Online Appendix:

Fighting Propaganda with Censorship: A Study of the Ukrainian Ban on Russian Social Media

Contents

Appendix A: Descriptive statistics

Table A1: Descriptive statistics

Users sub-sample and proportion of VKontakte friends in Russia

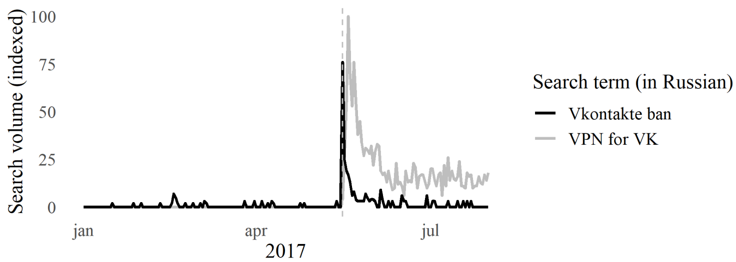
Mainland Ukraine				
Group	N users	Mean prop. of friends	Median prop. of friends	SD
Pro-Russian	528	0.140	0.077	0.168
Pro-Ukrainian	509	0.081	0.051	0.093
Random	396	0.069	0.044	0.081
All	1,433	0.099	0.057	0.128
Crimea				
Group	N users	Mean prop. of friends	Median prop. of friends	SD
Pro-Russian	539	0.211	0.158	0.172
Pro-Ukrainian	603	0.136	0.100	0.137
Random	449	0.200	0.151	0.171
All	1,591	0.179	0.135	0.163

Posting activity 90 days before and after the ban

Mainland Ukraine					
Group	N posts	Mean daily posts (before)	SD	Mean daily (after)	SD
Pro-Russian	17,424	0.272	0.486	0.116	0.268
Pro-Ukrainian	14,238	0.242	0.446	0.104	0.265
Random	15,072	0.233	0.456	0.112	0.283
All	46,734	0.251	0.464	0.111	0.271
Crimea					
Group	N posts	Mean daily posts (before)	SD	Mean daily (after)	SD
Pro-Russian	23,262	0.264	0.470	0.217	0.401
Pro-Ukrainian	17,745	0.184	0.373	0.149	0.367
Random	12,128	0.129	0.293	0.154	0.448
All	53,135	0.195	0.393	0.174	0.404

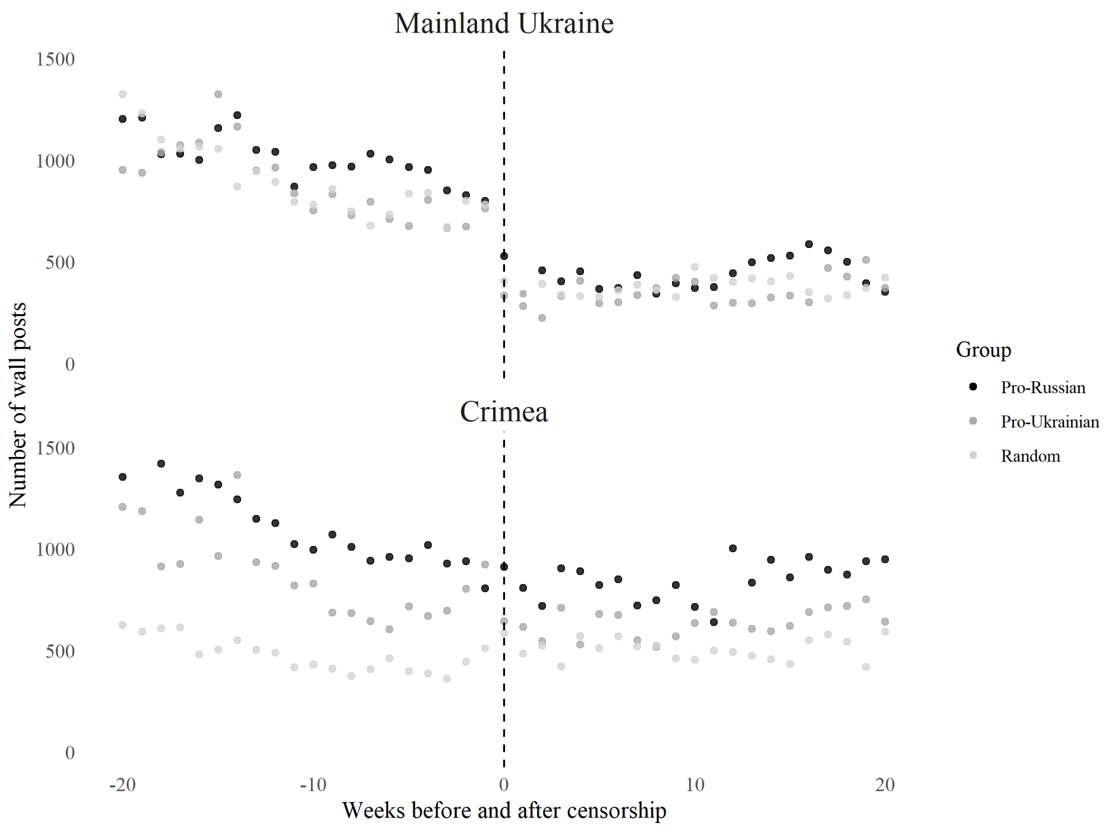
N users = 3,024; N posts (90 days bandwidth) =99,869

Figure A1: Indexed search volume on Google for 'VKontakte ban' and 'VPN for VK' in Ukraine.



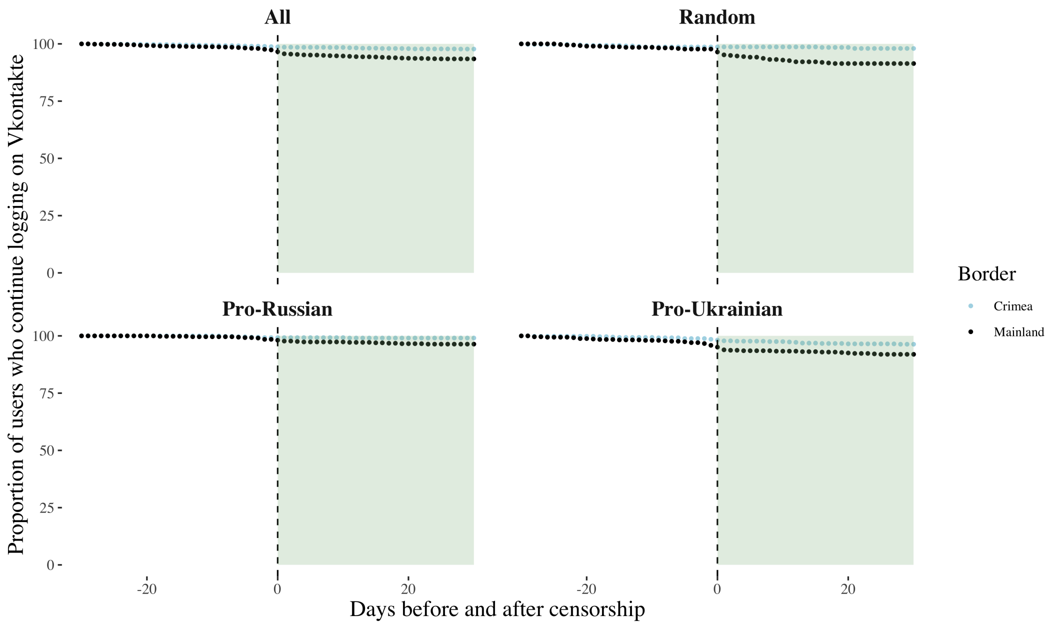
The dashed line marks when the ban was first announced in the media (16th of May 2017). The date with the highest number of search queries during the first half of 2017 is the base date (19th of May = 100). Original search keywords are in Russian. Source: Google Trends.

Figure A2: Posting activity per week before and after the ban



The first 7 days of the ban implementation (18th to 24 May 2017) are standardized as 0. The graph is based on two separate median values for Crimea and Mainland respectively.

Figure A3: Proportion of users who have access to their VKontakte profile (based on last login date)



The users are considered to have access to VKontakte until their very last login date. The majority of the users continued accessing VKontakte after the ban. The figure is based on the assumption that all of the 3,024 accounts existed throughout the entire period before the censorship. I confirm this by examining their oldest wall posts. All of them have posted at least once before the ban. 2989 (98.84%) of these accounts wrote their oldest post at least 30 days before the post. The remaining 35 users may have existed prior to the period, however, this cannot be confirmed through wall posts. This is equivalent to only 1.16 % of all the users and does not change the main results. Note that the decline begins before the ban. This is due to the fact that the figure is based on last logins: i.e. some users login for the last time not knowing that they will be locked out by the sudden ban.

Appendix B: Operationalizing political attitudes

Political attitudes toward Russia and Ukraine are operationalized in three steps, using information about which public VKontakte communities the users follow. The communities consist of either groups or pages dedicated to a topic, similar to Facebook. First, I identify the top 10,000 most frequently liked pages and groups among the sampled users on each side of the border. The annotation guidelines for this step are described in Codebook A. Secondly, I extensively evaluate *only* the name of each page, while keeping only those pages, whose names are thematically related to politics, news, patriotism, or government institutions. In the last phase, I follow the guidelines in Codeook B to evaluate whether the relevant pages are neutral, pro-Russian, or pro-Ukrainian, based on a close reading of their names, descriptions, cover pictures, and their 20 most recent wall posts. By doing so, I identify 122 public VKontakte communities that are either pro-Russian (N = 74) or pro-Ukrainian (N = 48). The list of these groups and pages and the codebooks with annotation guidelines are available in the chapters below.

Appendix C. Codebook A: Coding VKontakte community names

You are asked to identify pro-Russian and pro-Ukrainian Communities in VKontakte. Please read this guide carefully before proceeding. Below you will find a definition of pro-Russian and pro-Ukrainian communities and a step-by-step guide for identifying and categorizing the groups.

Categories

1. UNRELATED. Refers VKontakte community names that are unrelated to politics or news. See the categories below.

2. POLITICS. Assign these categories to community names that are related to local, national or global politics. This includes communities that are focused on political figures, movements and events.

3. PATRIOTISM. Use this category if the name directly indicates that this a community for patriots or nationalists. This category should be used both in cases where the name indicates praise or pride in ones own country or ethnic background or resentfulness towards other countries, nationalities or ethnicities.

4. NEWS. This category should be used in cases where the group name is related to either local, national or global news.

5. GOVERNMENT INSTITUTIONS. Use this category in cases where the community name indicates that the main topic is centered on a specific government institution such as the police, military, municipality, universities etc.

6. UNCLEAR. This category should be used in instances, where the category is unclear based on the name alone. Use this category if the community has to be manually examined in an browser in order to be assigned one of the previous categories

Examples:

UNRELATED:

- ” Гороскоп на каждый день” : The main topic of the groups is horoscope, which is why it is unrelated to the “politics” , “patriotism” and “news” categories.

- “Подростки | Весна | Любовь” the main topic of this group appears to be love, spring and teenage hood, which is why it is unrelated to the abovementioned categories.

- ” Хитрости жизни” . The precise topic of this group is unclear. However, it does not appear to be related to politics, patriotism or news.

- “Лайфхакерша | Женские Хитрости” – The main topic of behind the group name appears to be lifehacking for women, which is why it is unrelated to the abovementioned categories.

POLITICS:

- “Партия регионов” The main topic appears to be the Ukrainian “Party of Regions” .

- “*Путин*” : The community name indicates that the main topic is focused on a political figure

- “Международная Военная Политика МВП” : The name indicates that international and military politics are a central topic in this group.

- “ Мы против Евромаидана!: The main topic is the political movement, Euromaidan.

PATRIOTISM

- “Бессмертный полк России” The community is dedicated to a patriotic movement, whose goal is to commemorate Russia’s victory of Nazi Germany

- “Наше українське | АТО” the group name expresses a strong sentiment of Ukrainian sentiment, where the central topic is the discussion of ” our Ukraine” in the ” anti terrorist operation" against Russian separatists in Donbass

- " Верим в Россию! | Патриоты | Новости | История" is a patriotic group, because it is dedicated to supporting Russia.

- "Стильный Патриот Украины" is a patriotic group, because it is dedicated to supporting Ukraine through patriotic fashion.

NEWS

- "ВЕСТИ.ru | РОССИЯ 24" is an online community dedicated to a news program.

- " НТВ" : the name indicates that the community is dedicated to the Russian TV channel, NTV, and is therefore also the source of news that are broadcasted on the channel.

- "Типичный Славянск" : The name indicates that the community is dedicated to sharing local news about Slavyansk. This kind of "typical" groups exist for many cities in Ukraine.

- "Сводки от ополчения Новороссии" is categorized as " news" , because the name indicates that the group is dedicated to latest news updates on the war in Ukraine (from the perspective of Russian separatists).

GOVERNMENT INSTITUTIONS.

- "МИД России" The name suggests that the groups is related to the Russian Ministry of Foreign Affairs

UNCLEAR

- " Наша страна Россия" is unclear, because the name could suggest that this is a patriotic Russian community, a source of news about Russia or both.

- "Не только факты" : The name could potentially indicate that the community is dedicated to news sharing, however, the ambiguous name could also refer to other topics.

Appendix D: Codebook B - Coding VKontakte communities

You are asked to identify pro-Russian and pro-Ukrainian Communities in VKontakte. Please read this guide carefully before proceeding. Below you will find a definition of pro-Russian and pro-Ukrainian communities and a step-by-step guide for identifying and categorizing the groups.

Categories

1. PRO-RUSSIAN. Refers to VKontakte community pages that directly support Russia, the Russian government, President Putin, Russian patriotic groups, Russian separatist forces in Ukraine or media that are loyal to Kremlin. This includes groups and movements that affiliate themselves of the separatist forces or the Russian government or the Russian separatist cause in Ukraine. The category should be applied if the pro-Russian alignment is indicated in the following parts of the VK community page:

   (a) name

   (b) description

   (c) cover picture

   (d) At least 1 of the 20 most recent messages posted on the community wall by the admin(s)

   The community page's use of the Russian language is in itself not sufficient to indicate pro-Russian alignment, since Russian can also be used to mobilize pro-Ukrainian sentiment. Similarly, the use of the Russianf lag alone is not enough for the group to qualify as "Pro-Russian" , since the flag can be used to simply indicate that the group is Russian, without showing an active pro-Ukrainian position. The "pro-Russian" category does not apply to communities that are centered on sport clubs. The annotator should evaluate all 4 elements as a whole, including the visual symbols in the cover picture.

2. PRO-UKRAINE. Refers to VKontakte community pages that directly support Ukraine, The Ukrainian government, Ukrainian Army, Ukrainian patriotic groups and Ukrainian media that propagate Ukrainian sovereignty. The category should be applied if the pro-Ukrainian alignment is indicated in the following parts of the VK community page:

   (a) name

   (b) description

   (c) cover picture

   (d) the top 20 messages posted on the community wall by the admin(s)

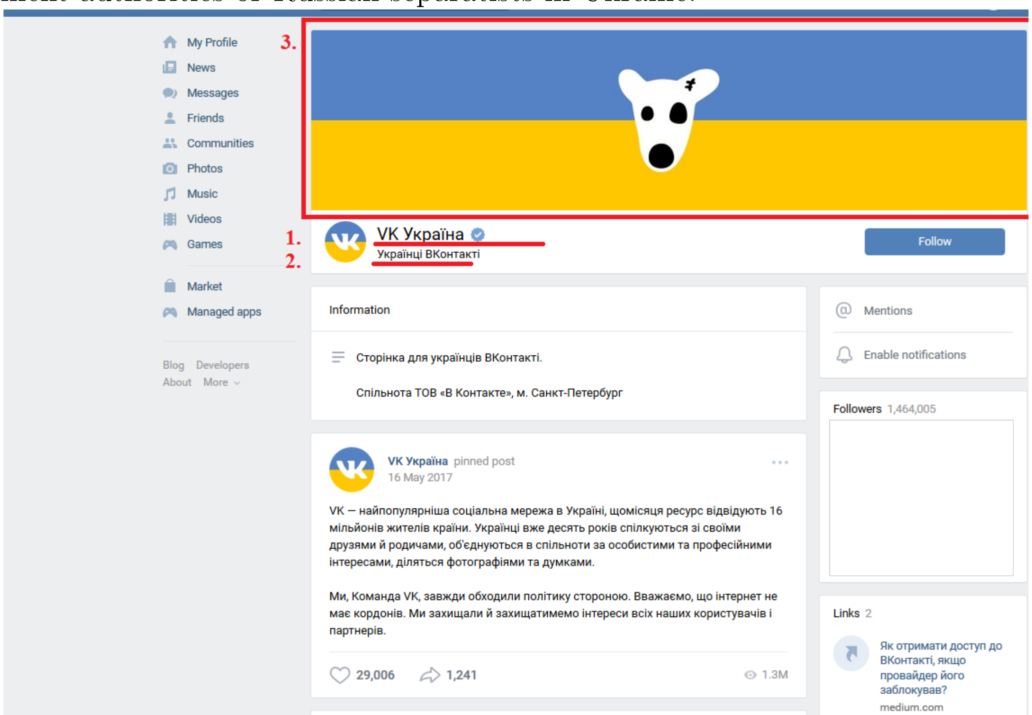
   The community page's use of the Ukrainian language is in itself not sufficient to indicate pro-Ukrainian alignment, since Ukrainian language can also be used to mobilize pro-Russian or pro-separatist sentiment. Similarly, the use of Ukrainian flag alone is not enough for the group to qualify as "Pro-Ukrainian" , since the flag can be used to simply indicate that the group is Ukrainian, without showing an active pro-Ukrainian position. However, there are instances where a flag can be used to signal a pro-Ukrainian position in the countries conflict between pro-Russian and pro-Ukrainian citizens. The annotator should evaluate the use of the symbol closely in context of other content on the page.
   The category should not be applied to communities that are focused on sport clubs. The annotator should evaluate all 4 elements as a whole, including the visual symbols in the cover picture.

3. UNCLEAR. Please use this category if you unsure whether the group is Pro-Russian, Pro-Ukrainian or unrelated.

4. UNRELATED. Use this category if the Community page is neither Pro-Russian or Pro-Ukrainian.

Example: Pro-Kremlin Alignment

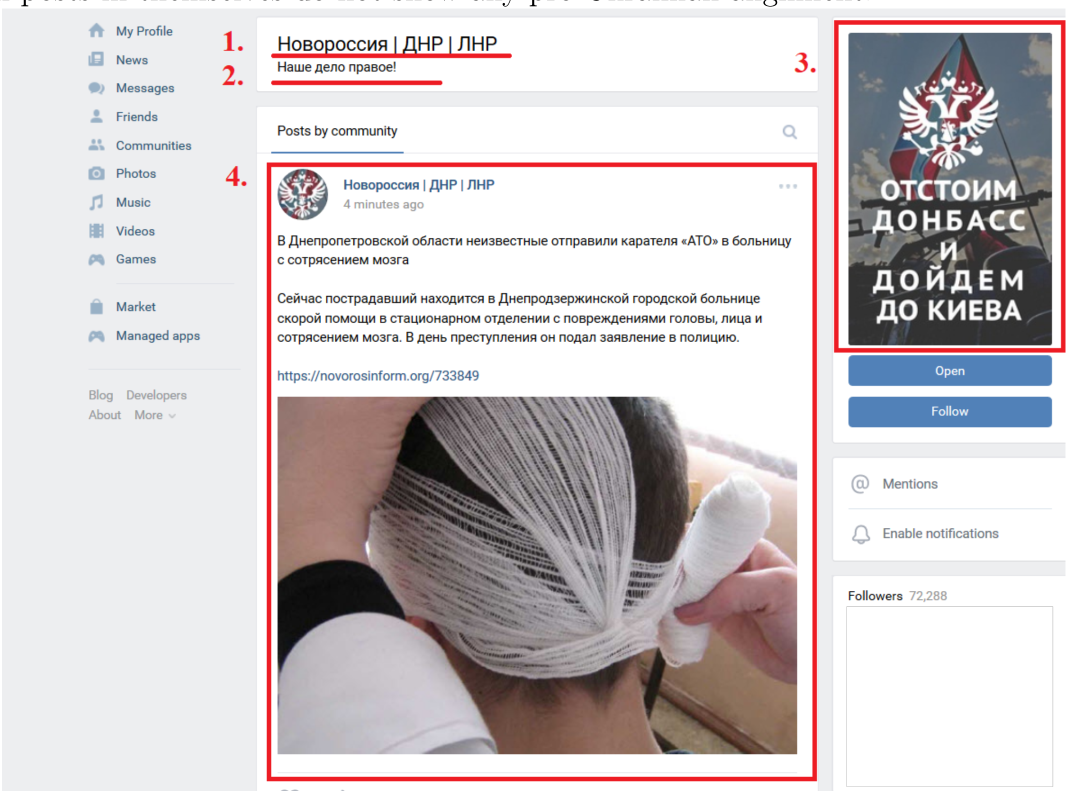
The VK community in the example bellow can be categorized as "pro-Kremlin" because its name, description, cover picture show support towards Russian separatists in Ukraine. This interpretation is further supported by admin's latest post, which describes the Ukrainian forces using a derogatory term, "karatel" (punisher) frequently used either by Russian government authorities or Russian separatists in Ukraine.



User avatars have been removed from this image.

Example: Pro-Ukrainian Alignment

The community in the example below is pro-Kyiv, because it directly supports Ukrainian national unity. In this context, the Ukrainian flag can be seen as an act of support for the Ukrainian nation in the conflict with Russian separatists and Russia. Note that most recent wall posts in themselves do not show any pro-Ukrainian alignment.



User avatars have been removed from this image.

Appendix E: Pro-Russian and pro-Ukrainian VKontakte communities

Table E1: VKontakte community groups annotated as pro-Russian (N= 74)

<table>
<tr><th colspan="2">Group name</th></tr>
<tr><td>МРК Россия-Китай: главное</td><td>ПЕРВЫЙ РУССКИЙ</td></tr>
<tr><td>Родные места: Москва</td><td>История России</td></tr>
<tr><td>Армия и Оружие</td><td>Россия | Russian Federation</td></tr>
<tr><td>АрмиЯ — сила!</td><td>Kino1TV</td></tr>
<tr><td>Новый Русский</td><td>IZ.RU</td></tr>
<tr><td>Наша страна Россия</td><td>Однажды в России</td></tr>
<tr><td>АРМИЯ | СПЕЦНАЗ | ОРУЖИЕ</td><td>Портал Госуслуг РФ</td></tr>
<tr><td>Армейское Братство (Армия | Дембель)</td><td>Волонтёры Победы Республика Крым</td></tr>
<tr><td>Типичная Армия| Служба| Солдаты| Дембель</td><td>РОССИЯ</td></tr>
<tr><td>Военная Техника России и Мира</td><td>Это Россия, детка</td></tr>
<tr><td>РусскаЯ АрмиЯ</td><td>РОССИЯ</td></tr>
<tr><td>РИА Новости</td><td>Полиция Крыма | МВД по Республике Крым</td></tr>
<tr><td>LIFE | Новости</td><td>Военный вестник</td></tr>
<tr><td>Первый канал</td><td>ТНТ</td></tr>
<tr><td>Русское Радио</td><td>Новости | ДНР | ЛНР | МИР | NEWS-FRONT</td></tr>
<tr><td>РОССИЯ - ВЕЛИКАЯ НАША ДЕРЖАВА!</td><td>РЕН ТВ | Телеканал</td></tr>
<tr><td>ВЕДОМОСТИ</td><td>Необычная Россия | Факты| История |Армия|Новости</td></tr>
<tr><td>#ЯРОССИЯ</td><td>Лента.ру</td></tr>
<tr><td>АнтиМайдан</td><td>НТВ</td></tr>
<tr><td>Новости RT на русском</td><td>ВЕСТИ.ru | РОССИЯ 24</td></tr>
<tr><td>Русские Онлайн</td><td>РЕН ТВ | Новости</td></tr>
<tr><td>МБОУ Братский УВК</td><td>Новости РФ | Армия России</td></tr>
<tr><td>"Волонтёры Победы" Джанкой</td><td>Телеканал "ПЕРВЫЙ КРЫМСКИЙ"</td></tr>
<tr><td>ЕДИНАЯ РОССИЯ Красноперекопск</td><td>Совет министров Республики Крым</td></tr>
<tr><td>Новости Крым | Симферополь | Севастополь</td><td>Прямая линия с Владимиром Путиным</td></tr>
<tr><td>Телеканал «Северный Крым»</td><td>ЕДИНАЯ РОССИЯ</td></tr>
<tr><td>Я РУССКИЙ</td><td>Мы против внешней политики США / Новости - Сирия</td></tr>
<tr><td>Бессмертный полк России</td><td>Армия России | Оружие | Факты | Геополитика</td></tr>
<tr><td>Я РУССКИЙ</td><td>Правительство России</td></tr>
<tr><td>ГЕРОИ РОССИИ | СПЕЦНАЗ | ГРУ | ФСБ | ОРУЖИЕ</td><td>МОЙ ПУТИН</td></tr>
<tr><td>РОССИЯ - Моя гордость !</td><td>Партия "Родина". Город Джанкой.</td></tr>
<tr><td>ОМОН.СОБР.ФСБ.ФСВНГ.ФСИН.ОСН.ГРУ.ГЕРОИ РОССИИ</td><td>МИД России</td></tr>
<tr><td>РОССИЯ - ВЕЛИКАЯ СТРАНА</td><td>США — спонсор мирового террора (16+)</td></tr>
<tr><td>Минобороны России</td><td>РУССКОЕ ОБОЗРЕНИЕ || РОССИЯ || ПУТИН</td></tr>
<tr><td>РОССИЯ | Нашу страну не победить!</td><td>ЕДИНАЯ РОССИЯ КРЫМ</td></tr>
<tr><td>Россия везде, где есть русские!</td><td>Вежливые Люди | Армия | Россия</td></tr>
<tr><td>Холодная война 2.0</td><td>ПЕРВЫЙ КРЫМСКИЙ Крым на ладони.</td></tr>
</table>

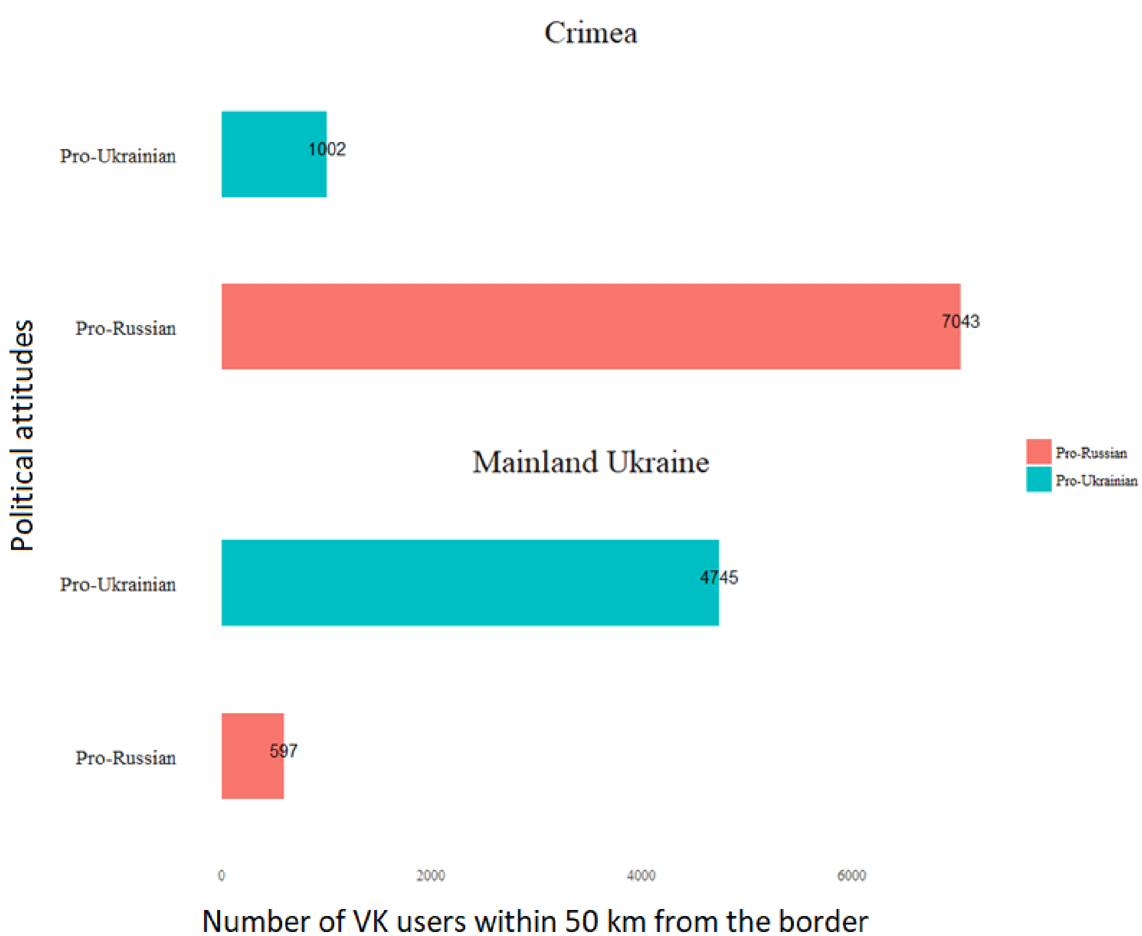
Table E2: VKontakte groups annotated as pro-Ukrainian (N= 48)

Group name	
РЕАЛЬНЫЙ КИЕВ | Київ	Типова Україна
Украина в шоке	Анекдоти UA
Країна У	VK Україна
Знайомся УКРАЇНА	Подслушано Чаплинка
Я + Ти = Україна	Чоткий Укропчик UA
Клуб Брутальних Українців (КБУ)	• Україна смієтьсЯ •
Де є Українці, там є приколи	Україна Online
Чоткий Українець!	Національна гвардія НОВИНИ ООС
Весела Україна	Жизнь мусульман Генического р-на(ДУМУ)
Телеканал СТБ	ДШВ | Десантно Штурмові Війська | Україна
Ржачна країна - наша Україна	Національна поліція України | Police UA
Скадовск 24	Чаплинка
Збройні Сили України	УНИАН | УНІАН
Про мову	Канал 1+1
АТО	ТСН
Український простір	Телеканал ICTV
StopFake	Слава Україні!
Телеканал ТЕТ	Я люблю Україну
-ПАТРІОТ- | НОВИНИ УКРАЇНИ | АТО | ООС	Україна для УКРАЇНЦІВ
У К Р А Ї Н К А	Патріот Ua
УКРАЇНА — ПОНАД УСЕ! - НОВИНИ	Справжній Українець
ВІЙНА ЗА УКРАЇНУ - НОВИНИ ДОНЕЦЬКА АТО	Патріот
УКРОП | НОВИНИ УКРАЇНИ | АТО / ООС	Свідок (HTH)
Суворий Бандерівець	СОБЫТИЯ КРЫМА информационный портал

Appendix F: Pro-Russian and Pro-Ukrainian users

I operationalize pro-Russian users as those who follow Pro-Russian communities and no pro-Ukrainian communities. Conversely, I categorize users as pro-Ukrainian, if they follow only follow pro-Ukrainian communities and not a single pro-Russian one. Using the method, I have classified 13,387 out of the 65,213 active users (see Methods section) in at least one of the two, mutually exclusive categories. Figure F1 below illustrates the distribution of these users on both-sides of the border. I did not download wall posts from all of these users due to VKontakte public API's restrictions on the data volume. As mentioned in the Methods section, I then download all 2,127,398 public posts from a smaller, stratified sub-sample of 3,024 of pro-Russian, pro-Ukrainian and Random users.

Figure F1: Number of pro-Russian and pro-Ukrainian users



Appendix G: Using Crimea as a control region

While locals in annexed Crimea are de facto under Russian rule today, they were subject to Ukrainian law, political system and mass media within just three years and two months prior the ban. This does not mean that the Autonomous Republic of Crimea is demographically or culturally the same as other administrative regions in Ukraine. This is also seen in the data: there are - not surprisingly - more pro-Russian VKontakte users and fewer pro-Ukrainian users in Crimea than in Mainland, as shown in Appendix F. However, Crimea is arguably more similar to other areas in Ukraine than regions in Russia that - despite shared communist past - never were subjected to Ukrainian law, educational system, draft in the Ukrainian army, Ukrainian political party campaigns or other institutions in the country. At the same time, Crimea offers a relatively clear border separating users exposed to censorship from those that are not, unlike the war zone in South-Eastern Ukraine, where parts of the front-line remain fluid and reliable information about individuals living close to the border may be difficult to obtain.

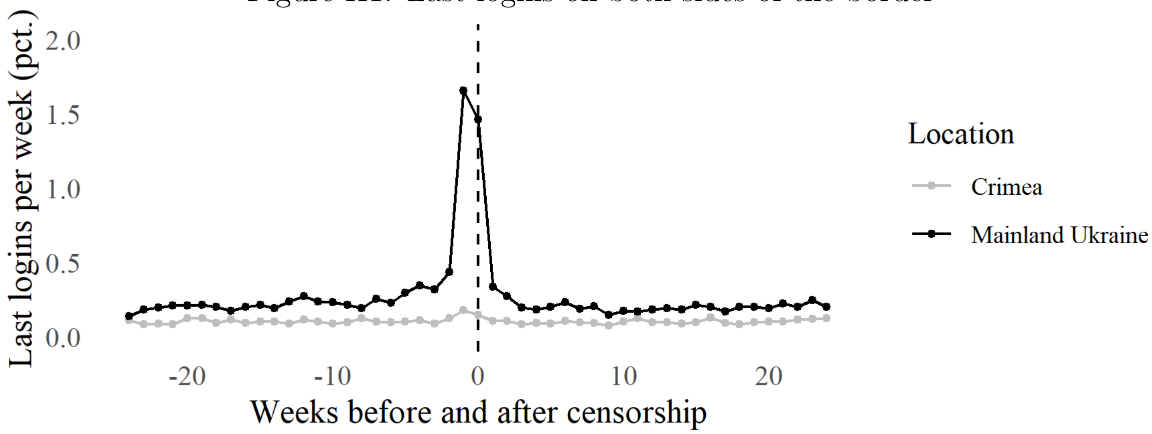
As mentioned in the article, the users in the "treatment" group are sampled from a region north of the Crimean border in Kherson Oblast or "Mainland". The region is demographically similar to northern Crimea. Both are relatively rural areas when comparing to the rest of Ukraine, with relatively similar age, education and gender distribution based on the 2001 national census (Ukrainian Census, 2001). Prior to the Russian annexation in 2014, both regions had a history of a relatively strong support for Ukrainian political candidates, who promoted a stronger integration with Russia. For example, both Kherson Oblast and Crimea (including Simferopol) showed a relatively high voter support for Viktor Yanukovich from Party of Regions during the 2004 and 2010 presidential elections – the president who fled to Russia following the Euromaidan in 2014. However, there are also differences between the two regions. Kherson Oblast has an even lower population density than Crimea (Ukrainian Census, 2001). Furthermore, Kherson Oblast has much lower average salaries than Crimea in 2013, ranking as one of the most impoverished regions in Ukraine (State Statistics, 2013).

Despite of political similarities, Kherson Oblast has shown less support for pro-Russian candidates while also consisting of more ethnic Ukrainians than Crimea (Ukrainian Census, 2001). Overall, Kherson Oblast represents a highly rural and economically strained part of Ukraine with moderate support for pro-Russian political candidates. Despite of this, the findings in this article also hold when rerunning the analysis on users from less pro-Russian, more urban, and economically well-off Kyiv (see Appendix N).

Appendix H: Last logins

What percentage of VKontakte users ceased logging back in, following the ban? Figure H1 illustrates the proportion of last logins per week, where the first 7 days following the earliest implementation of the ban (May 18, 2017) are standardized as 0. I compute the number by dividing the number of last logins by the total number of "active users" accounts on the respective sides of the border.

Figure H1: Last logins on both sides of the border



If the ban unanticipated by the individuals, one would expect some of the unprepared users to be locked out of the platform. The data is in line with this expectation: The number of last logins topped already a few days prior to the ban. This is likely caused by the fact that some users may log in prior to the ban without knowing that this may be the last time before being locked out by the respective ISP. Although the Ukrainian government gave the ISPs two weeks to implement the ban following the announcement of the executive order, the general public did not know that the unexpected ban would already be implemented two days after the government announcement. This gave ordinary citizens little time to prepare for the sudden shutdown. The exogenous nature of the ban is confirmed by the discontinuous decline in mean number of posts per day *after the ban* in Mainland.

The descriptive overview suggests that the sudden ban had a much greater effect on Mainland users than on those in the neighboring Crimea within just 50 km of the border. Of the 23,506 “active users” in Mainland, 3.12% have logged on to VKontakte for the last time in the week before and the week following the policy announcement (weeks - 1 and 0). This means that they did not use their respective VKontakte accounts a single time during the period of approximately 8 months between their last login and the data collection period. In comparison, only 0.33% of the 41.707 active Crimean users have logged in for the last time during this two-week period — one before and one following the implementation of the ban.

A part of the decline in online activity - in terms of last logins - is likely a result of "natural decay" - i.e. factors that are unrelated to the ban. This is likely to be the case, because accounts are regularly "abandoned" long before the ban and on both sides of the border. Nevertheless, the descriptive findings confirm the first hypothesis by showing that the censored Mainland region suffered a greater decline online activity - in term of last logins - than the non-censored Crimea. One must note, however, that this decline remains small, even in Mainland.

Appendix I: DD and DDD models with 30 and 90 days bandwidth

Table I1: Change in posting activity after the ban (for all users in the sub-sample)

	Bandwidth:	
	90 days	30 days
	(1)	(2)
Ban	0.0005	−0.006
	(0.006)	(0.010)
Mainland	0.063***	0.054***
	(0.004)	(0.007)
Pro-Russian	0.050***	0.046***
	(0.003)	(0.006)
Pro-Ukrainian	0.012***	0.006
	(0.003)	(0.006)
Friends in Russia	0.026***	0.023***
	(0.003)	(0.005)
Male	0.028***	0.029***
	(0.003)	(0.005)
Number of friends	0.0002***	0.0003***
	(0.00001)	(0.00002)
Weekend	0.016***	0.010
	(0.003)	(0.005)
Days after the ban	−0.0002***	−0.0002
	(0.0001)	(0.0003)
Ban*Mainland	−0.118***	−0.114***
	(0.005)	(0.009)
Constant	0.093***	0.101***
	(0.005)	(0.008)
Observations	547,344	184,464
Note:	**p<0.05; ***p<0.01	

The models are based on data from all 3,024 users (including those with political affiliations as well as random users).

Table I2: Change in posting activity after the ban (90 days bandwidth)

	Models:				
	DD			DDD	
	(1)	(2)	(3)	(4)	(5)
Ban	−0.012	−0.108***	−0.107***	−0.014	−0.006
	(0.007)	(0.009)	(0.009)	(0.009)	(0.009)
Mainland	0.038***			0.032***	0.063***
	(0.005)			(0.007)	(0.007)
Pro-Russian	0.043***	0.010**	0.018**	0.043***	0.074***
	(0.003)	(0.005)	(0.008)	(0.003)	(0.007)
Friends in Russia	0.015***	0.030***	0.023***	0.012	0.014***
	(0.003)	(0.008)	(0.005)	(0.007)	(0.003)
Male	0.043***	0.015***	0.015***	0.043***	0.043***
	(0.003)	(0.005)	(0.005)	(0.003)	(0.003)
Number of friends	0.0002***	0.0002***	0.0002***	0.0002***	0.0002***
	(0.00001)	(0.00002)	(0.00002)	(0.00001)	(0.00001)
Weekend	0.024***	0.037***	0.037***	0.024***	0.023***
	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)
Days after the ban	−0.0003***	−0.0004***	−0.0004***	−0.0003***	−0.0003***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Ban*Mainland	−0.107***			−0.097***	−0.104***
	(0.007)			(0.009)	(0.009)
Friends in Russia*Mainland				0.011	
				(0.010)	
Ban*Friends in Russia*Mainland				−0.019	
				(0.013)	
Ban*Friends in Russia		−0.015		0.004	
		(0.009)		(0.009)	
Pro-Russian*Mainland					−0.051***
					(0.010)
Ban*Pro-Russian*Mainland					−0.004
					(0.013)
Ban*Pro-Russian			−0.017		−0.013
			(0.009)		(0.009)
Constant	0.122***	0.171***	0.171***	0.123***	0.108***
	(0.005)	(0.007)	(0.007)	(0.006)	(0.006)
Observations	394,399	187,697	187,697	394,399	394,399

**p<0.05; ***p<0.01

The models are based on data from users with political affiliations (and not random users). Models 2-3 are limited to Mainland. The remaining models include users from both sides of the border.

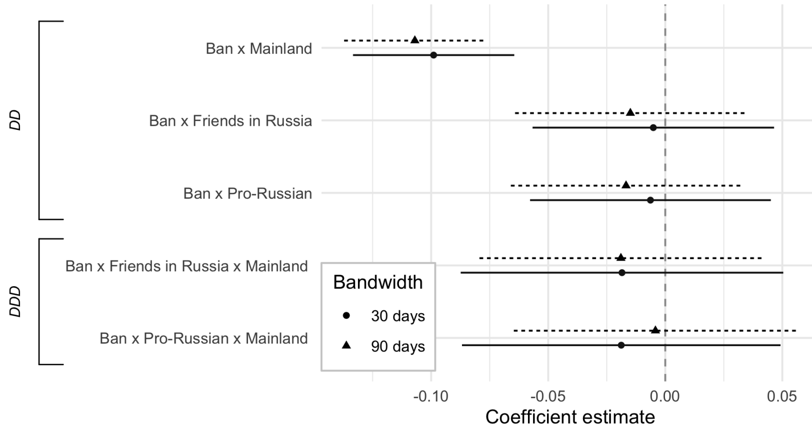
Table I3: Change in posting activity after the ban (30 days bandwidth)

	Models:				
	DD			DDD	
	(1)	(2)	(3)	(4)	(5)
Ban	−0.019	−0.102***	−0.101***	−0.026	−0.025
	(0.012)	(0.017)	(0.016)	(0.016)	(0.015)
Mainland	0.025***			0.012	0.039***
	(0.008)			(0.011)	(0.011)
Pro-Russian	0.045***	0.019**	0.022	0.045***	0.057***
	(0.006)	(0.008)	(0.013)	(0.006)	(0.011)
Friends in Russia	0.009	0.027**	0.024***	−0.006	0.009
	(0.006)	(0.013)	(0.008)	(0.011)	(0.006)
Male	0.048***	0.001	0.001	0.048***	0.047***
	(0.006)	(0.008)	(0.008)	(0.006)	(0.006)
Number of friends	0.0002***	0.0002***	0.0002***	0.0002***	0.0002***
	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)
Weekend	0.015**	0.021**	0.021**	0.015**	0.014**
	(0.006)	(0.009)	(0.009)	(0.006)	(0.006)
Days after the ban	−0.0003	−0.001	−0.001	−0.0003	−0.0003
	(0.0003)	(0.0004)	(0.0004)	(0.0003)	(0.0003)
Ban*Mainland	−0.099***			−0.090***	−0.090***
	(0.011)			(0.016)	(0.016)
Friends in Russia*Mainland				0.026	
				(0.017)	
Ban*Friends in Russia*Mainland				−0.018	
				(0.023)	
Ban*Friends in Russia		−0.005		0.013	
		(0.015)		(0.017)	
Pro-Russian*Mainland					−0.029
					(0.017)
Ban*Pro-Russian*Mainland					−0.019
					(0.022)
Ban*Pro-Russian			−0.006		0.012
			(0.015)		(0.016)
Constant	0.128***	0.174***	0.174***	0.136***	0.123***
	(0.009)	(0.012)	(0.012)	(0.011)	(0.011)
Observations	132,919	63,257	63,257	132,919	132,919

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

The models are based on data from users with political affiliations (and not random users). Models 2-3 are limited to Mainland. The remaining models include users from both sides of the border.

Figure I1: Change in mean number of posts/day after the ban (95% confidence intervals) with user-level Fixed Effects



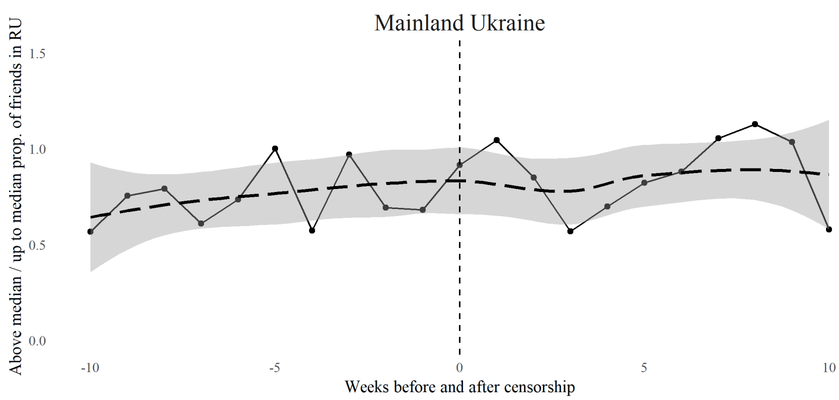
The models are the same as the ones in Table I2 and I3 but now also with user-level Fixed Effects with standard errors clustered by user ID.

Appendix J: Robustness and validity

This appendix gives an overview of the overall validation procedures and robustness tests. These relevant information for the different procedures are included in the subsequent appendices.

I have validated the main findings related to Hypotheses 1a and 1b through a series of steps. First, I have examined the ratio between online activity among users with different numbers of social ties to VKontakte friends in Russia. Using the sample of 23,506 "active users" from Mainland, I compute the ratio by dividing the number of last logins among users with above median proportion of VKontakte friends in Russia with those that have up to median proportion. The findings based on RD remain robust. As shown in Figure J1, there is no discontinuous shift in the ratio between online activity among users with above-median and up-to-median shares of VKontakte friends in Russia.

Figure J1: Ratio between number of last logins among users with above-median and up-to-median proportions of friends in Russia.



The smoothing line and confidence intervals (95%) are made using Loess regression.

Second, I test the DD and DDD models for robustness by replacing the variable for the proportion of VKontakte friends in Russia with different dichotomous variables, such as whether the user has 75% or more of their friends living in Russia. By doing so, I also ensure that the results hold, when examining the more "extreme" users with a very high proportion of friends in Russia. In addition to this, I ensure model robustness by reiterating the main analysis by using negative binomial regression models instead of OLS (Appendix R). I rerun the OLS models in a separate procedure by adding user-level Fixed Effects (Figure I1). Furthermore, I test whether the findings are persistent over time, by gradually expanding the post-ban time period from 10 to 400 days for the DD and DDD models used in the main analysis (results reported in Appendix K). Moreover, I test for tie strength by conducting a second wave of data collection in which I download all of the re-posts from the users. I repeat the DD and DDD models by measuring social ties to Russia using a dichotomous variable, indicating whether the user has at least one strong tie to a Russian user on VKontakte or not based on how often the respective users repost each other (see Appendix L for details). In all five instances, the overall findings reported in the results sections remains robust.

In the third step, I supplement the difference-in-differences models by using a Regression Discontinuity in Time (RDiT) in order to estimate the average effect of the ban on posting activity *around the cut-off date* that separates the pre-ban (control) and ban (treament) period. As shown in Appendix M, the RDiT models indicate that ban substantively reduced online activity among all of the groups in the sub-sample in Mainland. These results corroborate the findings based on DD and DDD models in the analysis.

As mentioned earlier, more than 90 % of the users in each political subgroup continue logging even within 30 days after the ban. In the fourth step, I ensure the findings based on DD and DDD in the previous sections remain robust when limiting the analysis only to these accounts. This is indeed the case: the ban has reduced online posting activity even among users who continue logging back in after the ban.

In the fifth step, I test whether the political affiliation variable captures meaningful differences between users. For instance, one would expect users who follow pro-Russian groups to be closely affiliated to Russia in other ways as well. This is indeed the case. Users labeled as "Pro-Russian" have 35.5 percent higher proportion of VKontakte friends in Russia than their "Pro-Ukrainian" counterparts in Crimea, and 42.6 percent higher in Mainland. Similarly, "Pro-Russian" profiles have 22.9 percent more VKontakte friends in Russia that they repost (on their own respective wall) than pro-Ukrainian users in Mainland. The number is as high as 66.3 percent in Crimea. Whereas pro-Russian users in both Mainland and Crimea upload on average 14.4 posts on their own walls with a link to a website with a Russian domain (".ru"), pro-Ukrainian users upload only 7.93 posts, i.e. 45.8 percent less.

In the sixth step, I examine the external validity of the findings related to political affiliations (Hypothesis 1a) by using DD to compare the change in posting activity among pro-Russian and pro-Ukrainian users from Kyiv. The results remain largely the same as those reported in the previous chapters, with no substantive or significant differences in terms of how pro-Russian and pro-Ukrainian users were affected by the ban (see Appendix N for details).

Lastly, I test whether the decline in online activity is caused by a "placebo" in the form a holiday or other reoccurring events rather than the actual censorship "treatment". As I show in Appendix S, online activity during the pre-censorship period *increased* on both sides of the Crimean border. This suggests that the sudden decline after the ban cannot be explained by re-occurring or seasonal events alone.

As mentioned previously, the study is limited to public VKontakte accounts and posts. For this reason, it is not possible to empirically exclude that the users made their posts private after the ban either out of fear of social stigma or legal reprisal, while keeping their overall activity (consisting of private and public posts) unchanged.

Similar to the discussion section, it is theoretically unlikely that the individuals in the sample shifted towards private wall posts out of fear of legal consequences, as their login date is visible to all users. Furthermore, private wall posts do not eliminate the risks of being socially stigmatized as unpatriotic by peers, since these posts are, in most cases, still visible to VKontakte friends. It is important to note that the pro-Russian users in the sample did not hide their affiliation to pro-Russian VKontakte communities, meaning that this information is available even to strangers outside of their immediate VKontakte "friendship" network. While this study is unique due to its focus on individual micro-level effects of the ban, multiple analytics firms have observed a major decline in the overall aggregate-level popularity or traffic towards VKontakte in Ukraine (Dek, 2019; Maria Shortgin and Roman Lebed, Maria Shortgin and Roman Lebed; Kantar, 2017). These results are also based on traffic from non-public accounts and are not limited to public wall posts. This corroborates the main findings, suggesting that the decline in online activity is more than a mere shift in privacy settings from public to private posts.

Appendix K: Long-term effects of the ban

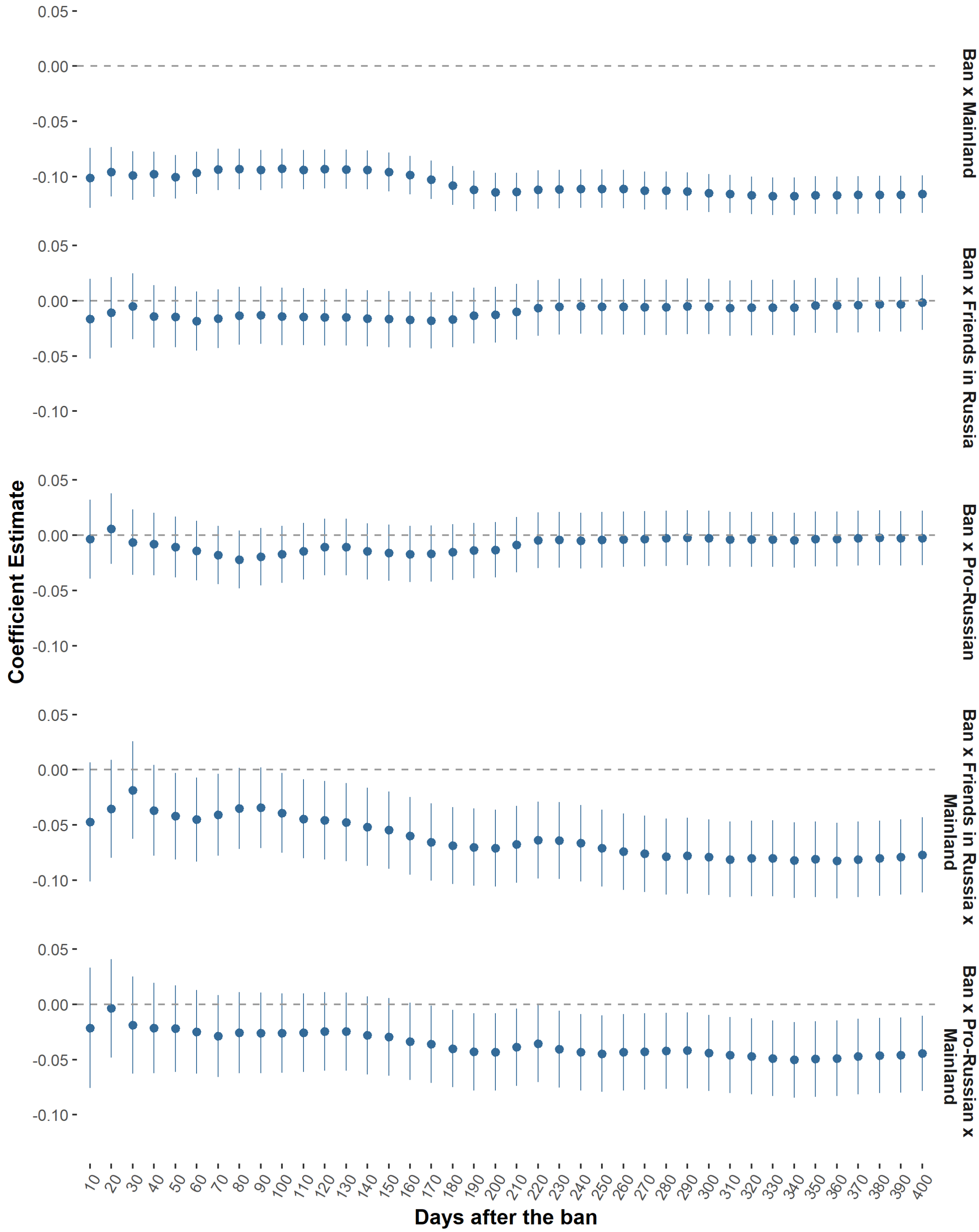
In this appendix, I will examine whether the findings based on the DD and DDD models in the analysis persistent over a longer period of time. Figure K1 show estimates from 200 DD and DDD models for users with a pro-Russian of pro-Ukrainian political affiliation. I gradually expand the post-ban period from 10 to 400 days, while keeping pre-ban period constant within a relatively narrow time range of 30 days (see Table K1 and Table K2 for a post ban period of 30 and 400 days). As reflected by the values for *Ban x Mainland* term in Figure K1, the effect of the ban among political users is both statistically significant, robust and long-term. This is the case even when comparing the daily mean number of posts in the 400 days post-ban period to the 30 days right before the ban. The substantively large effect size ranges between a decrease of 0.09 and 0.12 daily mean posts per user. This is equivalent to a 39.4% and 52.6% decline in the 30-days pre-ban activity in Mainland for the political users. The models becomes more vulnerable to confounding factors, when using a longer time range. Because of this, it is difficult to estimate the causal effect of the ban when using data from a period long after the censorship implementation date. However, these estimates do indicate that the online activity remained low among censored users even long after the ban.

I find no statistically significant difference in the effect size (*Ban x Friends in Russia*) among users with up to median and above median proportion of friends in Russia in the DD models based on users in Mainland. Similarly, the DD estimate based on users in Mainland show no statistically significant difference between pro-Russian and pro-Ukrainian users.

The difference between users with strong political and social affiliations to Russia becomes statistically significant when introducing Crimea as a control in the DDD model. The difference is not robust, since the estimates are only significant when using long-term data. Even in this case, however, the findings suggest that users with above-median proportion are *more* affected by the ban than those with fewer friends in Russia. The difference is equivalent to 0.039 mean posts per day ($t = -2.129, p < 0.05$) according to the most conservative and statistically significant estimate (100 days post-ban period). Similarly, the ban has reduced online activity by -0.036 posts per day ($t = -2.023, p < 0.05$) more among pro-Russian than pro-Ukrainian users according to the most conservative, statistically significant estimate (170 days post-ban period).

Overall, results based on both DD and DDD confirm the key findings in the analysis. The ban has reduced the overall posting activity in Mainland, as predicted by Hypothesis 1. The effect of the ban persists over a longer period of time. The findings show little support for Hypothesis 1a and 1b, which predict that users with either political or social affiliations to Russia are more affected by the anti-Kremlin shutdown. In case of a difference between the user groups, the findings suggest that those with stronger affiliations to Russia are more affected by the ban in longer term — contrary to the two hypotheses.

Figure K1: Change in mean number of posts per day after the ban (95% confidence intervals)



The models are based on data from users with political affiliations (and not random users).

Table K1: Change in posting activity (Timerange: from 30 days prior to 200 days after the ban)

	Models:				
		DD		DDD	
	(1)	(2)	(3)	(4)	(5)
Ban	-0.041^{***}	-0.122^{***}	-0.121^{***}	-0.070^{***}	-0.055^{***}
	(0.007)	(0.009)	(0.009)	(0.009)	(0.009)
Mainland	0.025^{***}			0.012	0.041^{***}
	(0.008)			(0.011)	(0.011)
Pro-Russian	0.047^{***}	0.010^{***}	0.022	0.047^{***}	0.054^{***}
	(0.003)	(0.004)	(0.012)	(0.003)	(0.011)
Friends in Russia	0.028^{***}	0.028^{**}	0.017^{***}	-0.005	0.028^{***}
	(0.003)	(0.012)	(0.004)	(0.011)	(0.003)
Male	0.040^{***}	0.016^{***}	0.016^{***}	0.040^{***}	0.039^{***}
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Number of friends	0.0002^{***}	0.0002^{***}	0.0002^{***}	0.0002^{***}	0.0002^{***}
	(0.00001)	(0.00002)	(0.00002)	(0.00001)	(0.00001)
Weekend	-0.001	0.034^{***}	0.034^{***}	-0.001	-0.002
	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)
Days after the ban	0.0003^{***}	0.0001	0.0001	0.0003^{***}	0.0003^{***}
	(0.00003)	(0.00003)	(0.00003)	(0.00003)	(0.00003)
Ban*Mainland	-0.114^{***}			-0.078^{***}	-0.093^{***}
	(0.009)			(0.012)	(0.012)
Friends in Russia*Mainland				0.026	
				(0.017)	
Ban*Friends in Russia*Mainland				-0.071^{***}	
				(0.018)	
Ban*Friends in Russia		-0.013		0.058^{***}	
		(0.013)		(0.012)	
Pro-Russian*Mainland					-0.031
					(0.017)
Ban*Pro-Russian*Mainland					-0.043^{**}
					(0.018)
Ban*Pro-Russian			-0.013		0.030^{**}
			(0.013)		(0.012)
Constant	0.131^{***}	0.167^{***}	0.167^{***}	0.148^{***}	0.129^{***}
	(0.006)	(0.009)	(0.009)	(0.008)	(0.008)

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

The models are based on data from users with political affiliations (and not random users). Models 2-3 are limited to Mainland. The remaining models include users from both sides of the border.

Table K2: Change in posting activity (Timerange: from 30 days prior to 400 days after the ban)

	Models:				
	DD			DDD	
	(1)	(2)	(3)	(4)	(5)
Ban	−0.005	−0.112***	−0.111***	−0.043***	−0.025***
	(0.006)	(0.008)	(0.009)	(0.009)	(0.008)
Mainland	0.026***			0.013	0.042***
	(0.008)			(0.011)	(0.011)
Pro-Russian	0.052***	0.014***	0.016	0.052***	0.048***
	(0.002)	(0.003)	(0.012)	(0.002)	(0.011)
Friends in Russia	0.042***	0.027**	0.026***	−0.007	0.041***
	(0.002)	(0.012)	(0.003)	(0.011)	(0.002)
Male	0.029***	0.001	0.001	0.029***	0.028***
	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
Number of friends	0.0003***	0.0003***	0.0003***	0.0003***	0.0003***
	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)
Weekend	0.014***	0.055***	0.055***	0.014***	0.013***
	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)
Days after the ban	0.00001	−0.00003***	−0.00003***	0.00001	0.00001
	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)
Ban*Mainland	−0.116***			−0.077***	−0.095***
	(0.009)			(0.012)	(0.012)
Friends in Russia*Mainland				0.026	
				(0.017)	
Ban*Friends in Russia*Mainland				−0.077***	
				(0.017)	
Ban*Friends in Russia		−0.001		0.076***	
		(0.013)		(0.012)	
Pro-Russian*Mainland					−0.031
					(0.017)
Ban*Pro-Russian*Mainland					−0.044**
					(0.017)
Ban*Pro-Russian			−0.003		0.042***
			(0.013)		(0.012)
Constant	0.107***	0.156***	0.156***	0.131***	0.110***
	(0.006)	(0.008)	(0.009)	(0.008)	(0.008)

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

The models are based on data from users with political affiliations (and not random users). Models 2-3 are limited to Mainland. The remaining models include users from both sides of the border.

Appendix L: Validation - The strength of social ties

Table L1: Change in posting activity after the ban (DD and DDD)

	Bandwidth:			
	30 days		90 days	
	(1)	(2)	(3)	(4)
Ban	-0.118^{***}	-0.033^{***}	-0.143^{***}	-0.041^{***}
	(0.008)	(0.009)	(0.008)	(0.009)
Friends in Russia	0.050^{***}	-0.026^{**}	0.028	-0.005
	(0.016)	(0.013)	(0.016)	(0.013)
Mainland		0.007		0.029^{***}
		(0.009)		(0.009)
Pro-Russian	0.024^{***}	0.049^{***}	0.015	0.046^{***}
	(0.008)	(0.006)	(0.008)	(0.006)
Male	0.002	0.048^{***}	0.016^{**}	0.043^{***}
	(0.008)	(0.006)	(0.008)	(0.006)
Number of friends	0.0001^{***}	0.0002^{***}	0.0002^{***}	0.0002^{***}
	(0.00002)	(0.00002)	(0.00002)	(0.00002)
Weekend	0.021^{**}	0.015^{***}	0.037^{***}	0.024^{***}
	(0.009)	(0.006)	(0.009)	(0.006)
Ban*Friends in Russia	-0.040^{**}	0.019	-0.017	0.006
	(0.019)	(0.019)	(0.019)	(0.019)
Ban*Mainland		-0.085^{***}		-0.101^{***}
		(0.013)		(0.013)
Friends in Russia*Mainland		0.074^{***}		0.034
		(0.021)		(0.021)
Ban*Friends in Russia*Mainland		-0.059^{**}		-0.023
		(0.027)		(0.027)
Constant	0.185^{***}	0.141^{***}	0.193^{***}	0.143^{***}
	(0.009)	(0.008)	(0.009)	(0.008)
Observations	63,257	132,919	187,697	394,399
Note:				**p<0.05; ***p<0.01

The variable "Friends in Russia" previously reflected whether the user had above median proportion of friends in Russia or not. In the table above, the variable reflects whether the user has a strong tie to at least one user living in Russia. All of the models include only pro-Russian and pro-Ukrainian users. I consider a social tie between two users to be "strong", if they are online friends *and* have re-posted each other at least once. I limit the analysis to outgoing social ties (i.e. outdegree), by examining who the users in the sub-sample re-post and not the re-posts done by their online friends. The DD Models 1 and 3 only contain users from Mainland. Model 2 and 3 are based on DDD and contain users from both sides of the border.

Appendix M: Regression Discontinuity (RDiT)

In this appendix, I use a regression discontinuity in time (RDiT) design to supplement the difference-in-differences (DD) and triple differences (DDD) approach and to strengthen the paper's causal argument. The two sections below specify the model as well as present the results based on user-day observations for the 1,433 users and a bandwidth of 90 days before and after the ban (181 days in total). Overall, these results corroborate the findings based on DD and DDD by showing that the ban substantively reduced online activity. The decline is discontinuous, statistically significant and substantively large for users in the pro-Russian, pro-Ukrainian and Random user split sample, as well as users with above median and up to median proportion of VKontakte friends in Russia.

Regression Discontinuity (RDiT) specification

I use RDiT (Hausman and Rapson, 2018) to estimate the average effect of the ban on mean number of posts in Mainland right around the threshold separating the period immediately before the ban (control) and after the ban (treatment). The implementation date (18th of May 2017) is standardized as 0.

Specifically, I estimate the local linear regression using the following equation:

$$y_t = \alpha + \beta_1 time_{it} + \beta_2 ban_{it} + \delta X + \varepsilon_{it} \tag{1}$$

In the equation above, y_{it} is the average number of wall posts for user i on day t, where $time_{it}$ is a running variable indicating the number of days before and after the ban. Here, the running variable is standardized, where the earliest feasible implementation date (May 18, 2017) has a value of 0. In this sense, $time_{it} = 0$ is a cutoff separating the period before and after censorship. ε_{it} indicates the error term, X reflects controls, and ban_{it} is a binary variable with a value of 1 if the post originates from after the ban and 0 otherwise. β_2 indicates the mean effect of the ban. In order to capture the change in the slope after the ban, I further expand the equation:

$$y_t = \alpha + \beta_1 time_{it} + \beta_2 ban_{it} + \beta_3 time_{it} * ban_t + \delta X + \varepsilon_{it} \quad (2)$$

The only difference between the two equations is the interaction term $\beta_3 time_{it} * ban_{it}$, which allows different slopes on both sides of the threshold. The RD in both equations is "sharp", where the individuals are seen either as exposed to censorship or not, depending on whether they post prior to the sharp cutoff date. Controls are unnecessary if the error term does not change discontinuously around the threshold (Anderson, 2014, 2777). Nevertheless, I add a binary control variable—indicating whether the post was written on Saturday or Sunday — in order to obtain more precise estimates and to control for a potential increase in user activity due to a "weekend effect" (Hausman and Rapson, 2018).

Regression Discontinuity (RDiT) results

In this section, I will use RDiT to estimate the mean effect of the ban on posting activity among a sub-sample of 1,433 profiles from Mainland Ukraine, including pro-Russian, pro-Ukrainian, and random users. Figure M1 shows the mean number of posts per user for each day. The figure illustrates the discontinuous reduction in the daily posting activity for the average user.

Results from the RD models reported in Table M1 indicate that the ban has a significant, negative effect on the average posting activity. Model 1 includes a running variable for the number of days before/after the cutoff and a dummy variable, *ban*, indicating whether the posts have been uploaded before or after the implementation of the ban. The estimates for *ban* indicate that the sudden shutdown has reduced the posting activity by 0.117 posts per day ($t = -15.214$) for the average user.

In addition to these variables, I have included an interaction term between the two variables in Model 2. The regression lines in Figure M1 reflect the estimates from this expanded equation. The effect remains largely similar, with an average decrease of -0.116

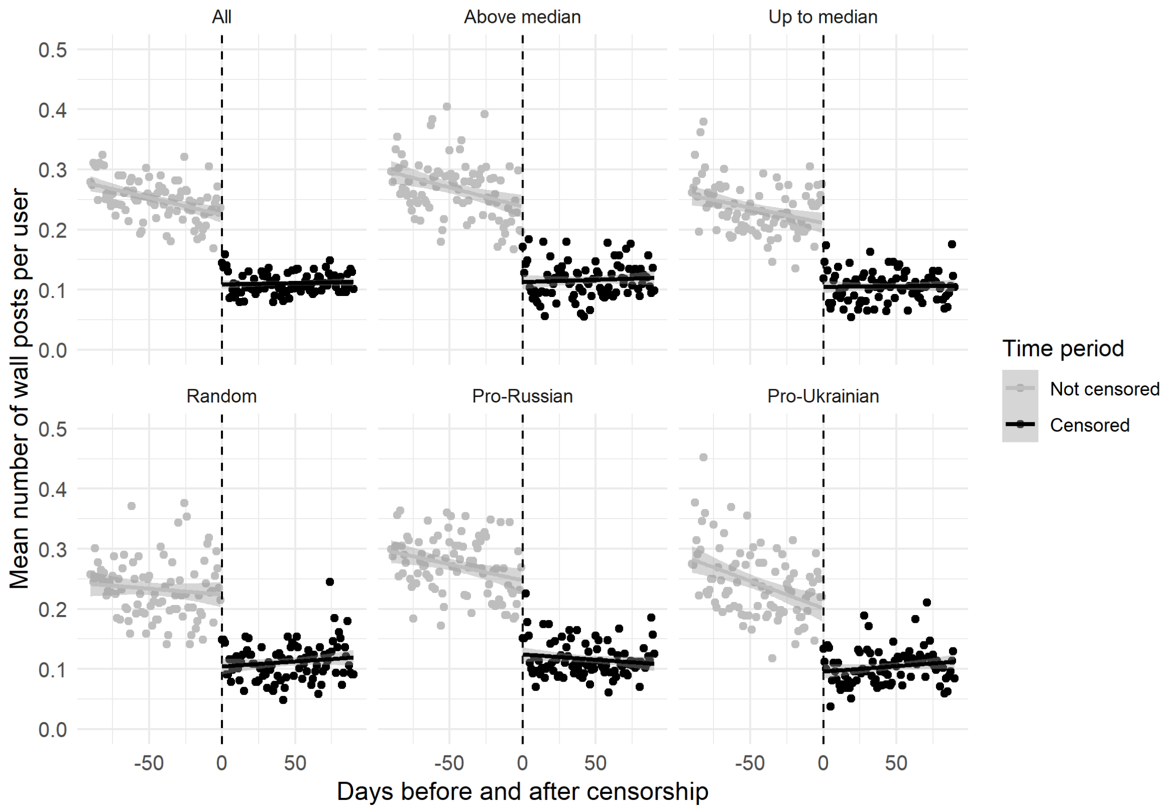


Figure M1: Regression discontinuity in posting activity 90 days before and after the ban (95% confidence interval)

The plot is based on models in table M1 and M1 for the users in the sub-sample.

posts per day ($t = -15.131$) or approximately 0.8 posts per week. This effect is statistically significant and substantively large. Users on the Mainland side of the border post on average 0.251 posts per day prior to the implementation of the censorship policy. The effect size of the ban is equivalent to a 46.2% decrease in the pre-ban average number of posts per day.

I then test whether the ban had an effect on the individual groups by running the local linear regression from Model 2 in table Table M1 in a series of split samples among users from Mainland. As shown in Table M1-M1, as well as Figure M1, the ban was followed by a substantive decline in daily mean number of posts among all of the groups in the split-sample. In other words, the ban had a negative effect on Random, pro-Russian and pro-Ukrainian users, as well as those with up to median- or above median proportion of VKontakte friends in Russia. This further validates the findings based on DD and DDD.

Table M1: Regression discontinuity in no. of posts in Mainland (90 days before and after the ban)

	User samples:			
	All		Above median	Up to median
	(1)	(2)	(3)	(4)
Days	-0.0003^***	-0.001^***	-0.001^***	-0.001^***
	(0.0001)	(0.0001)	(0.0002)	(0.0001)
Ban	-0.117^***	-0.116^***	-0.126^***	-0.106^***
	(0.008)	(0.008)	(0.012)	(0.010)
Weekend	0.021^***	0.021^***	0.065^***	-0.023^***
	(0.004)	(0.004)	(0.006)	(0.005)
Days*ban		0.001^***	0.001^***	0.001^***
		(0.0001)	(0.0002)	(0.0002)
Constant	0.232^***	0.218^***	0.218^***	0.218^***
	(0.004)	(0.006)	(0.009)	(0.007)
Observations	259,373	259,373	129,234	130,139
Note:				**p<0.05; ***p<0.01

Table M2: Regression discontinuity in no. of posts in Mainland (90 days before and after the ban)

	User samples:		
	Pro-Russian	Pro-Ukrainian	Random
	(1)	(2)	(3)
Days	-0.001^{***}	-0.001^{***}	-0.0002
	(0.0002)	(0.0002)	(0.0002)
Ban	-0.123^{***}	-0.106^{***}	-0.120^{***}
	(0.013)	(0.013)	(0.014)
Weekend	0.015^{**}	0.062^{***}	-0.023^{***}
	(0.007)	(0.007)	(0.008)
Days*ban	0.0004	0.001^{***}	0.0004
	(0.0003)	(0.0002)	(0.0003)
Constant	0.243^{***}	0.183^{***}	0.231^{***}
	(0.010)	(0.009)	(0.010)
Observations	95,568	92,129	71,676
Note:			**p<0.05; ***p<0.01

Appendix N: VKontakte groups used for sampling pro-Russian and pro-Ukrainian users from Kyiv

Figure N1: Pro-Ukrainian Group in VKontakte

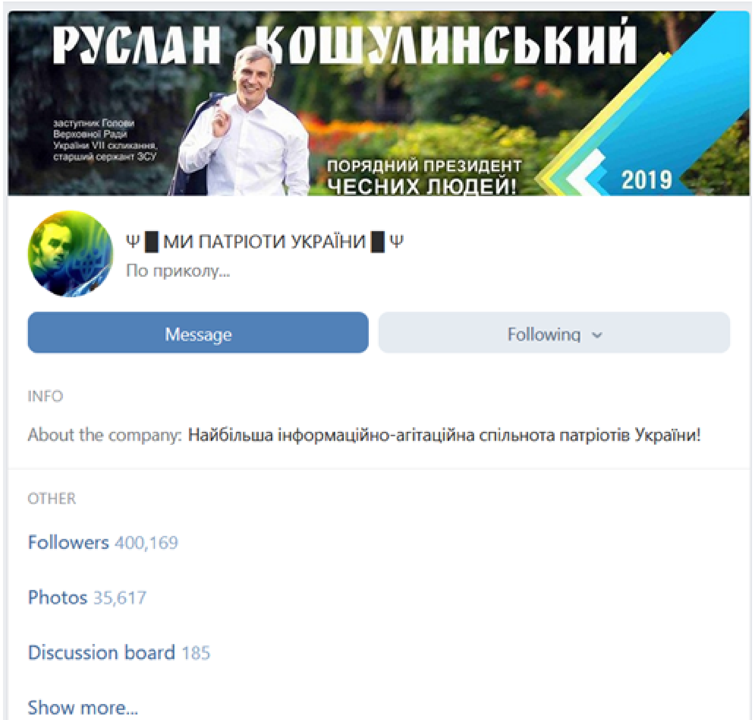
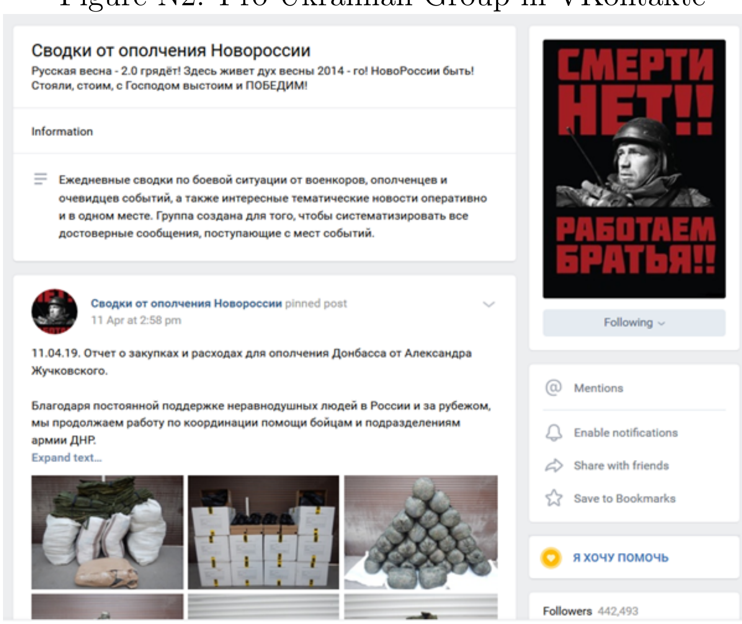


Figure N2: Pro-Ukrainian Group in VKontakte



Appendix O: Validation - Pro-Russian and Pro-Ukrainian users from Kyiv

I have collected all wall posts from a sub-sample of users living in Kyiv. The data collection took place 7 month after the initial collection of data from users living near the Crimea-Mainland border. I select Kyiv, because this is a region where nationalist and pro-Western political candidates are traditionally more popular than in the regions around the Crimean border. I proceed by selecting two additional VKontakte groups. One is highly pro-Russian and is used to crowdfund supplies and military gear for Russian separatists in Ukraine, while the other is heavily centered on Ukrainian nationalism and anti-Kremlin sentiment. I then download metadata for all of the group members and randomly select 300 users for each group among the users who living in Kyiv according to self-reported geolocation. I then use a DD approach to measure the difference between the two groups.

Table O1: DD results for users from Kyiv

	Bandwidth (before and after the ban):	
	30 days	90 days
	(1)	(2)
Ban	-0.003	-0.035^{***}
	(0.014)	(0.007)
Pro-Russian	0.022^{**}	0.004
	(0.009)	(0.005)
Weekend	-0.001	-0.001
	(0.008)	(0.004)
Ban*Pro-Russian	-0.004	0.007
	(0.024)	(0.011)
Constant	0.102^{***}	0.108^{***}
	(0.006)	(0.004)
Observations	32,818	97,378
Note:		**p<0.05; ***p<0.01

Appendix P: Telegram ban in Russia

After leaving Russia, VKontakte's Durov created Telegram, a end-to-end encrypted social media app. Russian authorities tried to block the app, because Durov refused to give them access to private messages and other sensitive data. The opposition and other activists responded to the ban with nation-wide protests (Burgess, 2018). While there is little academic research on this case, the Mediascope analytics firm has observed an increase in Russian traffic to the platform from 3,7 million daily users in April 2018 to 4 million in November (Sobolev, 2019) (see Appendix Q for a comparison with the VKontakte ban).[1]

Why did the VKontakte ban succeed, but not the Telegram ban in Russia? I advance at least two theoretical interpretations, which I will explain in further detail here. Following a "Streisand effect" logic, one possible explanation is that Telegram, with its relatively small audience and high novelty value at the time of the ban, was not as known as VKontakte (Sobolev, 2019). While VKontakte was already a mundane part of most online users' everyday life in both Russia and Ukraine, the censorship controversy may have attracted new users, who were curious to discover what the government wanted to hide from them. The novelty-value of the forbidden content, or the lack thereof, emerges as a potentially important factor for censors who wish to minimize the "Streisand effect".

A second explanation is that the Russian authorities did not succeed in technically implementing the Telegram ban. The failure is likely caused by indirect support from Western firms, which allowed Telegram to mask traffic as Google or Amazon in a technique called "domain fronting" (Burgess, 2018). Once the Russian authorities blocked a specific Google IP address, it was simply replaced with thousands of new ones in an endless circle of cat-and-mouse game. As a result, a large proportion of Russian users can easily access the website even without VPN, according to Telegram Analytics (TGStat, 2019). In contrast, Ukraine was relatively successful at implementing the technical aspect of the censorship policy. Major Ukrainian ISPs began effectively blocking access to the platform within just few days

[1] The numbers are based on individuals between 12–64 years old in cities above 100,000 inhabitants.

after the announcement (UNIAN, 2017). While this study does not empirically test the two interpretations, the findings may serve as a stepping stone for future research on why some large-scale censorship policies succeed and others do not.

Appendix Q: Repeating the analysis with users who did not post prior to the ban

This Appendix reiterates the main tables and figures from the analysis, while also including users who did *not* post anything prior to the ban. Whereas results in the paper are based on a sub-sample of up to 3,024 profiles, the expanded sub-sample below includes up to 3,502 profiles. The main findings remain the same.

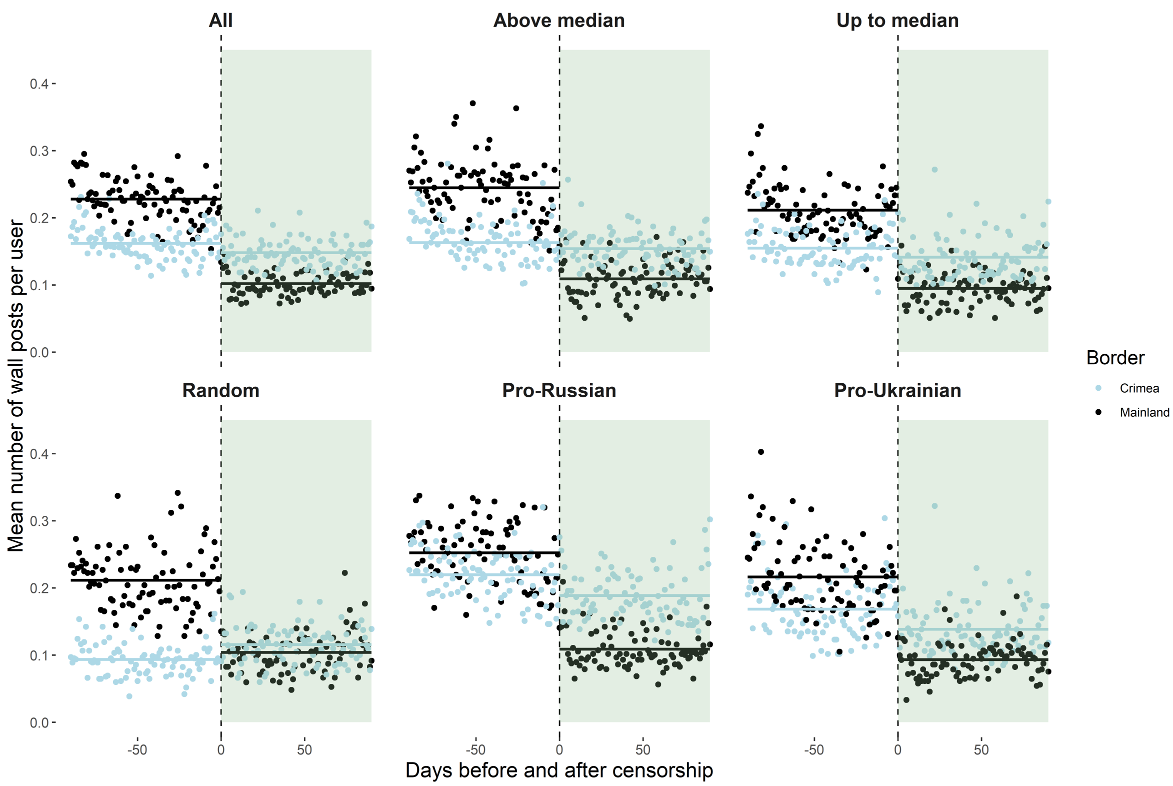


Figure Q1: Change in mean number of posts per day after the ban (for all users and the split-sample)

Table Q1: Descriptive statistics

Users sub-sample and proportion of VKontakte friends in Russia

Mainland Ukraine				
Group	N users	Mean prop. of friends	Median prop. of friends	SD
Pro-Russian	569	0.138	0.075	0.167
Pro-Ukrainian	571	0.079	0.050	0.092
Random	436	0.069	0.044	0.081
All	1,576	0.098	0.056	0.126
Crimea				
Group	N users	Mean prop. of friends	Median prop. of friends	SD
Pro-Russian	649	0.212	0.164	0.170
Pro-Ukrainian	658	0.140	0.100	0.143
Random	619	0.207	0.154	0.178
All	1,926	0.186	0.139	0.167

Posting activity 90 days before and after the ban

Mainland Ukraine					
Group	N posts	Mean daily posts (before)	SD	Mean daily (after)	SD
Pro-Russian	17,457	0.252	0.473	0.109	0.260
Pro-Ukrainian	14,275	0.216	0.428	0.093	0.252
Random	15,179	0.211	0.440	0.104	0.272
All	46,911	0.228	0.448	0.102	0.260
Crimea					
Group	N posts	Mean daily posts (before)	SD	Mean daily (after)	SD
Pro-Russian	23,739	0.220	0.439	0.189	0.391
Pro-Ukrainian	17,835	0.168	0.360	0.138	0.354
Random	12,342	0.093	0.256	0.115	0.389
All	53,916	0.161	0.365	0.148	0.379

N users = 3,502; N posts (90 days bandwidth) =100,827

Table Q2: Change in posting activity after the ban (90 days bandwidth)

	Models:				
	DD				
	(1)	(2)	(3)	(4)	(5)
Ban	−0.007	−0.097***	−0.094***	−0.009	−0.006
	(0.006)	(0.009)	(0.008)	(0.008)	(0.008)
Mainland	0.043***			0.036***	0.052***
	(0.005)			(0.006)	(0.006)
Pro-Russian	0.033***	0.013***	0.024***	0.033***	0.046***
	(0.003)	(0.004)	(0.007)	(0.003)	(0.006)
Friends in Russia	0.015***	0.029***	0.022***	0.011	0.015***
	(0.003)	(0.007)	(0.004)	(0.006)	(0.003)
Male	0.036***	0.010**	0.010**	0.036***	0.036***
	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)
Number of friends	0.0002***	0.0002***	0.0002***	0.0002***	0.0002***
	(0.00001)	(0.00002)	(0.00002)	(0.00001)	(0.00001)
Weekend	0.017***	0.032***	0.032***	0.017***	0.017***
	(0.003)	(0.005)	(0.005)	(0.003)	(0.003)
Days after the ban	−0.0003***	−0.0003***	−0.0003***	−0.0003***	−0.0003***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Ban*Mainland	−0.103***			−0.094***	−0.093***
	(0.006)			(0.008)	(0.008)
Friends in Russia*Mainland				0.015	
				(0.009)	
Ban*Friends in Russia*Mainland				−0.018	
				(0.012)	
Ban*Friends in Russia		−0.015		0.004	
		(0.008)		(0.008)	
Pro-Russian*Mainland					−0.018
					(0.009)
Ban*Pro-Russian*Mainland					−0.020
					(0.012)
Ban*Pro-Russian			−0.021**		−0.001
			(0.008)		(0.008)
Constant	0.106***	0.152***	0.150***	0.108***	0.099***
	(0.005)	(0.007)	(0.007)	(0.005)	(0.005)
Observations	442,907	206,340	206,340	442,907	442,907

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

The models are based on data from users with political affiliations (and not random users). Models 2-3 are limited to Mainland. The remaining models include users from both sides of the border.

Appendix R: Negative Binomial regressions

This appendix re-iterates the DD and DDD models using the same variables and data as in the main analysis, now using negative binomial regressions instead of a simple OLS. The findings remain robust: The ban reduced online activity in Mainland Ukraine. There is no statistically significant difference in the effect size among pro-Russian and pro-Ukrainian users. The pattern is the same when comparing users with many friends and few friends in Russia.

Table R1: Change in posting activity after the ban (90 days bandwidth)

	Models:				
	DD			DDD	
	(1)	(2)	(3)	(4)	(5)
Ban	-0.079^{**}	-0.750^{***}	-0.723^{***}	-0.096	-0.099
	(0.040)	(0.058)	(0.057)	(0.051)	(0.052)
Mainland	0.185^{***}			0.184^{***}	0.294^{***}
	(0.022)			(0.032)	(0.034)
Pro-Russian	0.212^{***}	0.067^{**}	0.061	0.212^{***}	0.301^{***}
	(0.018)	(0.028)	(0.032)	(0.018)	(0.031)
Friends in Russia	0.086^{***}	0.098^{***}	0.128^{***}	0.066^{**}	0.084^{***}
	(0.018)	(0.031)	(0.027)	(0.031)	(0.018)
Male	0.216^{***}	0.077^{***}	0.078^{***}	0.216^{***}	0.212^{***}
	(0.018)	(0.027)	(0.027)	(0.018)	(0.018)
Number of friends	0.001^{***}	0.001^{***}	0.001^{***}	0.001^{***}	0.001^{***}
	(0.00002)	(0.00005)	(0.00005)	(0.00002)	(0.00002)
Weekend	0.107^{***}	0.155^{***}	0.154^{***}	0.108^{***}	0.100^{***}
	(0.019)	(0.028)	(0.028)	(0.019)	(0.019)
Days after the ban	-0.001^{***}	-0.001^{***}	-0.001^{***}	-0.001^{***}	-0.001^{***}
	(0.0004)	(0.001)	(0.001)	(0.0004)	(0.0004)
Ban*Mainland	-0.655^{***}			-0.664^{***}	-0.641^{***}
	(0.037)			(0.053)	(0.057)
Friends in Russia*Mainland				0.002	
				(0.044)	
Ban*Friends in Russia*Mainland				0.018	
				(0.074)	
Ban*Friends in Russia		0.067		0.034	
		(0.055)		(0.050)	
Pro-Russian*Mainland					-0.218^{***}
					(0.044)
Ban*Pro-Russian*Mainland					-0.022
					(0.074)
Ban*Pro-Russian			0.013		0.037
			(0.055)		(0.050)
Constant	-2.064^{***}	-1.820^{***}	-1.832^{***}	-2.054^{***}	-2.104^{***}
	(0.029)	(0.039)	(0.039)	(0.033)	(0.034)
Observations	394,399	187,697	187,697	394,399	394,399

$^{**}p<0.05$; $^{***}p<0.01$

The models are based on data from users with political affiliations (and not random users). Models 2-3 are limited to Mainland. The remaining models include users from both sides of the border.

Table R2: Change in posting activity after the ban (30 days bandwidth)

	Models:				
	DD			DDD	
	(1)	(2)	(3)	(4)	(5)
Ban	−0.122	−0.748***	−0.737***	−0.163	−0.189**
	(0.069)	(0.109)	(0.098)	(0.087)	(0.088)
Mainland	0.152***			0.100	0.199***
	(0.038)			(0.055)	(0.058)
Pro-Russian	0.247***	0.147***	0.101	0.248***	0.233***
	(0.032)	(0.050)	(0.055)	(0.031)	(0.054)
Friends in Russia	0.069**	0.098	0.158***	−0.023	0.069**
	(0.032)	(0.055)	(0.049)	(0.054)	(0.032)
Male	0.259***	0.031	0.034	0.258***	0.258***
	(0.031)	(0.048)	(0.048)	(0.031)	(0.032)
Number of friends	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.00003)	(0.0001)	(0.0001)	(0.00003)	(0.00003)
Weekend	0.061	0.083	0.081	0.063	0.057
	(0.033)	(0.050)	(0.050)	(0.033)	(0.033)
Days after the ban	−0.001	−0.003	−0.003	−0.001	−0.001
	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
Ban*Mainland	−0.642***			−0.654***	−0.620***
	(0.066)			(0.097)	(0.100)
Friends in Russia*Mainland				0.103	
				(0.077)	
Ban*Friends in Russia*Mainland				0.017	
				(0.131)	
Ban*Friends in Russia		0.131		0.086	
		(0.098)		(0.087)	
Pro-Russian*Mainland					−0.091
					(0.076)
Ban*Pro-Russian*Mainland					−0.046
					(0.132)
Ban*Pro-Russian			0.099		0.138
			(0.099)		(0.088)
Constant	−2.056***	−1.845***	−1.849***	−2.012***	−2.048***
	(0.053)	(0.069)	(0.067)	(0.059)	(0.061)
Observations	132,919	63,257	63,257	132,919	132,919

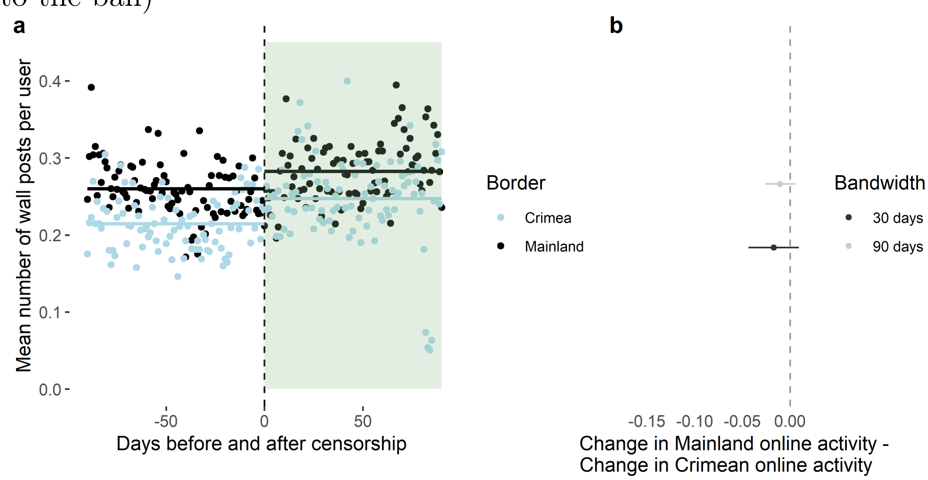
Note: **p<0.05; ***p<0.01

The models are based on data from users with political affiliations (and not random users). Models 2-3 are limited to Mainland. The remaining models include users from both sides of the border.

Appendix S: Placebo test

What if the decrease in activity on VKontakte is caused by annual holidays or other reoccurring events that only take place during May in Mainland and not Crimea? In this case, the drastic fall in online activity in Mainland would be the result of the time period of the ban, or a "placebo", rather than the active treatment (i.e. the ban) itself. In order to take this into account, I conduct a placebo test by rerunning the same DD models (described in the analysis) using wall posts from the same users from a time period of exactly 1 year prior to the ban. Figure S1 shows the observed means as well as DD estimates for the test, where 0 reflects May 18, 2016, instead of 2017. I observe no decline in the daily mean number of posts during this period. On the contrary, the activity moderately *increased* on both sides of the border, following the "placebo treatment". The activity increased slightly more in Crimea. The DD estimates, illustrated in Figure S1 **b** suggest that the difference is neither substantively large nor statistically significant. This validates the assumption that the changes in online activity on both sides of the border follow a parallel trend prior to the ban.

Figure S1: Placebo test: Change in online behavior before and after 18th of May 2016 (1 year prior to the ban)



The plots are based on all 3,024 users in the sub-sample. Plot **a** shows mean number of posts per individual user, where horizontal lines reflect overall means for the respective sides of the border. Plot **b** shows coefficient estimates for $BanxMainland$ interaction term in the DD placebo models.

Appendix T: Data overview

The table below gives an overview of the data collected about VKontakte users through the website's public API.

Data collected using VKontakte's API
Last login date
Gender (self-reported)
Town (self-reported)
Country (self-reported)
VKontakte friends
VKontakte communities
VKontakte posts on own wall*

* All of the data has been collected in August 2018, apart from the wall posts, which were collected both at the end of August 2018 and in September 2018.

References

Anderson, M. L. (2014). Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion. *American Economic Review 104*(9), 2763–96.

Burgess, M. (2018). This is why Russia's attempts to block Telegram have failed. *Wired*. Accessed October 26, 2020. https://www.wired.co.uk/article/telegram-in-russia-blocked-web-app-ban-facebook-twitter-google.

Dek, Anton, K. K. T. M. (2019). The Effects of Banning the Social Network VK in Ukraine. *NATO StratCom*. Accessed October 26, 2020. https://www.ceeol.com/search/viewpdf?id=791422.

Hausman, C. and D. S. Rapson (2018). Regression discontinuity in time: Considerations for empirical applications. *Annual Review of Resource Economics 10*, 533–552.

Kantar (2017). Reiting populyarnih saytiv za veresen 2017. Accessed October 26, 2020. https://tns-ua.com/news/reyting-populyarnih-saytiv-za-veresen-2017.

Maria Shortgin and Roman Lebed. Upalo Vsyo, kak ukraintsi prozhili god bez Vkontakte. *BBC News Ukraina*. Accessed October 26, 2020. https://www.bbc.com/ukrainian/features-russian-44250660.

Sobolev, S. (2019). Krizisnij messendzher - Kak blokirovka Telegram povliyala na regulirovanie Runeta. *RBC*. Accessed October 26, 2020. https://bit.ly/3drXbTw.

State Statistics (2013). Serednya zarobitn plata za regionami za misyats u 2013 rotsi. *State Statistics Service of Ukraine*. Accessed October 26, 2020. http://www.ukrstat.gov.ua/operativ/operativ2013/gdn/reg_zp_m/reg_zpm13_u.html.

TGStat (2019). Isledovaniye auditorii Telegram 2019. *TGStat*. Accessed October 26, 2020. https://tgstat.ru/research.

Ukrainian Census (2001). All-Ukrainian population census 2001. *State Statistics Committee of Ukraine*. Accessed October 26, 2020. http://2001.ukrcensus.gov.ua/eng/results/general/.

UNIAN (2017). "kievstar" uzhe blokiruet vkontakte i odnoklassniki (foto). *UNIAN*. Accessed October 26, 2020. https://bit.ly/33GPGDA.