of government control, we run a counterfactual simulation where the GC news outlet have ideological positions similar to the independent news outlets, and find that in this case a market share of the GC news outlets would increase

by 6%. By doing simple back-of-the-envelope calculation we show that this 6% increase approximately correspond to \$6.5 million dollars of annual advertising revenue of GC news outlet. For comparison, the government’s subsidies to the GC news outlet are \$1.2 billion dollars, suggesting that a loss of advertising revenue due to the government’s control is not very substantial for the GC news outlets.

In addition, model estimates reveal several empirical facts about the demand for news. First, our results suggest that ideological preferences of frequent and infrequent news consumers are different, with the frequent news consumer having more distaste for the pro-government bias. The frequent news consumers are a relatively small group of news readers (8.8%), but they are responsible for the majority (70%) of news consumption in our data. Thus, news outlets are incentivized to fit their coverage to the preferences of the frequent news consumers and ignore the preferences of the rest 91% the market, which questions the ability of an unregulated news market to inform average voter. Second, we find that 40% of the frequent news readers navigate to more ideologically diverse news outlets when the amount of sensitive news in the market increase, suggesting that they are “conscientious” consumers. Finally, we find that some consumers prefer both the pro-Russia and the pro-Ukraine slant in the Ukraine crisis news coverage, and infrequent consumers are more likely to have such preferences. Given that both the pro-Russia and the pro-Ukraine slant in the Ukraine crisis news coverage uses very emotional language, results suggest that consumer preferences for slant can be driven by pure entertainment, which to our knowledge is not considered by the prior literature.

This paper is the first to use a structural demand model to estimate consumer preferences for pro- and anti-government slant in autocracies.⁴ We propose a new method to

4. Gentzkow and Shapiro (2015) discuss a related demand model for online news consumption. Other work examined the consumer response to an increase in the pro-government bias in the news, with Durante and Knight (2012) documenting viewers response to the ideological change in TV programming of public television due to Berlusconi’s victory in the national elections in Italy, and Knight and Tribin (2016) documenting viewers response to the airing of cadenas, government propaganda on Venezuela channels.

measure media slant (Groseclose and Milyo 2005, Gentzkow and Shapiro 2010, Gentzkow et al. 2016), which to our knowledge is the first measure that splits media slant into multiple dimensions.⁵ We use a new identification strategy to estimate consumer preferences for slant in online news, which contributes to the empirical literature on media slant (Gentzkow and Shapiro 2010, Martin and Yurukoglu 2015) and online news markets (Gentzkow and Shapiro 2011, Gentzkow and Shapiro 2015, Sen and Yildirim 2016, Athey et al. 2017, Cage et al. 2017). Our demand estimation results inform theoretical literature on the demand-side slant (Mullainathan and Shleifer 2005, Xiang and Sarvary 2007, Zhu and Dukes 2014). Finally, our work is related to the theoretical (Besley and Prat 2006, Prat and Strömberg 2013, Gehlbach and Sonin 2014) and empirical (Durante and Knight 2012, Enikolopov et al. 2011, Bai et al. 2014, Roberts 2014, Garcia-Arenas 2016, Knight and Tribin 2016) literature on the effect of government control of the news on consumers.

The rest of the paper is organized as follows. Section 2 describes the Russian online news market and data sources. In Section 3, we find government-sensitive news topics and characterize the reporting of news outlets on these topics. Section 4 presents descriptive evidence. We build a demand model in Section 5 and present estimation results in Section 6. Section 7 studies counterfactual simulations and measures the effectiveness of government control. In Section 8 we extend the basic demand model to account for multiple consumptions of consumers within a day. Section 9 concludes.

5. Perego and Yuksel (2016) discuss the separate decision of news outlets on agenda setting and slant in the news in a theoretical framework, and Pan and Xu (2017) examine if Chinese ideological spectrum is multi-dimensional.

CHAPTER 2

DATA

2.1 Online News Market in Russia

Despite relatively high government control over the offline news market starting in 2000, online news outlets in Russia enjoyed relative freedom up until the 2013. A large number of independent players existed in the online news media landscape. Since the beginning of 2013, political pressure has forced a number of top online news outlets to change their chief editors.¹ The most prominent examples include dissolution of RIA Novosti, a state news agency known for balanced news coverage under its editor-in-chief Svetlana Mironyuk, in December 2013² and the layoff of Galina Timchenko, editor-in-chief of one of the top online news outlets in Russia, lenta.ru, in March 2014.³ Government control intensified in February of 2014 with the Ukrainian crisis and the annexation of Crimea, with the government blocking websites of some opposition leaders in March 2014⁴ and developing a law to limit the foreign ownership of Russian news outlets to 20% starting in January 2016.⁵

In the end of 2014, the online news media landscape in Russia still included both groups of GC and independent news outlets. Table 2.1 presents the top 48 Russian-language news

1. ura.ru - November 30, 2012, change of chief editor, <http://www.vesti.ru/doc.html?id=972487> (Russian); gazeta.ru - September 7, 2013, change of chief editor, http://slon.ru/russia/grekh_unyniya_kreaklov_gazete_ru_naznachili_novyy_glavnyy_redaktor-988192.xhtml (Russian); ria.ru - December 10, 2013, liquidation, chief editor fired, <http://www.rg.ru/2013/12/09/ykaz-dok.html> (Russian); lenta.ru - March 12, 2014, change of chief editor, <http://lenta.ru/news/2014/03/12/goreslavsky/> (Russian)

2. <http://www.telegraph.co.uk/news/worldnews/europe/russia/10505386/Vladimir-Putin-dissolves-Russias-RIA-Novosti.html>

3. <http://www.bbc.com/news/world-europe-26543464>

4. <http://www.bbc.com/news/technology-26578264>

5. <http://www.rg.ru/2014/10/17/smi.html> (Russian); http://www.squirepattonboggs.com/~media/files/insights/publications/2014/10/russia-moves-to-limit-foreign-ownership-in-the-2_/files/russiamovestolimitforeignownership/fileattachment/russiamovestolimitforeignownership.pdf

Table 2.1: Russian-language online news media by the type of influence in December 2014

International	Independent	Possibly Influenced (oligarchic)	Government (controlled)	Ukrainian (subset)	
bbc	newsru	bfm	fontanka	1tv	korrespondent
svoboda	newtimes	echo	gazeta	aif	liga
meduza	novayagazeta	interfax	lifenews	dni	unian
dw	rbc	mk	izvestia	ntv	
reuters	slon	znak	kommersant	rg	
	tvrain	ng	kp	ria	
	vedomosti	polit	lenta	rt	
	forbes	regnum		vesti	
	snob	ridus		vz	
	the-village	rosbalt		tass	
		sobesednik			
		utro			
	trud				

Table presents the simplified domain names; for example, 1tv stands for www.1tv.ru. Most domains have the www.*.ru structure, with some exceptions. Groups are created based on open information and interviews with media professionals.

outlets,⁶ including 40 Russian news outlets, five international news outlets that offer news stories in Russian, and three large Ukrainian outlets with popular Russian-language sections.

Russian news outlets are organized into four groups: (1) independent and not influenced, (2) independent but possibly influenced, (3) possibly influenced news media owned by oligarchs close to Kremlin, and (4) GC outlets. Classification is based on the interviews with media professionals who prefer to remain anonymous. The ownership structure and media reports support this classification. For example, news outlets classified as GC are owned by the government (7 out of 10 news outlets) or the state company Gazprom (1 news outlet), or were founded by a member of the current incumbent party and a strong supporter of Vladimir Putin (2 news outlets). Appendix A contains detailed information on the ownership structure and open information about the news outlets.

6. Top outlets are defined both by their market share observed on liveinternet.ru and by the frequency of citations by news aggregators.

Functionally, group (1) of independent and not-influenced news outlets does not face any direct government influence, but might be subject to self-censorship given that the Russian government can potentially punish the news outlets. Groups (2) and (3) of independent but potentially influenced news media and oligarchic media have formally independent news coverage, but are reported to be indirectly influenced by the government. The nature of government control in these groups is very similar, so we group these outlets together and call them “influenced” news outlets. Finally, group (4) contains GC news outlets that have news agenda directly controlled by the Kremlin. The majority of these news outlets are owned by the government and receive government subsidies.

2.2 Supply Data

For 48 outlets described above, we collect information on publications for the period starting April 1, 2013, and ending March 31, 2015. The data are collected directly from archives on news-outlet websites and from the media archives *medialogia.ru* and *public.ru*. The resulting panel contains 3.9 million publications. For each article, we collect the title, text,⁷ URL link, and timestamp. Table 2.2 presents the number of articles per type of news outlet.

Table 2.2: Number of articles by type of news outlet

Type	Articles
GC	1,168,569
Independent	494,087
Influenced	1,848,556
International	75,596
Ukrainian	315,927
Total	3,902,735

7. For five news outlets (“meduza”, “newtimes”, “ridus”, “snob”, “the-village”), text was not collected for technical reasons. We keep these outlets in parts of the textual analysis and use titles instead. We drop these news outlets for the descriptive analysis and demand estimation because without information on article texts, we could not get a reliable measure of slant.

2.3 Demand Data

To measure the demand for news, we use the Internet Explorer (IE) Toolbar browsing data, which include complete browsing histories for a subset of IE users. The users included in the IE Toolbar data have installed a plug-in on their IE and opted-in for the data collection.⁸ IE Toolbar data contain information about each webpage consumers visited (URL), websites where consumers came from (referral URL), timestamp of the visit, number of seconds spent, browsing session ID, user ID, language of the browser, country of the user, and other information. We focus the analysis on Toolbar users who specified Russian as the language of their browser.

Although IE Toolbar data are collected for several years, the unique user IDs are kept only for one and a half years. By the time the data collection was conducted, the earliest available browsing data with user IDs were from November 15, 2013. We thus collect the browsing data between November 15, 2013, and March 31, 2015⁹ for all users with the IE language set to Russian.

The resulting panel consists of 2.17 million users. Among these users, 284,574 navigated to a news website at least once over the sample period, which is only 13% of users with IE in Russian. At the same time, these users are the most active online whose browsing corresponds to 77.8% of all browsing of users who set their IE language to Russian. In total, our sample contains 26.54 million page views of the 48 news-outlet websites defined above.

2.3.1 *IE Toolbar Representativeness*

Online news consumers in the IE Toolbar are a subset of all online news consumers: they are consumers who use IE as their browser and also have the toolbar plug-in installed on

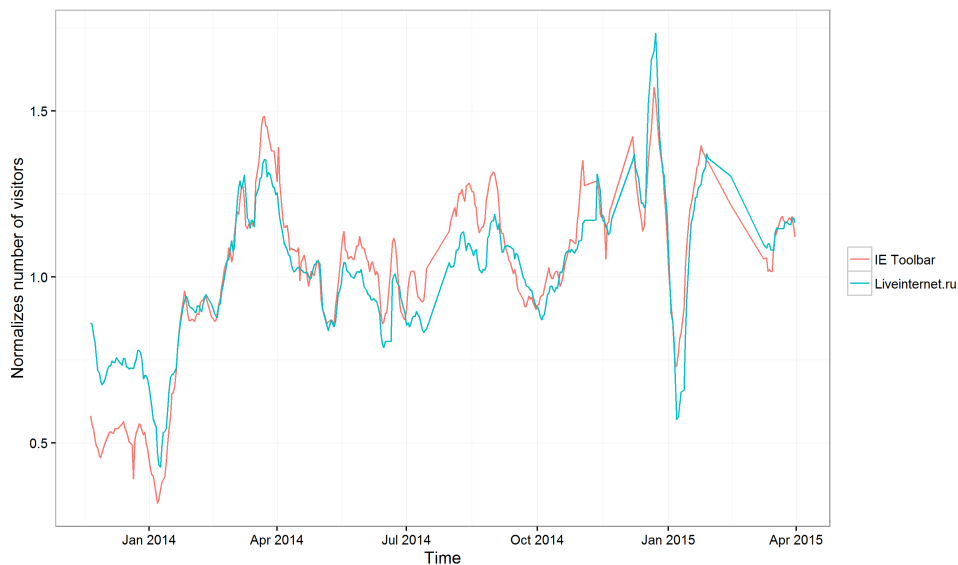
8. Based on Microsoft records, around 75% of users who installed the plug-in opted-in to the data collection.

9. Data for the period between April 1, 2013, and November 15, 2013, are available with scrubbed (deleted) user IDs.

their computer. We examine whether the news-consumption patterns of this small subset is different from the patterns of the overall population of news consumers in Russian. We collect the data on the overall number of visits of a subset of news outlets in Russia from `liveinternet.ru` (LI), a website that tracks statistics for the Russian internet. To collect historical data, we use the digital archive Wayback Machine to scrape the information about the usage rankings. Due to the layout of the website ranking on LI, we have reliable records of usage over the period of time studied for only the top 30 websites on the Russian internet.¹⁰ We collect information on the number of users navigating to these websites during the week.

Seven news websites frequently appear in top 30 websites. We compute the average traffic ranking across these websites, and compare it to the ranking of these websites in the Toolbar data. We find that five out of seven top news outlets in LI are also top news outlets in the IE Toolbar data, with `lenta.ru` being in position 8 and `ria.ru` being in position 14. Table B.1 in Appendix B presents the websites and rankings.

Figure 2.1: Normalized number of weekly visitors to `rbc.ru`, IE Toolbar and LI data



The average overall traffic is normalized to one; IE Toolbar data are corrected for the churn rate.

10. The top page includes only the top 30 websites; Wayback Machine does not have frequent records for the other pages.

We also examine if news consumption changes in a similar way over time. Figure 2.1 shows weekly news consumption for the market leader rbc.ru based on IE Toolbar and LI data. The average overall traffic is normalized to one, and IE Toolbar data are corrected for the churn rate.¹¹ The changes in news consumption in the IE Toolbar tracks the overall changes in news consumption of rbc.ru relatively well.¹² We conclude that news consumption by IE Toolbar users can be used to approximate the overall Russian news consumption.

11. Figure C.1 in Appendix C contains information about the number of weekly users of the IE Toolbar.

12. Appendix B contains similar figures for the other six top news websites.

CHAPTER 3

GOVERNMENT CONTROL AND SENSITIVE NEWS

3.1 Types of Government Control

In general, researchers acknowledge two broad types of news bias induced by governments: censorship and slant. Censorship of the news occurs when the government removes a certain topic from the news coverage of its controlled outlets, or allows reports of only certain facts. For example, a government instructing a news outlet not to cover a story about a corruption scheme organized by some government officials, or instructing a news outlet to omit certain facts about the involvement of government officials in the scheme classifies as censorship. The media economics literature refers to censorship as “issue and fact bias” (Prat and Strömberg 2013) or as “filtering or selection of news” (Gentzkow et al. 2015). Censorship works through the effects of agenda setting (McCombs and Shaw 1972) and priming (Iyengar and Kinder 1987). Cohen (1963) summarized the idea of agenda setting by arguing that the press “may not be successful much of the time in telling people what to think, but it is stunningly successful in telling its readers what to think about.”

Apart from censorship, government can control the news by adding slant to news reporting. Slant in the news coverage is reporting that uses language that favors one of the parties described in the news. Gentzkow and Shapiro (2010) provide multiple examples of slanted language used by the members of the US Congress, such as describing the Iraq war as “war on terror” (republicans) versus “war in Iraq” (democrats). In the media economics literature, slant corresponds to “framing and ideological stand bias” (Prat and Strömberg 2013) and “distortion of news” (Gentzkow et al. 2015).

3.2 Internal-Sensitive News

We identify the first set of government-sensitive news using the idea of censorship. Given that censorship is the omission of facts, we examine proper nouns. We take this approach because proper nouns mainly correspond to facts in the news; for example, they represent the actors in the news and the places where the news happened.¹ We consider all words starting with a capital letter as proper nouns except for the first words in the sentences. If the facts corresponding to a topic are censored or underreported, proper nouns related to the topic will be underused in the news outlet’s publications.

We then compare the usage of proper nouns by GC and independent news outlets. The goal is to find proper nouns that are used less by the *majority* of GC news outlets compared to the *majority* of independent news outlets. We use the following procedure:

1. For each outlet j and proper noun v , we compute usage of v by j by computing the fraction of occurrences of v in j ’s reporting (across all proper nouns used), sh_{vj} .
2. For each v , we rank usage sh_{vj} across outlets, sh_{v1}, \dots, sh_{v48} , where the outlet with the highest usage share receives a rank equal to 1, $rank_{vj'} = 1$.
3. For each v , we compute the average rank of usage for GC and independent news outlets, $Rank_{vInd}$ and $Rank_{vGov}$.
4. Proper nouns with the lowest difference between the two average ranks, $Rank_{vInd} - Rank_{vGov}$, are the ones that are underused by the majority of GC news outlets.

The benefit of this procedure as opposed to a simple comparison of shares of usage is twofold. First, news outlets differ in news volume; using shares of words instead of counts

1. For example, the title of one of top news stories on the day when this paragraph was written, “Panama Paper: David Cameron’s worst week as Prime Minister,” contains proper nouns “Panama Papers,” “David Cameron,” and “Prime Minister,” that summarize the topic of the news article, but does not capture the sentiment of this topic (captured by the word “worst”).

allows normalizing the size of news outlets. Second, some outlets might specialize in a particular topic (e.g., corruption scandals) and have a large share of usage of particular proper nouns. Using an ordinal-rank measure instead of cardinal-share measure for usage of proper nouns allows us to limit the effect of such outliers.

We apply this procedure to all proper nouns that appear more than 200 times in the publications.² Figure D.1 in Appendix D presents the histogram of the distribution of the rank differences, $Rank_{vInd} - Rank_{vGov}$. Figure D.2 in Appendix D shows the list of 100 proper nouns with the lowest rank differences. The top 10 proper nouns underused by GC outlets include “Khodorkovsky”³ (a former political prisoner and one of the most well-known people to oppose the government), “Rotenberg”⁴ and “Timchenko”⁵ (close confidants of Putin), “Roskomnadzor”⁶ (a Russian federal executive body overseeing the media), “Slon”⁷ and “Dozhd”⁸ (independent online news outlets in Russia), and other public figures criticizing the government. Thus, the most underused proper nouns correspond to facts that are anecdotally known to be sensitive for the government. We take this as evidence that proper nouns with the lowest $Rank_{vInd} - Rank_{vGov}$ rank difference represent government-sensitive news topics.

Using the result above, we define the first set of government-sensitive news topics. We take a subset of the top 100 proper nouns underused by GC outlets and define a news article as “internal sensitive” if it contains one of these proper names.⁹ We add the label

2. Using only frequent proper nouns ensures that we focus only on the influential facts. Proper nouns should appear 200 times across all publications (all journals, all days). We arbitrarily chose the threshold.

3. https://en.wikipedia.org/wiki/Mikhail_Khodorkovsky

4. https://en.wikipedia.org/wiki/Arkady_Rotenberg

5. https://en.wikipedia.org/wiki/Gennady_Timchenko

6. https://en.wikipedia.org/wiki/Federal_Service_for_Supervision_in_the_Sphere_of_Telecom,_Information_Technologies_and_Mass_Communications

7. https://ru.wikipedia.org/wiki/Slon_Magazine

8. <https://en.wikipedia.org/wiki/Dozhd>

9. We remove the proper names that correspond to the names of news outlets (e.g., Slon), editors of news

“internal” because most of these proper nouns correspond to an internal-sensitive issues such as government opposition and corruption.

Censorship is the mechanism with which we allocated internal-sensitive topics. However, we do not know if it is the only mechanism that government uses to bias these news stories. GC news outlets might report internal-sensitive news with a slant different from independent news outlet. To find this slant, we examine the difference in word usage between GC and independent outlets in articles about internal-sensitive news. We apply the ranking procedure described above, and examine words (excluding proper nouns) that are underused by GC news outlets in the articles about sensitive news. The top hits include “blocked” (referred to website), “prisoner” (referred to arrested activists), “meeting” (referred to unauthorized demonstrations), “corruption,” and other words referring to sensitive issues for the government. We use a subset of these words as indication of anti-government slant of the news outlet,¹⁰ and label a news article about internal-sensitive news “slanted” if it contains at least one of these words.¹¹

Knowing the identities of articles about internal-sensitive news and slant, we can characterize the reporting of news outlets. Figure 3.1 describes the reporting on internal-sensitive topics by various Russian news outlets. Subfigure (a) presents an example of articles containing the proper noun “Khodorkovsky,” a political prisoner who was released from jail in December 2013. The figure shows that news outlet forbes.ru and tvrain.ru reported the most about Khodorkovsky: around 4.7% of all of their publications contain his last name. Subfigure (b) combines the share of usage of all proper nouns on internal-sensitive topics, with

outlets (e.g. Venediktov), news aggregators (e.g., yandex, rambler), and names of some government officials and companies that appear in the news for a wide spectrum of reasons and thus create a lot of noise (e.g., Sberbank, Medvedev). The resulting list of proper names is presented in Figure D.2 in Appendix D.

10. We remove proper nouns and incidental words such as words related to the media. Appendix E contains more information on the selection of words.

11. We note that the allocated “slant” might contain some factual information; for example, the topic the independent media described using the word “corruption” could be covered by GC news outlets as “overspending the budget”, or could be not mentioned at all.

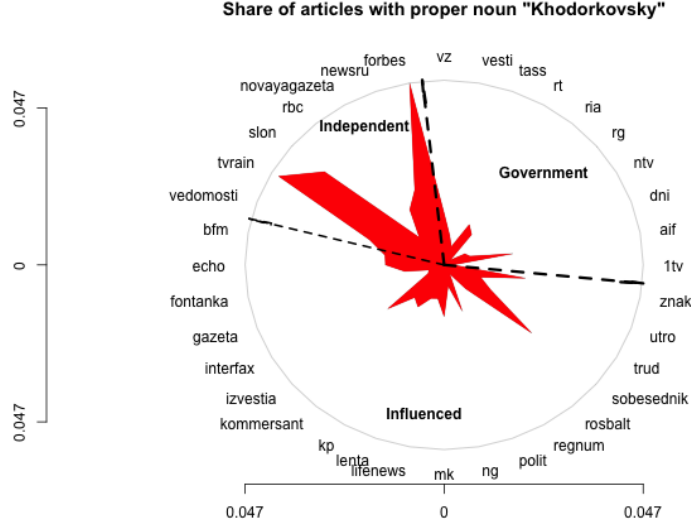
the darker red color corresponding to higher usage of more proper nouns. We can see that potentially influenced outlets differ in the coverage of internal-sensitive news: some news outlets (e.g., sobesednik.ru or izvestia.ru) are closer in their reporting to the independent news outlets, and some (e.g., gazeta.ru or regnum.ru) are closer in their reporting to the GC news outlets.

Figure 3.2 presents the average levels of news outlets' reporting and anti-government slant on internal-sensitive news. By construction, independent outlets have higher reporting and slant than GC outlets. Influenced outlets are between GC and independent outlets, with some influenced news outlets being closer to the GC and some being closer to the independent. Ukrainian news outlets report little about internal-sensitive news, and international news outlets' reporting is close to the reporting of independent news outlets.

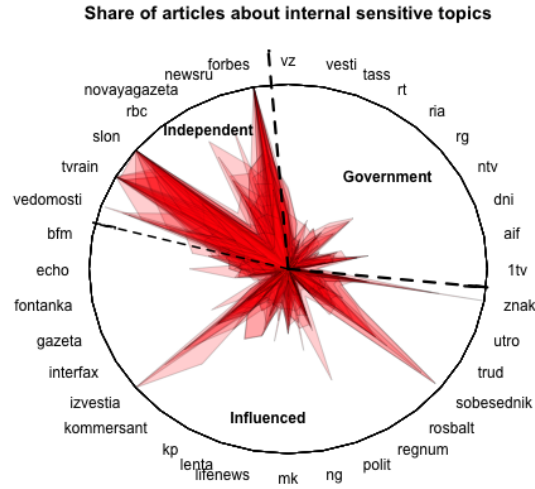
In addition to characterizing the reporting of news outlets, we use the share of articles about internal-sensitive news overall in the market on a given day to measure the relative importance of internal-sensitive news on that day. This approach provides a proxy for the number of sensitive news events that happened on a given day: on the days when nothing sensitive happened, news outlets will have nothing to report about, whereas on the days with a lot of sensitive news, the share of reporting will be higher.

We find that GC news outlets censor more on the days with more internal-sensitive news. For this finding, we compute the difference in reporting (share of articles about internal-sensitive news) of GC and independent news outlets on each day, and regress the difference in reporting on the overall share of internal-sensitive news in the market. We find a significant positive relationship: with a 1-percent-point increase in the amount of internal-sensitive news, the difference in reporting increases by 0.54 percentage points (standard error of 0.055). Thus, the difference in reporting in GC and independent news outlets is higher on the days with more internal-sensitive news.

Figure 3.1: Reporting about internal-sensitive news by the Russian news outlets



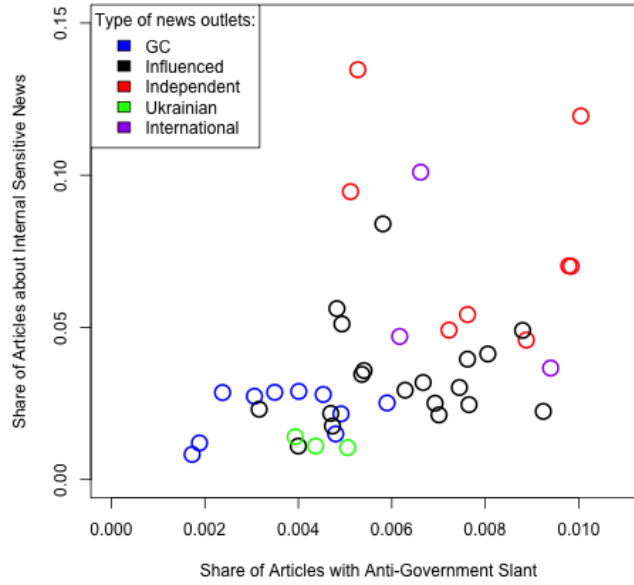
(a) Example of articles about Khodorkovsky



(b) Usage of all internal-sensitive words

Red radial lines correspond to the share of reporting about a topic by a news outlet outside of the circle. Subfigure (b) plots the share of usage of all proper nouns classified as internally sensitive. We normalize the maximum reporting of each word to reach the circle. Darker color corresponds to larger number of sensitive words.

Figure 3.2: Levels of reporting and slant in internal-sensitive news, by news outlets



Each dot represents a position of a news outlet.

3.3 Government-Sensitive News about the Ukraine Crisis

We now turn to another government-sensitive news topic: the Ukraine crisis of 2013-2015.¹² The conflict was widely covered in the Russian news media, and was reported to be heavily slanted by news outlets controlled by the Russian government.¹³

Figure 3.3 presents the share of news articles that contain the word “Ukraine” that were published in the independent, government-influenced, and GC outlets over time.¹⁴ With the beginning of the Ukraine conflict, the figure shows that all news outlets increase their reporting about Ukraine, but GC outlets increase it more than independent and influenced outlets. This increase is almost the opposite of censorship: GC news outlets report significantly more

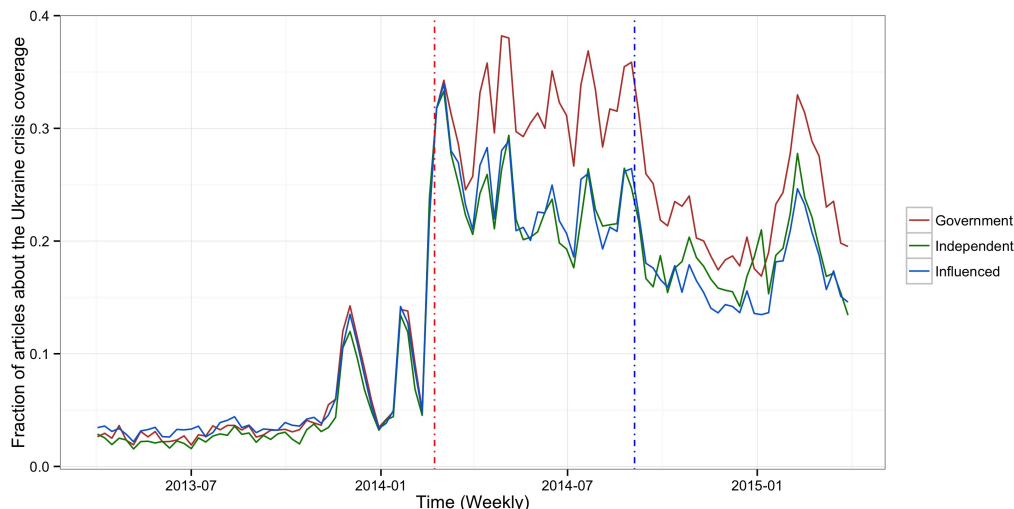
12. For a broad overview, please see https://en.wikipedia.org/wiki/Ukrainian_crisis.

13. For an overview, please see https://en.wikipedia.org/wiki/Media_portrayal_of_the_Ukrainian_crisis#Media_in_Russia.

14. Having the word “Ukraine” in the news coverage is a proxy for an article being about the Ukraine crisis.

on the Ukraine crisis than independent and influenced news outlets.¹⁵

Figure 3.3: Share of articles containing the word “Ukraine” in the weekly coverage of news outlets, by types



Red line corresponds to the GC media, green line - to the independent media, blue line - to the government-influenced media. Red dotted line corresponds to February 22, 2014, day when the former president Yanukovich fled Ukraine as a result of a revolution. Blue dotted line corresponds to the first Minsk Peace agreement, September 4, 2014.

We next look for evidence of the usage of slant by GC news outlets. An important feature of media coverage of the Ukrainian crisis is that journalists collected abundant observational evidence about both pro-Russia and pro-Ukraine slant in the news coverage. We compile the vocabulary of pro-Russia (slanted in favor of rebels from eastern Ukraine regions and against Ukraine’s new government) and pro-Ukraine slant (slanted against rebels from eastern Ukraine regions) using the materials on stopfake.org, a fact-checking website launched in March 2014.¹⁶ A key aspect of pro-Russia/anti-Ukraine slant is the framing of the new Ukrainian government as fascist: a pro-Russia media slant is described with the

15. We check the degree of censorship using the ranking method and find little evidence of such occasions. The only exception is information on killed Russian soldiers in Ukraine, with a prominent example of Leonid Kichatkin, killed in Ukraine at the end of August, news that was widely reported by the independent media but ignored by the government media. Appendix F contains the list of the top underused proper names by GC outlets.

16. The website is supported by faculty and alumni of the Mohyla School of Journalism and students from the Digital Future of Journalism program, Kyiv, Ukraine.

words “junta” (referring to a new Ukraine’s government), “punitive” (referring to actions of the Ukrainian army against rebels), and “banderovtsy” (referring to Ukrainian nationalists). Key words describing a pro-Ukraine slant are “terrorists,” “hitman,” “separatists” (all referring to rebels), “occupied” (referring to the territory of eastern Ukraine), and “aggressor” (referring to Russia). We denote articles that contain slanted words of either type as slanted articles. Appendix G contains a complete list of slanted words and describes the procedure of selecting these words.

Figure 3.4 describes the usage of pro-Russia slant by the news outlets. Subfigure (a) shows the share of articles about the Ukraine crisis using the word “punitive.” We can see that GC news outlets slant their reporting about the Ukraine crisis more than independent news outlets. Subfigure (b) shows the overall usage of slant in the Ukraine-crisis articles.

Figure 3.5 describes the usage of pro-Ukraine slant by the news outlets. Subfigure (a) shows the share of articles about the Ukraine crisis using the word “occupied.” Subfigure (b) shows the overall usage of slant in the Ukraine-crisis articles. Ukrainian news outlets use pro-Ukraine slant the most. At the same time, all other types of news outlets also use some of the words that are considered to be pro-Ukraine slant.

Figure 3.6 present the average levels of pro-Russia and pro-Ukraine slant in the Ukraine-crisis news by outlets. GC news outlets have relatively high levels of pro-Russia slant and average levels of pro-Ukraine slant. Independent news outlets have relatively low levels of pro-Russia slant and average to high levels of pro-Ukraine slant. Some influenced news outlets look more like GC, and some look more like independents. Finally, Ukrainian and international news outlets have low levels of pro-Russia slant and high levels of pro-Ukraine slant.

Figure 3.4: Amount of the pro-Russia slant in the reporting about the Ukraine Crisis



(a) Example of usage of word “punitive”



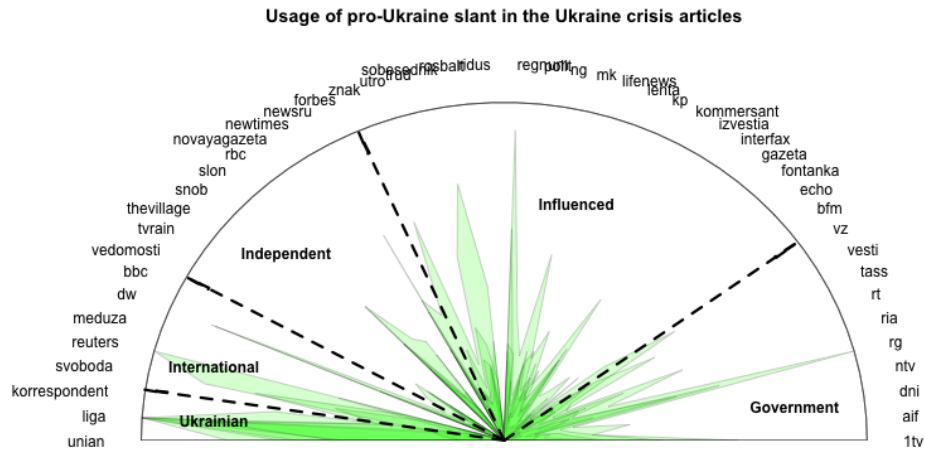
(b) Usage of all pro-Russia slanted words

Red radial lines correspond to the share of pro-Russia slanted reporting about the Ukraine crisis by a news outlets outside of the circle. In the Subfigure (b), we normalize the maximum reporting of each word to reach the circle. Darker color corresponds to larger number of sensitive words.

Figure 3.5: Amount of the pro-Ukraine slant in the reporting about the Ukraine crisis



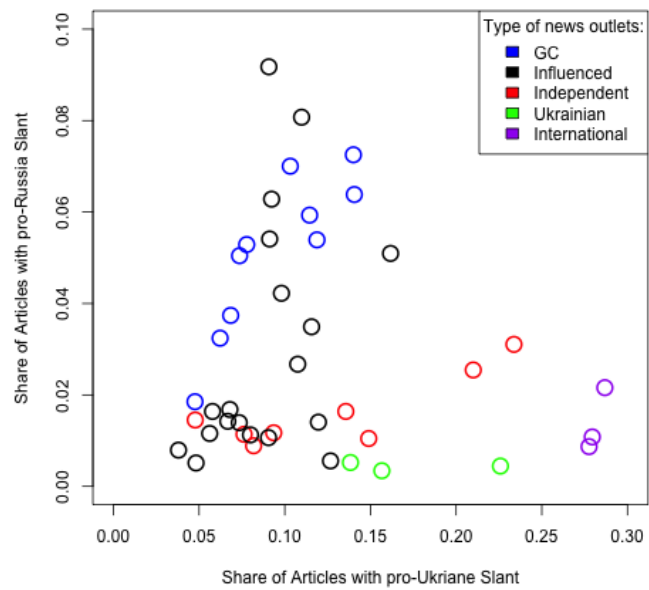
(a) Example of usage of word “occupied”



(b) Usage of all pro-Ukraine slanted words

Green radial lines correspond to the share of pro-Ukraine slanted reporting about the Ukraine crisis by a news outlets outside of the circle. In the Subfigure (b), we normalize the maximum reporting of each word to reach the circle. Darker color corresponds to larger number of sensitive words.

Figure 3.6: Level of the pro-Russia and pro-Ukraine slant in the Ukraine-crisis news coverage by each news outlet



Each dot represents a position of a news outlet.

CHAPTER 4

DESCRIPTIVE EVIDENCE

Before building a formal model of demand for news, we present some descriptive evidence on the role of government control in consumer demand. In particular, we examine the relationship of news outlets' market shares and the amount of sensitive news in the market. The ideological position of news outlets is more important on the days with more sensitive news. If consumers prefer the pro-government bias, on the days with more sensitive news market shares of the pro-government biased news outlets should grow more compared to the market shares of the news outlets with no pro-government bias.

We define news consumption of an outlet as navigation to at least one of its news articles on a given day.¹ We treat only visits to news articles' web pages as news consumption, because news outlets provide information other than news on their websites (e.g., financial analytics, industry analysis, photos and videos, etc.). To allocate consumption of news articles, we find URLs in the demand data that either match the article URLs in the supply data or have a similar structure.² Appendix H breaks down website URLs by page types, and presents summary statistics of consumers' browsing of these types.

We define news consumers as users who have read news in our sample at least once. If consumers are online on a given day but do not navigate to any news articles, we record them as choosing not to consume the news. Thus, market share of news outlets can grow from news consumers switching from other news outlets (intrinsic margin) and deciding to start consuming the news (extrinsic margin). We define the market share of outlet j on day t as the number of consumption occasions of j on day t divided by the total number of

1. We thus ignore the intensity of consumption; we examine the results when re-defining consumption at different levels of intensity (e.g., at least five news article web pages, at least two minutes spent on the webpages, etc.) and find similar results.

2. That is, if for some outlet a typical news article has the structure `http://website/news/year/month/day/article-name` we will consider all URLs with the structure `http://website/news/year/month/day/*` as a news article.

consumption occasions in the market on day t .

To examine the change in consumption due to an increase in the amount of sensitive news, we regress the market shares of news outlet j on the amount of internal-sensitive news events and Ukraine-crisis news events on day t :

$$\text{sh}_{jt} = b_{0j} + b_j^{Int} F_t^{Int} + b_j^{Ukr} F_t^{Ukr} + \xi_{jt}$$

where F_t^{Int} and F_t^{Ukr} correspond to the share of articles about internal-sensitive news and Ukraine-crisis news, respectively. The slope coefficients b_j^{Int} and b_j^{Ukr} correspond to the change in the market shares due to the change in the amount of sensitive news in the market.

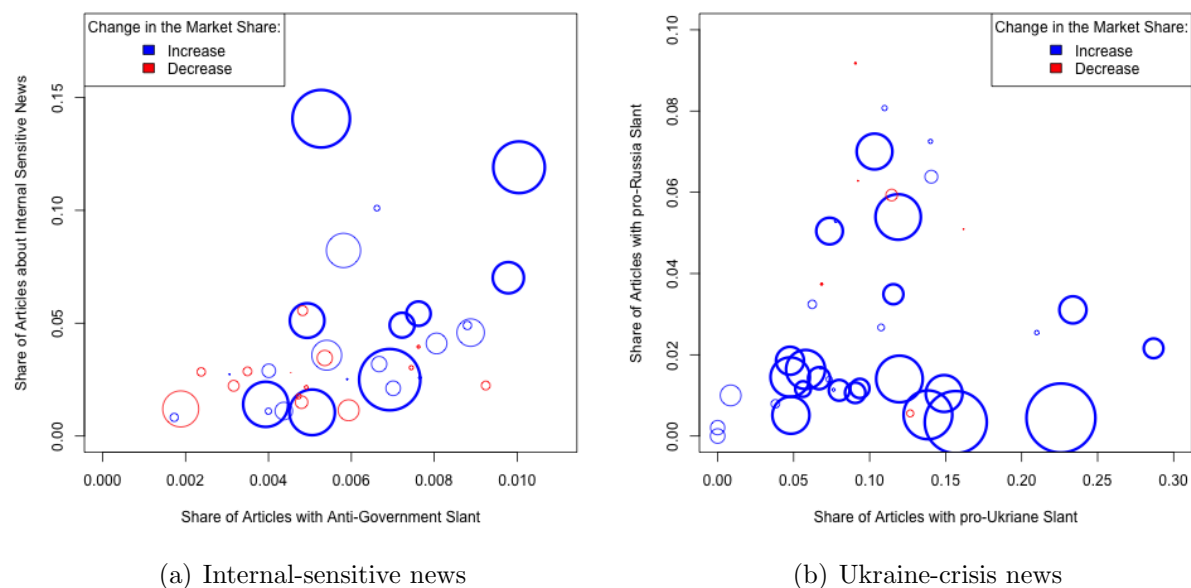
Figure 4.1 summarizes the results of the regression. Each point represents a news outlet. The size of the points represents the degree of change of the market share of the news outlets, measured as a percent of average market shares of this news outlet, $\beta_j^{Int}/\bar{\text{sh}}_{jt}$ and $\beta_j^{Ukr}/\bar{\text{sh}}_{jt}$. Blue color corresponds to the increase in the market shares, and red color corresponds to the decrease in the market share. Bold borders of the points correspond to significance of the change in the market share.³

Subfigure (a) presents the relationship of the market shares and the amount of internal-sensitive news. The vertical axis corresponds to the amount of reporting about internal-sensitive news by the outlet, while the horizontal axis corresponds to the amount of anti-government slant in the coverage of the news outlets, similar to the Figure 3.2. We find that ten news outlets gain significant increase in their market shares as there is more internal-sensitive news in the market. The news outlet that report more on internal-sensitive news and have more anti-government slant are more likely to get an increase in their market shares. Appendix I provides more details on the relationship.

Subfigure (b) presents the relationship of the market shares and the amount of news

3. Increase is significant at the 5% level, standard errors are corrected for autocorrelation.

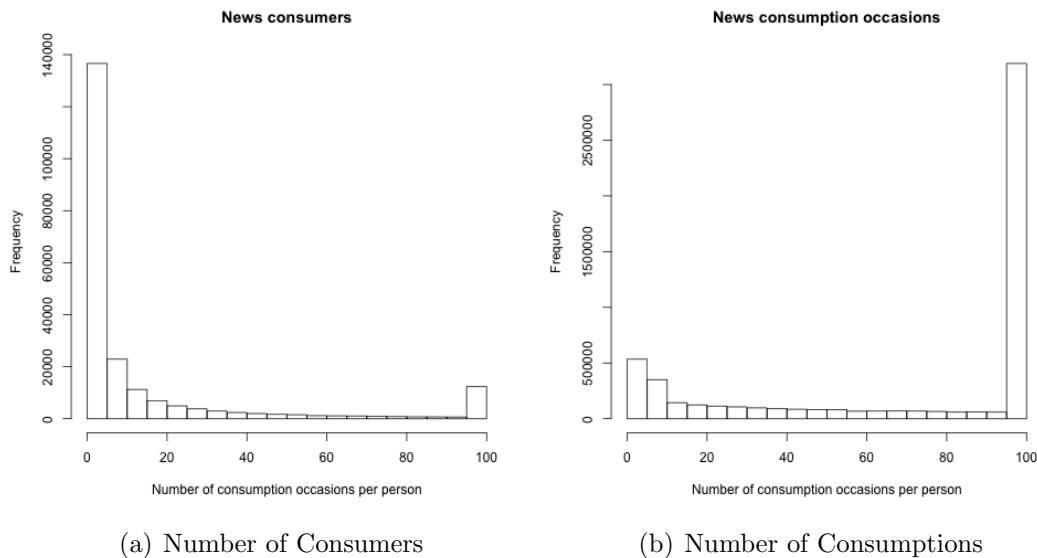
Figure 4.1: Predicted changes in the news outlets' market shares with the change in the amount of sensitive news, by news outlet



Each point represents a news outlet. The size of the points represents the degree of change of the market share of news outlets, measured as a percent of average market shares of this news outlet. Blue color corresponds to the increase in the market shares, and red color corresponds to the decrease in the market share. Bold borders of the points correspond to significance of the change in the market share.

about the Ukraine crisis. The vertical axis corresponds to the amount of pro-Russia slant in the news, while the horizontal axis corresponds to the amount of pro-Ukraine slant, similar to the Figure 3.6. We find that 20 news outlets gain significant increase in their market shares as there is more Ukraine-crisis news in the market. The news outlet that use less pro-Russia slant are more likely to get an increase in their market shares. Appendix I provides more details on the relationship. In addition, the share of consumers choosing the outside option (not to read the news on a given day) decreases on the days with more Ukraine-crisis news, which implies that part of the increase in the market shares of the news outlet comes from the extrinsic margin.

Figure 4.2: Histograms of number of consumers and number of consumptions by types of consumers

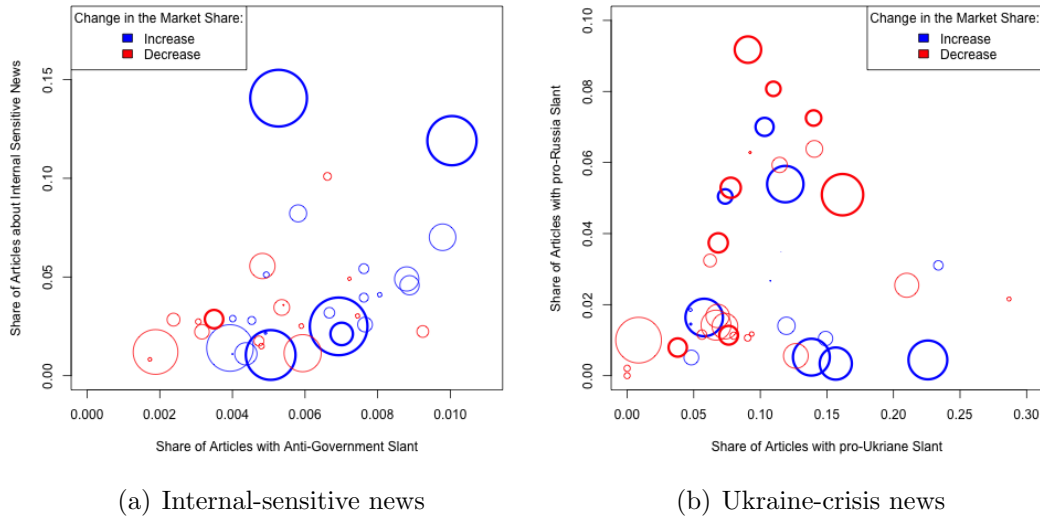


Types of consumers are measured with number of consumptions in the sample period. Figure presents the scale truncated at 100 consumptions occasions per individual.

If consumers have similar preferences for news coverage, the above results suggest that an average consumer prefers news on government-sensitive news topics to non-sensitive news, prefer the coverage of internal-sensitive news with more anti-government slant and the coverage of the Ukraine-crisis with less pro-Russia slant. However, potentially there is heterogeneity of consumer preferences, in particular between frequent and infrequent consumers.

Figure 4.2 presents the histograms of number of consumers and number of consumption occasions by different types of consumers, grouped by the frequency of consumption. Subfigure (a) shows that majority of consumers in this market rarely visit the news outlets, with only 9.7% of consumers having more than 50 occasions during the sample period of 502 days. However, these 9.7% of consumers are responsible for 68.7% of the news consumption, which is shown in subfigure (b). This implies that preferences of a small group of frequent news consumer have large effect on the market shares. If frequent and infrequent news consumers have different news preferences, changes in the market share might not represent the preferences of an average news consumer.

Figure 4.3: Predicted changes in the news outlets' market shares from the infrequent news consumers with the change in the amount of sensitive news, by news outlet



Only individuals with less than 5 news consumption occasions are used. Each point represents a news outlet. The size of the points represents the degree of change of the market share of news outlets, measured as a percent of average market shares of this news outlet. Blue color corresponds to the increase in the market shares, and red color corresponds to the decrease in the market share. Bold borders of the points correspond to significance of the change in the market share.

To examine the preferences of infrequent news consumers, we regress their consumption on the amount of sensitive news in the market, using the similar procedure as summarized

in figure 4.1. We define infrequent news consumers as the individuals who have less than 5 consumption occasions during the sample period.⁴ Figure 4.3 presents the results. Results in the subfigure (a) suggest that, similar to the rest of the market, infrequent news consumers prefer the outlets with more internal-sensitive news coverage and with more anti-government slant on the days with more internal-sensitive news. However, subfigure (b) reveals that infrequent news consumers are less interested in the news about the Ukraine crisis. In addition, we cannot conclude that these consumers prefer news with less pro-Russia slant, as the relationship between the change in their market shares with more Ukraine-crisis news and degree of pro-Russia slant of the news outlet is not significant. To measure the demand for bias in the news and to examine the consumer heterogeneity more deeply, we build and estimate a model of demand for news.

4. The threshold is chosen arbitrarily.

CHAPTER 5

MODEL

5.1 Nature

There is a set of possible news events S . Each event can be characterized by one of three news topics it covers: non-sensitive news for the government, internal-sensitive news, and news about the Ukraine crisis. At every time period t , nature produces news about a subset of these events, S_t . We denote the number of events that happens about a given topic as N_t^x , where x corresponds to non-sensitive, internal-sensitive or Ukraine-crisis news, $x = \{Non, Int, Ukr\}$. We assume that the news-market participants take the underlying production process as given.

5.2 News Outlets

The market contains J news outlets. Each news outlet j is given its type $type_j$: independent, influenced, GC, or Ukrainian, $type = \{ind, inf, gov, ukr\}$.¹ News outlet j chooses its quality α_j , the level of anti-government slant in the reporting of internal-sensitive news sl_j^{Int} , and level of pro-Russia and pro-Ukraine slant in the reporting of Ukraine-crisis news, sl_j^R and sl_j^U . News outlet j also chooses the number of news articles about each type of news on day t , N_{jt}^x , given $N_t^x \forall x$ and their relative size. As researchers, we do not observe N_t^x . Instead, we measure the relative importance of topic x on day t with a fraction of news articles about topic x in the market (across all news outlets), $F_t^x = \frac{\sum_j N_{jt}^x}{\sum_x \sum_j N_{jt}^x}$.

We would like to measure how informative the reporting of a particular news outlet on topic x is given the relative importance of news topics on this day F_t^x . One natural measure of the amount of information is the number of articles this news outlet wrote about

1. We do not include international websites, because the data for most of them are sparse and their market share is low.

this topic on a given day, N_{jt}^x . This measure has a couple of downsides: the number of articles depends on the size of the news outlet, and the length of news articles can vary. An alternative measure is the fraction of articles on topic x in news outlet j 's reporting on day t , $F_{jt}^x = \frac{N_{jt}^x}{\sum_x N_{jt}^x}$. Such a measure allows normalizing the size of the news outlet j . A downside of this measure is that it is not necessarily related to the relative importance of a news topic x on a given day: a particular news outlet might always allocate 60% of its space to the news topic, even when nothing serious happen, perhaps because they specialize in the topic. To penalize the news outlets' reporting for such specialization, we define the measure of reporting as

$$\text{rep}_{jt}^x = \begin{cases} F_t^x \left(\frac{F_{jt}^x}{F_t^x} \right) = F_{jt}^x & \text{if } F_{jt}^x \leq F_t^x \\ F_t^x \left(1 + \frac{F_{jt}^x - F_t^x}{1 - F_t^x} \right) & \text{if } F_{jt}^x > F_t^x. \end{cases}$$

If news outlet j has fewer articles on topic x than the average in the market on day t , $F_{jt}^x \leq F_t^x$, its reporting on topic x is just the fraction of news articles it has on this topic, $\text{rep}_{jt}^x = F_{jt}^x$. If news outlet j has more articles on topic x than the average in the market on day t , $F_{jt}^x > F_t^x$, we define its reporting on topic x as F_t^x plus F_t^x multiplied by an extra % of space news outlet j allocated to topic x , $\frac{F_{jt}^x - F_t^x}{1 - F_t^x}$. Such a measure penalizes news outlets for over-reporting on news topic x when the relative importance of this topic is low.

For example, consider two news outlets: outlet A always matches the relative importance of news topic x with the fraction of articles it allocates to this topic, and outlet B always allocates 60% of space to the topic x (e.g., it specializes in this topic). On the days with the relative importance of topic x of 5%, the fraction of articles news outlets A and B have on topic x will differ by a magnitude of 12: outlet A will have $F_{At}^x = 0.05$, and outlet B will have $F_{Bt}^x = 0.6$. The difference in the measures of reporting we have constructed will be much smaller: outlet A will have $\text{rep}_{At}^x = 0.05$, and outlet B will have $\text{rep}_{Bt}^x = 0.05(1 + \frac{0.6-0.05}{1-0.05}) = 0.079$. In this way, the reporting of news outlet B is linked to the relative importance of

topic x , and is more comparable to news outlet A .²

5.3 Demand

There are I consumers in the market. We assume that consumers are in the market for online news on the days when they are browsing online. On each consumption occasion τ on day t , consumer i can choose one news outlet, or choose an outside option of not consuming any news. We treat a consumption occasion as reading news for a short period of time, which makes navigating to more than one news outlet impractical for consumers (Gentzkow and Shapiro 2015). Consumers observe the slant of news outlets, $sl_{j'}$, and the amount of reporting on news topics $rep_{j't}$, for all j' . Consumer i has preferences for news topics, β_i , ideological slant $\gamma_i = \{\gamma_i^{Int}, \gamma_i^R, \gamma_i^U\}$ and news outlets, $\alpha_i = \{\alpha_{i1}, \dots, \alpha_{iJ}\}$. We assume consumer preferences are fixed over time, and denote them as $\theta_i = \{\beta_i, \gamma_i, \alpha_i\}$.

Consumer i ' utility from visiting news outlet j on day t on the consumption occasion τ is

$$u_{ijt\tau} = \alpha_{ij} + rep_{jt}^{Int}(\beta_i^{Int} + sl_j^{Int}\gamma_i^{Int}) + rep_{jt}^{Ukr}(\beta_i^{Ukr} + sl_j^R\gamma_i^R + sl_j^U\gamma_i^U) + \epsilon_{ijt\tau}, \quad (5.1)$$

where $\epsilon_{ijt\tau}$ is an idiosyncratic shock to the utility distributed iid type-1 extreme value. On the days with no sensitive news, rep_{jt}^{Int} and rep_{jt}^{Ukr} are zero, so the utility of consumer i from reading news outlet j comes primarily from α_{ij} . On the days when sensitive news is more important, the relative preference of consumers for sensitive news, β_i^{Int} and β_i^{Ukr} , and their preferences for slant γ_i influence the utility they get from news outlets. The utility of the outside option of being online but not consuming the news is normalized to $u_{i0t\tau} = \epsilon_{i0t\tau}$. The consumer chooses outlet j at choice occasion τ such that $u_{ijt\tau} \geq u_{ij't\tau} \forall j' \in \{0, \dots, J\} : j' \neq j$.

2. We acknowledge that in defining the reporting of news outlets this way, we impose a specific functional form on the reporting. We intend to relax this functional form in future work.

Denote consumers' choices as y . The probability that consumer i chooses news outlet j at on day t on the consumption occasion τ is

$$\pi(y_{it\tau} = j|\theta_i) = \frac{\exp(\alpha_{ij} + \text{rep}_{jt}^{Int}(\beta_i^{Int} + \text{sl}_j^{Int}\gamma_i^{Int}) + \text{rep}_{jt}^{Ukr}(\beta_i^{Ukr} + \text{sl}_j^R\gamma_i^R + \text{sl}_j^U\gamma_i^U))}{1 + \sum_{j'} \exp(\alpha_{ij'} + \text{rep}_{j't}^{Int}(\beta_i^{Int} + \text{sl}_{j'}^{Int}\gamma_i^{Int}) + \text{rep}_{j't}^{Ukr}(\beta_i^{Ukr} + \text{sl}_{j'}^R\gamma_i^R + \text{sl}_{j'}^U\gamma_i^U))}.$$

The likelihood of θ_i observing a sequence of choices y_i is

$$L(\theta_i|y_i) = \prod_t \prod_\tau \prod_j \pi(y_{it\tau} = j|\theta_i)^{I(y_{it\tau}=j)}.$$

A number of assumptions underlie the specified model. First, we assume the discrete-choice structure of the news consumption and independence of consumers' choices over time. These assumptions are of particular importance on the days when consumer i navigates to more than one news outlet. Although such occasions are rare in the data (77% of browsing sessions of consumers have visited only one news outlet), visiting multiple news outlets might be more important on the days with more sensitive news. In this case, misspecification of the model will affect the estimates of consumer preferences. In section 8, we relax this assumption by allowing interdependence of consumer choices within a day, which partly solves the problem of visits to multiple news outlets, but does not capture any further interdependences of choices (across days). We plan to relax this assumption further in future work.

Second, we assume consumers know their preferences for all news outlets α_i , as well as the slant and reporting of outlets on news topics, when they make consumption decision. Matching values α_i and the slant of news outlets are fixed over time, so in this case, it is sufficient that consumers have learned about their preferences and ideological positions of news outlets some time in the past. Reporting of news outlets changes over time due to changes in the relative importance of sensitive news; we assume consumers are aware of the relative importance of sensitive news on a given day and have information about the share

of reporting of news outlets. Consumers can get such information online (from search, social networks, etc.) before deciding to navigate to news articles. This assumption is supported by common indirect navigation of consumers to news articles: at least 46% of consumers in our sample navigate to their first news article during the browsing session using links on news aggregators, search, or social media, with only 19% navigating from other web page on news websites.³ Media reports has also found indirect navigation being very common for news consumption.⁴

Third, we assume consumers' preferences are fixed over time. This assumption contradicts the behavior of government, which controls the news to influence consumers' political beliefs, which would potentially affect consumers' ideological preferences. We treat this behavior of government as a long-run investment in consumers' ideology, and we assume that in the short run, such behavior will not affect consumers' preferences. At the same time, we acknowledge that adding consumer learning and dynamics in consumers' preferences is more realistic and will enrich the model.

Finally, the model does not allow news outlet to differ in the quality of reporting on sensitive news. Thus, if some difference in the quality of reporting across news outlets is not captured by the amount of reporting rep_{jt}^x and is correlated with the amount of slant news outlets use in the reporting, consumers' preferences for slant might be biased. For example, if news outlets that have higher-quality reporting on Ukraine-crisis news use less slant (and consumers prefer higher-quality reporting), the model will underestimate consumers' preferences for slant. In future work, we plan to address this problem by creating

3. The referral website for the remaining 35% of news article visits was not recorded. Browsing session is defined using the session ID in the IE Toolbar data.

4. In January 2015, a survey of internet users found that 67% navigate to news through search and news aggregators, 36% through the social media, and only 17% through direct navigation to the websites. Source: <http://rusability.ru/news/fom-rasskazal-o-potreblenii-novostnogo-kontenta-v-rossii/>, <http://fom.ru/SMI-i-internet/12491>. However, 67% and 36% can be overestimates, because these percentages include people who consume news from search and social media without navigating to news outlets.

a more flexible measure of news outlets’ reporting, as well as creating an observed measure of quality of reporting on news topics and incorporating this measure into the demand model.⁵

5.4 Government Control

Government control affects the reporting and slant of news outlets. The censorship constraint affects the reporting of news outlets: under censorship, the government determines the fraction of news articles of outlet j on sensitive topic x , F_{jt}^x (and the fraction is below the optimal amount of reporting on censored topics). The propaganda constraint affects the slant levels of news outlets: under propaganda, the government determines the slant of outlet j , sl_j^* (and it is more “pro-government” than optimal for the news outlet).

To determine the optimal levels of reporting slant, we need to specify the supply-side model of the market. Appendix J presents a basic supply-side model in the spirit of Gentzkow and Shapiro (2015). We leave the refinement of the model and its application for future work.

5. Multiple potential proxies exist for the quality of reporting: number of facts (proper nouns) reported, readability of the text, etc.

CHAPTER 6

ESTIMATION AND RESULTS

6.1 Demand Estimation

We estimate the distribution of θ_i using a Bayesian hierarchical model. The first-stage prior on θ_i is a linear combination of consumer characteristics z_i and a random variable u_i :

$$\theta_i = \Delta' z_i + u_i.$$

The prior distribution of Δ is normal:

$$\text{vec}(\Delta) \sim N(\bar{\delta}, A_{\delta}^{-1}).$$

The prior distribution of u_i is a mixture of K normal distributions, with the Dirichlet prior distribution over the components, the normal distribution over the mixture components' means, and the inverse Wishart prior over the mixture components' covariance matrix:

$$u_i \sim N(\mu_{\text{ind}_i}, \Sigma_{\text{ind}_i})$$

$$\text{ind} \sim \text{Multinomial}_K(\text{pvec})$$

$$\mu_k \sim N(\bar{\mu}, \Sigma_k \otimes a_{\mu}^{-1})$$

$$\Sigma_k \sim IW(\nu, V).$$

We estimate the distribution of the parameters θ by simulating from the posterior distribution using an MCMC hybrid sampler. We pick the standard tuning parameters based on the discussion in Rossi, Allenby, and McCulloch (2005).

For the individual-level estimation, we define the consumer choice at the level of the type

of outlet, allowing for four types: GC, government-influenced, independent, and Ukrainian. This transformation reduces the number of consumer choices to five alternatives, which significantly reduces the computation time. Estimation is done on a sample of 50,000 news consumers.¹

6.2 Estimation Results

Figures 6.1 and 6.2 present the pointwise posterior means and 90% credibility region of the marginal densities of consumers' preferences for news coverage, β_i and γ_i , defined in equation (5.1)². The scale of variables is normalized to represent the effect of a one-standard-deviation change in the underlying variable on consumers' utility, which allows us to compare the relative importance of news characteristics.

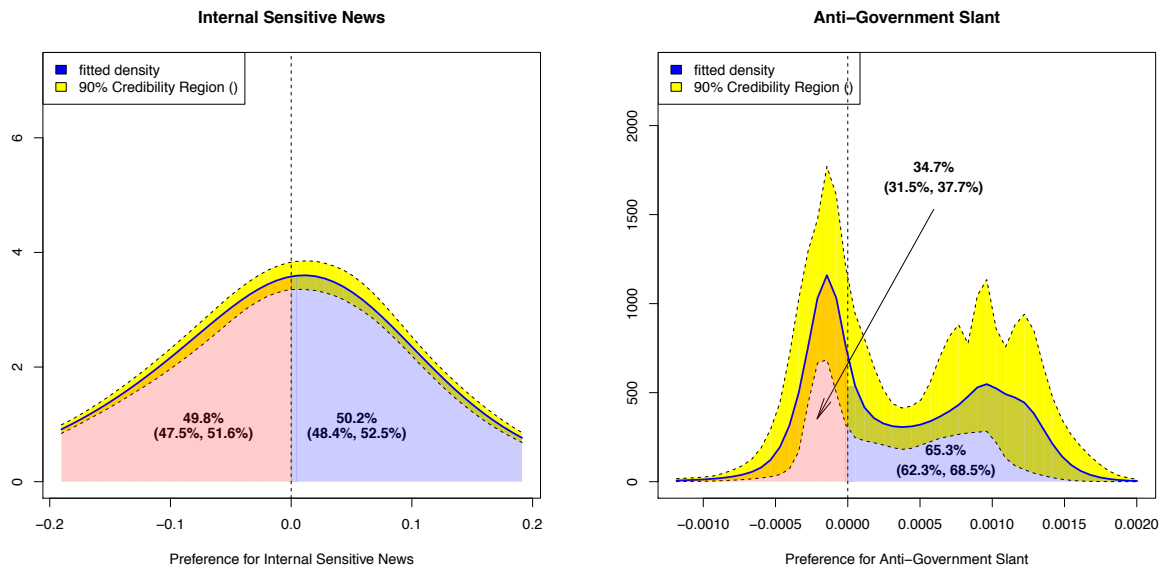
Figure 6.1 presents estimates of consumer preferences for internal-sensitive news. Subfigure (a) shows that consumers' preferences for the internal-sensitive news, β_i^{Int} , are split: 50.2% of consumers prefer reading about the internal-sensitive news, with a 90% credibility interval of (48.4%, 52.5%), whereas the rest prefer non-sensitive news. This indicates the heterogeneity of consumers' preferences for news topics. An average consumer is indifferent between the internal-sensitive and non-sensitive news, with the average β_i^{Int} of -0.003, as shown in Table 6.1.

Subfigure (b) shows the distribution of consumer preferences for the anti-government slant in the internal-sensitive news, γ_i^{Int} . On average, consumers prefer reading the internal-sensitive news with the anti-government slant: a one-standard-deviation increase in slant leads to a 4e-4 (3e-4; 5e-4) increase in consumer utility. The majority of consumer, 65.3% (62.3%, 68.5%), prefer the anti-government slant in the internal-sensitive news. However,

1. Sampling done for computational reasons. We simulate 1 million draws, keeping every 500th draw for the reasons of data storage and use the first 0.5 million draws as a burn-in period. We are currently re-estimating the model at the news-outlet level.

2. Distributions of matching values α_{ij} are available in Figure K.1 in Appendix K.

Figure 6.1: Posterior distributions of preferences for internal-sensitive news and corresponding slant



(a) Preference for Internal-sensitive News

(b) Preference for Anti-Government Slant

We normalize the scale of variables: values on the x axis represent effect of a one-standard-deviation increase in the corresponding variable on consumers' utilities. Blue line is a posterior mean; yellow shaded regions are 90% credibility regions. Numbers represent a percentage of probability mass of the distributions below and above zero, and 90% credibility intervals of these percentage.

Table 6.1: Posterior mean point estimates of β and γ parameters: unweighted and weighted by number of news-consumption occasions

	Unweighted		Weighted	
	Mean	S.D.	Mean	S.D.
Internal Sensitive:				
$\hat{\beta}^{Int}$	-0.003	0.121	0.021	0.132
	(-0.007, 0.001)	(0.114, 0.127)	(0.018, 0.024)	(0.128, 0.136)
$\hat{\gamma}^{Int}$	4e-04	5e-04	3e-04	6e-04
	(3e-04, 5e-04)	(2e-04, 7e-04)	(1e-04, 4e-04)	(5e-04, 7e-04)
Ukraine Crisis:				
$\hat{\beta}^{Ukr}$	-0.075	0.505	0.117	0.541
	(-0.087, -0.065)	(0.462, 0.533)	(0.105, 0.128)	(0.524, 0.557)
$\hat{\gamma}^R$	0.015	0.031	0.001	0.045
	(0.011, 0.02)	(0.015, 0.04)	(-0.004, 0.007)	(0.038, 0.051)
$\hat{\gamma}^U$	0.162	0.195	0.144	0.289
	(0.15, 0.176)	(0.156, 0.223)	(0.131, 0.157)	(0.264, 0.311)

Consumers are weighted based on the number of occasions of news consumption.

from the scale of $\hat{\gamma}_i^{Int}$ we can conclude that consumers' preferences for the anti-government slant is less important than consumers' preferences for the internal-sensitive news, β_i^{Int} : the effect of a one-standard-deviation change in the amount of the internal-sensitive news coverage on consumers' utilities is of two orders of magnitude higher than the effect of a one-standard-deviation change in the amount of the anti-government slant in this coverage. This implies that for the internal-sensitive news, censorship has a stronger effect on consumption than choice of language used to describe the internal-sensitive news.

To measure consumer preferences for the coverage of GC news outlets versus independent news outlets, we compare the utility consumers get from the internal-sensitive news coverage of these news outlets. We assume that the coverage of the independent news outlets is “unbiased”, and refer to the deviation from this coverage by GC news outlets as “pro-government biased”. Thus, we record that consumer i dislikes the pro-government bias in internal-sensitive news on day t if

$$\text{rep}_{\text{Govt}}^{Int}(\hat{\beta}_i^{Int} + sl_{\text{Gov}}^{Int}\hat{\gamma}_i^{Int}) < \text{rep}_{\text{Indt}}^{Int}(\hat{\beta}_i^{Int} + sl_{\text{Ind}}^{Int}\hat{\gamma}_i^{Int})$$

. We find that on the day with an average amount of internal-sensitive news, 51.3% (49.6%; 53.5%) of consumers get more utility from the coverage of the independent news outlets. If we weight the consumers by the estimated probability to visit GC news websites (putting more weight on the consumers of GC news outlets), results is almost unchanged: 51.5% (50.3%, 53%) of consumers of GC news outlets dislike the pro-government bias. This implies that on the days with an average volume of the internal-sensitive news, around 50% of the consumers of GC news outlets would prefer the internal-sensitive news coverage of independent news outlets, so they navigate to GC news websites despite the control. On the days with a large internal-sensitive news event, the share of GC consumers who prefer independent coverage increase to 67.7% (66.5%, 68.9%)³.

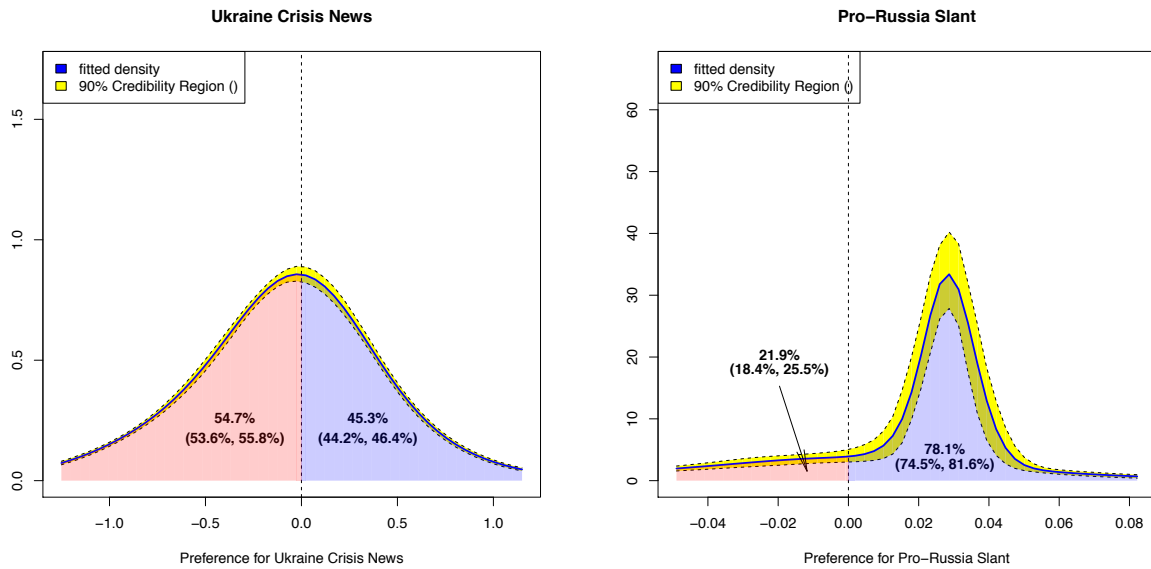
Figure 6.2 presents estimates of consumer preferences for the Ukraine-crisis news. Subfigure (a) shows that the majority of consumers, 54.7% (53.6%, 55.8%), have negative estimates of β_i^{Ukr} , which means that they prefer non-sensitive news to the Ukraine-crisis news. This supports the finding that there is heterogeneity in consumers' preferences for news topics, shown in the figure 6.1⁴. An average consumer prefers non-sensitive news to the Ukraine-crisis news, with an average β_i^{Ukr} of -0.075, as shown in Table 6.1.

Subfigure (b) and (c) show the distribution of consumer preferences for the pro-Russia and pro-Ukraine slant in the Ukraine-crisis news, γ_i^R and γ_i^U , respectively. The majority of consumers prefer the pro-Russia slant (78.1%), and the majority of consumers prefer the pro-Ukraine slant (86.9%). This is surprising: in the context of Ukraine crisis pro-Russia slant ("Ukraine are fascists") and pro-Ukraine slant ("Russia is aggressor") can be viewed as opposing each other, in which case we would expect consumers to support one side of the conflict versus the other. If this is true, we would expect consumers to have unidimensional

3. We estimate this for a day with the highest observed volume of the internal-sensitive news in the sample period.

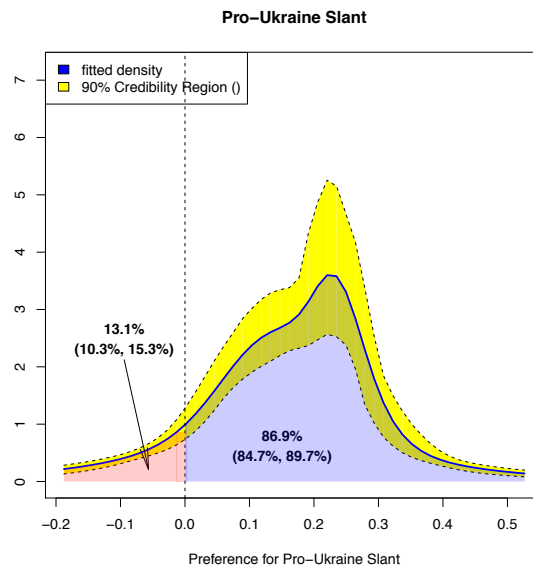
4. We also find that consumers preferences for internal-sensitive news (β_i^{Int}) and Ukraine-crisis news (β_i^{Ukr}) are positively correlated: 0.243 (0.206, 0.282).

Figure 6.2: Posterior distribution of preference for the Ukraine-crisis news and corresponding slant



(a) Preference for Ukraine-crisis News

(b) Preference for a pro-Russia Slant

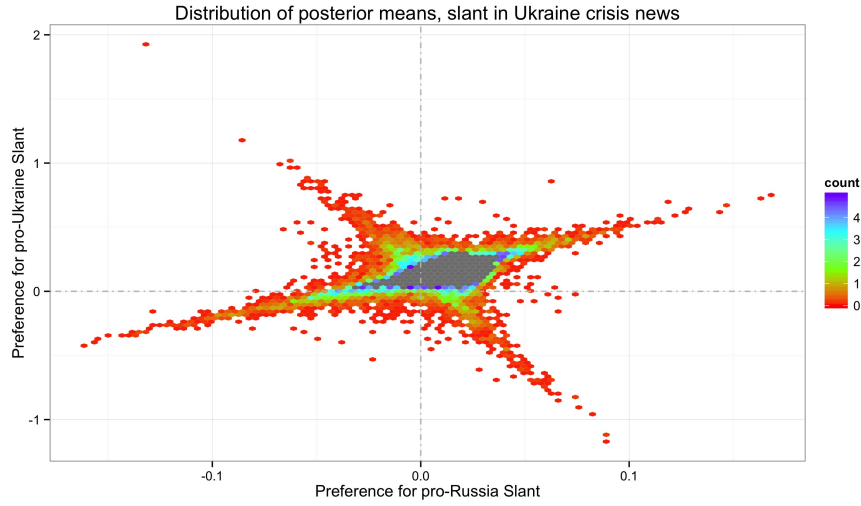


(c) Preference for a pro-Ukraine Slant

We normalize the scale of variables: values on the x axis represent effect of a one-standard-deviation increase in the corresponding variable on consumers' utilities. Blue line is a posterior mean; yellow shaded regions are 90% credibility regions. Numbers represent a percentage of probability mass of the distributions below and above zero, and 90% credibility intervals of these percentage.

preferences for ideological slant, which would result in γ_i^R and γ_i^U being negatively correlated across consumers. However, alternative theory is that consumers prefer the slant in the news in general. For example, consumers might be entertained by news articles about “terrorist” as well as by news articles about “fascists”. In this case, we would expect positive correlation between γ_i^R and γ_i^U .

Figure 6.3: Joint distribution of posterior means of preferences for pro-Russia and pro-Ukraine slant



Each dot represents preferences of a consumer. With the increase in the number of dots in the same location, color changes from red to dark-blue, as noted in the color code to the right of the figure.

Figure 6.3 plots the mean posterior preferences of consumers for pro-Russia and pro-Ukraine slants. As we saw before, the majority of consumer prefer news about the Ukraine crisis with both pro-Russia and pro-Ukraine slants. However, consumers tastes differ widely: some consumers prefer a less pro-Ukraine slant and a more pro-Russia slant (bottom-right corner), some consumers prefer a less pro-Russia slant and a more pro-Ukraine slant (top-left corner), and some consumers dislike both a pro-Russia or a pro-Ukraine slant (bottom-left corner). Thus, the figure suggests a non-linear relationship between consumer preferences. However, the overall relationship is positive, with the correlation between preferences across consumers of 0.206 (0.024, 0.328), which rejects the unidimensional preferences of consumers

for slant in Ukraine-crisis news.

Estimates of γ_i^R and γ_i^U reveal that the magnitude of the effect of news with pro-Russia and pro-Ukraine slant on consumers utility is different. Based on the results in Table 6.1, consumers get an order of magnitude more utility from the pro-Ukraine slant (average consumer gets 0.162 utils from a one-standard-deviation increase in the degree of slant) than from the pro-Russia slant (average consumer gets 0.015 utils). This suggests a nuanced role of government control in case of the Ukraine-crisis news. On the one hand, GC news outlets have more pro-Russia slant, which provides utility to the majority of consumers. On the other hand, GC news outlets have less pro-Ukraine slant, which leads to the disutility for the majority of consumers. Given that the pro-Ukraine slant delivers more utility to consumers, we would expect them to prefer the coverage of independent news outlets (which has more of the pro-Ukraine slant) even though GC news outlets have more of the pro-Russia slant. We test for this by comparing the utility individuals get from the Ukraine-crisis coverage by independent and GC news outlets. We record that consumer i dislikes the pro-government bias in Ukraine-crisis news on day t if

$$sl_{\text{Gov}}^R \hat{\gamma}_i^R + sl_{\text{Gov}}^U \hat{\gamma}_i^U < sl_{\text{Ind}}^R \hat{\gamma}_i^R + sl_{\text{Ind}}^U \hat{\gamma}_i^U$$

.

We find that on the day with an average amount of the Ukraine-crisis news, 82.3% (78%; 86%) of consumers get more utility from the coverage of the independent news outlets. If we weight the consumers by the estimated probability to visit GC news websites, the result is almost unchanged: 86.5% (84.1%, 89.1%) of consumers of GC news outlets dislike the pro-government bias. This implies that on the days with an average volume of the Ukraine-crisis news, 86.5% of the consumers of GC news outlets would prefer the Ukraine-crisis news coverage of independent news outlets, so they navigate to GC news websites despite the control. On the days with a large Ukraine-crisis news event, the share of GC consumers who

prefer independent coverage decreases to 78.1% (74.7%, 81.9%). We conclude that consumers navigate to GC news outlets despite government control: the majority of consumers of GC news outlets prefer the coverage of independent news outlets on the Ukraine-crisis news.

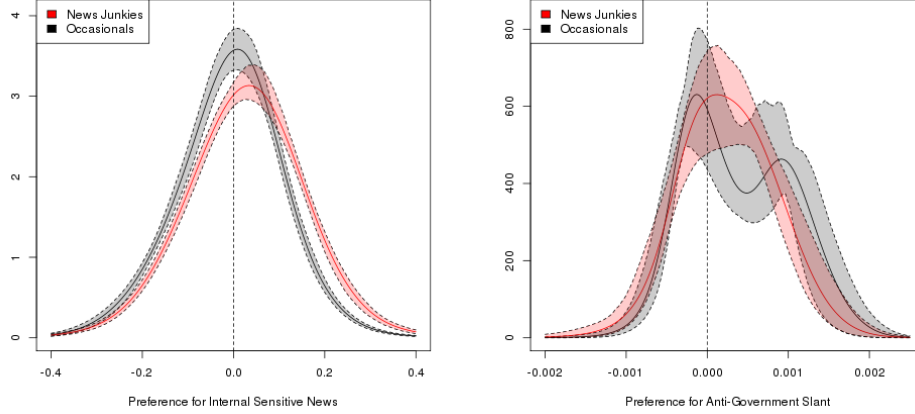
The preferences of consumers for the Ukraine-crisis news and the pro-Russia/pro-Ukraine slant are different from the preferences suggested by the descriptive analysis. In particular, we find that the majority of consumers prefer non-sensitive news to the Ukraine-crisis news (while Figures 4.1 and I.2 suggest that consumers prefer the Ukraine-crisis news) and that the majority of consumers prefer more pro-Russia slant in the news (while Figures 4.1 and I.2 suggest the opposite). These results are driven by the difference in preferences of frequent and infrequent news consumers. First, if we compute average preferences in the market weighting consumers by the number of consumption occasions, we find that an average consumer prefers the Ukraine-crisis news to non-sensitive news and does not like the pro-Russia slant in the coverage (columns 3 and 4 of Table 6.1). Second, we can split the sample of consumers to frequent (have more than 50 consumption occasions during the sample period⁵) and infrequent (everyone else) and examine the posterior distributions of the preference parameters.⁶ We label frequent news consumers as “news junkies” and everyone else as “occasional consumers”.

Figures 6.4 and 6.5 compare posterior distributions of preferences of the news junkies and the occasional consumers. Figure 6.4 reveals that news junkies have higher preference for internal-sensitive news (Subfigure a) and similar preference for anti-government slant (Subfigure b) compared to the occasional consumers. Figure 6.5 reveals that news junkies have higher preference for the Ukraine-crisis news (Subfigure a), lower preference for the pro-Russia slant (Subfigure b) and lower preference for the pro-Ukraine slant (Subfigure c)

5. Threshold chosen arbitrarily.

6. Frequency of consumption is an endogenous variable, which might reflect interest in the particular news topic and not preference for the news in general. To check the results, we define frequent news consumers as consumers with high average estimate of α_i across j , $\bar{\alpha}_i$, and find similar results.

Figure 6.4: Posterior distributions of preferences for internal-sensitive news and corresponding slant, by frequency of consumption: News junkies and occasionals



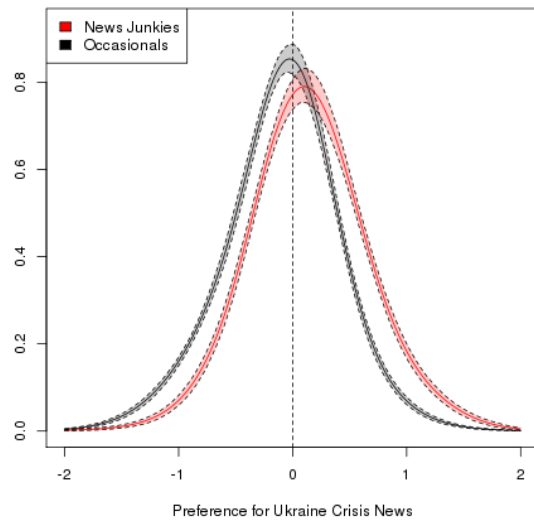
(a) Preference for Internal Sensitive News (b) Preference for Anti-Government Slant

We normalize the scale of variables: values on the x axis represent effect of a one-standard-deviation increase in the corresponding variable on consumers' utility. Black line correspond to occasional consumers, red line corresponds to news junkies. Shaded regions correspond to a 90% credibility region.

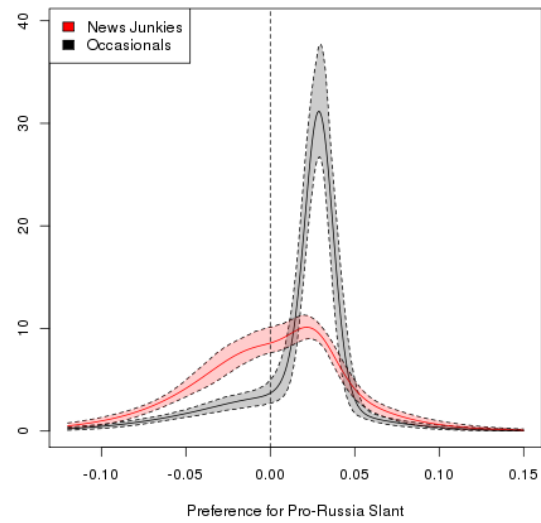
compared to the occasional consumers. The difference in preferences is especially large in the case of the pro-Russia slant: 47.6% (41.1%, 53.2%) of frequent news consumers get disutility from the pro-Russia slant ($\hat{\gamma}^R < 0$), compared to only 19.4% (16.1%, 23%) of the occasional consumers.

The difference in preferences of frequent and infrequent news consumers has important implications for the news markets in general. As we have seen in Figure 4.2, frequent news consumers represent only around 9% of all consumers, but they are responsible for around 70% of the consumption. Thus, frequent news consumers are very important for the online news outlets, given that the majority of their revenue comes from display advertisement, which is often priced as cost-per-impression. This implies that news outlets are incentivized to bias their coverage towards preferences of the frequent news consumers. However, this news coverage might not represent the preferences of the other 91% consumers, who are infrequent visitors. This might discourage the rest of the consumers from reading the news,

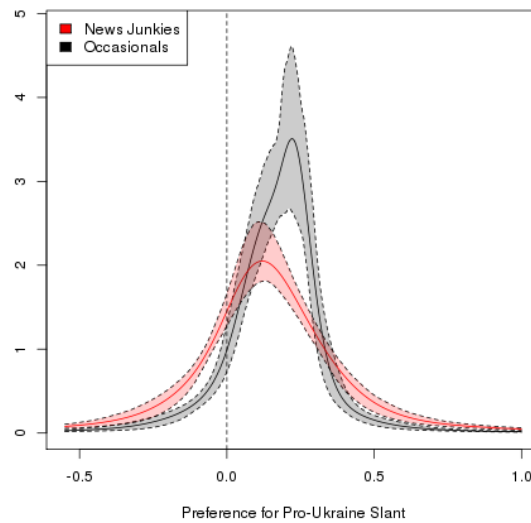
Figure 6.5: Posterior distributions of preferences for Ukraine-crisis news and corresponding slant, by frequency of consumption: News junkies and occasionals



(a) Preference for Ukraine Crisis News



(b) Preference for pro-Russia Slant



(c) Preference for pro-Ukraine Slant

We normalize the scale of variables: values on the x axis represent effect of a one-standard-deviation increase in the corresponding variable on consumers' utility. Black line correspond to occasional consumers, red line corresponds to news junkies. Shaded regions correspond to a 90% credibility region.

and they will be not informed. From the perspective of a benevolent government, which has a goal of providing information to all citizen so that they can make informed political decisions, this would be a bad outcome.

CHAPTER 7

COUNTERFACTUALS

Demand estimates reveal that consumers have heterogeneous preferences for the news coverage. This implies that the role of government control in this market is still ambiguous: even though we have found that the majority of consumers prefer the coverage of independent news outlets about the internal-sensitive news, a fraction of consumers has a preference for the pro-government bias in the news, so having the pro-government bias in the coverage might be optimal for some news outlets. To understand whether government control is a binding constraint for the news outlets, we need to examine their performance in the absence of control. In this section, we conduct the counterfactual simulations in which we change the level of control in the market and examine the market outcomes.

One approach to understanding the role of government control in the news market is to formalize the control constraints on the supply side and simulate the game between the news outlets once these constraints are removed. If the control constraints are indeed binding, such approach would allow to measure the reactions of independent news outlets to the higher levels of competition between the firms. This requires specifying the assumptions on the production function of the news outlets and the nature of the game they are playing. Estimating the supply-side of the model is still work in progress, and we plan to follow this approach in the later drafts of this paper.

Another approach is to examine the short-term effect of changing the coverage of the controlled news outlets. In the previous section, we have defined the pro-government bias as the difference in coverage between the GC and independent news outlets. Thus, we have assumed that independent news outlets are “unbiased”. To understand the role of government control, we match the reporting of independent news outlets by the controlled news outlets, thus making them also “unbiased”. If the formerly-controlled news outlets have better performance under this new coverage, we can conclude that government control

is a binding constraint: by exactly matching the coverage of independent news outlets, we create the highest level of competition possible, so if this is beneficial for the controlled news outlets, choosing their coverage optimally would be even more beneficial, and they will be better off in the absence of control.

We examine the effect of removing control both from the GC and influenced news outlets. The majority of GC news outlets are owned by the government, which Gehlbach and Sonin (2014) defined as direct control. At the same time influenced news outlets are controlled indirectly (Gehlbach and Sonin 2014), e.g. through encouraging the private owner to bias the news. By matching the coverage of GC and influenced news outlets to the independent news outlets, we can measure the role of direct and indirect control in this market. In particular, for the internal-sensitive news we match the volume of coverage and anti-government slant, rep_{jt}^{Int} and sl_j^{Int} , to remove the effect of censorship and propaganda; for the Ukraine-crisis news we match the level of pro-Russia and pro-Ukraine slant, sl_j^R and ls_j^U , to remove the effect of propaganda.

Table 7.1 presents the simulated market shares under different levels of government control. The market shares are averaged across the period of study. Column (1) presents the market shares under the current level of control. Column (2) presents the market shares under no direct government control: GC news outlets have the coverage of independent news outlets about sensitive news. The market share of GC news outlets increase by 0.231 percentage points, which is 5.9% of their current market share. This increase comes from the intrinsic (influenced and independent news outlets lose 0.048 and 0.03 percentage points of the market shares, respectively) and from the extrinsic margins (the share of consumers not visiting any news outlets decrease by 0.152). Column (3) present the market shares under no indirect control: influenced news outlets have the coverage of independent news outlets about sensitive news. Now the market share of the influenced news outlets increase by 0.261 percentage points (5.3% of their market share), which is driven by consumers switching from

GC (0.034) and independent news outlets (0.034), as well as the outside option (0.194). Column (4) present the market shares under no direct and indirect control. Both GC and influenced news outlets gain in the market shares (0.19 and 0.207, respectively), with the increase coming from the independent news outlets (0.061) and the outside option (0.339).

Table 7.1: Simulated market shares for different levels of government control.

	(1)	(2)	(3)	(4)	(5)
	$\bar{s}h$	Δ	Δ	Δ	Δ
	Current shares	No direct control	No indirect control	No control	More control
$\bar{s}h_{Gov}$	3.94 (3.91, 3.95)	+0.23 (0.2, 0.26)	-0.03 (-0.04, -0.03)	+0.2 (0.16, 0.22)	+0.02 (0.01, 0.02)
$\bar{s}h_{Inf}$	4.92 (4.89, 4.94)	-0.05 (-0.06, -0.04)	+0.26 (0.24, 0.28)	+0.21 (0.19, 0.22)	+0.03 (0.02, 0.03)
$\bar{s}h_{Ind}$	3.37 (3.45, 3.5)	-0.03 (-0.034, -0.027)	-0.034 (-0.04, -0.030)	-0.06 (-0.068, -0.053)	-0.16 (-0.183, -0.139)
$\bar{s}h_{Ukr}$	0.5 (0.49, 0.52)	+0.0005 (0.0001, 0.001)	+0.001 (0.0004, 0.002)	+0.002 (0.001, 0.003)	-0.001 (-0.002, -0.001)
$\bar{s}h_{Out}$	87.18 (87.11, 87.24)	-0.15 (-0.18, -0.13)	-0.194 (-0.211, -0.179)	-0.34 (-0.37, -0.3)	+0.12 (0.1, 0.13)

The market share are in % of the entire market.

Using the results in columns (1)-(4) of Table 7.1, we can conclude that both direct and indirect control constrain the news outlets. Without the direct control, GC news outlets would gain higher market shares, both with and without the indirect control. Same holds for the influenced news outlets: without the indirect control, they would also gain higher market shares. Given that GC and influenced news outlets might decide not to match the independent news outlets exactly once control is removed, an actual increase in their market shares might be even higher. Interestingly, the major part of this increase comes from the extrinsic margin, which implies that control conditions reduce the overall size of the advertising market.

We can do a back-of-the-envelope calculation of the loss in the advertising revenue of GC and influenced news outlets due to the direct and indirect control constraints. We could not find information on the overall size of advertising market for the online news

outlets in Russia. However, the majority of advertising revenue for these news outlets come from display advertising. The total advertising expenditure in Russian internet on display advertising in 2014 was 19.1 billion rubles¹, which is around 360 million USD using the exchange rate of 60 rubles for a dollar of the end of 2014.² Even if we assume that the online news outlets deliver all display advertising in Russia, the amount of revenue GC and influenced news outlets lose due to the control is tiny: if we remove direct control, GC news outlets would gain 0.231 percentage points in their market shares, which is around \$6.5 millions, while if we remove the indirect control, influenced news outlets would gain 0.261 percentage points, which is around \$7.3 millions.³ Thus, an annual loss of all the controlled news outlets due to government control is around \$13.8 millions. For comparison, government subsidies to mass media in Russia in 2015 were 72.6 billion rubles (\$1.21 billion),⁴ which is 88 times more than the loss of online news outlets in the advertising revenue due to the government control.

Alternative counterfactual of interest is the online news market under *more* government control. By the middle of 2016, online news market in Russia did not become less controlled; in fact, the market became more controlled: several independent news outlets had to change their ownership due to a new law⁵ and rbc, one of the top online news outlets in Russia, had to change the editorial team due to the government pressure⁶. With these developments, it

1. http://www.akarussia.ru/knowledge/market_size

2. The exchange rate of ruble to dollar has changed dramatically in 2014, increasing for 33 to 60 rubles for a dollar during the year. We use the exchange rate of 60 rubles per dollar given that it stayed at this level at least for one and a half year after the end of 2014. Source: <http://www.xe.com/currencycharts/?from=USD&to=RUB&view=5Y>

3. With the current levels of control, served market is $100\% - 87.175\% = 12.825\%$. If we assume that each consumption occasion is equally valuable for news outlets, 0.235 percentage points is $\$360 \text{ million} / 12.825\% = \2.77 billion , and 0.272 percentage points is $\$360 \text{ million} / 12.825\% = \2.77 billion .

4. Source: <http://www.rbc.ru/politics/29/06/2015/55912ffa9a7947453982cda9>. Same exchange rate used. The sum of 72.6 billions rubles includes subsidies to the television and print.

5. <https://rg.ru/2016/01/01/smi-site-anons.html>

6. http://www.bbc.com/russian/news/2016/05/160513_rbc_badanin

is interesting to study what would happen if all independent news outlets become indirectly controlled by the government. To examine this, we match the sensitive topics coverage of independent news outlets to the influenced ones in the counterfactual simulation. Column (5) of Table 7.1 presents the average market shares under this change. Independent news outlets lose 0.159 percentage points of their market share, which is around 4.6% of their market shares. The majority of these consumers leave the market (0.115 percentage points), and some switch to GC (0.017) and influenced (0.028) news outlets. Using the same back-of-the-envelope calculations as above, we get that independent news outlet will lose at most \$4.5 million in annual advertising revenues. Thus, if independent news outlet only care about the advertising revenue, it is relatively cheap for the government to bribe them.

7.1 Effectiveness of Government Control

Changes in consumption due to changes in the level of control allows us to understand its effectiveness. For now, we define the effectiveness of government control by the fraction of news readers who stay with the news outlet once control is imposed.⁷ E.g. if all consumers of GC and influenced news websites keep consuming these news outlets once they become controlled, government control is 100% effective. If all of the consumers switch away from these news outlets, control is completely not effective. This way, effectiveness of control depends on the rigidity of consumers' choices of news outlets, with more rigidity implying more effective control.

Using the results of Table 7.1, we can conclude that government control is highly effective. Under the control, GC news outlet lose only 5.9% of their visitors, and influenced news outlets lose only 5.3% of their visitors. Thus, controlled news outlets are able to maintain around 94% of consumers, so control is 94% effective. Moreover, the majority of consumers who

7. A better measure of effectiveness of government control should incorporate (1) preferences of consumers for news coverage and (2) where consumers decide to switch to. We are currently working on such measure.

switch away from the controlled news outlets stop reading the news, which implies that they do not get the unbiased coverage on the independent news outlet. Only a small fraction of consumers navigate to the independent news outlet and read the unbiased coverage.

Results in Table 7.1 are averages across the days with different levels of sensitive news. However, effectiveness of control on the days with large sensitive events is more important for the government, as on these days it is more important to prevent people from getting the unbiased coverage of the news. Tables L.1 and L.2 in the Appendix L present the simulated market shares for the days with large internal-sensitive and Ukraine-crisis news events, respectively. On the days with a large internal-sensitive news event, GC news outlets lose 0.491 percentage points of their market shares (12.2%) once direct control is imposed; the change in the market share of influenced news outlets is not significant once the indirect control is imposed. On the days with a large Ukraine-crisis news event, GC news outlets lose 0.639 percentage points of their market shares (11%) once direct control is imposed; influenced news outlets also lose 0.639 percentage points of their market share (9.7%) once indirect control is imposed. Overall, on these days control is less effective, but the controlled news outlet can still maintain around 90% of their customers.

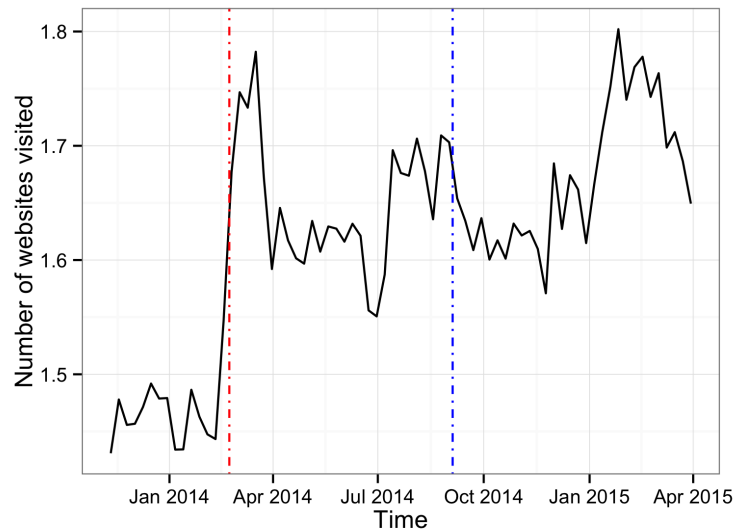
Even though we find that government control is effective, results suggest that it is limited: some consumers switch away from the GC news outlets. In this way, our findings support conclusions of Knight and Tribin (2016) that consumers actively switch away from the news sources to avoid propaganda.

CHAPTER 8

NEWS CONSUMPTION FOR INFORMATION

In this section, we extend the demand model to allow for the interdependence of consumers' choices within a day. Typically, news consumers navigate to one news outlet on a given day: 77% of consumers' browsing sessions include a visit to only one news outlet. However, news consumers sometimes prefer to read multiple news outlets. This behavior might be especially important on the days with a lot of sensitive news: if consumers want to read about government-sensitive topics, they might be more likely to navigate to multiple news outlets. Figure 8.1 plots an average number of news outlets consumers have visited, conditional on reading the news on a given day. On average, consumers navigate to 1.61 news outlets. The number of news outlets consumed increases at the beginning of the Ukraine-crisis, suggesting consumption of multiple news outlets is more important on the days with more sensitive news.

Figure 8.1: Average number of websites consumers visit on a day when they consume news



Red line corresponds to GC media, green line - to independent media, blue line - to government-influenced media. Red dotted line corresponds to February 22, 2014, day when the former president Yanukovich fled Ukraine as a result of revolution. Blue dotted line corresponds to the first Minsk Peace agreement on September 4, 2014

Allowing interdependence of consumers’ choices within a day is important for two reasons. First, misspecification of the model might lead to biased estimates of consumers’ preferences. For example, if consumers navigate to ideologically more diverse news outlets on the days with a large volume of sensitive news in order to become better informed about the news, estimates of consumers’ preferences for slant will be biased.

Second, the question of why consumers read multiple news outlets is interesting by itself. Two explanations are possible. One is that consumers are trying to get multiple points of view about the topic to filter out the bias. In other words, consumers are “sophisticated” news readers, and they know bias exists.¹ In this case, on the days with more sensitive news events we would expect news consumers to visit websites that are on the opposite sides of the ideological spectrum. A different explanation is that consumers are trying to get more information simply because they are interested in the topic, while still being subject to confirmation bias. In other words, consumers are “naive” to the existence of bias in the news, and on the days with more sensitive news events we should see people navigating to websites that are ideologically similar.

We extend the model to allow for such interdependencies. On each day t , consumer i can have multiple news consumptions $T_{it} = \{1, \dots, M+1\}$, where M is the total number of news outlets in the market. We define news consumption of consumer i in the news outlet space, and treat each news consumption event as a discrete-choice problem with the same number of choices available.² That is, at each day t , consumer i chooses the outlets sequentially on occasions $\tau = \{1, \dots, T_{it}\}$, where in the last choice occasion T_{it} he chooses an outside

1. Mullainathan and Shleifer (2005) and Xiang and Sarvary (2007) refer to sophisticated news consumers as “conscientious readers”.

2. If consumer i chose GC news outlet on day t , there are 9 more government controlled news outlets which he did not consume, so he can still choose a GC outlet on the next consumption occasion on the same day. We do not observe any consumer exhausting the number of news outlets of some type on a given day.

option. The consumer's utility is defined as

$$\begin{aligned}
u_{ijt\tau} = & \alpha_{ij} + \text{rep}_{jt}^{Int} \left(\beta_i^{Int} + sl_j^{Int} \gamma_i^{Int} + |sl_j^{Int} - sl_{y_{i\tau-1}}^{Int}| I(\tau > 1) \rho_i^{Int} \right) + \\
& + \text{rep}_{jt}^{Ukr} \left(\beta_i^{Ukr} + sl_j^R \gamma_i^R + sl_j^U \gamma_i^U + |sl_j^{Ukr} - sl_{y_{i\tau-1}}^{Ukr}| I(\tau > 1) \rho_i^{Ukr} \right) + \\
& + |sl_{jt}^{Int} - sl_{y_{i\tau-1}}^{Int}| I(\tau > 1) \zeta_i^{Int} + |sl_{jt}^{Ukr} - sl_{y_{i\tau-1}}^{Ukr}| I(\tau > 1) \zeta_i^{Ukr} + s_{it\tau} \eta_i + \epsilon_{ijt\tau},
\end{aligned} \tag{8.1}$$

where $sl_{jt}^{Ukr} = sl_{jt}^R - sl_{jt}^U$, the overall difference in the slant in the Ukraine-crisis coverage. The measure $|sl_j^x - sl_{y_{i\tau-1}}^x|$ captures the ideological distance between consumer i 's current and previous choice on day t . Coefficient ζ_i^x captures people's taste for different slants relative to their previous choice on day t . Coefficient ρ_i^x captures the change in consumers' preferences for different slants in the news on the days with more sensitive news. State variable $s_{it\tau}$ is an indicator variable equal to 1 if type j was consumed on day t on one of the previous choice occasions $1, \dots, \tau - 1$. Coefficient η captures both the decrease in utility consumer gets from news of a similar type because he was exposed to this information, and any potential correlation in the unobserved shock for news outlets of a similar type.

To test whether consumers are “sophisticated” or “naive” regarding slant in the news, we focus on the ρ_i^x coefficient. Negative ρ_i^x implies consumers are more likely to read news of a similar ideological slant on the days with more sensitive news, suggesting they are naive. Positive ρ_i^x implies consumers prefer more ideologically diverse news on the days with more sensitive news, suggesting they are sophisticated. We take the model to the data by estimating a Bayesian hierarchical multinomial logit model with heterogeneity approximated with the mixture of normals. We focus our estimation on the subset of the “news junkies”, for whom multiple news consumption occasions within a day are more common: we use people who have consumed news 50-100 times within the observation period.³

To measure the importance of misspecification of the original model on the estimates,

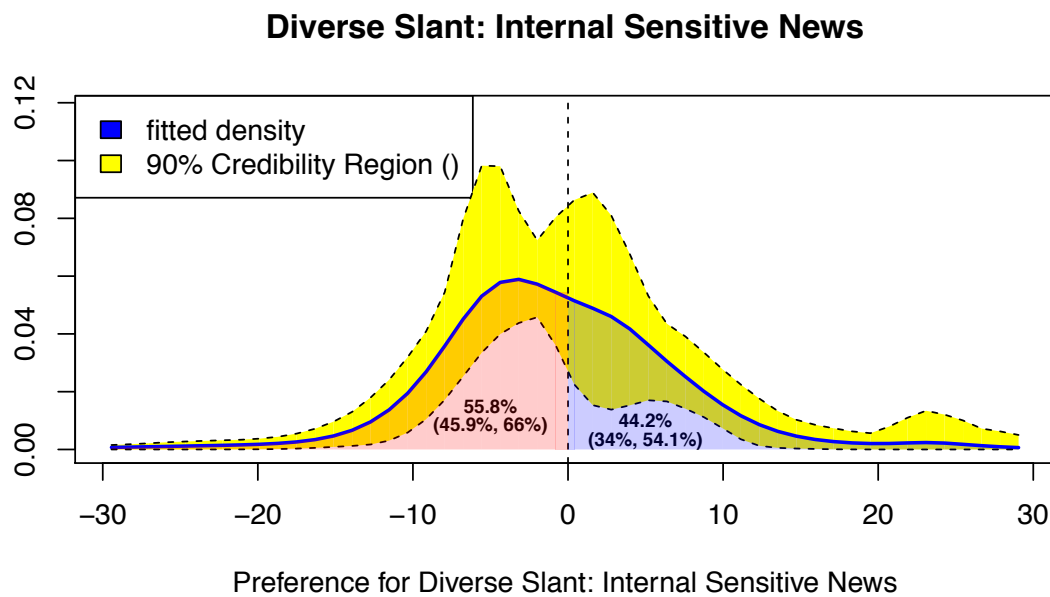
3. Estimation with a more diverse subset of consumers is in progress.

we re-estimate the original model only for the news junkies, and compare the two sets of estimates. Appendix M presents the estimates from the base and extended model. Allowing for the interdependence of consumers' choices within the day does not change the overall conclusions: the majority of news junkies prefer sensitive news to non-sensitive news, and prefer less pro-government slant in the news. However, the estimates of consumers' preferences for slant are significantly different in the extended model: more consumers dislike the anti-government slant in the internal-sensitive news and dislike a pro-Ukraine slant in Ukraine-crisis news. We conclude that extending the basic model of demand to allow for interdependencies of consumers' choices across time is important. We plan to perform this extension in future work.

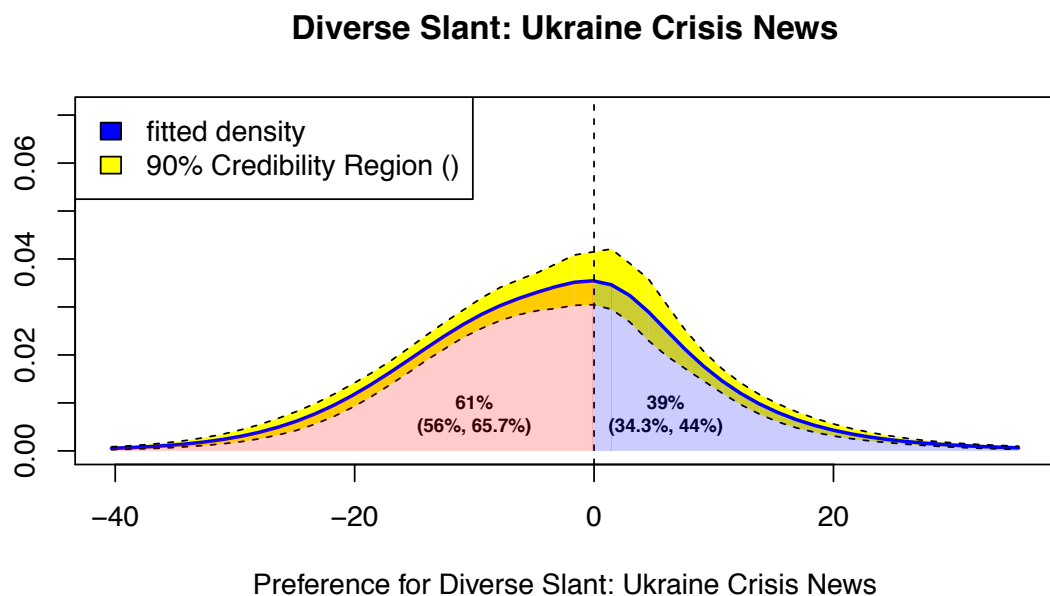
We now examine whether news junkies behave more like sophisticated or naive types. To do so, we focus on ρ_i^x estimates. Figure 8.2 presents the posterior distributions of ρ_i^{Int} and ρ_i^{Ukr} . The majority of consumers have negative ρ_i^x estimates: 55.8% (45.9%, 66%) of consumers prefer ideologically more similar news on the days with more internal-sensitive news, and 61% (56%, 65.7%) of consumers prefer ideologically more similar news on the days with more Ukraine-crisis news. We conclude the majority of news consumer behave more like naive types, whereas some news consumers behave like sophisticated types.

If positive ρ_i^x coefficients indeed signal that consumers are sophisticated, such consumers should be more likely to search for more diverse information both on days with more internal-sensitive news and on days with more Ukraine-crisis news. Thus, we would expect that ρ_i^{Int} and ρ_i^{Ukr} are positively correlated across consumers. Figure 8.3 shows individual-level posterior means of ρ_i^{Int} and ρ_i^{Ukr} . A strong positive correlation in the preferences exists (0.495, with a 90% credibility interval of 0.366,0.632), suggesting news consumers who are sophisticated readers of internal-sensitive news are more likely to be sophisticated readers of news about the Ukraine crisis.

Figure 8.2: Posterior distribution of consumers' preference for ideologically diverse news (ρ_i^x)



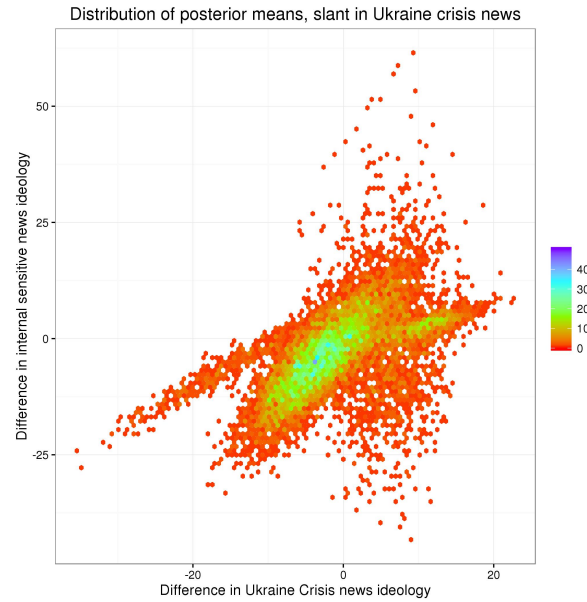
(a) Internal-sensitive News



(b) Ukraine-crisis News

We do not normalize the scale of variables. Blue line is a posterior mean; yellow shaded regions are 90% credibility regions. Numbers represent a percentage of probability mass of the distributions below and above zero, and 90% credibility intervals of these percentage.

Figure 8.3: Individual-level posterior means of preferences for ideological diversity in government-sensitive news (ρ_i^{Int} and ρ_i^{Ukr})



Each dot represent preferences of a consumer. With the increase in number of dots in the same location, color changes from red to dark-blue, as noted in the color code to the right of the figure.

CHAPTER 9

CONCLUSION

In this paper, we explore consumers' preferences for news coverage in the controlled online news market. Using a case study of a Russian online news market, we show that the majority of consumers visit the controlled news outlets despite the pro-government bias. Counterfactual simulations show that the controlled news outlets will have higher market shares without the control constraint. Assuming online news outlets are responsible for all display advertising revenues in the Russian internet, their market shares increase corresponds only to \$13.8 million increase in the annual advertising revenues. In reality, we expect this number to be lower, given that the share of online news outlets in the overall display advertising in Russia is less than 100%. For comparison, annual government subsidies to the controlled news outlets are more than \$1.2 billion, which implies that news outlets lose only a small fraction of their budgets due to the government control.

There is substantial heterogeneity in consumer preferences. In particular, we find that frequent news consumers get more disutility from the biased news. The difference in preferences of frequent and infrequent news consumers has important implications for the news markets in general. In this market, frequent news consumers are only 9% of all consumers, but they are responsible for around 70% of the consumption. Thus, they are important for news outlets, which get the majority of revenue from display advertisement, priced as cost-per-impression. This implies that news outlets are incentivized to bias their coverage towards preferences of the frequent news consumers, and are less incentives to tailor the news coverage to preferences of the rest of consumers. The news coverage fitted to the frequent news consumers might discourage the rest of consumers from reading the news, keeping them not informed. From the perspective of a benevolent government, which has a goal of providing information to all citizen so that they can make informed political decisions, this is a bad outcome.

This shows that business incentives of the news outlets in the free market economy and political incentives of the benevolent government are not well aligned. However, if online advertising is more effective for the infrequent news consumers and an appropriate advertising pricing schedule is cost per unique visitor, business and political incentives are more aligned.¹ In short, understanding the effectiveness of online advertising is crucial for understanding how well a free online news market will provide information to the citizens.

We show that consumers’ ideological preferences are not necessarily unidimensional. Using the example of slant in the news about the Ukraine crisis, we show that consumers who prefer the pro-Russia bias in the news (“Ukraine are fascists”) are more likely to prefer the pro-Ukraine bias (“Russia is in aggressor”) in the news. This suggests that consumers prefer biased news not only because of the confirmation bias (Mullainathan and Shleifer 2005, Gentzkow and Shapiro 2010), but also for other reasons, such as entertainment.

Finally, we test if frequent news consumers read the news with the pro-government bias because they prefer the news with such bias (which we call “naive” consumers), or because they would like to learn the extent of the pro-government bias, comparing it to the unbiased version of the news (which we call “sophisticated” consumers). We find that around 40% of consumers behave like sophisticated types, and such behavior is consistent across the sensitive news topics.

We believe this paper makes several contributions to the literature. To the best of our knowledge, we are the first to estimate a structural model of demand for news in the online news market under government control. To do this, we use two novel data sets (publication records and individual-level browsing data) and use a novel identification strategy to disentangle consumers preferences for news outlets versus their coverage on government-sensitive topics. Our results contribute to the theoretical and empirical literature about the demand

1. Advertising might be more effective for infrequent news consumers e.g. because for them the informational effects of advertising (Akerberg 2001, Blake et al. 2015) might be stronger, given that they are less informed almost by definition.

for bias in the news (Mullainathan and Shleifer 2002, Mullainathan and Shleifer 2005, Xiang and Sarvary 2007, Gentzkow and Shapiro 2010, Zhu and Dukes 2014, etc.) as well as to the literature on government control of the media (Durante and Knight 2012, Enikolopov et al. 2011, Gehlbach and Sonin 2014, Prat 2014, Garcia-Arenas 2016, Knight and Tribin 2016, etc.). In particular, our results suggest that consumers demand biased news not only because of the confirmation bias, assumed to be the main driver of demand for biased news in the recent works.² We also find the evidence of “conscientious readers” (Mullainathan and Shleifer 2005, Xiang and Sarvary 2007), consumers who visit multiple news outlets to discount the bias. Finally, our findings support conclusions of Knight and Tribin (2016) that consumers actively switch away from the news sources to avoid propaganda.

Our work has a number of important limitations and potential improvements. On the conceptual side, we analyze the consumption of news on the internet. Adding other news mediums (such as TV) to the choice set of consumers will provide a more complete picture of the news consumption. If consumers complement TV consumption with reading news online, we might underestimate consumers’ preferences for pro-government bias, given that all three biggest TV channels in Russia are controlled by the government. In addition, the current draft does not contain a supply-side model. Accounting for the supply-side is important for measuring the competitive effects in the counterfactual simulations, as well as for understanding the effect of various advertising pricing schedules on the degree of competition in this market and on the effectiveness of government control.

The methodological side allows for a number of potential improvements of our work. First, currently, we ignore the information about the topics of news articles consumers click on, and the amount of time people spend on these news articles. Incorporating this information should allow us more precise estimates of preferences for news topics. Second, in this work, we assume people’s preferences are fixed over time. At the same time, we know people’s

2. Mullainathan and Shleifer (2002) is one of the few exceptions, with authors discussing the difference between media slant (confirming the beliefs of consumers) and media spin (creating a memorable story).

preferences for different topics can change over time. Incorporating this fact in the model will allow us to examine whether consumers' preferences get more segregated as a large news topic (e.g., the Ukraine crisis) evolves. Finally, we measure only the relative importance of different news topics, because we base it on the average reporting in the news market. Collection of the timeline of facts happening in the conflict would allow us to examine the supply-side decision of news outlets in more detail.

Although we study a particular case of news consumption, we believe the question we study has implications outside of the media economics literature. We estimate consumers' preferences for bias in information. Biased information exists not only in the news; consumers often search for information about a large variety products, and companies need to decide on the most effective ways to deliver this information to customers. For example, if consumers prefer biased information, companies might want to introduce some degree of bias to their advertising, making it more subjective, whereas if consumers dislike the bias, companies might decide to make the advertising message more objective. We hope this work will encourage more research on the role of bias in the demand for information in different areas of economics and marketing.

REFERENCES

- Ackerberg, Daniel (2001): “Empirically Distinguishing Informative and Prestige Effects of Advertising,” *RAND Journal of Economics*, 32(2), 316-333.
- Athey, Susan, Markus Mobius and Jenő Pal (2017): “The Impact of Aggregators on Internet News Consumption,” Working paper.
- Blake, Thomas, Chris Nosko and Steven Tadelis (2015): “Consumer Heterogeneity and Paid Search Effectiveness: A Large-scale Field Experiment,” *Econometrica*, 1-26.
- Anderson, Simon, David Strömberg and Joel Waldfogel (2016): *Handbook of Media Economics* 1st Edition, North Holland.
- Bai, Jie, Mikhail Golosov, Nancy Qian and Yan Kai (2014): “Understanding the Influence of Government Controlled Media: Evidence from Air Pollution in China,” Working Paper.
- Baron, David (2006): “Persistent Media Bias,” *Journal of Public Economics*, 90: 1-36.
- Besley, Timothy and Robin Burgess (2002): “The Political Economy of Government Responsiveness: Theory and Evidence from India,” *Quarterly Journal of Economics*, 117: 1415-51.
- Besley, Timothy and Andrea Prat (2006): “Handcuffs for the Grabbing Hand? Media Capture and Government Accountability,” *American Economic Review*, 96: 720-736.
- Cagé, Julia, Nicolas Hervé and Marie-Luce Viaud (2017): “The Production of Information in an Online World: Is Copy Right?,” Working paper.
- Della Vigna, Stefano and Ethan Kaplan (2007): “The Fox News Effect: Media Bias and Voting,” *Quarterly Journal of Economics* 122: 1187-1234.

- Dubé, Jean-Pierre, Günter Hitsch and Peter Rossi (2010): “State Dependence and Alternative Explanations for Consumer Inertia,” *RAND Journal of Economics*, 41: 417-445.
- Durante, Ruben and Brian Knight (2012): “Partisan control, media bias, and viewer responses: Evidence from Berlusconi’s Italy,” *Journal of the European Economic Association*, 10: 451-481.
- Enikolopov, Ruben, Maria Petrova and Ekaterina Zhuravskaya (2011): “Media and Political Persuasion: Evidence from Russia,” *American Economic Review*, 101: 3253-85.
- Garcia-Arenas, Javier (2016): “The impact of free media on regime change: Evidence from Russia,” Working Paper.
- Gehlbach, Scott and Konstantin Sonin (2014): “Government Control of the Media,” *Journal of Public Economics*, 118: 163-171.
- Gentzkow, Matthew, Michael Sinkinson, and Jesse M. Shapiro (2011): “The Effect of Newspaper Entry and Exit on Electoral Politics,” *American Economic Review* 101: 2980-3018.
- Gentzkow, Matthew and Jesse M. Shapiro (2006): “Media Bias and Reputation,” *Journal of Political Economy*, 114: 280-316.
- Gentzkow, Matthew and Jesse M. Shapiro (2010): “What Drives Media Slant? Evidence from U.S. Daily Newspapers,” *Econometrica* 78: 35-71.
- Gentzkow, Matthew and Jesse M. Shapiro (2015): “Ideology and Online News,” NBER Working Paper 19675.
- Gentzkow, Matthew, Jesse M. Shapiro and Matt Taddy (2016): “Measuring Polarization in High-Dimensional Data: Method and Application to Congressional Speech,” NBER Working Paper 22423.

- Knight, Brian and Ana Tribin (2016): “The Limits of Propaganda: Evidence from Chavez’s Venezuela” NBER Working Paper 22055.
- Martin, Gregory J. and Ali Yurukoglu (2015): “Bias in Cable News: Persuasion and Polarization,” Working paper.
- McCombs Maxwell E. , and Donald L. Shaw (1972): “The Agenda-Setting Function of Mass Media,” *Public Opinion Quarterly* 36: 176-187.
- Mullainathan, Sendhil and Andrei Shleifer (2002): “Media Bias,” NBER Working Paper 9295.
- Mullainathan, Sendhil and Andrei Shleifer (2005): “The Market for News,” *American Economic Review* 95: 1031-1053.
- Pan, Jennifer and Yiqing Xu (2017): “China’s Ideological Spectrum,” *The Journal of Politics*, forthcoming
- Perego, Jacopo and Sevgi Yuksel (2016): “Media Competition and the Source of Disagreement,” Working paper
- Prat, Andrea and David Strömberg (2013): “The Political Economy of Mass Media,” *Advances in Economics and Econometrics: Volume 2, Applied Economics*.
- Prat, Andrea (2014): “Media Power,” CEPR Discussion Paper N. 10094.
- Roberts, Margaret E. (2014): “Fear or Friction? How Censorship Slows the Spread of Information in the Digital Age,” Working Paper.
- Rossi, Peter (2014): “Bayesian Non- and Semi-parametric Methods and Applications,” Princeton University Press.
- Rossi, Peter, Greg Allenby and Robert McCulloch (2005): “Bayesian Statistics and Marketing,” Wiley Series in Probability and Statistics.

Sen, Ananya and Pinar Yildirim (2016): “Clicks Bias in Editorial Decisions: How does Popularity Shape Online News Coverage?,” Working paper.

Xiang, Yi and Miklos Sarvary (2007): “News Consumption and Media Bias,” *Marketing Science*, 26(5): 611-628.

Zhu, Yi and Anthony Dukes (2015): “Selective Reporting of Factual Content by Commercial Media,” *Journal of Marketing Research*, 52(1): 56-76.

APPENDIX A

NEWS OUTLETS CLASSIFICATION

Classification is done based on interviews with media professionals who preferred to stay anonymous. At the same time, groups are supported by the ownership structure and existing media reports.

Government-controlled news outlets:

- *vesti*, *1tv*, *tass*, *rg*, *rt* and *ria* are owned by the government.
- *aif* is owned by Moscow city hall.
- *ntv* is owned by Gazprom, a state-owned gas monopolist.
- *vz* and *dni* were founded by Konstantin Rykov, a member of United Russia (incumbent political party) who led the political campaigns in support of Vladimir Putin in 2007. *vz* is owned by the Institute of Socio-Economics and Political Research, which is managed by Dmitry Badovsky, a former deputy chief of the Presidential Administration of Russia (2012).

Oligarchic news outlets:

- *lenta* and *gazeta* are owned by Alexander Mamut. Both were considered independent at the beginning of 2013. *Gazeta* changed an independent editor-in-chief to a more government-loyal editor-in-chief in September 2013; *lenta* got a similar change in March of 2014¹.
- *izvestia* is owned by Yuri Kovalchuk through the National Media Group (NMG). Yuri Kovalchuk is a close friend of Vladimir Putin.

1. <https://meduza.io/feature/2016/05/17/12-redaktsiy-za-pyat-let>

- *lifenews* is owned by Aram Gabrelyanov, a manager of NMG².
- *kommersant* is owned by Alisher Usmanov, one of the richest Russian oligarchs³.
- *kp* is owned by Grigory Berezkin, who is on board of directors of state-owned RZD⁴.
- *fontanka* is owned by “Azur-Media”.

Potentially government-influenced news outlets:

- *bfm* is owned by Rumedia, a company of Russian steel tycoon Vladimir Lisin⁵.
- *echo* is jointly owned by journalists of *echo* (34%) and a state-owned gas monopolist Gazprom (66%). One of the most famous Russian independent media, it is reported to be influenced by the government. Reported to publish paid articles⁶.
- *interfax*’s beneficiary is not disclosed. It is argued to be top-management⁷.
- *mk* is owned by Pavel Gusev, a confidant of Vladimir Putin. There are examples of *mk* removing published articles about government-sensitive topics⁸.
- *znak* was formerly *ura.ru*, It had to change its name due to government pressure⁹. It’s likely to be independent¹⁰.

2. <http://www.kommersant.ru/doc/2311510>

3. <https://lenta.ru/lib/14164974/>

4. <http://www.forbes.ru/profile/grigorii-berezkin>

5. https://en.wikipedia.org/wiki/Vladimir_Lisin

6. <https://tjournal.ru/p/media-denim>

7. https://www.vedomosti.ru/business/articles/2012/01/19/lgota_dlya_smi

8. <http://www.rbc.ru/politics/27/12/2013/897386.shtml>

9. <http://www.forbes.ru/news/227994-redaktsiya-uraru-budet-vypuskat-internet-gazetu-znakcom>

10. <http://theins.ru/politika/5463>

- *ng* is owned by Konstantin Remchukov. It's reported to publish articles which are paid for by the government¹¹.
- *polit's*, *utro's* and *ridus's* ownerships are unclear.
- *regnum* is reported to have been purchased by Gazprom media¹². It's reported to publish paid articles¹³.
- *rosbalt*, *sobesednik* and *trud* are reported to publish paid articles¹⁴.

Independent news outlets:

- *newsru* is owned by Vladimir Gusinsky, a tycoon opposing the incumbent Russian government since 2001.
- *newtimes* is owned by a non-profit fund The New Times Foundation.
- *novayagazeta* is owned by journalists (76%), Alexander Lebedev (14%) and Mikhail Gorbachev (10%).
- *rbc* and *snob* are owned by Mikhail Prokhorov, a Russian billionaire and politician. He participated in the presidential elections of 2012. RBC.ru stayed independent till May 2016, when the top managers were fired due to political pressure¹⁵.
- *slon* and *tvrain* are owned by Alexander Vinokurov and Natalia Sidneeva. *tvrain's* TV channel was taken off air by major TV providers after covering the street protests of 2011. It's website operates based on subscriptions.

11. <http://theins.ru/politika/6015>

12. <https://lenta.ru/news/2014/06/20/media/>

13. <https://tjournal.ru/p/media-denim>

14. <https://tjournal.ru/p/media-denim>

15. <http://www.newyorker.com/news/news-desk/the-demise-of-rbc-and-investigative-reporting-in-russia>

- *vedomosti* was jointly owned by Sanoma Independent Media (33%), Financial Times (33%) and The Wall Street Journal (33%) before the end of 2015. It was sold to Demyan Kudryavsev in November 2015 due to the a new law limiting foreign ownership of media to 20% starting in 2016.
- *forbes* was owned by Axel Springer before the end of 2015. It was sold to Alexander Fedotov in October 2015 due to a new law limiting foreign ownership of media to 20% starting in 2016.
- *the-village* is owned by Look at Media, which is registered in The Netherlands.

International News Outlets:

- *bbc* is a Russian version of BBC.
- *svoboda* is Radio Liberty.
- *meduza* is a news outlet founded in Latvia by a former journalists of lenta.ru, who were fired in March 2014 due to Ukraine Crisis coverage.
- *dw* is a Russian version of Deutsche Welle.
- *reuters* is a Russian version of Reuters.

A small subset of Ukrainian news outlets: *korrespondent*, *liga* and *unian*.

APPENDIX B

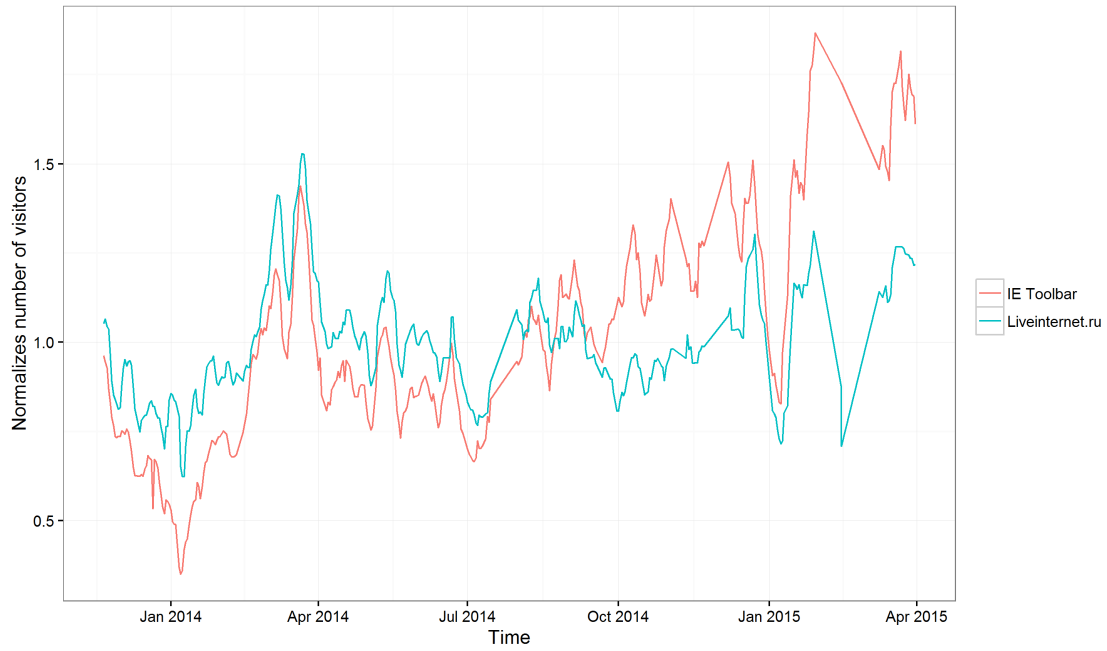
COMPARISON OF IE TOOLBAR AND LIVEINTERNET.COM

DATA

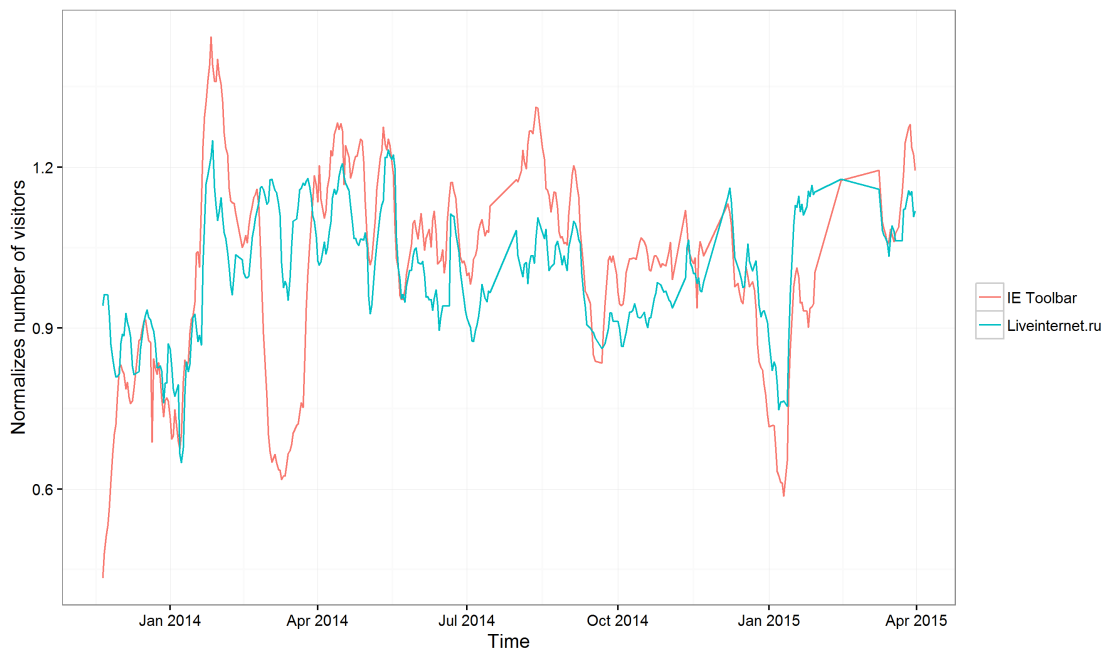
Table B.1: Comparison of website rankings in IE Toolbar and LI.ru

Website	Ranking in LI.ru	Ranking in IE Toolbar
rbc	1	1
ria	2	14
vesti	3	2
kp	4	4
lenta	5	8
gazeta	6	5
rg	7	3

Figure B.1: Normalized number of weekly visitors to top second and third news websites, IE Toolbar and Liveinternet.ru

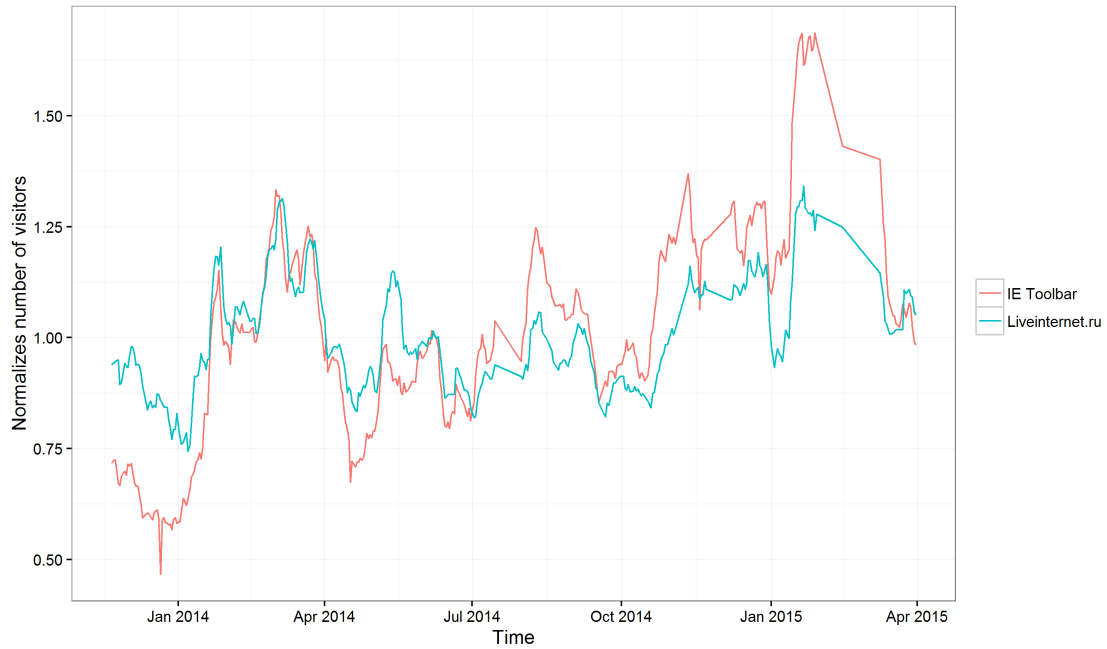


(a) ria.ru

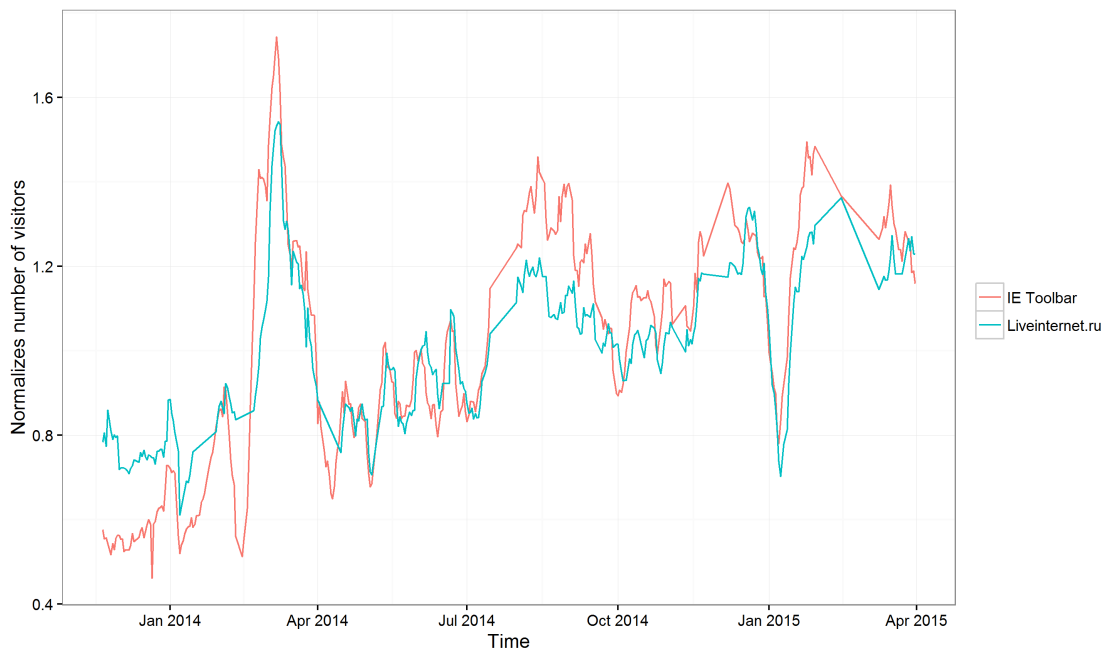


(b) vesti.ru

Figure B.2: Normalized number of weekly visitors to top fourth and fifth news websites, IE Toolbar and Liveinternet.ru

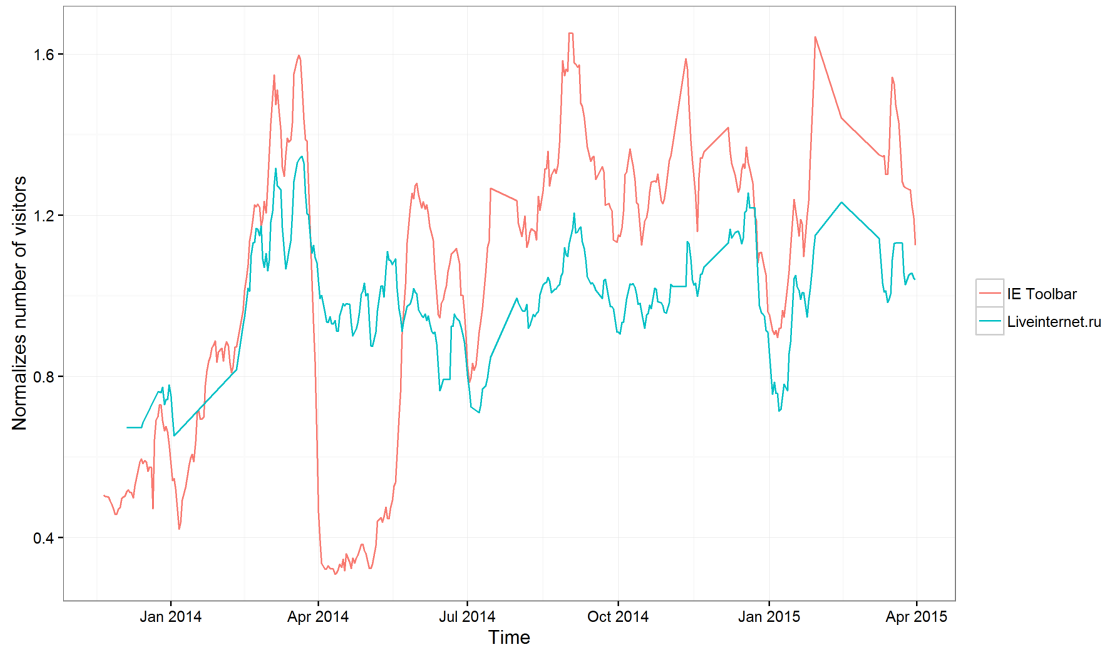


(a) kp.ru

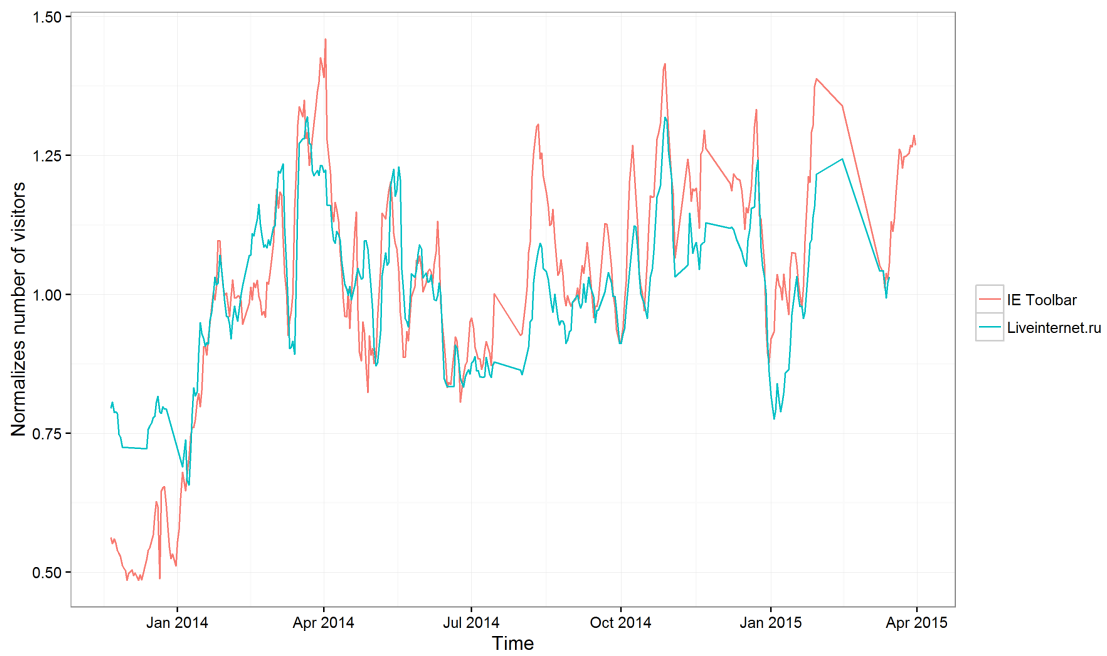


(b) lenta.ru

Figure B.3: Normalized number of weekly visitors to top sixth and seventh news websites, IE Toolbar and Liveinternet.ru



(a) gazeta.ru

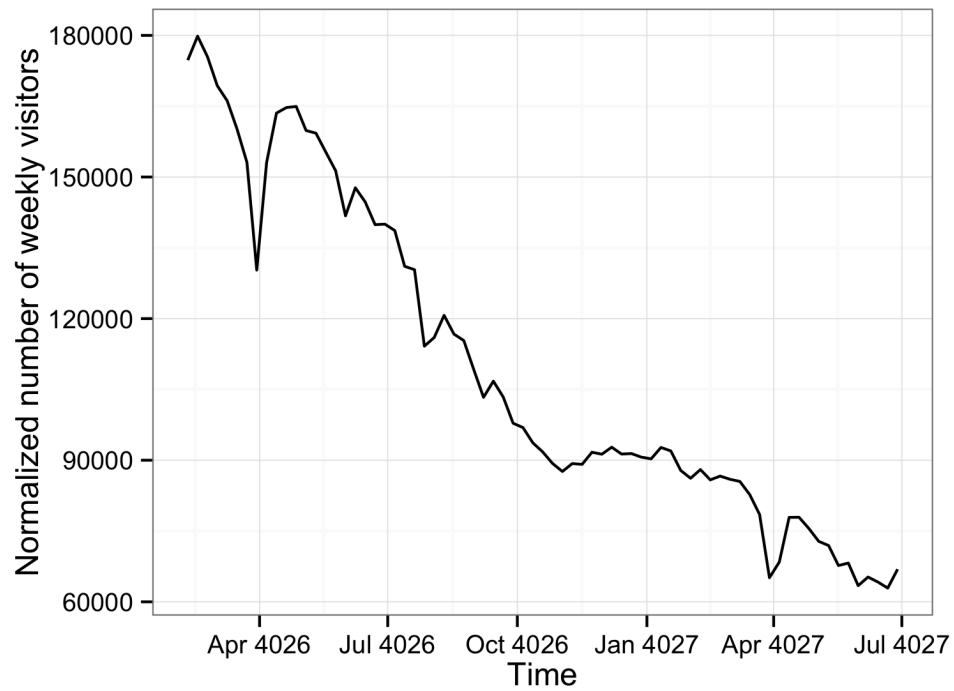


(b) rg.ru

APPENDIX C

ATTRITION OF USERS IN IE TOOLBAR

Figure C.1: Normalized number of weekly visitors of IE Toolbar data



APPENDIX D

CENSORED PROPER NOUNS

Figure D.1: Distribution of the differences in usage rankings of proper names between government-controlled and independent outlets

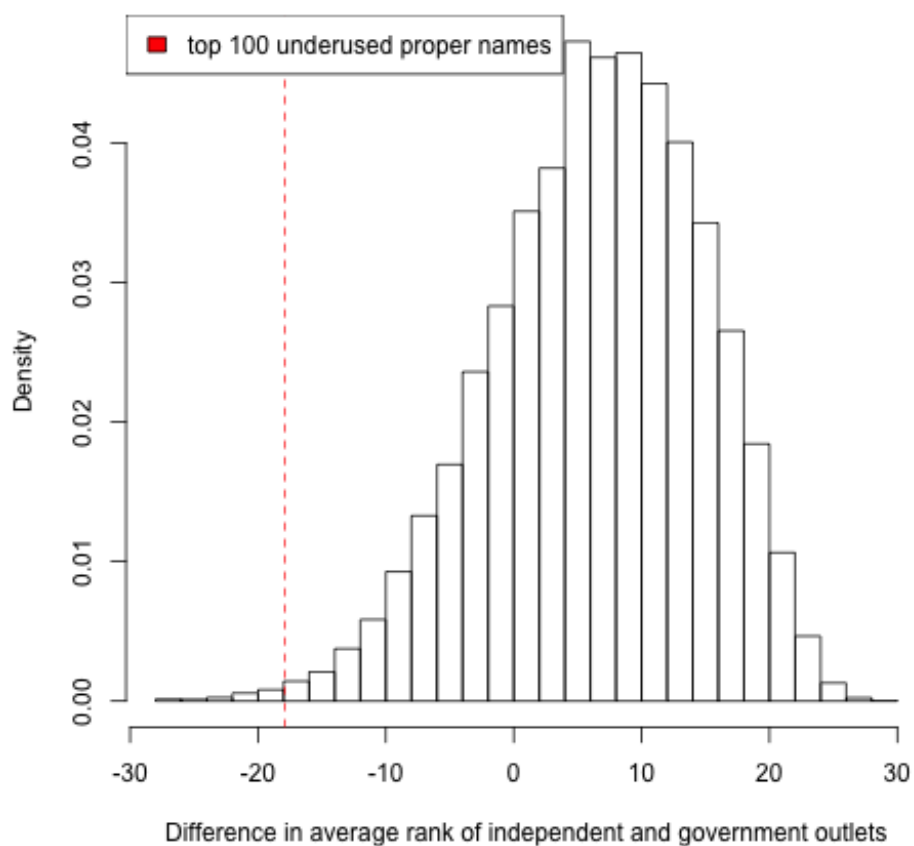


Figure D.2: Top 100 proper names that are under-used by the government-controlled media, and selected subset used to define the internal sensitive news topics

[1]	"венедиктов"	"дожд"	"дзядк"	"ревзин"	"ротенберг"	"тимченк"
[7]	"Slon"	"ходорковск"	"роскомнадзор"	"косенк"	"эх"	"кудрин"
[13]	"лукойл"	"сечин"	"дур"	"лебедев"	"навальн"	"прохоров"
[19]	"плющев"	"бизнес"	"яндекс"	"пархоменк"	"буковск"	"гельма"
[25]	"кисел"	"рамблер"	"болотн"	"лент"	"гур"	"сбербанк"
[31]	"лужков"	"чик"	"шлосберг"	"Riot"	"венедикт"	"Zar"
[37]	"Puss"	"мед"	"малофеев"	"деньг"	"афиш"	"максимовск"
[43]	"кашин"	"киселев"	"аркад"	"зубов"	"медведев"	"Gunvor"
[49]	"синдеев"	"акименков"	"New"	"инф"	"рош"	"зимин"
[55]	"ив"	"марш"	"давыдов"	"габрелянов"	"связн"	"алексашенк"
[61]	"роснефт"	"тиньков"	"парфен"	"Villag"	"усман"	"плох"
[67]	"олевск"	"алексе"	"мамут"	"евросет"	"перекрестк"	"билайн"
[73]	"сурков"	"зуб"	"плющ"	"ведом"	"сахаров"	"кичаткин"
[79]	"agor"	"сноб"	"манежн"	"волож"	"гражданкин"	"герма"
[85]	"Esquir"	"фельгенгауэр"	"кадыров"	"лесин"	"бершидск"	"мемория"
[91]	"маркв"	"богатенков"	"греф"	"гражданск"	"чичваркин"	"Rambler"
[97]	"албуров"	"акунин"	"предпринимател"	"капк"		

(a) All

[1]	"ревзин"	"ротенберг"	"тимченк"	"ходорковск"	"роскомнадзор"	"косенк"	"кудрин"
[8]	"сечин"	"дур"	"лебедев"	"навальн"	"прохоров"	"гельма"	"болотн"
[15]	"гур"	"шлосберг"	"Riot"	"Puss"	"малофеев"	"Gunvor"	"акименков"
[22]	"рош"	"ив"	"габрелянов"	"алексашенк"	"парфен"	"кичаткин"	"гражданкин"
[29]	"лесин"	"мемория"	"чичваркин"	"албуров"	"ройзма"		

(b) Selected

APPENDIX E

SLANT IN THE INTERNAL-SENSITIVE NEWS

To find slant in internal-sensitive news, we first allocate articles which mention government-sensitive proper nouns. We use two groups of proper nouns: (1) related to the actions of the opposition and (2) related to corruption. Proper nouns selected as opposition-related are Khodorkovsky, Bolotnaya, Pussy Riot, Navalny, and Royzman. Corruption-related proper nouns are Timchenko and Rottenberg. We then allocate the top 50 words underused by GC outlets in the articles which contain these sensitive proper names. Figure E.1 contain the top 50 words allocated as slant for both topics and a subset of words which was selected after removing incidental words and proper names.

Figure E.1: Slant related to sensitive news

заблокирова роскомнадзор узник сход спецприемник интерв страниц допрос тимченк сечин
коррупц 15 экстремизм антикризисн задержан народн вконтакт бутырк кремл дзядк
инф фбк выписк остав антикоррупцион роснефт объяв блокировк арест еспч
автозак внепланов ярмыш slop обвинен facebook косенк блог объясн запуга
опубликова едет блокирова капк обыск суток коммерсант лент ссылок запрошен

(a) Opposition, Top 50

**заблокирова узник сход спецприемник допрос
коррупц экстремизм антикризисн задержан народн
антикоррупцион блокировк арест автозак обвинен
запуга блокирова обыск суток**

(b) Opposition, Selected

ходорковск сын шамалов строительств зимин кооператив увол ротенберг увольнен расследован
роснефт тимченк релиз уск анонимн проект генподрядчик швейцарск подрядчик ведом
рыб конкурс геннад сумм редакцион запуск ведет заказчик подтвержда закупочн
смет бизнес мостотрест зят газпром бенефициар источник meduz питьев агентств
колпак субподрядчик дачн limited специализир галин протокол акулинин труб залог

(c) Corruption, Top 50

**строительств кооператив увол увольнен расследован
проект генподрядчик швейцарск подрядчик конкурс
заказчик закупочн смет**

(d) Corruption, Selected

We then compute the frequency of usage of selected slant words in the publications

about sensitive topics across all outlets. We compute these frequencies for both opposition and corruption topics and use average frequencies across these two topics to construct an index of slant for each outlet. This index is used in Figure 3.2.

APPENDIX F

CENSORSHIP IN UKRAINE-CRISIS NEWS

We look for underused proper nouns in the Ukraine-crisis news coverage. The top proper noun which appears is a name of the Russian soldier who allegedly was killed in Ukraine while fighting for the pro-Russian opposition. We do not find much incidents of censorship with proper nouns and conclude that censorship might not be the main mechanism of control in the Ukraine-crisis news coverage.

Figure F.1: Top 50 proper names that are underused by the government in articles about the Ukraine crisis

[1]	"кичаткин"	"кремл"	"коммерсант"	"Slon"	"федорин"
[6]	"передрук"	"васюнин"	"зубов"	"зуб"	"выбут"
[11]	"богатенков"	"шлосберг"	"ходорковск"	"олевск"	"киселев"
[16]	"гражданкин"	"сноб"	"кривенк"	"теленедел"	"афиш"
[21]	"лент"	"барабан"	"зеркальн"	"манежн"	"волож"
[26]	"инф"	"дожд"	"каныгин"	"алексашенк"	"бендукидз"
[31]	"сурк"	"Meduz"	"клю"	"горл"	"солоп"
[36]	"мальдон"	"глазьев"	"друг"	"мамут"	"белковск"
[41]	"путин"	"фридман"	"Radisson"	"роскомнадзор"	"левад"
[46]	"горлов"	"сурков"	"сов"	"ведом"	"джемилев"

APPENDIX G

SLANT IN UKRAINE-CRISIS NEWS

Figure G.1 includes all slanted words (both pro-Russia and pro-Ukraine slants) and figure G.2 contains selected words which describe pro-Russia (top) and pro-Ukraine (bottom) slants.

Figure G.1: List of slanted words used related to the Ukrainian conflict

[1]	"свобод"	"колон"	"нацист"	"оккупац"
[5]	"народн"	"сектор"	"террорист"	"боевик"
[9]	"ат"	"нацизм"	"хунт"	"карательн"
[13]	"фашист"	"сепаратист"	"оккупирова"	"стрелк"
[17]	"предател"	"самооборон"	"фашизм"	"агрессор"
[21]	"ополченц"	"неофашист"	"колорад"	"карател"
[25]	"головорез"	"бандеровц"	"ляшк"	"ярош"
[29]	"загнива"	"бендеровц"	"днр"	"новоросс"
[33]	"гиркин"	"ватник"	"лнр"	"праворадикализм"
[37]	"неофашизм"	"вышиватник"	"ультранацизм"	

Figure G.2: Pro-Russia and pro-Ukraine slants in the Ukraine conflict

[1]	"хунт"	"карательн"	"карател"	"бандеровц"	"бендеровц"
(a) Pro-Russia slant					
[1]	"террорист"	"боевик"	"ат"	"сепаратист"	"оккупирова"
[6]	"агрессор"				
(b) Pro-Ukraine slant					

APPENDIX H

SUMMARY OF BROWSING BEHAVIOR

Each news website consists of 4 different types of pages: the main page, news articles pages, news subdirectories, and other pages¹. News articles account for most page views on news websites. Other webpages are visited half as often as news articles. The main directory and news subdirectories are also each visited only half as often as news articles. Table H.1 shows statistics of browsing of the webpage types. While some consumers read news from the headlines, most of the time the main pages and news subdirectories help readers to navigate to the news articles. This also includes navigation to the non-news content in the ‘other’ sections. Thus, we only use navigation to news articles as records of news consumption.

Table H.1: Summary of browsing behavior

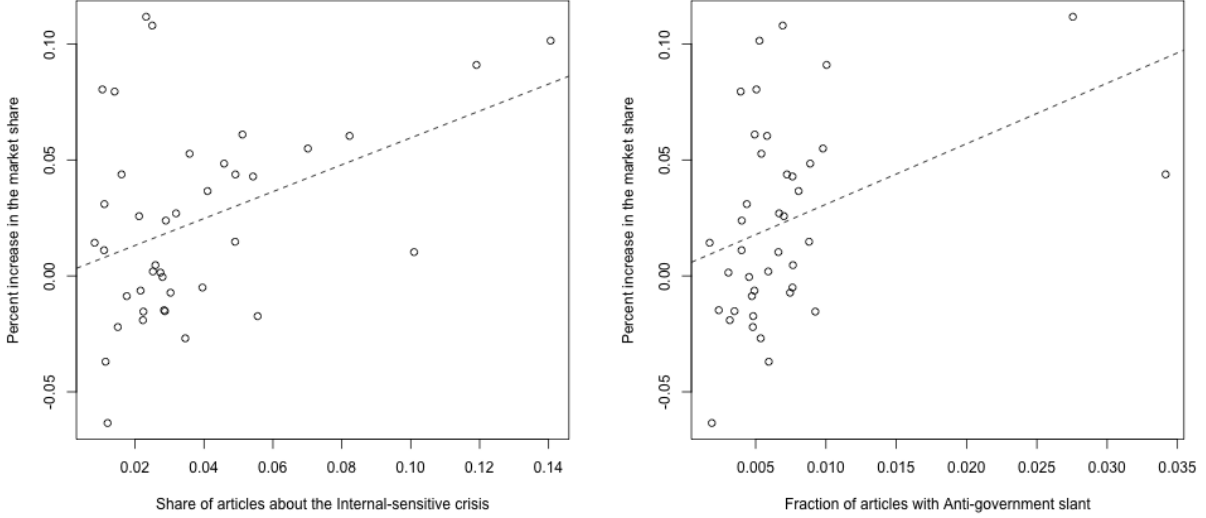
	Page views	Visits (Sessions)	Seconds spent	
			Mean	Median
Main page	5,344,041	1917206	128	42
News articles	1,042,0780	4240831	186	86
News subdirectories	4,225,221	1484410	263	90
Other	6,584,713	2,389,635	145	44
Total	26,537,267	6,630,400	176	64

1. News subdirectories are the directories that list news on a given day or topic; other links correspond to the specialized content which is published by the news website, for example, financial analytics, industry analysis, pictures, etc.

APPENDIX I

DESCRIPTIVE EVIDENCE FIGURES

Figure I.1: Changes in the market shares with the increase in the amount of the internal-sensitive news

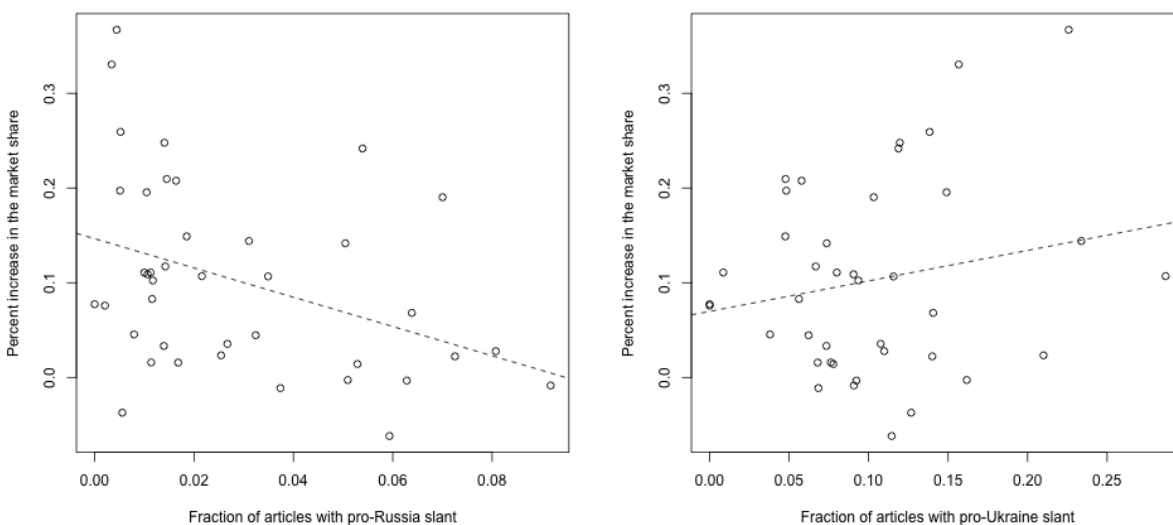


(a) By amount of coverage

(b) By amount of anti-government slant

Each point represents a news outlet. Vertical axis corresponds to the percent change in the market shares as amount of internal-sensitive news increase by 1 standard deviation. Both relationships are statistically significant at 5% level.

Figure I.2: Changes in the market shares with the increase in the amount of the Ukraine-crisis news



(a) By amount of pro-Russia slant

(b) By amount of pro-Ukraine slant

Each point represents a news outlet. Vertical axis corresponds to the percent change in the market shares as amount of internal-sensitive news increase by 1 standard deviation. The relationships of the market shares with the amount of coverage and level of pro-Russia slant are statistically significant at 5% level. The relationship between the market shares and the level of pro-Ukraine slant is positive but not statistically significant.

APPENDIX J

BASIC SUPPLY MODEL

News outlet j profits from advertisements that depend on the total number of visitors, the number of unique visitors, and the number of exclusive visitors (Gentzkow and Shapiro, 2015). For simplicity we assume that advertisement revenue on a given day only depends on the number of visitors of website j , V_{jt} :

$$Profit_{jt} = f(V_{jt}) - g(\alpha_j)$$

We also assume a linear relationship between the number of visitors and profits, e.g. $f(V_{jt}) = aV_{jt}$, and a cost function which depends on the quality of the news outlet but not on the allocation of topics reported and ideology.¹ This implies that $g(\cdot)$ is strictly increasing in α_j and does not depend on rep_{jt}^x and sl_{jt}^x :

$$\frac{\partial g(\cdot)}{\partial \alpha_j} > 0$$

$$\frac{\partial g(\cdot)}{\partial rep_{jt}^x} = 0$$

$$\frac{\partial g(\cdot)}{\partial sl_{jt}^x} = 0$$

The number of visitors going to website j at time period t is

$$V_{jt} = I \int \pi(y_{it} = j | \theta_i) \phi(\theta_i) d\theta_i$$

Let us assume that the observed characteristics of news outlets $\{rep_{jt}^s, \alpha_j, sl_j\}$ are the result of a Nash equilibrium game with all players moving simultaneously. The objective

1. This set-up is similar to Gentzkow and Shapiro (2015) where firms can choose slant along the ideological plane for free.

of news outlet j is to maximize profit at each time period t . The corresponding first-order conditions are

$$^a \frac{\partial V_{jt}}{\partial \alpha_j} = \frac{\partial g(\cdot)}{\partial \alpha_j}$$

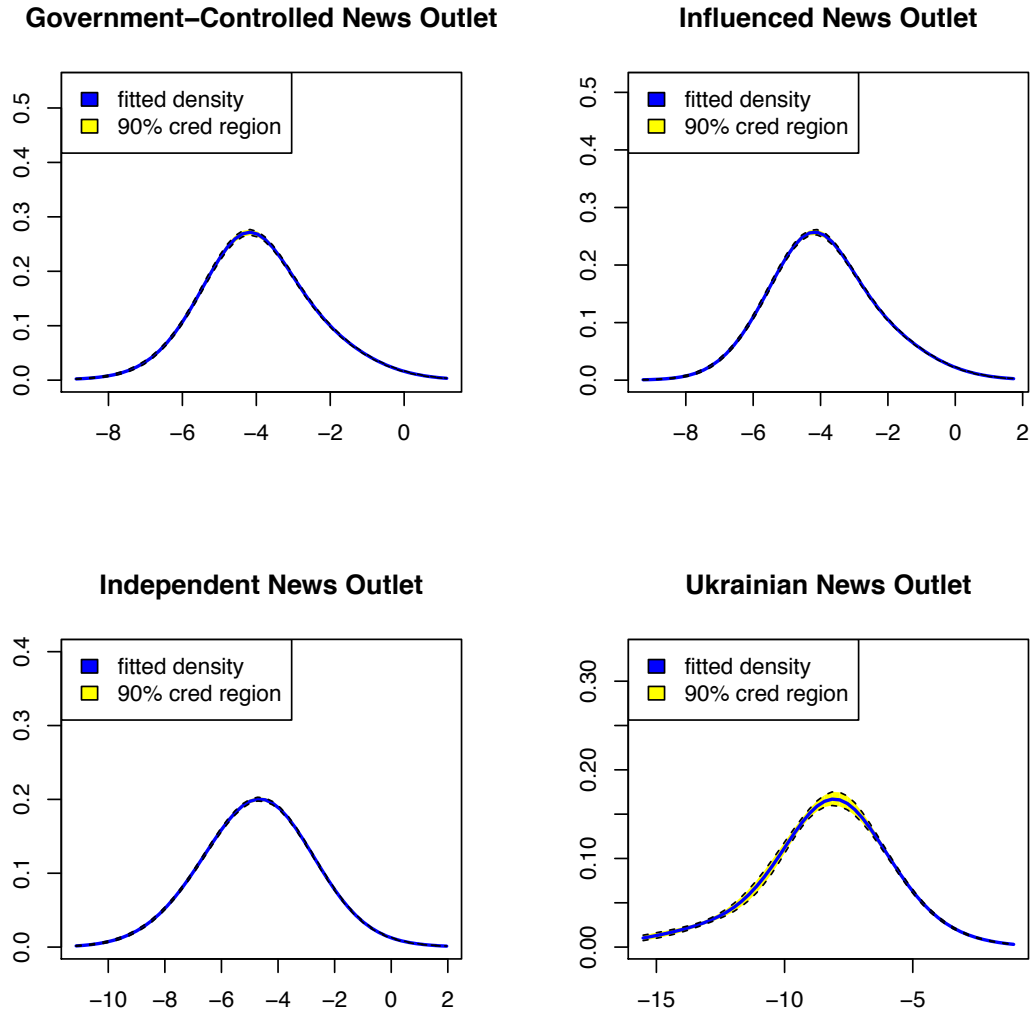
$$^a \frac{\partial V_{jt}}{\partial rep_{jt}^x} = 0$$

$$^a \frac{\partial V_{jt}}{\partial sl_j} = 0$$

APPENDIX K

ESTIMATION RESULTS: DISTRIBUTIONS OF MATCHING VALUES α_{IJ}

Figure K.1: Posterior distributions of demand parameters: matching values



APPENDIX L

SIMULATED MARKET SHARE, LARGE VOLUME OF SENSITIVE NEWS

Table L.1: Simulated shares, day with a large internal-sensitive news event.

	(1)	(2)	(3)	(4)	(5)
	$\bar{s}h$	Δ	Δ	Δ	Δ
	Current shares	No direct control	No indirect control	No control	More control
$\bar{s}h_{Gov}$	4.031 (3.982, 4.079)	+0.491 (0.448, 0.534)	+0.013 (0.006, 0.022)	+0.505 (0.463, 0.546)	-0.008 (-0.013, -0.004)
$\bar{s}h_{Inf}$	5.826 (5.750, 5.938)	-0.148 (-0.162, -0.136)	+0.031 (-0.008, 0.061)	-0.116 (-0.151, -0.086)	-0.024 (-0.033, -0.016)
$\bar{s}h_{Ind}$	4.193 (4.133, 4.247)	-0.087 (-0.094, -0.080)	+0.015 (0.006, 0.023)	-0.072 (-0.084, -0.060)	+0.008 (-0.017, 0.038)
$\bar{s}h_{Ukr}$	0.462 (0.445, 0.477)	-0.001 (-0.002, -0.0005)	+0.002 (0.001, 0.003)	+0.001 (-0.0001, 0.003)	-0.002 (-0.002, -0.001)
$\bar{s}h_{Out}$	85.489 (85.310, 85.649)	-0.255 (-0.282, -0.227)	-0.062 (-0.083, -0.037)	-0.319 (-0.358, -0.282)	+0.025 (0.008, 0.041)

Table L.2: Simulated shares, day with a large Ukraine-crisis news event.

	(1) \bar{sh}	(2) Δ	(3) Δ	(4) Δ	(5) Δ
	Current shares	No direct control	No indirect control	No control	More control
\bar{sh}_{Gov}	5.809 (5.733, 5.899)	+0.639 (0.497, 0.776)	-0.060 (-0.087, -0.028)	+0.535 (0.403, 0.652)	+0.028 (0.013, 0.044)
\bar{sh}_{Inf}	6.478 (6.408, 6.545)	-0.169 (-0.216, -0.116)	+0.639 (0.557, 0.719)	+0.431 (0.355, 0.487)	+0.071 (0.055, 0.089)
\bar{sh}_{Ind}	4.768 (4.708, 4.835)	-0.106 (-0.128, -0.083)	-0.114 (-0.130, -0.097)	-0.191 (-0.226, -0.155)	-0.425 (-0.494, -0.356)
\bar{sh}_{Ukr}	1.379 (1.329, 1.433)	+0.007 (-0.001, 0.014)	+0.014 (0.007, 0.019)	+0.031 (0.017, 0.043)	-0.011 (-0.015, -0.007)
\bar{sh}_{Out}	81.566 (81.406, 81.738)	-0.370 (-0.442, -0.299)	-0.480 (-0.526, -0.436)	-0.807 (-0.906, -0.707)	+0.336 (0.279, 0.384)

The market share are in % of the entire market. We use days with the highest volume of sensitive news observed in the sample period.

APPENDIX M

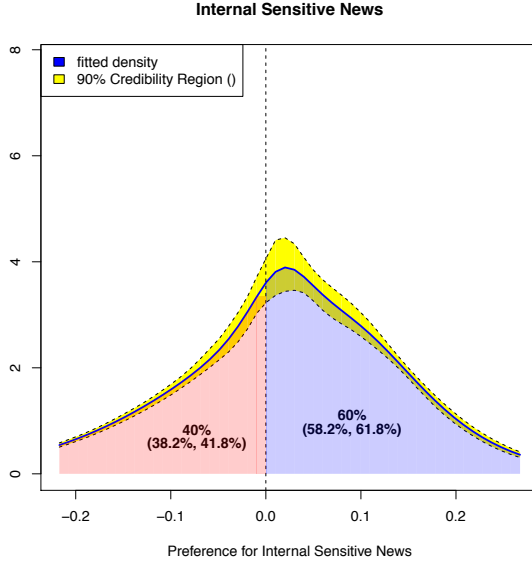
PREFERENCES OF NEWS JUNKIES

We run the estimation procedure on choices of people who consume news 50-100 times in the sample, using both the base and the extended model. Figures M.1 and M.2 present the pointwise posterior means and the 90% credibility region of the marginal densities of consumers' preferences. Figure M.1 presents estimates of consumers' preferences for internal-sensitive news. Based on the results of both models, the majority of consumers prefer internal-sensitive news to non-sensitive news: 60% (58.2%, 61.8%) in the case of the base model and 62.4% (60.8%, 63.9%) in the case of the extended model. Estimates of consumers' preferences for an anti-government slant are significantly different: 89.1% (84.6%, 93.6%) of consumers prefer an anti-government slant based on the results of the base model, while only 60.6% (56.7%, 64.6%) of consumers prefer such a slant based on the extended model.

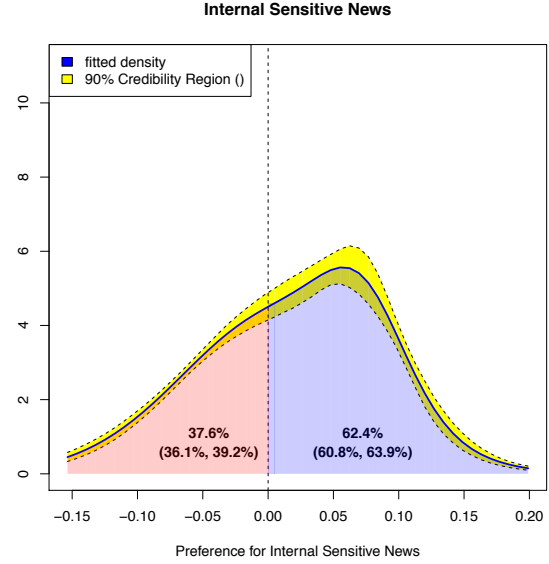
Figures M.2 and M.3 presents estimates of consumer preferences for Ukraine-crisis news. Based on the results of both models, majority of consumers prefer Ukraine-crisis news to non-sensitive news: 62.9% (61.4%, 64.3%) in the case of the base model and 62.3% (60.5%, 64.2%) in the case of the extended model. Estimates of consumers preferences for a pro-Russia and a pro-Ukraine slant are significantly different: 31.5% (26.2%, 36.6%) of consumers prefer a pro-Russia slant based on the results of the base model, while only 21.7% (16.8%, 25.8%) of consumers prefer such slant based on the results of extended model, and 70% (65.2%, 73.9%) of consumers prefer a pro-Ukraine slant based on the results of the base model, while only 57.3% (52.7%, 63.1%) of consumers prefer such slant based on the extended model.

Based on these findings, there are two things that we need to adjust in the current model and estimation. First, the original estimation based on 50,000 consumers who are randomly drawn from the sample finds that the majority of 'news junkies' prefer a pro-Russia slant in the Ukraine-crisis news (their preference is not as strong as the one of the 'occasional' customers, but still positive). The estimation based on only 'news junkies' reveals that the

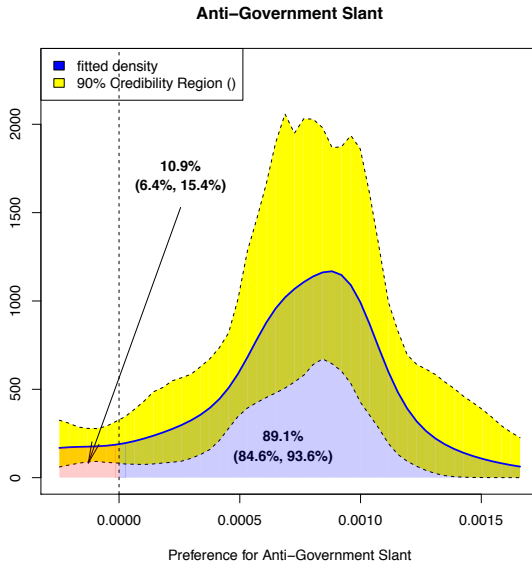
Figure M.1: Posterior distribution of preference for Internal-sensitive news and corresponding slant: Model with and without the interdependence of choices within a day



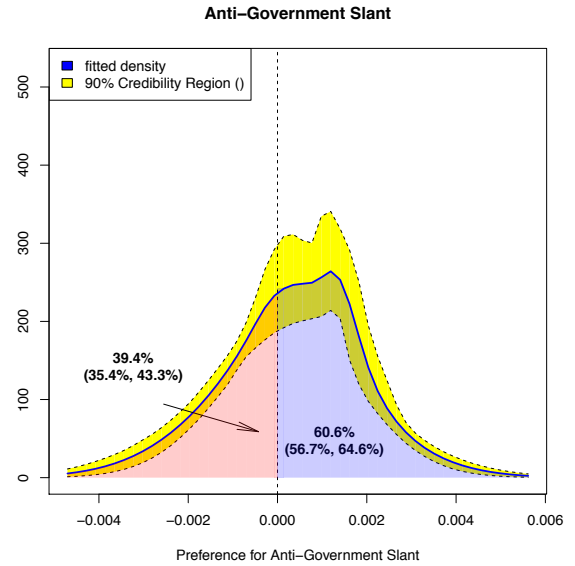
(a) Internal Sensitive News: Base Model



(b) Internal Sensitive News: Extended Model



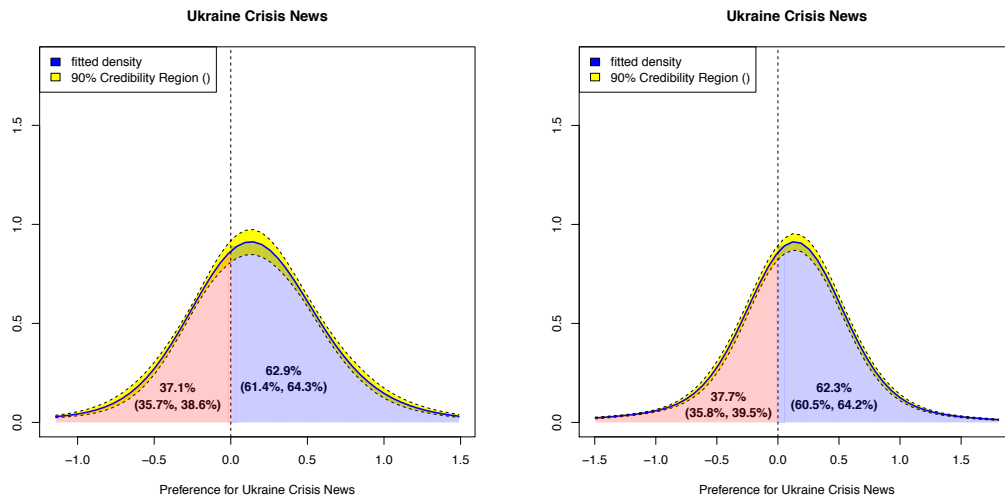
(c) Anti-Government Slant: Base Model



(d) Anti-Government Slant: Extended Model

The scale of variables is normalized: values on the x axis represent effect of 1 standard deviation increase in the corresponding variable on consumers' utilities. Blue line is posterior mean; yellow shaded regions are 90% credibility regions. Numbers represent percent of probability mass of distributions below and above zero, and 90% credibility intervals of these percents.

Figure M.2: Posterior distribution of preference for the Ukraine Crisis news



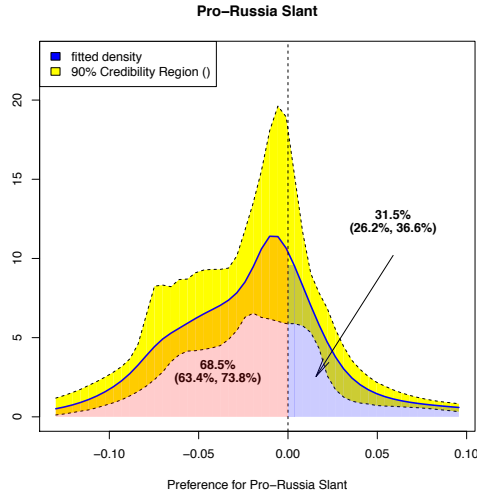
(a) Ukraine-Crisis News: Base Model (b) Ukraine-Crisis News: Extended Model

Please refer to the description of the previous figure.

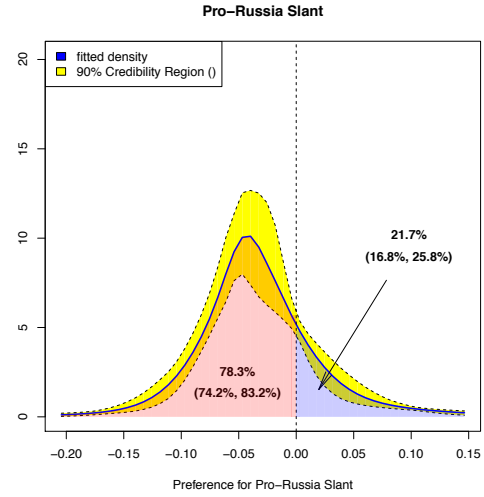
majority of them have a negative preference for a pro-Russia slant. This suggests that the estimation based on 50,000 consumers has shrunk the estimates too much, which implies we need to allow for more than 5 mixture components. We are currently re-running the estimation with 10 mixtures, and test a more flexible Dirichlet Process.

Second, results show that ignoring the interdependence of consumer preferences significantly affects the estimates. While the conclusion for consumers' preferences for the pro-government bias might not be affected (the majority of 'news junkies' still prefers less pro-government bias in the news and navigates to GC news outlets despite the pro-government bias), the counterfactual estimates might not be precise. We plan to re-estimate the model allowing for interdependencies of consumers' choices in the future work.

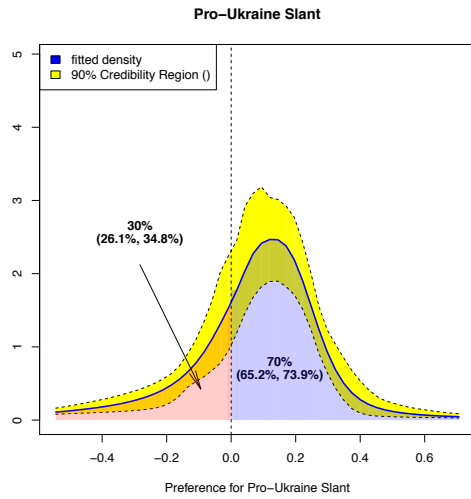
Figure M.3: Posterior distribution of preference for slant in the Ukraine Crisis



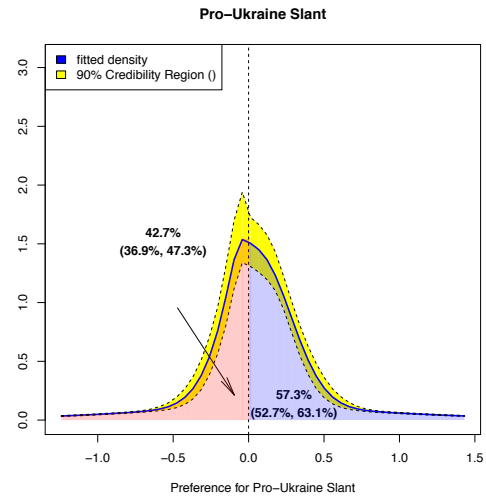
(a) Pro-Russia Slant: Base Model



(b) Pro-Russia Slant: Extended Model



(c) Pro-Ukraine Slant: Base Model



(d) Pro-Ukraine Slant: Extended Model

Please refer to the description of the previous figure.