Social Media and Protest Participation: Evidence from Russia[1]

Ruben Enikolopov,[a,b,c,d] Alexey Makarin,[e] and Maria Petrova,[a,b,c,d]
[a] ICREA-Barcelona Institute of Political Economy and Governance
[b] Universitat Pompeu Fabra
[c] Barcelona Graduate School of Economics
[d] New Economic School, Moscow
[e] Northwestern University

PRELIMINARY. PLEASE DO NOT CITE WITHOUT PERMISSION.

July 2015

In 2011-2012, Russia experienced the first large-scale political protests since the fall of the Soviet Union. We provide evidence that social media played an important role in these events. Using data on the dominant Russian online social network, VK, we show that higher penetration of the social network across cities increased the likelihood of protest occurrence and the number of participants in these protests. To identify the effect, we use data on the number of students from different cities who happen to study together with the VK founder, as compared with the number of students from the same cities who studied at the same university several years earlier or later, thus using intention-to-treat research design. Corresponding IV estimates imply that a 10% increase in VK penetration increases the probability of a protest by 3.7%, and the size of a protest by 15%. Placebo tests confirm that VK penetration was not correlated with protest activity or voting behavior before the emergence of social media. Analysis of the effect of fractionalization of network users between VK and Facebook and its evolution over time suggests that reducing the costs of coordination is a mechanism behind social media influence.

[1] We are grateful to Sergey Chernov, Nikolai Klemashev, Natalya Naumenko, and Alexey Romanov for invaluable help with data collection. We thank the Center for the Study of New Media and Society for financial and organizational support. We also thank Lori Beaman, Sam Greene, Kosuke Imai, Kirabo Jackson, John Londregan, Eliana La Ferrara, Ricardo Perez-Truglia, Gautam Rao, Tom Romer, Jake Shapiro, Gaurav Sood, Adam Szeidl, Josh Tucker, Glen Weyl, Katia Zhuravskaya, and seminar participants at Higher School of Economics, Central European University, Bocconi, Mannheim, Microsoft Research, Princeton University, Univsersitat Pompeu Fabra, New York University, ISNIE Annual Meeting, 11th Workshop in Media Economics in Tel Aviv, 6th Workshop in Applied Economics in Petralia, "Social Media and Political Participation" conference in Florence, and "Social Media and Social Movements" conference in St Petersburg for helpful discussions. The research support of the Ministry of Education and Science of the Russian Federation, grant No. 14.U04.31.0002 (for Ruben Enikolopov and Maria Petrova) is gratefully acknowledged.

1. Introduction

In both economics and politics collective action is often important to achieve socially beneficial outcomes. Ability to overcome collective action problems depends on people's information environment and opportunities to communicate. The emergence of online social media, such as Facebook or Twitter, changed these opportunities by allowing users to connect and communicate directly, thus potentially reducing the costs of coordination. However, so far there has been no systematic evidence on whether social media actually improves people's ability to overcome collective action problems. This paper fills the gap by looking at the effect of the penetration of the most popular online social network in Russia on political protests in 2011-2012.

Estimating the causal impact of social media is methodologically challenging, as social media usage is endogenous to individual and community characteristics. Studying the effect of social media on political protests, as an example of collective action, poses additional problems, as protests are typically concentrated in one or several primary locations, as was the case for Tahrir Square in Egypt or Maidan in Ukraine. Overtime variation in protest intensity can provide evidence on the association between social media usage and protests (Acemoglu et al. 2015),[2] but does not allow studying the causal effect of social media penetration.

We focus on the case of Russia in 2011-2012 because it is particularly suitable for empirical analysis of the effects of online social media on protest participation. First, there was a substantial geographic and time variation in the intensity of protest activities in Russian cities and in the penetration of major online social networks. Second, particularities of the early stage development of VK, the most popular social network in Russia, allow us to exploit quasi-random variation in the penetration of this network across cities to identify the causal effect of network penetration on political protests.

The unexpected wave of protests was triggered by electoral fraud in December 2011 parliamentary elections.[3] Active protests lasted until May 2012. This was the first large-scale political protest movement in Russia since the collapse of the Soviet Union. Among 626 cities in our sample, 133 witnessed at least one protest demonstration during that period. In the analysis

[2] See also Hassanpour (2014) and Tufekci and Wilson (2013).
[3] Electoral fraud was e.g. documented in Enikolopov, Korovkin, Petrova, Sonin, and Zakharov (2013) and Klimek, Yegorov, Hanel, and Thurner (2012).

we mostly focus on protest activity during the first weekend after the elections, though in some parts of the analysis we also use overtime variation.

Our identification is based on the particularities of the early stages of development of VK ("*VKontakte*"), the dominant social network in Russia.[4] It was launched in October 2006 by Pavel Durov, who graduated the same year from the Saint Petersburg State University (SPbSU). Through November 2006, to open a new account users had to send a request in response to the invitation made by Durov in the online forum of the university students (also owned by Durov), and all the requests had to be approved by Durov personally. Thus, the first users of the network were mostly students who studied at SPbSU at the same time as Durov. The distribution of early users in terms of their cities of origin turns out to be important for the future penetration of the network. Although many of these students never come back to their city of origin, their networks of friends and relatives helped faster development of VK communities in these cities and led to persistent difference in VK penetration. At the same time, the origin of the students from the same university several years older or younger did not affect the early development of VK network, and, eventually, subsequent VK penetration. As a result, fluctuations in the distribution of SPbSU students in terms of their city of origins had long-lasting effect on VK penetration.

In the empirical analysis we use the origin of the students who come to study to SPbSU in the same 5-year cohort as the VK founder as an instrument for VK penetration in 2011, conditional on the origin of the students who come to study to SPbSU in older and younger cohorts. We show that the number of students from a city who studied together with the VK founder, conditional on similar numbers for older and younger cohorts, was strongly related to VK penetration in summer 2011, shortly before the start of the protests, and to early VK penetration in November 2006, when the access to network was still by invitation only. Notably, we find that the number of students from a given city was important for VK penetration (both early and later) *only* for the VK founder cohort, but not for the older or the younger cohorts. Thus, we use intention-to-treat methodology, and the variation that we exploit is based on year-to-year fluctuations in student flows between various Russian cities and SPbSU.[5] Our identification is, therefore, based on the assumption that these fluctuations were quasi-random and not related to unobserved city characteristics.

[4] Russia is one of a few countries in the world in which Facebook is not dominant online social network.
[5] In the analysis we exclude Moscow and Saint Petersburg from the sample as important outliers.

In the reduced form analysis, we find that the number of students from a city in the VK founder cohort has a positive and significant effect on protest participation, while there was no such effect for students from older or younger cohorts. The corresponding IV estimates indicate that the magnitude of the effect is sizable – a 10% increase in the number of VK users in a city led to a 15% increase in the number of protest participants and 3.7 percentage points higher probability of having a protest.

We perform a number of additional tests to ensure that our results are not driven by unobserved heterogeneity. First, we show that VK penetration in 2011 was not significantly related to either protest participation in the Soviet Union or social protests in 2005, which suggests that our results are not driven by time-invariant unobserved preferences for protests. Second, we show that Facebook penetration at the time of the protests was closely related to contemporaneous VK penetration, but not the early VK penetration in 2006. This pattern suggests that VK penetration in in 2011 was mostly driven by general demand for social media in cities, which was also driving Facebook penetration. The distribution of early VK users, however, was not primarily determined by general demand for social media across cities. Third, we show that VK penetration in 2011 is not related to voting outcomes before the creation of VK. We also note that none of the early VK users have ever participated in any of the VK protest groups created in 2011, so that political preferences of these users are unlikely to have a direct effect on protest participation. Finally, the tests in the spirit of Altonji, Elder, and Taber (2005) indicate that unobservables that are positively correlated with observables are unlikely to drive our results. Overall, these results suggest that our results are unlikely driven by unobserved heterogeneity.

We explore potential mechanisms for the effect of social media. Social media can have an impact on protests through information channel, coordination channel, and social pressure channel. The information channel reflects the fact that online social media can serve as an important source of information, both on the fundamental issues that cause protests (e.g. electoral fraud) and on the protests themselves. We expect this effect to be especially strong in countries with government-controlled traditional media, like Russia. At the same time, both the coordination channel and social pressure channel rely on the fact that users not only consume, but also produce information. In particular, social media not only allows users to coordinate the

logistics of the protests (coordination channel), but also introduces social motivation if users' friends openly announce that they are joining the protest (social pressure channel).

To provide evidence on the mechanisms behind the effect, we first show that that fractionalization[6] of users between VK and Facebook, conditional on the total number of users in the two networks, had a negative impact on protest participation. This finding is consistent with both coordination and social pressure channels. This effect is stronger in larger cities with population over 100,000. In addition, we find that the effect of VK penetration on both protest participation and protest incidence was getting smaller over time, which is consistent with both information and coordination channels. Taken together, these results are consistent with reduction of the costs of coordination being an important mechanism of social media influence.

Overall, our results indicate that social media penetration facilitates participation in political protests. They also suggest that lowering the coordination costs is the mechanism through which social media promotes this form of collective action. This effect have been predicted by the theoretical literature (e.g. Edmond 2013, Little 2015) and widely discussed in the popular press, but so far there has been no systematic empirical evidence to support this prediction. Thus, availability of social media can have important consequences, as political protests can play an important role in within-regime power sharing agreements and related economic and political outcomes (Madestam, Shoag, Veuger, and Yanagizawa-Drott, 2013; Pasarelli and Tabellini 2013; Acemoglu, Hassan, and Tahouc 2015; Aidt and Frank 2015). The broader implication of our results is that social media has a potential to reduce the costs of collective action in various circumstances. In addition, our paper also proposes the methodology to use to study the effects of technology adoption that has a potential to be used in a wide range of circumstances.

Our paper is related to the literature on the impact of ICTs and traditional media on political preferences and policy outcomes. A number of recent works identifies the impact of broadband penetration on various outcomes (e.g. Czernich et al (2011), Falck, Gold, and Heblich (2013), Campante, Durante, and Sobbrio (2014), Bhuller, Havnes, Leuven, and Mogstad (2013), Gavazza et al. (2015)); however, they do not provide evidence if this effect is due to the accessibility of online newspapers, search engines, email and Skype communications, or social media. Recent works have shown that traditional media has an impact on voting behaviour

(DellaVigna and Kaplan 2007, Enikolopov et al. 2011, Gentzkow et al. 2011, Knight and Chiang 2011), violence (Yanagizawa-Drott 2014, Adena et al. forthcoming), or policy outcomes (Strömberg 2004, Eisensee and Strömberg 2007, Snyder and Strömberg 2010).

Our paper is closely related to Acemoglu et al. (2015) who study the impact of Tahrir protest participation and Twitter posts on the expected future rents of politically connected firms in Egypt. They find that the protests were associated with lower future abnormal returns of politically connected firms, i.e. protests predicted smaller future rents for firms connected to the incumbent regime. They also show that the activity in Twitter precedes the actual protest activity. Our analysis is different from theirs in several respects. First, we focus on studying the causal impact of social media penetration across cities, rather than the activity in the already existing social media. Thus, we consider the counterfactual of not having social media, rather than having no protest-related content in social media. Second, we look not only at the number of protesters, but also at the probability of protests, i.e. at the extensive margin of the effect. Finally, we provide suggestive evidence on potential mechanisms behind the impact of social media.

The rest of the paper is organized as follows. Section 2 provides background information about the environment that we study. Section 3 contains theoretical background and outlines our main empirical hypotheses. Section 4 describes our data and its sources. Section 5 discusses our identification strategy. Section 6 shows the empirical results. Section 7 concludes.

2. Background

Internet and social media in Russia.

In 2011, approximately half of the population in Russia had access to the Internet at home,[7] which made Russia the largest Internet market in Europe, surpassing Germany in terms of the number of users (50.8m vs. 50.1m)[8] and accounting for about 15% of the European Internet users. Although more than 80 countries in the world had higher Internet penetration at the time, in Russia it was growing at an incredible 23% yearly rate (averaged over the 2007-2011 period).

Social media was already very popular in Russia at that time. According to ComScore, in

6 We define fractionalization as the probability that two randomly picked individuals will be users of different social media sites. We correct our measure for potential overlap between social media, allowing individuals to be users of both Facebook and VK, and it does not change our results.

[7] http://www.internetlivestats.com/internet-users/russia/

[8] http://www.comscore.com/Insights/Press-Releases/2011/11/comScore-Releases-Overview-of-European-Internet-Usage-in-September-2011

2010 Russians spent on average 9.8 hours per month on social media sites - more than any other nation in the world.[9] Meanwhile, social media penetration in Russia was comparable with the most developed European countries at the time with 88% of Russian Internet users having at least one social media account, compared, for instance, to 93% in Italy and 91% in Germany. Although Russians lost the title of the most social media addicted nation to Israel in October 2011, they have remained the third with 10.4 hours per visitor.[10]

Despite the fact that social media had been becoming increasingly popular, Russia remained one of the very few markets in which Facebook was not the most popular social network. In fact, there was only one other country, where Facebook couldn't even get the second place for reasons other than censorship - South Korea. Russian social media market was (and still is) dominated by the homegrown networks *VKontakte* (VK) and *Odnoklassniki* (OK). As of August 2011, VK was the biggest player with a daily audience of 23.4m unique visitors (54.2% of online population), OK was catching up with 16.5m unique visitors (38.1%), leaving Facebook trailing with 10.7m unique visitors (24.7%).[11]

The reason for such an unusual pattern to emerge was a relatively late entry by Facebook into the Russia market. By the time it introduced Russian interface in mid-2008, both VK and OK already had more close to 20m registered users.[12] Besides, VK and OK could offer services that Facebook couldn't offer either for legal reasons (e.g. an opportunity to listen to music and watch movies for free), or because of a different marketing strategy (e.g. users were attracted by lower amount of advertising in both VK and OK).

While VK started off as a student- and youth-oriented site, OK specifically targeted older people. "Odnoklassiki" translates as "classmates" from Russian, and the main targeting message of OK was to help people find their former classmates and friends from their past. On the contrary, "VKontakte" translates as "in contact", and the targeting message of VK was to help students stay in touch later in life, and their target audience was similar to the target audience of Facebook. According to a marketing study performed in 2010, the average age of VK users was

---

[9] http://www.comscore.com/Insights/Presentations-and-Whitepapers/2011/it_is_a_social_world_top_10_need-to-knows_about_social_networking

[10] http://thenextweb.com/socialmedia/2011/12/22/israelis-are-now-the-worlds-biggest-social-network-addicts-says-new-report/

[11] http://www.dreamgrow.com/social-media-in-russia/

[12] https://vk.com/blog?id=92 and http://www.dni.ru/society/2008/11/1/152405.html

3 years less than the average age of OK users.[13] Moreover, VK was more widespread in large cities than OK. In turn, Facebook became especially popular among those who would want to communicate with their foreign friends and was most popular in large cities, especially in Moscow and St. Petersburg.[14]

As of December 2011 the Internet in general and social media in particular enjoyed relative freedom in Russia, as serious attempts to control online content began occurring only in 2012, i.e., after the period we focus on and, to a large extent, as a consequence of the protests examined in the paper. This freedom has made social media outlets an important avenue for transmitting information and enhancing political debate, which was gradually phased out of Russian TV and major newspapers after Vladimir Putin came to power in 2000.[15]

Protest movement of 2011-2012.

The wave of protest demonstrations in 2011-2012 was triggered by electoral fraud in December 2011 parliamentary elections, and it was the first large-scale political protest movement in Russia since the collapse of the Soviet Union. Similar to many other protest events in authoritarian countries (Kuran, 1989), Russian protests of December 2011 surprised everyone, including their leaders.[16] Russian society was politically dormant in the 2000s, with rapid economic growth softening any criticism of Putin's regime, so that electoral fraud of allegedly similar magnitude in 2007 did not trigger any serious protests. Finally, traditional media was under almost full government control at the time, so there was no chance any information about electoral fraud would leak through the main TV channels, which served as the primary source of information for nearly 80% of Russians at the time.[17]

Parliamentary elections were held on December 4, 2011. During the course of that day, reports on election fraud were quickly growing in numbers, documented both by independent observers and by regular voters. In a vast majority of cases, fraud was done in favor of the incumbent party *United Russia*. Videos with apparent cases of ballot staffing and `*carousel*'

[13] http://www.teleskop-journal.spb.ru/files/dir_1/article_content1327333874323807file.pdf
[14] http://corp.mail.ru/media/files/issledovanie-auditorij-sotcialnykh-setej.pdf
[15] Since 2009, Freedom House has ranked mass media as "not free," and Reporters Without Borders has classified Russia as a country with a "difficult situation" in terms of freedom of the press.
[16] For instance, a day before the first protest gathered over 5,000 participants its organizers were debating whether a threshold of 500 people will be surpassed.
[17] For instance, see Levada survey in 2011 (http://www.levada.ru/sites/default/files/2012_eng.pdf, p. 135) or VTSIOM study in 2011 (http://www.wciom.com/index.php?id=61&uid=31)

voting (when the same voters vote on multiple poll stations) began to circulate in the web and social media. Startling differences between exit polls and actual results emerged, with. some exit polls predicting a 23.6% vote for *United Russia* in Moscow, 20% lower than the officially reported result.[18] Some scholars confirm that the fraud was sizable using up-to-date research methodology. For instance, Enikolopov et al. (2013) show that the presence of randomly assigned independent observers decreased *United Russia's* results by at least 11% (from 47% to 36%). Apparent evidence of electoral fraud together with the absence of its acknowledgement by the government outraged some people and urged them to take to the streets.

On December 5, five to six thousand people appeared at the rally in the center of Moscow that intended to gather at most a thousand. The rally was followed by minor clashes with the police and detention of several opposition leaders. Although the number of protesters was not particularly large, this rally set up an example for the future, more massive ones. The rallies on December 10 and 24 had record attendance both in Moscow (near 100,000 participants on both dates) and across the country (more than 100 cities in total). Later waves of protests were less popular and involved fewer cities. Moscow and St. Petersburg, however, still hosted at least two major rallies per month. The last peak of the movement happened on May 6, 2012, a few days before Vladimir Putin's inauguration. Whereas all previous demonstrations were peaceful and non-violent, this march broke out in serious clashes with police forces. Within a few days more than 30 activists were picked and prosecuted for allegedly igniting mass riots, with many getting 3-4 years of prison later. This trial, together with the absence of tangible achievements, decreased the popularity of protests as a form of political activity.

History of VKontakte.

VKontakte (or VK) is a social network with very similar functionality to Facebook – one can build up his or her profile, add friends, chat with them, create events to make coordination easier, write posts, share information (in audio and video form as well), etc. It was launched in October 2006. The core of VK team was more or less stable until 2012; it consisted of Pavel Durov (philology major at SPbSU at the time), his brother Nikolay Durov (physics PhD student at SPbSU at the time, winner of the world programming and math competitions), and his fellow students. VK gained its popularity quickly. At first the network was advertised by Pavel Durov at

[18] Note that exit polls were later deleted from those polling agency's websites.

the SPbSU students' forum (which he created earlier). The first several thousands of users were created through a referral system, i.e. to get registered in VK one needed to have a friend or acquaintance who was already registered. One person could issue only a limited number of invitations and all the new accounts were personally approved by Pavel Durov. After making the registration open in the end of November 2006, the number of users skyrocketed–from 50 thousand users in January 2007 to 3 million in November 2007, to 100 million in November 2010. Figure A1 in Appendix shows the growth in the number of VK users overtime. VK.com became the most visited web site in Russia by early 2008.

VK leadership had a strong position against any censorship. During the 2011-2012 protests Pavel Durov was approached by Federal Security Agency (FSB) and asked to block opposition-minded online communities, as well as pages of protest events, some of which had 30,000 subscribed (Kononov, 2012). Durov refused to do it, arguing that it would lead to a large number of people switching to its foreign competitors, such as Facebook; others pointed to his libertarian views as well. VK policies regarding freedom of speech remained unchanged until he was forced to sell his share of VK and lost control of the firm in 2014.[19]

3. Hypotheses

Protest participation is a form of political participation with potential social benefits and individual costs, similar to voting in elections. The key difference is that these costs are larger than for the case of voting, and these costs also depend on the total number of participants, so that coordination is more important for the organization of a successful protest.

There are several channels through which social media can affect protests. Social media can provide its users with uncensored information about the government, facilitate coordination, or strengthen social pressure. To fix the ideas, consider a simple model of an individual decision to participate in a protest. The number of protesters is $P=p*N$, where p is the propensity to protests, and N is the number of people in a city.

Each individual first decides whether or not s/he thinks the regime is good or bad. Then for those individuals who do not like the regime and think it should be changed, the decision to

[19] Durov's share was bought out by Mail.ru after the same issue, but two years forward in September 2014, when he refused to block groups and accounts of Ukranian revolutionaries. He left VK for his new start-up Telegram, an application resembling WhatsApp in functions. He left the country too, after obtaining citizenship of Saint Kitts and Nevis.

participate in protests is determined by the combination of costs and benefits. In other words, these individuals choose to participate if and only if their benefits are higher than their costs:

$$B_i * \Delta Prob_i(success|P) + costs(P) - benefits(P) - \varepsilon_i > 0 \qquad (1)$$

where B_i is the benefit for individual *i* if the protest is successful, $\Delta Prob_i(success|P)$ is the marginal increase in the probability that the protest is successful if person *i* decides to participate, *benefits(P)* is the sum of an individual intrinsic value from expressing the opinion (that does not depend on *P*) and a social pressure component (that can potentially depend on the number of protesters *P*), and costs*(P)* is the expected cost of protesting (also depends on *P*, because of coordination, or, for example, because of the "safety in numbers" phenomenon[20]). In a large country $\Delta Prob_i(success|P) = 0$ for an individual participant.

Assuming a simple functional form, similar to the calculus of voting, the utility of a potential protester can be written as

$$benefits - costs = \gamma_i + \pi(s,P)\alpha_i - c_1(s,P) - \varepsilon_i \qquad (2)$$

Here γ_i is individual intrinsic value from participation; α_i is the utility of being seen in a protest by peers (can potentially be positive or negative); $\pi(s,P)$ is the probability of being seen (can be amplified by social media), where *s* is the social media penetration; $c_1(s,P)$ is the cost of learning the logistics of a protest, which is a diminishing function of the social media penetration, together with the expected cost of participation, which is a decreasing function of the size of a protest P, with some potential complementarities.

Consider partial equilibrium, i.e. the situation in which an individual decides whether to participate in a protest taking *P* as given and not internalizing the effect of his/her decision on the equilibrium outcome. Then the expected propensity to participate in a protest is

$$I(\theta < 0|s) * F\big(\gamma_i + \pi(s,P)\alpha_i - c_1(s,P)\big) \qquad (3)$$

where θ is the individual's opinion about the regime; $I(\theta < 0|s)$ is an indicator function for whether individual *i* dislikes the regime, which can be a function of social media penetration; *F* is a distribution function of individual idiosyncratic cost ε_i. Denote the argument of the function *F* in (3) as *X*. If we differentiate the protest propensity with respect to *s* in partial equilibrium, we get

[20] See e.g. Lohmann (1994).

$$\frac{\partial p}{\partial s} = \frac{\partial (P/N)}{\partial s} = \underbrace{\frac{\partial I(\theta<0|s)}{\partial s} * F(X)}_{information} + \left[f(X) * \left(\underbrace{\frac{\partial \alpha_i \pi(s,P)}{\partial s}}_{social\ pressure} - \underbrace{\frac{\partial c_1(s,P)}{\partial s}}_{coordination} \right) \right] \quad (4)$$

This suggests that the impact of social media penetration potentially consists of information effect, social pressure effect, and coordination effect, which theoretically can have different signs.

In the environment that we study we expect all three mechanisms to increase propensity to protest. In a situation in which traditional media is suppressed, information effect is expected to be positive (e.g. Reuter and Szakonyi (2015) find that social media users are more likely to know about electoral fraud). Enikolopov et al. (2015) suggest that social motivation was an important driver of protest participation in 2011-2012.[21] Finally, suggestive evidence and the theoretical literature (Little, 2015) imply that the coordination of protests might be easier with availability of social media. Thus, our basic hypothesis is

Hypothesis 1. Social media penetration increases the propensity to participate in political protests.

Similar logic holds for the extensive margin, i.e. for the probability of having a protest. Consider that there is a fixed cost of staging the protest, at least for some organizers, and the protest happens only if $pN > FixedCosts + \eta_i$, where η_i is a noise component. Then we expect the incidence of protests to be an increasing function of *s*, as long as propensity to protest, *p*, is an increasing function of *s*, and the distribution of η_i has continuous support (density of η_i is positive). Our second hypothesis, therefore, is:

Hypothesis 2. Social media penetration increases the probability of observing a political protest.

4. Data

We use several sources of data. Our sample consist of all the 626 Russian cities with population over 20,000 according to the 2010 Census, with Moscow and Saint Petersburg excluded. We take population figures from Rosstat, Russian statistical agency. We also check that our results are robust to the changes in population thresholds.

[21] Note, however, that this result seems to be context specific and might not hold for protest participants in other environments and time periods.

First, for all the users of VK with public accounts[22] we collect information about their city, whether the user joined VK before the autumn of 2011, and whether the user joined VK before December 2006. Based on this information we compute two city-level statistics: (i) the number of users in each city as of the summer of 2011, before the parliamentary elections were scheduled and electoral campaign began; (ii) the number of users for each city who joined VK before December 2006, i.e. who were among first 5,000 users of the network who opened their accounts while VK was still by-invitation-only. We also use the number of VK users in 2013 in some specifications. More details about data collection are available in the Online Appendix.

Second, we use hand-collected data on protests between December 2011 and May 2012 When the protests started in December 2011, we hired a graduate-level research assistant who monitored Internet, social media, and newspaper databases to find information about protests in as many Russian cities as possible. The research assistant repeated this process every week before the protests started to subside in summer 2012. For each event we recorded the number of protesters as reported by the police, by the organizers, and the news source that reported on the protest.[23] We believe that, as a result, we have the most comprehensive city-level database on Russian protests in 2011-2012. The list of our data sources is available in the corresponding section of the Online Appendix. For each week of the protests we construct two variables: a dummy for the existence of a protest in a given city and the number of protesters (computed by taking the average number of protesters as reported by the police, organizers, and the news source).[24]

Third, we use data on the number of users in Facebook by city in 2011 and 2013. Data on Facebook availability in 2011 is taken from Zapuskalov (2012). Data on Facebook penetration in 2013 was collected manually for each of the cities in the sample.[25]

Fourth, we collected the data on VK protest communities where the information about protests in a city was posted. This information was collected through manual searches in VK, and Yandex (the most popular Russian search engine).

[22] Public accounts have the basic information about the users, including their home city, available to any Internet user. The timing of the creation of an account can be inferred from the account ID number. Note that more that 90% of accounts in VK are public

[23] We have all the three estimates in only 9.5% of the cases and in 64% of the cases only one estimate is available.

[24] Our estimates remain practically unchanged if we use median value of available estimates instead of the mean.

[25] Missing numbers for 2011 were imputed using data on Facebook availability in 2013, VK availability in 2011, and VK availability in 2013 (see the Online Appendix for more details).

Fifth, we use data on protests in the late Soviet Union from Beissinger (2002). The data was compiled by the author based on 150 different news sources in 1987-1992 and contain information on 6,663 protest demonstrations. We aggregate this information at the city level for all the cities in our sample. In the analysis we exploit two different measures: the maximum number of protesters in a city and a dummy for at least one protest in a city.

Sixth, we use information on social protests in 2005 from the website of a youth communist organization (http://trudoros.narod.ru/).

Finally, we use a number of control variables collected from different sources. The data on electoral outcomes is coming from the Central Election Commission of the Russian Federation. The data socio-demographic characteristics of the cities (average age, education, population, average wage, and location of administrative centers) comes from Rosstat. Additional city characteristics (latitude, longitude, years of city establishment) comes from the Big Russian Encyclopedia. Summary statistics are presented in Table 1.

5. Identification Strategy

Basic approach

Our first hypothesis is that social media penetration (specifically, VK, penetration) had a positive effect on the number of protesters. Following (4) and taking logs, we estimate the following model:

$$\log(1 + protesters_i) = \beta_0 + \beta_1 VKpenetration_i + \beta_2 \mathbf{X_i} + \varepsilon_{it} \quad (5)$$

where $protesters_i$ is the number of protesters in city i in the first week of the protests, $VKpenetration_i$ is the measure of social media penetration, $\mathbf{X_i}$ is the vector of controls that includes fifth-order polynomial of population, dummy for being a regional or rayon administrative center, average wage in the region, the number of city residents in different age cohorts, distance to Moscow and Saint Petersburg, dummy for the existence of a university in a city, percentage of population with higher education, and internet penetration. In some specifications it also includes the number of protest participants in 1987-1992 and voting outcomes in past elections. Standard errors in the regressions are clustered at the regional level.

However, the OLS estimate of equation (5) is likely to be biased, since unobserved characteristics that make people from certain cities more likely to become VK users can at the same time make them more likely to participate in protests. To deal with this problem, we need a

source of exogenous variation that is related to social media penetration, but that does not have any other independent effect on protest participation.

To deal with this problem we exploit information about the origin of the students who studied at SPbSU at the same time as the VK founder, as well as the origin of the students who studied in the same university earlier or later. Specifically, we use five-year student cohorts, to match Census definition of cohorts, i.e. we look at the students who have studied at the same year as Durov, as well as the students who studies one or two years earlier and later. The identifying assumption is that conditional on population, education, and other controls, fluctuations in student flows in SPbSU in 2000-2010 are orthogonal to other determinants of protest participation.

Table 2 presents the full distribution of the data for student cohorts that we use. Note that in all but one cases the number of students is less than 35 students per cohort. Thus, the numbers are sufficiently small to allow for random fluctuations in the distribution of students across cities, that justifies our identification strategy.

Determinants of VK penetration

For our identification strategy to work, we need to show that our instrument is relevant. Panel A of Table 3 provides evidence on the determinants of VK penetration across Russian cities in 2011, and, in particular, on the effect of the number of SPbSU students in different cohorts. The results indicate that once population controls are included, only the VK founder five-year cohort is positively and significantly (at 1% level) correlated with the later VK penetration, in contrast to younger and older cohorts. The coefficient for the number of SPbSU students from VK founder cohort is stable across the specifications (2)-(7). In particular, it does not depend on the age distribution in a city, as we control for the number of people in each of the five-year age cohorts over 20 years of age. The magnitude of the effect implies that a 10% increase in the size of VK founder cohort leads to a 15% increase in the number of VK users in a given city in 2011. The coefficients for other cohorts are much smaller in magnitudes and not significant, and the coefficient for the size of younger cohort is consistently negative across specifications (2)-(7). These results are also summarized in graphical form in Figure 1A.

There are several other takeaway points from the results presented in Panel A of Table 3. First, as expected, VK penetration is significantly related to urbanization controls, such as flexible polynomial of population and a dummy for a regional center, and age distribution.

Second, VK penetration is positively and significantly related to average wage, and richer cities are more likely to have larger VK communities. Finally, VK penetration was not determined by pre-existing voting preferences, because voting controls from either 1995 or 1999 or 2003 elections were not statistically significant in the corresponding specifications (5)-(7).[26]

Early VK penetration

To better demonstrate the mechanism that makes our instrument to work, we also use the fact that initially VK was by invitation only, and the people who got invitation were friends or classmates of the VK founder, Pavel Durov. We look at the origin of the first 5,000 users of VK who have joint VK before December 2011, when the subscription to the network became open, and relate them to the size of the SPbSU student cohorts and all the other controls from Panel A of Table 3. Note that the early users of VK were unlikely to have a direct effect on political protests as we know from the protest group membership data that none of them registered for the protest events or participated in any of the protest groups. In fact, they were rather apolitical and rarely mention politics or list political interests in their profiles.

The results for the determinants of early penetration are shown in Pane B of Table 3. The results are similar to those for the VK penetration in 2011 reported in Panel A. The size of VK founder SPbSU student cohort is positively and significantly related to early VK penetration, whereas the effect of the older or younger students cohorts is not statistically significant. The magnitude of the effect is smaller as compared with Pane A of Table 3, so that a 10% increase in the size of student cohort was associated with a 5% larger VK penetration in November 2006. Cohort coefficients, together with their confidence intervals, are also shown graphically in Figure 1B.

Overall, results in Tables 3 are consistent with the story that the differences in VK penetration are explained by year-to-year fluctuations in student flows from different cities, and small initial differences have large long-term consequences on the penetration of the social network. In subsequent sections we use additional tests to ensure that our results are not driven by other type of unobserved heterogeneity.

[26] In addition, in Table A1 in the Online Appendix we show that student cohort coefficients do not alter the relationship between these other determinants and the later VK penetration, and the coefficients for other controls remain practically unchanged.

6. Empirical results

Reduced form estimation

We start by presenting the results of the reduced form estimation. Table 4 shows how the size of the protests (columns (1)-(4)) and protest occurrence (columns (5)-(8)) are related to the number of students in SPbSU student cohorts. We find that the size of the VK founder cohort has a positive and significant effect on both the size of protests and the incidence of protests, with the coefficients for other cohorts being much smaller and not statistically significant, with the sign of coefficient for older cohort being consistently negative across specifications. Figures 2A and 2B show these results graphically. Overall, the results presented in Figure 2 and Table 4 suggest that VK founder student cohort is positively and significantly associated with protest participation, in contrast to older and younger cohorts, and VK penetration is the only reasonable channel through which this effect can work.

There are some other important results in Table 4. First, though protest participation is related to urbanization controls, it is not significantly related to income in these cities (the coefficient for average wage is consistently negative in all the specifications). Second, voting controls from 1995, 1999, or 2003 predict protest participation, but not VK penetration, consistent with idea that protests were results of political grievances. Third, protest participation was positively and significantly related to the presence of a university in a city, in contrast to VK penetration. Finally, Soviet protest participation is positively related to modern protest participation in some specifications. Overall, most of the observable determinants of protest participation in 2011 were different from determinants of VK penetration, thus alleviating some potential remaining concerns.

Baseline IV results

Reduced form results provide us with evidence that SPbSU student cohorts, through their impact on VK penetration, affected protest activity in 2011. However, they do not allow us to quantify the magnitudes of the effect of social media penetration on protests. As a next step, we estimate equation (5) using the number of SPbSU students in VK founder cohort as an instrument for VK penetration in summer 2011.These results are shown in columns (1)-(4) of

Panel A of Table 5. The results indicate that social media penetration had a quantitatively large and statistically significant effect on protest participation. A 10% increase in the number of VK users leads to a 15-16% increase in the number of protesters, depending on the specification. Although the effect appears large, it is important to have in mind that VK users constituted a reasonably large share of city population (the average penetration in 2011 was 15%) while protest participants were a tiny fraction of city population, with only 0.4% of city population participating in protests in cities in which the protests occurred.

Weak instrument is potentially an important concern in our estimation. The traditional Stock and Yogo (2005) thresholds, derived for the case of homoscedastic errors cannot be used in our case, as we use clustered standard errors. Instead, we use recently developed methodology of Olea Montiel and Pflueger (2013) who derived weak instrument thresholds for 10% potential bias and 5% significance level similar to Stock and Yogo (2005) for the case of heteroskedastic and clustered standard errors. Adjusted F-statistics calculated according to their methodology exceed 270 in all specifications, which is much higher than the required threshold level of 23. To be conservative, we conduct weak instrument test using their methodology after each IV specification that we employ. We also checked that our instruments are not considered weak if we use more traditional Cragg-Donald F-statistics and compare it with Stock and Yogo (2005) thresholds under assumption about homoscedasticity.

For comparison, we show OLS coefficients for the equation (5)in columns (5)-(8) of Table 5. The coefficients are still highly significant, but are much smaller in magnitude that the corresponding IV estimates. An explanation for the difference between OLS and IV is the negative selection, e.g. based on unobserved income component. If people with higher unobserved income are more likely to become VK users, but are less likely to participate in a protest (which is consistent with the results in Tables 3 and 4) OLS estimate of the impact of VK penetration on protest participation will be downward biased.

Panel B of Table 5 shows the results if we use protest participation in late Soviet Union or participation in social protests in 2005 as a placebo outcomes. The results indicate that there is no association between VK penetration in 2011 and protests in either 1987-1992 or 2005, which is consistent with the assumption of no time-invariant unobserved taste-for-protest heterogeneity driving our results. Note, however, that because of large standard errors, we cannot reject tests

for equality of coefficients for VK penetration on the 2011 outcomes and on the placebo outcomes.

Extensive margin

Our next step is to test the hypothesis that protests are more likely to occur if social media penetration is higher. The results in Table 5 mix together two related, but distinct processes: the incidence of protests and the number of people joining once the protests get organized. Table 6 looks specifically at the extensive margin of this effect. To be able to combine IV estimation with clustered standard errors and weak instrument tests, we use linear probability model. The results suggest that VK penetration increases the probability of protest being organized (significant at 5% level). A10% increase in the number of VK users led to a 3.7-3.8% higher probability of a protest being organized.

Panel B of Table 6 shows that, similar to the results in Panel B of Table 5, there is no evidence that VK penetration was related to the incidence of Soviet protests or social protests in 2005, i.e. there is no time-invariant unobserved heterogeneity in city inhabitants' preferences for protesting that is driving our results. Note that for the pro-democracy protests in 1987-1992 (80% of all protests of that time) and for social protests in 2005, we can reject the equality of coefficients for VK penetration on protests in 2011 and on placebo outcomes at 10% level.

Other online social networks

If there is unobserved heterogeneity that makes some people more likely to join online social networks, we can expect VK penetration to be closely associated with penetration of Facebook, its main competitor on the Russian market, and *Odnoklassniki*, which, however, was targeted toward a somewhat different audience. Table 7 shows how early and later VK penetration is associated with Facebook and *Odnoklassniki* penetration, conditional on all other controls. Unobserved heterogeneity would suggest positive and significant relationship between penetration of VK and other networks. The results in columns (2)-(4) show that it is indeed the case for later VK penetration. Once the network size reached 50,000 users, its city-level penetration was positively and significantly associated with Facebook penetration in 2013. However, early VK penetration is not significantly related to 2013 Facebook penetration. Moreover, the coefficient for early VK penetration is negative rather than positive. These results imply that even if there is an unobserved heterogeneity that drives penetration of both networks

in later periods, it is likely to be orthogonal to early VK penetration. As a side note, quantitatively large and highly significant positive coefficient for the relationship between later VK and Facebook penetration suggest they indeed have similar audience. Columns (5)-(8) show that similar relationship does not hold for *Odnoklassniki*, which confirms that *Odnoklassniki* is unlikely to be a close substitute for VK.

Placebo tests

To further ensure that our results are not driven by unobserved heterogeneity we also conduct another set of placebo tests. In particular, we replicate the results in Table 5 (instrumental variables) by using different pre-2006 voting outcomes as placebo dependent variables. Voting outcomes can capture pre-existing political preferences, and Table 4 suggests that they are collectively important to predict protest activity in 2011. Table 8 summarizes the results of our placebo tests. Each cell in this table represents the coefficient for 2011 VK penetration in IV regression similar to equation (5), but with different voting outcomes, outlined in the column name in the year outlined in the row name. In sum, we find that out of 30 regression coefficients, 1 was significant at 1% level, 3 were significant at 5% level, and 8 turn out to be significant at 10% level. These numbers are slightly higher than suggested by random generation of these numbers, but they largely support our argument. In addition, almost all the significant effects show negative effect of VK on the support of opposition parties (e.g. *Yabloko* and its leader Grigory Yavlinsky) and the vote choice, traditionally associated with protest vote (vote 'Against all'). To deal with potential bias, we control for pre-existing voting preferences in all our tests. Note also that though pre-existing voting preferences significantly affect protest participation (Table 4), they are not related to VK penetration (Panel A of Table 3).

Altonji-Elder-Taber tests

Table 8 shows a version of Altonji-Elder-Taber (2005) test, similar to the one used in DellaVigna et al (2014) and Adena et al (forthcoming). Specifically, we predict the size of a VK founder student cohort with observable city characteristics. Then we regress protest participation on predicted independent variable and direct controls (flexible polynomial of population and a dummy for a regional center). If the positive significant relationship between VK penetration and political protest is driven by unobservables, that are positively correlated with observables, we should expect to see a positive and significant relationship between the predicted VK penetration

and the dependent variable. However, we do not observe such a relationship: the coefficients for the predicted variables are not statistically significant and, occasionally, flip their sign. Thus, our results are unlikely to be driven by unobservables that are positively correlated with observables.

Channels: fractionalization

Equation (4) outlines several potential channels through which social media can affect protest participation, including information channel, social pressure channel, and coordination channel. Disentangling these channels is not a trivial task. To tackle this question we take advantage of the fact that Facebook was close competitor of VK and was also used in protest activities.[27] We compute fractionalization index $fract_i = 1 - \sum_j s_{ij}$ where s_{ij} is share of users in network j in city i in total number of online social network in city i. We then estimate the following specification:[28]

$$\log(protesters)_i = \beta_0 + \beta_1 fractionalization_i + \beta_1 \log(totalusers)_i + \beta_3 \mathbf{X_i} + \varepsilon_{it} \qquad (6)$$

Coordination channel would imply that the coefficient for fractionalization is negative, as it is more difficult to coordinate if the size of both networks is roughly equal. Information channel would imply zero coefficient for fractionalization, as for information it is important how many users in total can spread the information, not their distribution in two networks, as long as there is at least one connection between them. Finally, for social pressure the expected sign is ambiguous, as fractionalization may mean (and may not) that the number of online friends is higher. Note that we are forced to use OLS for this specification, as we do not have a good instrument for fractionalization.

The results of the estimation are presented in columns (1)-(4) in Panel A of Table 10, imply that consistent with the coordination mechanism, fractionalization is negative associated with protest participation (significant at 10% level in all specification). Columns (5)-(8) show that these results are higher in magnitude and statistically significant at 5% level for cities with population over 100,000, in which the problems of coordination are likely to be more important.

Panel B of Table 10 shows that similar results hold for the extensive margin, i.e. for the incidence of protests. Fractionalization is not significantly related to the dummy for protests in

[27] In contrast to the *Odnoklassniki* network, which was not actively employed in the protest movement (Reuter and Szakonyi, 2015).

[28] Note that we check that our results are robust to change in fractionalization index that allows for partial overlap between users from different networks, with results being significant up to 70% overlap.

cities with population below 100,000, but becomes very important for cities with population over 100,000.

Overall, these findings are consistent with hypothesis that social media penetration affects protest participation by facilitating coordination, although it is also consistent with the social pressure mechanism.

Channels: variation over time.

Another exercise that helps us provide additional evidence on the mechanisms behind the effect is the estimation of social media effect over time. Protests in Russia took place regularly before presidential elections took place in March 2012 and then gradually subsided. Different channels suggest different predictions regarding the change in the effect of social media over time. Under assumption that protests were triggered by the evidence of electoral fraud, the information channel suggests that social media should become less important over time, as this information becomes more well known. If protest participation remains socially acceptable activity, as Enikolopov et al. (2015) results suggest, we expect that social pressure channel remains constant or increases over time. Finally, coordination channel suggests that the effect should decrease over time, as coordination links are important to establish initially, but after that coordination can proceed with more traditional offline means. To separate these channels, we estimate the following equation:

$$\log(protesters)_{it} = \beta_0 + \beta_1 VKpenetration_i * t + \beta_2 \mathbf{X_i} * t + \theta_i + t[+t^2 + \cdots] + \varepsilon_{it} \qquad (7)$$

where all the controls and $VKpenetration_i$ are interacted with time trend t, and the list of controls additionally includes a flexible polynomial function of t. An important benefit of using panel data is that we can use city fixed effects to control for time-invariant unobserved heterogeneity, and we control for overtime common protest trends by including either linear time term or the 5th polynomial of time.

Table 11 presents the results of this estimation. The coefficients for the interaction between VK penetration and time trend are negative and significant at 1% level in all specifications. The magnitude of the effect is such that for the city with the mean number of VK users, the social media effect on protests fully disappears after 30 weeks. Panel B of Table 11 confirms that similar results hold for the extensive margin. In other words, social media penetration becomes less important not only for the size of protest demonstrations, but also for the likelihood of protests.

Overall, the results of Table 11 are consistent with information and coordination channel, but are not consistent with social pressure channel. Taken together with the results of Table 10, these results provide an important argument for social media reducing the costs of collective action by lowering the costs of coordination.

7. Conclusion

This paper shows that social media penetration had a causal impact on both the occurrence and the size of protest demonstrations in Russia in 2011-2012. Our results suggest that social media can change the ability of people to overcome collective action problem by reducing the costs of coordination. To the best of our knowledge, it is the first study to show the causal impact of social media penetration on socio-economic outcomes, and we believe that our methodology can be used for studying the impact of social media penetration in other circumstances.

However, there are several limitations of our study worth mentioning. First, we do not know if these results could be extended to other types of collective action, e.g. volunteer activity, monitoring public officials, or political campaigns. Second, we do not know if these results hold for other times and places. Based on the hypotheses outlined in the theoretical section of the paper, we expect social media to have a positive effect on protest participation if all potential channels work in the same direction, as it was the case for Russia 2011-2012. However, if people are afraid of being seen as participants of the protests, and the regime uses a repressive machine to prevent protest activities, the overall impact of social media becomes ambiguous. Third, though our results on the channels behind social media influence are only suggestive, as they are based on the auxiliary tests and are not well identified, so that, unfortunately, we cannot quantify the relative influence of these channels for the overall social media effect.

In sum, our results are the first step in finding the ways through which social media change the opportunities for societies. More research is needed to know more about social media role in collective action in various circumstances.

References

Acemoglu, D., T. Hassan, and A. Tahoun (2015) "The Power of the Street: Evidence from Egypt's Arab Spring," *Working paper*

Adena, M., R. Enikolopov, M. Petrova, V. Santarosa, and E. Zhuravskaya (2015) "Radio and the Rise of Nazis in pre- War Germany," forthcoming in *Quarterly Journal of Economics*

Aidt, T., and R. Franck (forthcoming) "Democratization Under The Threat Of Revolution: Evidence from the Great Reform Act of 1832," *Econometrica*.

Antonji Joseph, Todd Elder and Christopher Taber (2005) "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 113(1), 151-184.

Beissinger, Mark R. (2002). *Nationalist mobilization and the collapse of the Soviet State*. Cambridge University Press.

Bhuller, M., Havnes, T., Leuven, E., and M. Mogstad (2013) "Broadband Internet: An Information Superhighway to Sex Crime?" *Review of Economic Studies*, 80 (4):1237-1266.

Campante, F., R. Durante, and F. Sobbrio (2014) "Politics 2.0: The Multifaceted Effect of Broadband Internet on Political Participation," *Working paper*.

Czernich, Falck, Kretschmer, & Woessmann (2011) Broadband infrastructure and economic growth. *The Economic Journal*, *121*(552): 505-532.

DellaVigna, Stefano and Ethan Kaplan (2007) "The Fox News Effect: Media Bias and Voting," *Quarterly Journal of Economics,* 122(3), 807–860.

Edmond, C. (2013) "Information manipulation, coordination, and regime change," *Review of Economic Studies*, 80(4): 1422–1458.

Eisensee, T. and D. Strömberg. 2007. "News Droughts, News Floods, and U.S. Disaster Relief." *The Quarterly Journal of Economics* 122(2):693-728.

Enikolopov, R., Korovkin, V., Petrova, M., Sonin, K., & Zakharov, A. (2013) "Field experiment estimate of electoral fraud in Russian parliamentary elections," *Proceedings of the National Academy of Sciences*, *110*(2), 448-452.

Enikolopov, R., A. Makarin, M. Petrova, and L. Polichshuk (2015) "Social networks, peer effects, and protest participation," working paper

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

Falck, Oliver, Robert Gold, and Stephan Heblich (2014) "E-Lections: Voting Behavior and the Internet", *American Economic Review,* 104 (7): 2238-2265

Gentzkow, Matthew, Jesse Shapiro, and Michael Sinkinson (2011) "The Effect of Newspaper Entry and Exit on Electoral Politics," *American Economic Review*, 101(7), 2980-3018.

Hassanpour, Navid (2014). Media Disruption and Revolutionary Unrest: Evidence From Mubarak's Quasi-Experiment. *Political Communication*, *31*(1), 1-24.

Klimek, P, Y. Yegorov, R. Hanel, and S. Thurner (2012) "Statistical detection of systematic election irregularities," *Proceedings of National Academy of Sciences*, 109 (41):16469-16473.

Knight, Brian and Chun-Fang Chiang (2011) "Media bias and influence: Evidence from newspaper endorsements," *Review of Economic Studies*, 78(3), 795–820.

Kononov, Nikolay (2012) *Kod Durovai*, Mann, Ivanov and Ferber: Moscow [*in Russian*].

Kuran, Timur (1991) "Now out of never: The element of surprise in the East European revolution of 1989," *World politics*, *44*(1), 7-48.

Little, Andrew (2015) "Communication Technology and Protest," *Working paper*.

Madestam, A., D. Shoag, S. Veuger, and D. Yanagizawa-Drott (2013) "Do Political Protests Matter? Evidence from the Tea Party Movement," *Quarterly Journal of Economics*, 128(4): 1633-1685.

Montiel Olea, José, Luis and Carolin Pflueger (2013) "A robust test for weak instruments," *Journal of Business & Economic Statistics*, 31(3): 358-369.

Passarelli, F., and G. Tabellini (2014) "Emotions and Political Unrest," *Working paper*.

Reuter, Ora John, and David Szakonyi (2015) "Online Social Media and Political Awareness in Authoritarian Regimes." *British Journal of Political Science*, 45(1): 29-51.

Snyder, James and David Strömberg (2010) "Press Coverage and Political Accountability," *Journal of Political Economy,* 118(2), 355–408.

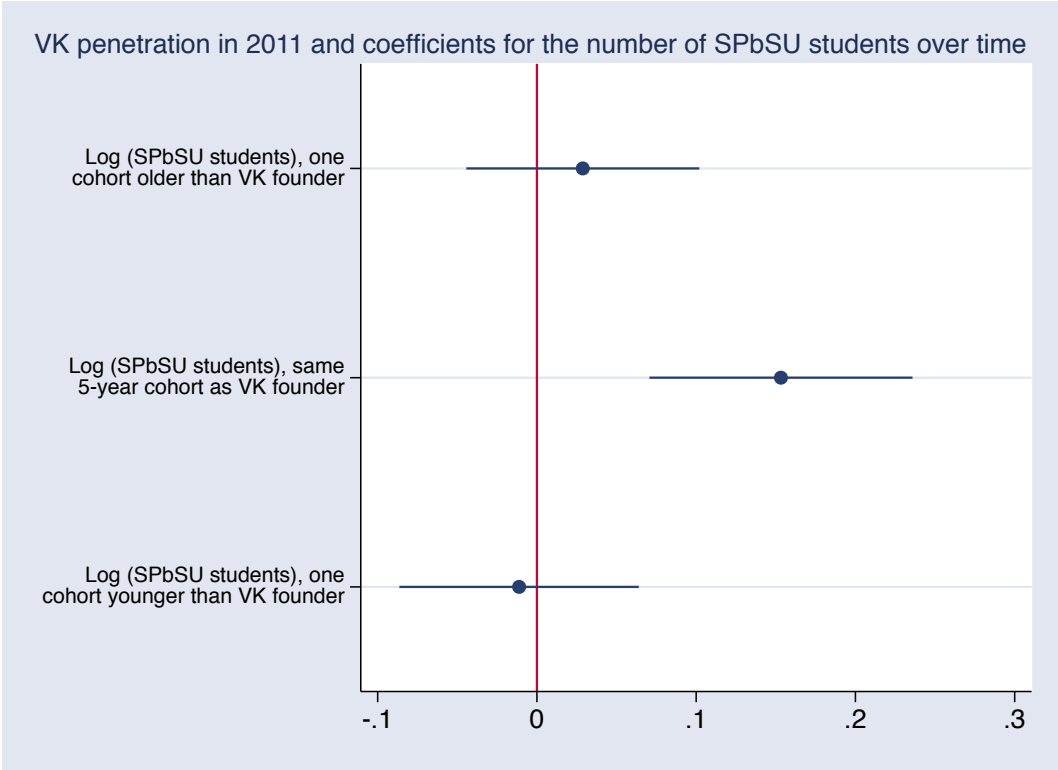
Stock, J., and Yogo, M. (2005), "Testing for Weak Instruments In Linear IV Regression," in *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, eds. D. W. Andrews and J. H. Stock. Cambridge: Cambridge University Press, pp. 80–108.

Strömberg, David (2004) "Radio's Impact on Public Spending," *Quarterly Journal of Economics,* 119(1), 189–221.

Tufekci, Zeynep and Christopher Wilson (2012) "Social Media and the Decision to Participate in Political Protest: Observations From Tahrir Square", *Journal of Communication,* 62(2), 363-379.

Yanagizawa-Drott, David (2014) "Propaganda and Conflict: Theory and Evidence from the Rwandan Genocide," *Quarterly Journal of Economics,* 129(4), pp.1947-1994

Figure 1. Social media penetration and SPbSU student cohorts.
A. SPbSU cohorts from different cities and VK Penetration in 2011



B. SPbSU cohorts from different cities and early (November 2006) VK penetration

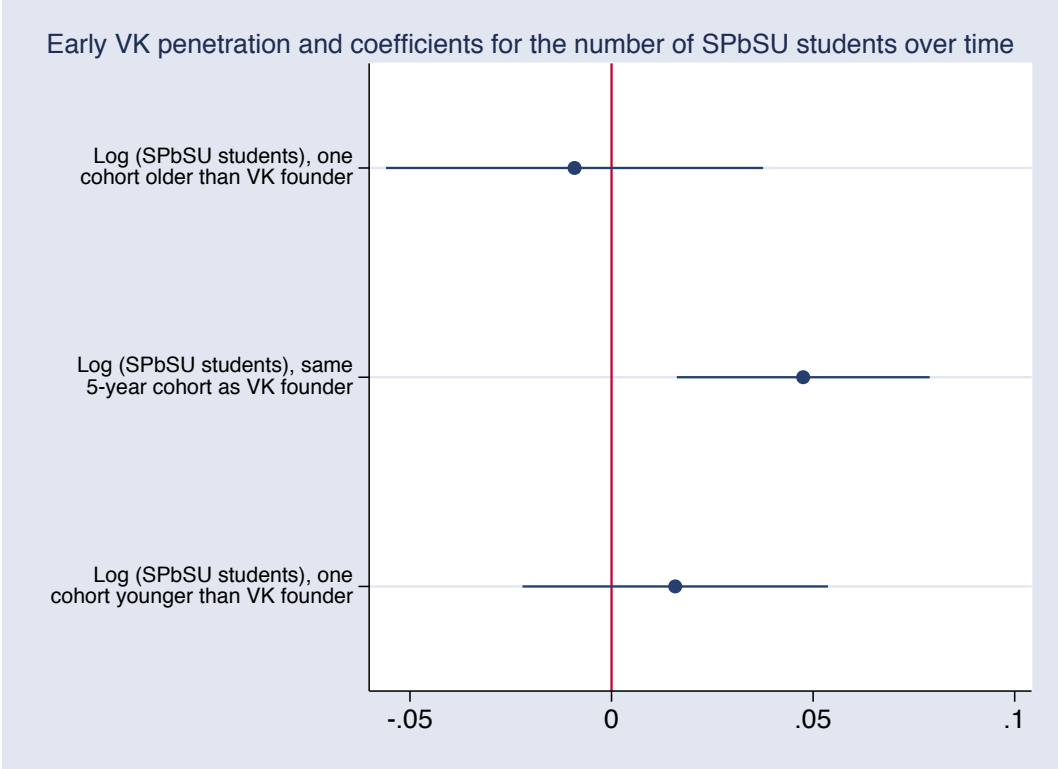
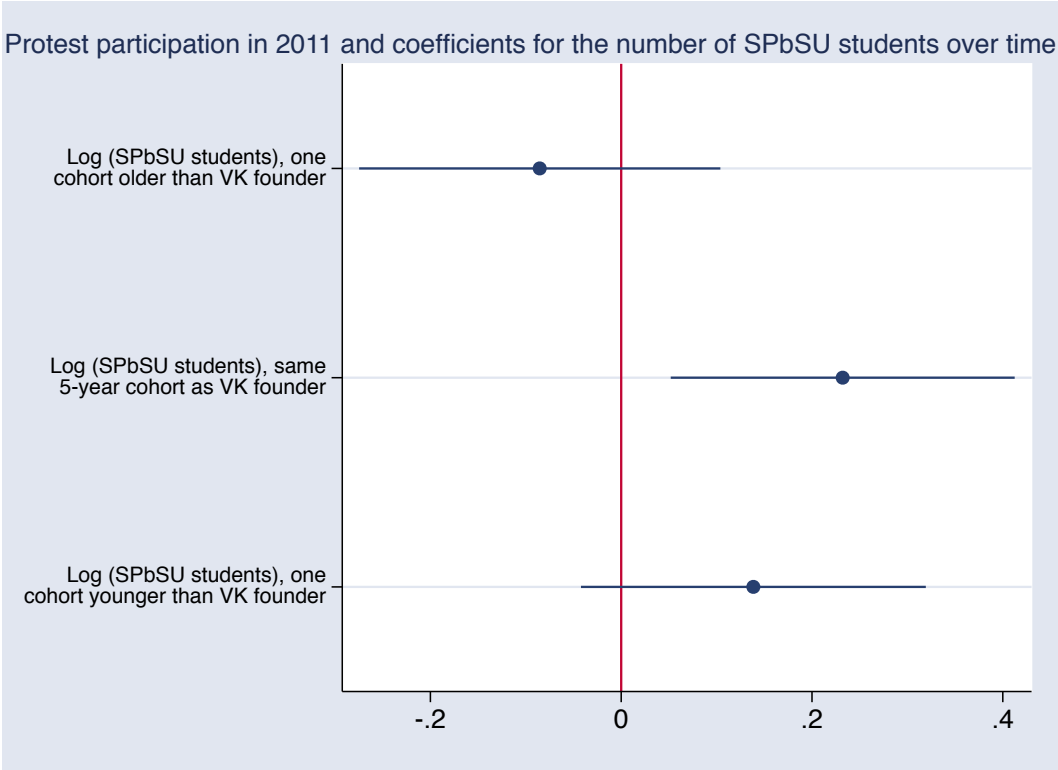


Figure 2. Protest acitivity and SPbSU student cohorts

A. SPbSU cohorts from different cities and protest participation



B. SPbSU cohorts from different cities and the incidence of protests

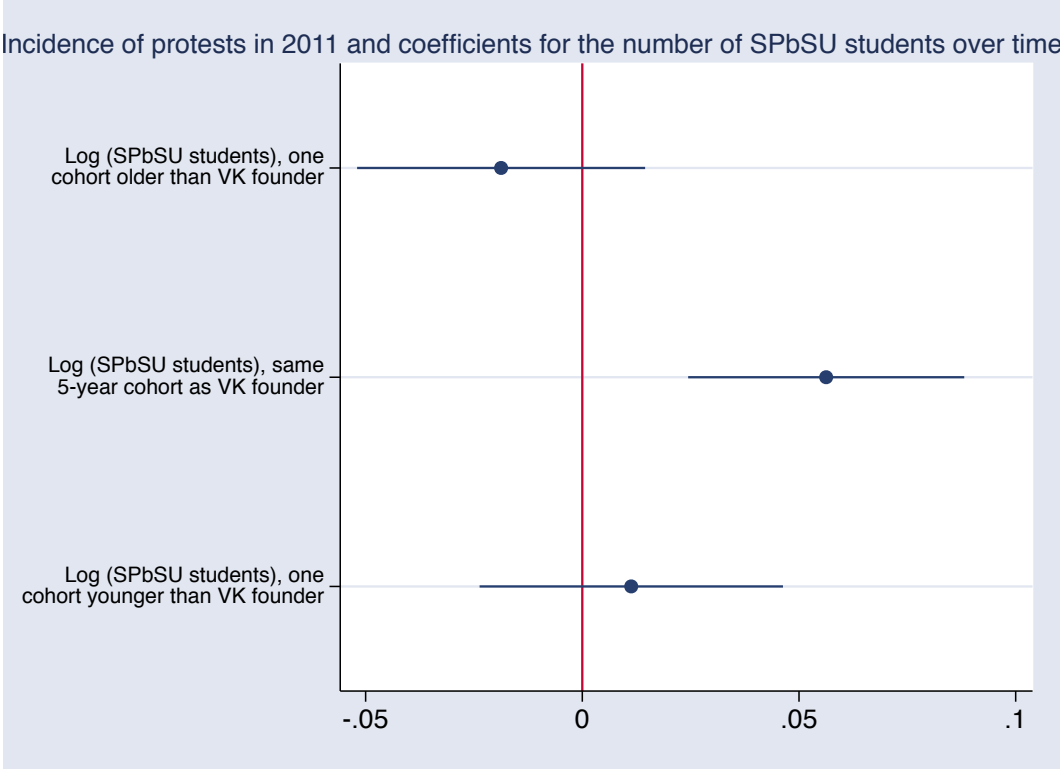


Table 1. Summary statistics.

	Observations	Mean	Standard deviation	Median	Min	Max
Log (protest participation in December 2011)	626	0.77	2.02	0	0	8.66
Incidence of protests in December 2011 (dummy)	626	0.13	0.34	0	0	1
Log(protest participation in USSR in 1987-1992)	626	1.41	2.76	0	0	12.99
Incidence of protests in USSR in 1987-1992 (dummy)	626	0.22	0.41	0	0	1
Log(participationin pro-democratic protests in USSR in 1987-1992)	626	1.38	3.08	0	0	13.93
Incidence of pro-democratic protests in USSR in 1987-1992 (dummy)	626	0.18	0.38	0	0	1
Log (VK users in 2011)	626	9.07	1.36	8.84	2.83	13.38
Log (early VK users, 11/2006)	626	0.07	0.3	0	0	3.5
Log (VK users in 2013)	626	9.74	1.38	9.54	3.33	13.95
Log (Odnoklassniki users 2014)	626	10.72	1.12	10.45	7.94	14.36
Log (Facebook users in 2013)	626	6.9	2.06	6.76	0	12.3
Population in 2010, in thousands	626	117.55	189.5	52.5	20	1393.5
Administrative center	626	0.12	0.32	0	0	1
Distance to Saint Petersburg	626	1480.54	839.17	1417	21.7	4646
Distance to Moscow	626	1151.38	875.95	1011	15.75	4174
Rayon center (county seat, dummy)	626	0.79	0.41	1	0	1
Log (average wage, 2011)	626	9.71	0.26	9.64	8.93	10.69
Log(number of people with age 20-24 in 2010	626	8.46	1.04	8.22	4.57	11.73
Log(number of people with age 25-29 in 2010	626	8.46	0.99	8.25	4.72	11.62
Log(number of people with age 30-34 in 2010	626	8.36	0.99	8.15	4.63	11.48
Log(number of people with age 35-39 in 2010	626	8.39	0.98	8.18	4.75	11.45
Log(number of people with age 40-44 in 2010	626	8.62	0.96	8.41	5.08	11.66
Log(number of people with age 45-49 in 2010	626	8.57	0.96	8.38	4.8	11.65
Log(number of people with age 50 and older in 2010	626	9.79	0.99	9.58	5.86	12.95
Presence of university in a city	626	0.15	0.35	0	0	1
Percentage with higher education	626	0.15	0.06	0.13	0.05	0.45
Internet penetration in 2011	626	0.27	0.17	0.22	0.01	0.63
Log (SPbSU students, same 5-year cohort as VK founder)	626	0.48	0.75	0	0	4.64
Log (SPbSU students, one cohort younger than VK founder)	626	0.4	0.63	0	0	2.77
Log (SPbSU students, one cohort older than VK founder)	626	0.43	0.7	0	0	3.53

Table 2. Distribution of size of SPbSU student cohorts

Number of SPbSU students from a city in VK founder's cohort	Frequency	Number of SPbSU students from a city one cohort older than VK founder	Frequency	Number of SPbSU students from a city one cohort yonger than VK founder	Frequency
0	389	0	404	0	412
1	96	1	106	1	85
2	50	2	32	2	48
3	19	3	16	3	31
4	15	4	19	4	14
5	12	5	7	5	13
6	9	6	11	6	6
7	1	7	8	7	7
8	5	8	5	8	2
9	10	9	1	9	2
10	3	10	3	10	1
11	4	11	2	12	1
12	2	12	2	13	1
13	2	13	4	14	2
14	1	14	1	15	1
15	1	20	2		
16	1	21	1		
17	1	29	1		
20	1	33	1		
23	1				
25	1				
29	1				
103	1				

Note: all the results in the paper are robust to exclusion of a city with 103 people in VK founder cohorts (if anything, results get stronger without this outlier).

Table 3. Determinants of VK penetration.

Panel A. Determinants of VK penetration in 2011. First stage regression.

	Log (VK users in August 2011)						
Log (SPbSU students), same 5-year cohort as VK founder	0.4881***	0.1609***	0.1531***	0.1508***	0.1555***	0.1574***	0.1603***
	[0.1444]	[0.0424]	[0.0442]	[0.0490]	[0.0493]	[0.0481]	[0.0486]
Log (SPbSU students), one cohort younger than VK founder	0.5785***	-0.0254	-0.0102	-0.0180	-0.0143	-0.0185	-0.0131
	[0.1063]	[0.0557]	[0.0449]	[0.0449]	[0.0452]	[0.0482]	[0.0471]
Log (SPbSU students), one cohort older than VK founder	0.3127*	0.0277	0.0248	0.0339	0.0332	0.0272	0.0309
	[0.1871]	[0.0526]	[0.0441]	[0.0446]	[0.0439]	[0.0440]	[0.0433]
Regional Center		0.2916***	0.4373***	0.4113**	0.3895**	0.4336***	0.4090***
		[0.0903]	[0.1099]	[0.1578]	[0.1552]	[0.1519]	[0.1505]
Distance to St.Petersburg			0.0001	0.0001	0.0001	0.0001	0.0000
			[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0001]
Distance to Moscow			-0.0002	-0.0001	-0.0002	-0.0002	-0.0001
			[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0001]
Rayon Center (county seat)			0.0535	0.0488	0.0501	0.0526	0.0437
			[0.0878]	[0.0887]	[0.0884]	[0.0903]	[0.0850]
Average wage, logged, region-level			0.3429**	0.3766**	0.3601**	0.3746**	0.2933*
			[0.1337]	[0.1429]	[0.1546]	[0.1573]	[0.1519]
Presence of university in the city				-0.0237	-0.0181	-0.0171	-0.0175
				[0.1599]	[0.1557]	[0.1556]	[0.1533]
% with higher education				-0.0913	-0.1581	-0.5135	-0.3295
				[1.1609]	[1.2097]	[1.1626]	[1.0623]
Internet penetration 2011, region-level				-0.0853	-0.0740	0.0404	-0.0454
				[0.2493]	[0.2400]	[0.2530]	[0.2387]
				0.0944	0.0997	0.0944	0.1253
				[0.0816]	[0.0822]	[0.0807]	[0.0776]
Log (protest participants), 1987-1992				0.0036	0.0064	0.0030	0.0008
				[0.0080]	[0.0084]	[0.0085]	[0.0080]
Observations	626	626	626	626	626	626	626
R-squared	0.3957	0.8056	0.8302	0.8303	0.8321	0.8326	0.8323
Population controls		Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Age cohort controls			Yes***	Yes***	Yes***	Yes***	Yes***
Electoral controls, 1995					Yes		
Electoral controls, 1999						Yes	
Electoral controls, 2003							Yes

Table 3. Determinants of VK penetration (continued).

Panel B. Determinants of early VK penetration.

	Log (early VK users in November 2006)						
Log (SPbSU students), same 5-year cohort as VK founder	0.0897***	0.0469**	0.0545***	0.0479**	0.0495**	0.0465**	0.0493***
	[0.0219]	[0.0200]	[0.0185]	[0.0190]	[0.0189]	[0.0193]	[0.0185]
Log (SPbSU students), one cohort younger than VK founder	0.0768**	0.0156	0.0137	0.0169	0.0210	0.0154	0.0184
	[0.0379]	[0.0246]	[0.0229]	[0.0235]	[0.0243]	[0.0239]	[0.0239]
Log (SPbSU students), one cohort older than VK founder	0.0310	-0.0162	-0.0025	-0.0100	-0.0115	-0.0122	-0.0124
	[0.0338]	[0.0230]	[0.0277]	[0.0280]	[0.0276]	[0.0280]	[0.0276]
Regional Center		-0.0768	-0.0786	-0.1271	-0.1170	-0.1234	-0.1318
		[0.0624]	[0.0662]	[0.0873]	[0.0880]	[0.0857]	[0.0867]
Distance to St.Petersburg			0.0001*	0.0001	0.0001	0.0001	0.0000
			[0.0000]	[0.0000]	[0.0000]	[0.0001]	[0.0000]
Distance to Moscow			0.0000	0.0000	0.0000	0.0000	0.0000
			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Rayon Center (county seat)			-0.0258**	-0.0219*	-0.0261*	-0.0280**	-0.0191
			[0.0126]	[0.0122]	[0.0145]	[0.0120]	[0.0131]
Average wage, logged, region-level			0.0454	0.0228	0.0006	0.0134	0.0321
			[0.0372]	[0.0387]	[0.0449]	[0.0430]	[0.0385]
Presence of university in the city				0.0501	0.0577	0.0503	0.0534
				[0.0797]	[0.0807]	[0.0789]	[0.0788]
% with higher education				0.5264***	0.1792	0.3671**	0.3774**
				[0.1682]	[0.1802]	[0.1601]	[0.1839]
Internet penetration 2011, region-level				-0.0159	-0.0265	-0.0279	-0.0325
				[0.0484]	[0.0452]	[0.0469]	[0.0500]
Log (Odnoklassniki users in 2014)				-0.0153	-0.0049	-0.0112	-0.0076
				[0.0159]	[0.0144]	[0.0166]	[0.0150]
Log (protest participants), 1987-1992				-0.0005	-0.0015	-0.0015	-0.0014
				[0.0041]	[0.0042]	[0.0042]	[0.0044]
Observations	626	626	626	626	626	626	626
R-squared	0.1737	0.5246	0.5343	0.5397	0.5464	0.5432	0.5455
Population controls		Yes	Yes	Yes	Yes	Yes	Yes
Age cohort controls			Yes	Yes***	Yes	Yes	Yes
Electoral controls, 1995					Yes		
Electoral controls, 1999						Yes	
Electoral controls, 2003							Yes*

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets are adjusted by clusters within regions. Unit of observation is a city. When "Yes" is added to indiciate inclusion of a group of controls, a significance level is reported immediately after for this group of controls. Flexible controls for population (5th polynomial) are included in all specifications. Age cohort controls include the number of people aged 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50 and older years, in each city according to 2010 Russian Census. Electoral controls include vote for Yabloko party, Communist Party (KPRF), LDPR party, the ruling party (Our Home is Russia in 1995, Unity in 1999, United Russia in 2003), vote against all, and electoral turnout for a corresponding year.

Table 4. Student cohorts and protest participation in 2011. Reduced form estimation.

	Log (protesters in 2011)				Incidence of protests in December 2011 (dummy)			
Log (SPbSU students), same 5-year cohort as VK founder	0.228**	0.246**	0.245**	0.260**	0.055***	0.058***	0.058***	0.061***
	[0.107]	[0.106]	[0.108]	[0.106]	[0.019]	[0.018]	[0.019]	[0.019]
Log (SPbSU students), one cohort younger than VK founder	0.129	0.126	0.120	0.143	0.008	0.007	0.006	0.010
	[0.106]	[0.107]	[0.107]	[0.107]	[0.021]	[0.021]	[0.021]	[0.021]
Log (SPbSU students), one cohort older than VK founder	-0.074	-0.081	-0.094	-0.077	-0.016	-0.017	-0.020	-0.016
	[0.115]	[0.115]	[0.113]	[0.115]	[0.020]	[0.020]	[0.020]	[0.021]
Regional Center	0.313	0.320	0.350	0.327	-0.002	-0.000	0.004	0.002
	[0.491]	[0.486]	[0.479]	[0.486]	[0.102]	[0.101]	[0.098]	[0.101]
Distance to St.Petersburg	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Distance to Moscow	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Rayon Center (county seat)	0.011	0.009	-0.020	-0.043	0.002	0.002	-0.004	-0.008
	[0.043]	[0.050]	[0.052]	[0.054]	[0.008]	[0.009]	[0.010]	[0.010]
Average wage, logged, region-level	0.001	-0.041	-0.161	-0.175	0.002	-0.002	-0.026	-0.031
	[0.162]	[0.167]	[0.171]	[0.180]	[0.033]	[0.035]	[0.035]	[0.034]
Presence of university in the city	0.898**	0.917**	0.896**	0.914**	0.213*	0.216*	0.213*	0.215*
	[0.448]	[0.449]	[0.444]	[0.446]	[0.109]	[0.109]	[0.107]	[0.108]
% with higher education	-1.048	-1.530	-2.687**	-2.491**	-0.133	-0.195	-0.400*	-0.372*
	[1.088]	[1.353]	[1.261]	[1.134]	[0.213]	[0.252]	[0.234]	[0.219]
Internet penetration 2011, region-level	0.260	0.291	0.213	0.229	0.027	0.043	0.017	0.024
	[0.259]	[0.259]	[0.294]	[0.274]	[0.049]	[0.050]	[0.057]	[0.053]
Log (Odnoklassniki users in 2014)	0.146	0.132	0.195	0.157	0.042**	0.038*	0.052**	0.041**
	[0.102]	[0.111]	[0.120]	[0.114]	[0.017]	[0.019]	[0.022]	[0.020]
Log (protest participants), 1987-1992	0.041	0.047*	0.037	0.039				
	[0.026]	[0.027]	[0.027]	[0.026]				
Incidence of protests in 1987-1992 (dummy)					0.059*	0.066*	0.053	0.058
					[0.036]	[0.035]	[0.036]	[0.035]
Observations	626	626	626	626	626	626	626	626
R-squared	0.818	0.822	0.823	0.823	0.767	0.772	0.771	0.773
Population controls	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Age cohort controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes*	Yes*
Electoral controls, 1995		Yes**				Yes**		
Electoral controls, 1999			Yes**				Yes*	
Electoral controls, 2003				Yes**				Yes**

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets are adjusted by clusters within regions. Unit of observation is a city. When "Yes" is added to indiciate inclusion of a group of controls, a significance level is reported immediately after for this group of controls. Flexible controls for population (5th polynomial) are included in all specifications. Age cohort controls include the number of people aged 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50 and older years, in each city according to 2010 Russian Census. Electoral controls include vote for Yabloko party, Communist Party (KPRF), LDPR party, the ruling party (Our Home is Russia in 1995, Unity in 1999, United Russia in 2003), vote against all, and electoral turnout for a corresponding year.

Table 5. VK penetration and protest participation

Panel A. IV and OLS estimates. December 2011

	Log (protesters in December 2011)							
	IV	IV	IV	IV	OLS	OLS	OLS	OLS
Log (VK users in 2011)	1.511**	1.579**	1.555**	1.623**	0.185***	0.173***	0.170***	0.184***
	[0.749]	[0.754]	[0.750]	[0.722]	[0.062]	[0.062]	[0.061]	[0.064]
Log (SPbSU students), one cohort younger than VK founder	0.156	0.149	0.148	0.164	0.191*	0.192*	0.185	0.211*
	[0.117]	[0.121]	[0.121]	[0.120]	[0.111]	[0.113]	[0.112]	[0.110]
Log (SPbSU students), one cohort older than VK founder	-0.125	-0.133	-0.137	-0.127	0.006	0.009	-0.005	0.016
	[0.143]	[0.145]	[0.137]	[0.143]	[0.093]	[0.092]	[0.090]	[0.094]
Population controls	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Age cohort controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Electoral controls, 1995		Yes**				Yes**		
Electoral controls, 1999			Yes*				Yes*	
Electoral controls, 2003				Yes**				Yes**
Observations	626	626	626	626	626	626	626	626
Effective F-statistics (Olea Montiel and Pflueger 2013)	281	278.1	278.1	278.1				

Panel B. Placebo results. Soviet protests.

	Log (protesters in 1987-1992)				Log (participants in pro-democracy protests in 1987-1992)			
Log (VK users in 2011)	0.770	0.700	0.478	0.810	0.178	0.053	0.028	0.251
	[1.605]	[1.679]	[1.573]	[1.625]	[1.374]	[1.379]	[1.355]	[1.391]
	Log (participants in social protests in 2005)							
Log (VK users in 2011)	-0.443	-0.468	-0.422	-0.224				
	[1.420]	[1.337]	[1.336]	[1.342]				

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets are adjusted by clusters within regions. Unit of observation is a city. When "Yes" is added to indiciate inclusion of a group of controls, a significance level is reported immediately after for this group of controls. Flexible controls for population (5th polynomial) are included in all specifications. Age cohort controls include the number of people aged 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50 and older years, in each city according to 2010 Russian Census. Electoral controls include vote for Yabloko party, Communist Party (KPRF), LDPR party, the ruling party (Our Home is Russia in 1995, Unity in 1999, United Russia in 2003), vote against all, and electoral turnout for a corresponding year. Other controls include dummy for regional and county centers , distances to Moscow and St Peterburg, log (average wage), % with higher education, internet penetration in 2011, log (Odnoklassniki users in 2014), and, in panel A, protest participation in 1987-1992. F-stat for equality of coefficients in main results and placebos do not suggest that the difference is significance.

Table 6. VK penetration and protest participation in 2011. Extensive margin.

Panel A. Actual results. December 2011

	Incidence (protests in December 2011), dummy							
	IV	IV	IV	IV	OLS	OLS	OLS	OLS
Log (VK users in 2011)	0.365**	0.370**	0.368**	0.381***	0.033***	0.031**	0.031**	0.033***
	[0.146]	[0.146]	[0.143]	[0.141]	[0.012]	[0.012]	[0.012]	[0.012]
Log (SPbSU students), one cohort younger than VK founder	0.014	0.012	0.012	0.015	0.023	0.022	0.021	0.026
	[0.024]	[0.024]	[0.025]	[0.024]	[0.021]	[0.022]	[0.022]	[0.021]
Log (SPbSU students), one cohort older than VK founder	-0.028	-0.029	-0.029	-0.027	0.005	0.005	0.002	0.007
	[0.028]	[0.028]	[0.027]	[0.028]	[0.018]	[0.017]	[0.017]	[0.018]
Population controls	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Age cohort controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Electoral controls, 1995		Yes				Yes		
Electoral controls, 1999			Yes				Yes	
Electoral controls, 2003				Yes				Yes
Observations	626	626	626	626	626	626	626	626
Effective F-stat (Montiel Olea and Pflueger 2013)	281	278.1	278.1	278.1				

Panel B. Placebo results. Soviet protests.

	Log (protesters in 1987-1992)				Log (participants in pro-democracy protests in 1987-			
Log (VK users in 2011)	0.048	0.047	0.013	0.066	-0.005	-0.009	-0.019	0.021
	[0.230]	[0.237]	[0.221]	[0.230]	[0.173]	[0.173]	[0.169]	[0.172]
	Log (participants in social protests in 2005)							
Log (VK users in 2011)	-0.081	-0.089	-0.080	-0.049				
	[0.204]	[0.195]	[0.194]	[0.195]				

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets are adjusted by clusters within regions. Unit of observation is a city. When "Yes" is added to indiciate inclusion of a group of controls, a significance level is reported immediately after for this group of controls. Flexible controls for population (5th polynomial) are included in all specifications. Age cohort controls include the number of people aged 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50 and older years, in each city according to 2010 Russian Census. Electoral controls include vote for Yabloko party, Communist Party (KPRF), LDPR party, the ruling party (Our Home is Russia in 1995, Unity in 1999, United Russia in 2003), vote against all, and electoral turnout for a corresponding year. Other controls include dummy for regional and county centers , distances to Moscow and St Peterburg, log (average wage), % with higher education, internet penetration in 2011, log (Odnoklassniki users in 2014), and, in panel A, incidence of protests in 1987-1992. Equality of VK coefficients for 2011 protests and pro-democracy protests in 1987-1992 is rejected at 10% level in columns 5-8 (using 3sls framework, for column 7 with electoral controls of 1999 at 5% level).

Table 7. VK and penetration of other popular online social networks

	Log (Facebook users, 2013)				Log (Odnoklassniki users, 2014)			
Log (early VK users, id<5k)	-0.013				-0.042			
	[0.173]				[0.053]			
Log (early VK users, id<50k)		0.298*				-0.025		
		[0.153]				[0.035]		
Log (early VK users, id<100k)			0.235*				0.005	
			[0.136]				[0.035]	
Log (VK users, 2011)				0.446***				0.046
				[0.137]				[0.046]
Regional Center	0.053	0.023	0.011	-0.125	0.225**	0.234**	0.230**	0.212**
	[0.267]	[0.277]	[0.280]	[0.247]	[0.112]	[0.109]	[0.109]	[0.102]
Distance to St.Petersburg	-0.000*	-0.000	-0.000	-0.000*	0.000***	0.000***	0.000***	0.000***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Distance to Moscow	0.000	0.000	0.000	0.000*	-0.000***	-0.000***	-0.000***	-0.000***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Rayon Center (county seat)	0.084	0.097	0.096	0.054	0.028	0.027	0.029	0.025
	[0.259]	[0.253]	[0.256]	[0.236]	[0.073]	[0.073]	[0.073]	[0.072]
Average wage, logged, region-level	0.386	0.333	0.305	0.226	0.080	0.083	0.077	0.062
	[0.279]	[0.264]	[0.269]	[0.246]	[0.121]	[0.122]	[0.122]	[0.122]
Presence of university in the city	0.669**	0.578**	0.588*	0.656**	0.026	0.031	0.022	0.022
	[0.296]	[0.280]	[0.316]	[0.321]	[0.102]	[0.103]	[0.099]	[0.097]
% with higher education	4.321*	3.262	3.243	4.105*	1.116*	1.177*	1.068*	1.068*
	[2.519]	[2.364]	[2.288]	[2.194]	[0.588]	[0.633]	[0.613]	[0.560]
Internet penetration 2011, region-level	-1.024	-1.047	-1.090	-0.963	-0.307	-0.304	-0.308	-0.300
	[0.933]	[0.932]	[0.937]	[0.877]	[0.205]	[0.205]	[0.205]	[0.208]
Log (protest participants), 1987-1992	0.047*	0.041	0.041*	0.045*	0.003	0.004	0.003	0.003
	[0.027]	[0.025]	[0.024]	[0.025]	[0.010]	[0.010]	[0.010]	[0.009]
Population controls	Yes***	Yes**	Yes**	Yes	Yes***	Yes***	Yes***	Yes***
Age cohort controls	Yes**	Yes**	Yes*	Yes*	Yes***	Yes***	Yes***	Yes***
Observations	626	626	626	626	626	626	626	626
R-squared	0.558	0.559	0.559	0.571	0.881	0.881	0.881	0.882

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets are adjusted by clusters within regions. Unit of observation is a city. When "Yes" is added to indiciate inclusion of a group of controls, a significance level is reported immediately after for this group of controls. Flexible controls for population (5th polynomial) are included in all specifications. Age cohort controls include the number of people aged 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50 and older years, in each city according to 2010 Russian Census. These results are robust to inclusion of electoral controls, but corresponding specifications are not shown to save space.

Table 8. Placebo specifications.

Panel A. Pre-VK parliamentary elections

	Dependent variable					
	Yabloko vote share	Communists vote share	Pro-government party vote share	LDPR vote share	Turnout	Against all share
Voting results in 1995, IV with SPbSU cohorts	-0.026	0.107*	-0.003	0.015	0.028	-0.011
	[0.018]	[0.064]	[0.031]	[0.047]	[0.031]	[0.007]
Voting results in 1999, IV with SPbSU cohorts	-0.007	0.074*	0.053	-0.014	-0.050	-0.007
	[0.013]	[0.043]	[0.048]	[0.012]	[0.041]	[0.007]
Voting results in 2003 IV with SPbSU cohorts	-0.021**	0.003	-0.002	-0.014	0.020	-0.020*
	[0.010]	[0.024]	[0.003]	[0.025]	[0.044]	[0.012]

Panel B. Pre-VK presidential elections, 1996

	Dependent variable					
	Yeltsin vote share	Zyuganov vote share	Lebedev vote share	Yavlinksy vote share	Turnout	Against all share
Voting results, IV with SPbSU cohorts	-0.115*	0.153*	-0.030	0.002	0.018	-0.004
	[0.065]	[0.082]	[0.042]	[0.015]	[0.021]	[0.003]

Panel C. Pre-VK presidential elections, 2000

	Dependent variable					
	Putin vote share	Zyuganov vote share	Yavlinsky vote share	Tuleev vote share	Turnout	Against all share
Voting results, IV with SPbSU cohorts	0.123*	-0.024	-0.034**	-0.013	0.026	-0.014***
	[0.070]	[0.051]	[0.014]	[0.025]	[0.028]	[0.005]

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets are adjusted by clusters within regions. Each cell reports the coefficient for log (VK users) from IV regression with a standard set of controls (i.e. Table 5A, column 1) with various pre-2006 dependent variables, indicated in column titles.

Table 9. Student cohorts and protest participation in 2011. Altonji-Elder-Taber tests.

	Log (protesters in 2011)				Incidence of protests in December 2011 (dummy)			
Log (SPbSU students), same 5-year cohort as VK founder, predicted	0.341	0.147	0.163	-0.013	0.090	0.061	0.060	0.024
	[0.287]	[0.263]	[0.239]	[0.254]	[0.063]	[0.058]	[0.056]	[0.051]
Log (SPbSU students), one cohort younger than VK founder	0.092	0.156	0.151	0.208	-0.000	0.009	0.009	0.021
	[0.162]	[0.155]	[0.147]	[0.149]	[0.030]	[0.029]	[0.028]	[0.027]
Log (SPbSU students), one cohort older than VK founder	-0.033	0.072	0.064	0.159	-0.020	-0.005	-0.004	0.016
	[0.174]	[0.159]	[0.151]	[0.161]	[0.036]	[0.033]	[0.033]	[0.032]
Direct population controls	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Observations	626	626	626	626	626	626	626	626
R-squared	0.799	0.799	0.799	0.799	0.738	0.738	0.738	0.737
Electoral controls, used for prediction		1995	1999	2003		1995	1999	2003

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets are adjusted by clusters within regions. Unit of observation is a city. When "Yes" is added to indiciate inclusion of a group of controls, a significance level is reported immediately after for this group of controls. Flexible controls for population (5th polynomial) are included in all specifications. Age cohort controls include the number of people aged 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50 and older years, in each city according to 2010 Russian Census. Electoral controls include vote for Yabloko party, Communist Party (KPRF), LDPR party, the ruling party (Our Home is Russia in 1995, Unity in 1999, United Russia in 2003), vote against all, and electoral turnout for a corresponding year.

Table 10. Fractionalization of networks and protest participation. OLS estimates.

Panel A. Network fractionalization and protest participation.

	Log (protesters in December 2011)							
	Whole sample				Cities with more than 100 000 inhabitants			
Fractionalization of social media networks (Facebook+Vkontakte)	-1.194**	-1.334**	-1.262**	-1.215**	-3.390**	-3.905***	-3.282**	-4.029**
	[0.598]	[0.602]	[0.594]	[0.609]	[1.412]	[1.346]	[1.372]	[1.523]
Log (number of users in both networks)	1.791***	1.784***	1.717***	1.789***	1.535***	1.364***	1.511***	1.569***
	[0.334]	[0.332]	[0.332]	[0.326]	[0.514]	[0.465]	[0.511]	[0.513]
Population controls	Yes***	Yes***	Yes***	Yes***	Yes	Yes	Yes	Yes
Age cohort controls	Yes	Yes	Yes	Yes**	Yes	Yes	Yes	Yes
Electoral controls, 1995		Yes*				Yes**		
Electoral controls, 1999			Yes*				Yes**	
Electoral controls, 2003				Yes				Yes**
Observations	626	626	626	626	158	158	158	158
R-squared	0.831	0.835	0.835	0.834	0.790	0.814	0.818	0.822

Panel A. Network fractionalization and the incidence of protest

	Incidence of protests in December 2011 (dummy)							
	Whole sample				Cities with more than 100 000 inhabitants			
Fractionalization of social media networks (Facebook+Vkontakte)	-0.173	-0.196	-0.186	-0.174	-0.560*	-0.592**	-0.543*	-0.666**
	[0.127]	[0.125]	[0.125]	[0.130]	[0.302]	[0.265]	[0.297]	[0.333]
Log (number of users in both networks)	0.281***	0.281***	0.271***	0.279***	0.188*	0.169*	0.169*	0.202**
	[0.065]	[0.064]	[0.065]	[0.064]	[0.096]	[0.086]	[0.101]	[0.096]
Population controls	Yes***	Yes***	Yes***	Yes***	Yes	Yes	Yes	Yes
Age cohort controls	Yes*	Yes	Yes	Yes**	Yes	Yes	Yes	Yes***
Electoral controls, 1995		Yes*				Yes		
Electoral controls, 1999			Yes				Yes	
Electoral controls, 2003				Yes**				Yes**
Observations	626	626	626	626	158	158	158	158
R-squared	0.777	0.782	0.781	0.782	0.728	0.756	0.754	0.772

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets are adjusted by clusters within regions. Unit of observation is a city. When "Yes" is added to indiciate inclusion of a group of controls, a significance level is reported immediately after for this group of controls. Flexible controls for population (5th polynomial) are included in all specifications. Age cohort controls include the number of people aged 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50 and older years, in each city according to 2010 Russian Census. Electoral controls include vote for Yabloko party, Communist Party (KPRF), LDPR party, the ruling party (Our Home is Russia in 1995, Unity in 1999, United Russia in 2003), vote against all, and electoral turnout for a corresponding year. Other controls include dummy for regional and county centers , distances to Moscow and St Peterburg, log (average wage), % with higher education, internet penetration in 2011, log (Odnoklassniki users in 2014), and, in panel A, protest participation in 1987-1992.

Table 11. VK penetration effect over time, 2011-2012. City Fixed Effects estimation.

Panel A. Protest participation over Time.

	Log (protesters)							
Log (VK users in 2011)*Time	-0.0015***	-0.0015***	-0.0015***	-0.0015***	-0.0015***	-0.0015***	-0.0015***	-0.0015***
	[0.0005]	[0.0005]	[0.0005]	[0.0005]	[0.0005]	[0.0005]	[0.0005]	[0.0005]
Time	0.0403**	0.0369**	0.0324*	0.0316	-0.0450*	-0.0484**	-0.0529**	-0.0537**
	[0.0186]	[0.0184]	[0.0192]	[0.0198]	[0.0230]	[0.0232]	[0.0237]	[0.0238]
City Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls interacted with Time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5th polynomial of Time					Yes	Yes	Yes	Yes
Voting controls 1995, interacted with Time		Yes				Yes		
Voting controls 1999, interacted with Time			Yes				Yes	
Voting controls 2003, interacted with Time				Yes				Yes
Observations	26,292	26,292	26,292	26,292	26,292	26,292	26,292	26,292
R-squared	0.120	0.120	0.120	0.120	0.121	0.121	0.122	0.122

Panel B. Protest incidence over Time.

	Incidence of protest demonstrations (dummy)							
Log (VK users in 2011)*Time	-0.0003***	-0.0003***	-0.0003***	-0.0003***	-0.0003***	-0.0003***	-0.0003***	-0.0003***
	[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0001]
Time	0.0061*	0.0053	0.0044	0.0046	-0.0084*	-0.0092**	-0.0101**	-0.0099**
	[0.0033]	[0.0034]	[0.0035]	[0.0035]	[0.0043]	[0.0044]	[0.0045]	[0.0044]
City Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls interacted with Time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5th polynomial of Time					Yes	Yes	Yes	Yes
Voting controls 1995, interacted with Time		Yes				Yes		
Voting controls 1999, interacted with Time			Yes				Yes	
Voting controls 2003, interacted with Time				Yes				Yes
Observations	26,292	26,292	26,292	26,292	26,292	26,292	26,292	26,292
R-squared	0.121	0.122	0.122	0.122	0.123	0.123	0.123	0.123

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets are adjusted by clusters within regions. Unit of observation is a city. Flexible controls for population (5th polynomial), interacted with time, are included in all specifications. Age cohort controls include the number of people aged 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50 and older years, in each city according to 2010 Russian Census, all interacted with time. Electoral controls include vote for Yabloko party, Communist Party (KPRF), LDPR party, the ruling party (Our Home is Russia in 1995, Unity in 1999, United Russia in 2003), vote against all, and electoral turnout for a corresponding year, all interacted with time. Other controls include dummy for regional and county centers , distances to Moscow and St Peterburg, log (average wage), % with higher education, internet penetration in 2011, log (Odnoklassniki users in 2014), and, in panel A, protest participation in 1987-1992, or, in panel B, incidence of protests in 1987-1992, all interacted with time.

APPENDIX. Additional evidence.

Figure A1. VK penetration over time. Number of users (vertical axis) and the date of the first post (horizontal axis) are shown.

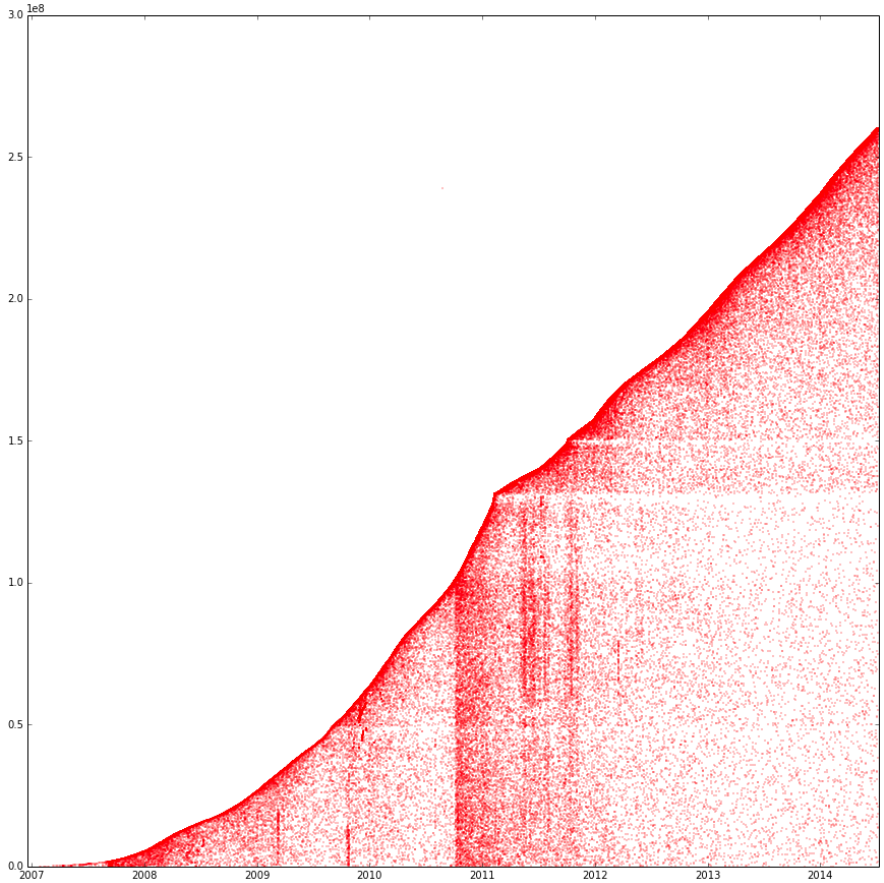


Table A1. Determinants of VK penetration in 2011, cohorts excluded.

VARIABLES	Log (VK users in 2011)						
Regional Center	0.336***	0.398***	0.462***	0.397**	0.374**	0.426**	0.403**
	[0.088]	[0.096]	[0.109]	[0.177]	[0.177]	[0.172]	[0.172]
Distance to St.Petersburg		-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
		[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Distance to Moscow		-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
		[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Rayon Center (county seat)		0.137	0.067	0.066	0.071	0.069	0.069
		[0.095]	[0.087]	[0.089]	[0.089]	[0.091]	[0.085]
Average wage, logged, region-level		0.390***	0.334**	0.358**	0.362**	0.376**	0.303*
		[0.143]	[0.141]	[0.147]	[0.157]	[0.158]	[0.155]
Presence of university in the city				0.029	0.032	0.029	0.032
				[0.176]	[0.174]	[0.173]	[0.175]
% with higher education				0.463	0.550	0.075	0.500
				[1.268]	[1.318]	[1.281]	[1.169]
Internet penetration 2011, region-level				-0.135	-0.104	0.018	-0.100
				[0.258]	[0.253]	[0.271]	[0.254]
Log (protest participants), 1987-1992				0.006	0.009	0.006	0.005
				[0.009]	[0.010]	[0.009]	[0.009]
Observations	626	626	626	626	626	626	626
R-squared	0.800	0.808	0.826	0.827	0.828	0.829	0.828
Population controls	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Age cohort controls			Yes***	Yes***	Yes***	Yes***	Yes***
Electoral controls, 1995					Yes		
Electoral controls, 1999						Yes	
Electoral controls, 2003							Yes

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets are adjusted by clusters within regions. Unit of observation is a city. When "Yes" is added to indiciate inclusion of a group of controls, a significance level is reported immediately after for this group of controls. Flexible controls for population (5th polynomial) are included in all specifications. Age cohort controls include the number of people aged 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50 and older years, in each city according to 2010 Russian Census. Electoral controls include vote for Yabloko party, Communist Party (KPRF), LDPR party, the ruling party (Our Home is Russia in 1995, Unity in 1999, United Russia in 2003), vote against all, and electoral turnout for a corresponding year.

Table A2. Sources of variables used in the analysis.

Variable	Description
Protest participation in December 2011	Number of people who participated in the protests against electoral fraud on the week of December 10-16, 2011. The first wave of massive protests happened on the first weekend after the Duma elections, which are December 10-11, 2011, and that is why we focused our analysis on that week. The data was gathered manually from the open sources in the internet. Where possible, three numbers were collected - an estimate from Ministry of Internal Affairs, an estimate from activists, and an estimate from journalists. Whenever more than one number was obtained, we used the average.
Incidence of protests in December 2011	1 = at least one protest occurred in the city on the week of December 10-16, 2011; 0 = no protests that week
Protest participation in USSR in 1987-1992	Number of people who participated in protests in USSR in the period from 1987 to 1992. The data was taken from Mark Beissenger's website (http://www.princeton.edu/~mbeissin/research1.htm). This variable does not distinguish between different protest agendas, e.g. both pro-democratic and pro-communist protests count equally. For protests with more than one estimate of participation, an average number was taken. For cities with multiple protests during that period, we use median participation in our calculations.
Incidence of protests in USSR in 1987-1992	1 = at least one protest occurred in the city in 1987-1992; 0 = no protests occurred in 1987-1992

Participation in pro-democratic protests in USSR in 1987-1992	Number of people who participated in anti-soviet or pro-democratic protests in USSR in the period from 1987 to 1992. The data was taken from Mark Beissenger's website (http://www.princeton.edu/~mbeissin/research1.htm). We identified 75 various demands in the dataset which we considered either anti-soviet or pro-democratic. Examples would include demands like "Against Party Priveleges", "Decentralize Economic Administration", "Democratization of Political institutions", etc. Full list of anti-soviet/pro-democratic demands is available upon request. For protests with more than one estimate of participation, an average number was taken. For cities with multiple protests during that period, we use median participation in our calculations.
Incidence of pro-democratic protests in USSR in 1987-1992	1 = at least one anti-soviet or pro-democratic protest occurred in the city in 1987-1992; 0 = no anti-soviet or pro-democratic protests occurred in 1987-1992
Number of VK users in 2013	Number of all registered VK users in 2013, who was then living in a given city. Manually collected data.
Number of VK users in 2011	Number of valid and active VK users in 2011, who picked a given city as their hometown. "Valid" means not blocked, while "active" means they were online at least once during June 21 - July 7, 2011. Collected by a professional programmer - full description of the gathering process can be found at http://habrahabr.ru/post/123856/ (in Russian).
Number of early VK users, 11/2006	Number of VK users with id<5,000, who picked a given city as their hometown. In other words, those were the first 5,000 users ever registered in VK. They all were registered within less than a month, November 2006.
Number of Odnoklassniki users in 2014	Number of all registered Odnoklassniki users in 2014, who was then living in a given city. Manually collected data.
Number of Facebook users in 2013	Number of all registered Facebook users in 2013, who was then living in a given city. Manually collected data.
Population in 2001, in thousands	Collected from mojgorod.ru, which itself gathered data from Russian Federal State Statistics Service.

Distance to Saint Petersburg	Spherical distance from a given city to Saint Petersburg, in km
Distance to Moscow	Spherical distance from a given city to Moscow, in km
Administrative center	1 = city is the administrative center of its region; 0 = not. Data collected from Wikipedia.
Rayon center (county seat, dummy)	1 = city is the administrative center of its district (rayon); 0 = not. Data collected from Wikipedia.
Average wage in 2001, region-level	Collected from mojgorod.ru, which itself gathered data from Russian Federal State Statistics Service.
Number of people with age xx-xx in 2002	As of Russian census in 2002. Collected from Russian Federal State Statistics Service website, http://std.gmcrosstata.ru/webapi/opendatabase?id=vpn2002
Presence of university in a city	1 = city has at least one university; 0 = not. Data collected from Wikipedia.
Percentage with higher education in 2002	Percentage of adults with higher education or better. As of Russian census in 2002. Collected from Russian Federal State Statistics Service website, http://std.gmcrosstata.ru/webapi/opendatabase?id=vpn2002
Internet penetration in 2011, region-level	Number of unique users from Russia divided by population according to the 2010 census. Data collected from liveinternet.com
Number of SPbSU students, same 5-year cohort as VK founder	Manually collected from ok.ru, as the number of Odnoklassiniki users who studied in Saint Petersburg State University, and is in the same age 5-year cohort as Pavel Durov, former CEO of VK: 28-32 years.
Number of SPbSU students, one cohort younger than VK founder	Manually collected from ok.ru, as the number of Odnoklassiniki users who studied in Saint Petersburg State University, and is one 5-year cohort younger than Pavel Durov, former CEO of VK: 23-27 years.
Number of SPbSU students, one cohort older than VK founder	Manually collected from ok.ru, as the number of Odnoklassiniki users who studied in Saint Petersburg State University, and is one 5-year cohort older than Pavel Durov, former CEO of VK: 33-37 years.