# Social Media and Protest Participation: Evidence from Russia*

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

April 2017

**Abstract**

Do new communication technologies, such as social media, reduce collective action problem? This paper provides evidence that penetration of VK, the dominant Russian online social network, affected protest activity during a wave of protests in Russia in 2011. As a source of exogenous variation in network penetration, we use information on the city of origin of the students who studied together with the founder of VK, controlling for the city of origin of the students who studied at the same university several years earlier or later. We find that a 10% increase in VK penetration increased the probability of a protest by 4.6%, and the number of protesters by 19%. At the same time, VK penetration increased pro-governmental support, with no evidence of increased polarization. Additional results suggest that social media has affected protest activity by reducing the costs of coordination, rather than by spreading information critical of the government. We find that cities with higher fractionalization of network users between VK and Facebook experienced fewer protests, and there is a critical mass of VK users necessary to jumpstart the protests. Finally, we provide suggestive evidence that municipalities with higher VK penetration received smaller transfers from the central government after the occurrence of protests.

*We are grateful to Sergey Chernov, Nikolai Klemashev, Aleksander Malairev, Natalya Naumenko, and Alexey Romanov for invaluable help with data collection, and to Tatiana Tsygankova and Aniket Panjwani for editorial help in preparing the manuscript. We thank the Center for the Study of New Media and Society for financial and organizational support. Ruben Enikolopov and Maria Petrova acknowledge financial support from the Spanish Ministry of Economy and Competitiveness (Grant BFU2011-12345) and the Ministry of Education and Science of the Russian Federation (Grant No. 14.U04.31.0002). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 638221). We are indebted to Daron Acemoglu, Sinan Aral, Lori Beaman, Matt Gentzkow, Sam Greene, Kosuke Imai, Kirabo Jackson, John Londregan, Eliana La Ferrara, Monica Martinez-Bravo, Samuel Norris, Ricardo Perez-Truglia, Gautam Rao, Tom Romer, Jake Shapiro, Jesse Shapiro, Gaurav Sood, Erik Snowberg, David Strömberg, Adam Szeidl, Josh Tucker, Glen Weyl, Noam Yuchtman, Katia Zhuravskaya, and seminar participants at Higher School of Economics, Central European University, Berkeley, Bocconi, CEMFI, CREI, Hebrew, Mannheim, Microsoft Research, Princeton University, Universitat Pompeu Fabra, Northwestern University, New York University, NBER Digitization Meeting, 11th Workshop in Media Economics in Tel Aviv, 6th Workshop in Applied Economics in Petralia, "Social Media and Political Participation" conference in Florence, "Social Media and Social Movements" conference in St Petersburg for, and Political Economy Conference in Vancouver for helpful discussions.

# 1 Introduction

Collective action problems have traditionally been seen as one of the major barriers to achieving socially beneficial outcomes (e.g., Olson, 1965; Hardin, 1982; Ostrom, 1990). People's ability to overcome collective action problems depends on their information environment and on their ability to communicate to each other. New horizontal information exchange technologies, such as Facebook, Twitter, and other social media allow users to converse directly without intermediaries at a very low cost, thus potentially enhancing the spread of information and weakening the obstacles to coordination. However, so far there has been no systematic evidence on whether social media indeed improves people's ability to overcome collective action problems. Our paper fills this gap in the literature by looking at the effect that the most popular online social network in Russia had on a particular form of collective action – political protests.

The rise in the usage of social media in recent years coincided with waves of political protests around the world. But did social media play any role in political participation, i.e. inciting the protests, or did its content just reflect the preferences of the population?[1] Recent theoretical works argue that social media may indeed increase the probability of political protests occurring (Edmond, 2013; Little, 2015; Barberà and Jackson, 2016). However, testing this hypothesis is methodologically challenging since social media usage is endogenous to individual and community characteristics. In addition, protests are typically concentrated in one or a few primary locations, as was the case for Tahrir Square in Egypt or Maidan in Ukraine. Hence, geographic variation in protests is often limited. Temporal variation in protest intensity can provide evidence on the association between activity in social media and subsequent protests (Acemoglu, Hassan, and Tahoun, 2017),[2] but it does not allow studying the causal effect of the availability of social media.

To understand whether social media indeed promotes protest participation, we study an unexpected wave of political protests in Russia triggered by electoral fraud in the December 2011 parliamentary elections.[3] Our empirical setting allows us to overcome the problems of previous studies for two reasons. First, there was substantial geographic and time variation in the protest activities and in the penetration of major online social networks across Russian cities. E.g., among 625 cities in our sample, 133 witnessed at least one protest demonstration after the December 2011

[1]While not based on systematic empirical evidence, previous popular and academic literature disagreed even about the direction of the potential effect of social media on protests. Some have argued that the effect must be positive, as social media promotes cooperation (Shirky, 2008), fosters a new generation of people critical of autocratic leaders (Lynch, 2011), and increases the international visibility of protests (Aday et al., 2010, 2012). Others, however, have noted that social media is either irrelevant or even helps to sustain authoritarian regimes by crowding out offline actions (Gladwell, 2010), allowing governments to better monitor and control dissent (Morozov, 2011), and spread misinformation (Esfandiari, 2010).

[2]See also Hassanpour (2014) and Tufekci and Wilson (2012) for survey-based evidence on temporal variation in protests in Egypt.

[3]Electoral fraud was documented, for instance, in Enikolopov, Korovkin, Petrova, Sonin, and Zakharov (2013) and Klimek, Yegorov, Hanel, and Thurner (2012).

elections. Second, particularities of the development of VKontakte (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.

Our identification is based on the information about the early stages of the development of VK. It was launched by Pavel Durov in October 2006, the same year he graduated from Saint Petersburg State University (SPbSU). Upon VK's creation, Durov issued an open invitation on an SPbSU online forum for students to apply for membership on VK. Interested students then requested access to VK, and Durov personally approved all accounts. Thus, the first users of the network were primarily students who studied at Saint Petersburg State University together with Durov. This, in turn, made their friends and relatives at home more likely to open an account, which sped up the development of VK in these cities. Network externalities magnified these effects and, as a result, the distribution of the home cities of Durov's classmates had a long-lasting effect on VK penetration. In particular, we find that the distribution of the home cities of the students who studied at SPbSU at the same time as Durov predicts the penetration of VK across cities in 2011, but the distribution of the home cities of the students who studied at SPbSU several years earlier and later does not.

We exploit this feature of the VK development in our empirical analysis by using the origin of the students who studied at SPbSU in the same five-year cohort as the VK founder as an instrument for VK penetration in summer 2011, controlling for the origin of the students who studied at SPbSU several years earlier and later. Thus, our identification is based on the assumption that temporal fluctuations in the number of students coming to SPbSU from different cities were quasi-random and were not related to unobserved city characteristics.

Using this instrument, we then test the impact of VK penetration on the incidence of protests and protest participation. In the reduced form analysis, we find that the number of students from a city in the VK founder cohort had a positive and significant effect on protest participation while there was no such effect for the number of 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 4.6 percentage points higher probability of having a protest and a 19% increase in the number of protest participants . These results indicate that VK penetration indeed had a causal impact on protest participation in Russian cities in December 2011.

We also study the impact of VK on electoral outcomes. If the effect of social media was driven by provision of information critical of the government, we would expect to see negative effect of social media on the support for the government. However, we did not find any evidence of overwhelmingly negative content in social media posts. Moreover, we find that higher VK penetration led to higher, not lower, pro-governmental vote shares in the presidential elections of 2008 and 2012 and in the parliamentary elections of 2011. We find similar results using data from a large-scale survey conducted weeks before the 2011 elections.

Respondents in cities with higher VK penetration expressed greater support for the government and were more likely to say that they are planning to vote for the pro-governmental party. There was no evidence of an increased disapproval of the government or of an increased support for the opposition, so we find no indication of a polarizing effect of social media. Moreover, respondents in cities with higher VK penetration were less likely to say that they are ready to participate in political protests. A potential explanation for these results is the prevalence of pro-government content posted in VK before the elections. In any case, these results indicate that social media has not increased the number of people dissatisfied with the government, at least before the 2011 elections, in contrast to common perception that social media sure erode the support of autocratic leaders.

We perform a number of placebo tests to ensure that our results are not driven by unobserved heterogeneity. First, we show that VK penetration in 2011 was not related to protest participation in the same cities before the creation of VK using three different protest measures: anti-government protests in the end of the Soviet Union (1987-1992), labor protests in 1997-2002, and social protests in 2005. Second, we show that VK penetration in 2011 was not related to voting outcomes before the creation of VK. These findings suggest that our results are not driven by time-invariant unobserved characteristics of the cities that affect protest activity or political preferences. We also replicate our first stage regressions using information on the cities of origin of the students who studied in more than 60 other Russian universities of comparable quality. We find that the coefficient for our instrument – VK founder's cohort at SPbSU – lies at the top end of the distribution of the coefficients for the same cohort in other universities while the coefficients for younger and older cohorts lie close to the medians of the corresponding distributions, which is consistent with our identifying assumptions. The tests in the spirit of Antonji, Taber, and Elder (2005) and Oster (2016) also indicate that unobservables that are positively correlated with observables do not drive our results.

Next, we explore potential mechanisms behind the effect. Social media can have an impact on protests through the information channel or through the collective action channel. The information channel reflects the fact that online social media can serve as an important source of information on the fundamental issues that cause protests (e.g. the quality of the government or electoral fraud). This effect is likely to be especially strong in countries with government-controlled traditional media, such as Russia. The collective action channel relies on the fact that users can not only consume, but also exchange information. In particular, social media may both allow users to coordinate the logistics of the protests (coordination) and also introduce social motivation if users and their online friends openly announce that they are joining the protest (social pressure). Thus, the information channel increases the number of people dissatisfied with the regime, whereas collective action channel increases the probability that dissatisfied people participate in protests.

There is an important difference between the roles social media plays in the two channels. Social media is only relevant in the information channel to the extent that it allows for freer protest related content provision than in state-controlled media. Thus, in principle, any free traditional media could have played the same role. However, the role of social media in the collective action channel reflects an inherent distinction between social media and other forms of media, in that social media can facilitate horizontal flows of information between the users. Note that our results for VK penetration and support of the government speak, to some extent, against the information channel.

We also show that fractionalization of users between VK and Facebook,[4] conditional on the total number of users in the two networks, had a negative impact on protest participation, though this effect becomes significant only for larger cities with population over 100,000. This finding is consistent with the collective action channel, which requires users to be in the same network, but not with the information channel, as information about electoral fraud was widely discussed in both networks. Taken together, these results are consistent with the reduction of the costs of collective action being an important mechanism of social media influence.

Overall, our results indicate that social media penetration facilitates participation in political protests, and the reduction in the costs of collective action is a mechanism behind this effect. The positive impact of social media penetration on collective action has been predicted by theoretical literature (e.g., Edmond, 2013; Little, 2015; Barberà and Jackson, 2016) and widely discussed in the popular press (e.g., Shirky, 2011), but so far there has been no systematic empirical evidence to support this prediction. Our results imply that the availability of social media may have important consequences as political protests can affect within-regime power-sharing agreements, as well as related economic and political outcomes (Madestam, Shoag, Veuger, and Yanagizawa-Drott, 2013; Aidt and Franck, 2015; Battaglini, 2017; Passarelli and Tabellini, 2017). A broader implication of our results is that social media has the potential to reduce the costs of collective action in various circumstances.

More generally, our paper speaks to the importance of horizontal information exchange on the ability of people to overcome collective action problems. Information technologies affect collective action potential by changing the opportunities for such exchange. In the past, technologies such as leaflets, telephones, or even coffeehouses (Pendergrast, 2010) were used to facilitate horizontal information flows. Our results imply that social media, as a new technology in this line, also promotes collective action potential by dramatically increasing the scale of horizontal information exchange. Development of this new technology can have far-reaching implications since collective

[4] We define fractionalization as the probability that two randomly picked social media users belong to different networks. 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.

action problems have traditionally been seen as one of the major barriers to achieving socially beneficial outcomes (e.g., Olson, 1965; Hardin, 1982; Ostrom, 1990).

Our paper is closely related to Acemoglu, Hassan, and Tahoun (2017) 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. They also show that the activity in Twitter preceded the protest activity on Tahrir Square, but did not have an independent impact on abnormal returns of connected companies. 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 looking at the changes in activity in already existing social media accounts over time. 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 the protests occurring, i.e. at the extensive margin of the effect. Finally, our results shed some light on potential mechanisms behind the impact of social media.

Our paper is also related to the literature on the impact of information and communication technologies (ICTs) and traditional media on political preferences and policy outcomes. A number of recent works identify the impact of broadband penetration on economic growth (e.g., Czernich, Falck, Kretschmer, and Woessmann, 2011), voting behavior (Falck, Gold, and Heblich, 2014; Campante, Durante, and Sobbrio, 2014), sexual crime rates (Bhuller, Havnes, Leuven, and Mogstad, 2013), and policy outcomes (Gavazza, Nardotto, and Valletti, 2015). However, these papers do not provide specific evidence about whether this effect is due to the accessibility of online newspapers, search engines, email, Skype communications, or social media.[5]

There are recent papers that study the association between social media usage and collective action outcomes. Qin, Strömberg, and Wu (2017) analyze the content of posts in Chinese microblogging platform Sina Weibo and show that Sina Weibo penetration was associated with the incidence of collective action events, without interpreting these results causally. Steinert-Threlkeld, Mocanu, Vespignani, and Fowler (2015) show that the content of Twitter messages was associated with subsequent protests in the Middle East and North Africa countries during the Arab Spring. Hendel, Lach, and Spiegel (2015) provide a detailed case study of a successful consumer boycott organized on Facebook.[6]

[5]There are also papers that study the impact of cellphone penetration on price arbitrage (Jensen, 2007) and civil conflict (Pierskalla and Hollenbach, 2013). In a similar vein, Manacorda and Tesei (2016) look at the impact of cellphone penetration on political mobilization and protest activity in Africa.

[6]Papers that are less directly related to collective action include Bond et al. (2012) who show that that political mobilization messages in Facebook increased turnout in the U.S. elections, Qin (2013) who shows that the spread of Sina Weibo led to improvement in drug quality in China, and Enikolopov, Petrova, and Sonin (2017b) who show that anti-corruption blog posts by a popular Russian civic activist had a negative impact on market returns of targeted companies and led to a subsequent improvement in corporate governance.

Recent works have also shown that traditional media has an impact on voting behavior (DellaVigna and Kaplan, 2007; Enikolopov, Petrova, and Zhuravskaya, 2011; Gentzkow, Shapiro, and Sinkinson, 2011; Gentzkow, Petek, Shapiro, and Sinkinson, 2015a; Chiang and Knight, 2011), violence and ethnic tensions (Yanagizawa-Drott, 2014; DellaVigna, Enikolopov, Mironova, Petrova, and Zhuravskaya, 2014; Adena, Enikolopov, Petrova, Santarosa, and Zhuravskaya, 2015), and policy outcomes (Strömberg, 2004; Eisensee and Strömberg, 2007; Snyder and Strömberg, 2010). A number of papers also study ideological segregation online (Gentzkow and Shapiro, 2011; Halberstam and Knight, 2016; Gentzkow, Shapiro, and Taddy, 2015b). Our paper is also related to the literature on the impact of technology adoption (e.g., Dittmar, 2011; Cantoni and Yuchtman, 2014).

The rest of the paper is organized as follows. Section 2 provides background information about the environment that we study. Section 3 presents a theoretical framework 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

## 2.1 Internet and Social Media in Russia

By 2011, approximately half of the Russian population had Internet access at home,[7] which made Russia the largest Internet market in Europe, accounting for about 15% of all European Internet users.[8] Although more than 80 countries enjoyed a higher Internet penetration rate than Russia at the time, Russia started catching up rapidly demonstrating a 23% average yearly growth rate between 2007-2011.

Social media was already popular in Russia by 2011. On average, Russians were spending 9.8 hours per month on social media websites in 2010 – more than any other nation in the world.[9] Social media penetration in Russia was comparable to that of the most developed European countries, with 88% of Russian Internet users having at least one social media account, compared to, for instance, 93% in Italy and 91% in Germany. Although Russians lost the title of most social-media-addicted nation to Israel in October 2011, they remained third with 10.4 hours per user.[10]

Despite the increasing popularity of social media, Russia remains one of the very few markets where Facebook is not dominant. In fact, there is only one other country where Facebook could not secure even the second largest share of the market for reasons other than censorship -– South Korea. Instead, homegrown networks VKontakte (VK) and Odnoklassniki were able to quickly

[7] http://bit.ly/2pilVDs
[8] According to comScore http://bit.ly/2oTnmfp
[9] According to comScore http://bit.ly/2oPqRDP
[10] http://bit.ly/2ofKbZf

take over the Russian social media market. As of August 2011, VK had the largest daily audience at 23.4m unique visitors (54.2% of the online population in Russia); Odnoklassniki was in second with 16.5m unique visitors (38.1%), leaving Facebook trailing in third place with 10.7m unique visitors (24.7%).[11]

Such an unusual pattern of market shares emerged because of the relatively late market entry by Facebook. By the time Facebook introduced the Russian language interface in mid-2008, both VK and Odnoklassniki already accumulated close to 20m registered users.[12] Besides, VK and Odnoklassniki could offer certain services that Facebook could not, either due to legal reasons (e.g. Facebook could not provide music and video streaming services because of copyright issues) or because of a different marketing strategy (e.g. users were attracted by a lower amount of advertising in both VK and Odnoklassniki).

VK started off as a student- and youth-oriented website. ”VKontakte” translates to ”in contact”, and the original mission of VK was to help current students stay in touch later in life, with its target audience similar to that of Facebook. As a result, VK was more widespread in large cities than Odnoklassniki.[13] In contrast, Facebook gained popularity among those who wanted to communicate with their foreign friends, and thus had a higher market share in large cities, especially in Moscow and St. Petersburg.[14]

As of December 2011, the Internet in general – social media in particular – enjoyed relative freedom in Russia, as there were no serious attempts to control online content up until 2012. Centralized censorship and content manipulation in social media began after the period we focus on and, to a large extent, were consequences of the protests examined in this paper. The relative freedom made social media websites an important channel for transmitting information and enhancing political debate, taking this role away from Russian TV and major newspapers.[15]

## 2.2 History of VK

VK is a social media website very similar to Facebook in its functionality. The user can create an individual profile, add friends, converse with them, create events, write blog posts, share information (in audio and video format as well), etc. VK was launched in October 2006. The core of the VK development team was more or less stable until 2012, consisting of Pavel Durov (philology major at SPbSU at the time), his brother Nikolai Durov (physics graduate student at SPbSU at the

[11] http://bit.ly/2nRJlif

[12] https://vk.com/blog?id=92 and http://bit.ly/2oTDIoi

[13] The original mission of Odnoklassniki was to help people find their former classmates and friends from their past, so the targeted different audience was older, on average. According to a marketing study performed in 2010, the average age of VK users was 3 years less than the average age of Odnoklassniki users (http://bit.ly/2nRL5b1)

[14] http://bit.ly/2oToKi7

[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.

time, winner of the world programming and math contests), and their fellow students. Upon VK's creation, Durov issued an open invitation on an SPbSU online forum for students to apply for membership on VK. Interested students then requested access to VK, and Durov personally approved all accounts. Registration in VK opened to the general public at the end of November 2006. Shortly after, the number of users skyrocketed from 5 thousand users to 50 thousand in January 2007, to 3 million in November 2007, to 100 million in November 2010 (see Figure A1 in Online Appendix). By early 2008, VK became the most visited website in Russia.

VK creators held a strong position against any forms of censorship. During the protests of 2011-2012 Pavel Durov was approached by the Federal Security Service (FSB) and was asked to start blocking the opposition-minded online communities, as well as protest events, some of which had more than 30,000 subscribers (Kononov, 2012). Durov refused, arguing that it would lead to a large number of people switching to VK's foreign competitors, such as Facebook. Some ascribed his actions to his libertarian views. VK policies regarding freedom of speech remained unchanged until Durov was forced to sell his share of VK and lost control of the firm in 2014.[16] Note that Durov himself, at least before 2013, was not involved in any political activity, and, in particular, did not advertise or create any politically related content in VK (Kononov, 2012).

## 2.3 Protest Movement of 2011-2012

A wave of protest demonstrations in 2011-2012 was triggered by electoral fraud in the parliamentary elections of December 2011. It was the first large-scale political protest movement in Russia since the collapse of the Soviet Union. Similar to other protest events in authoritarian countries, Russian protests of December 2011 surprised everyone, including their leaders (Kuran, 1991).[17] Russian society was politically inactive in the 2000s, with rapid economic growth softening any criticism of Putin's regime. For instance, electoral fraud of allegedly similar magnitude in 2007 parliamentary elections did not trigger any serious protests (Treisman, 2011). In addition, traditional media has been under heavy government control in 2007-2012 (and beyond), so there was hardly any chance information on electoral fraud could be transmitted through the main TV channels. The latter served as the primary source of information for nearly 80% of Russians at the time, ensuring steady control over the information flows.[18]

Parliamentary elections were held on December 4, 2011. During the course of that day, reports

[16] Durov was dismissed as the VK CEO after a similar incident two years later, in September 2014, when he refused to block groups and accounts of Ukrainian revolutionaries. He was forced to sell his shares of VK to Mail.ru earlier that year. He left VK for his new start-up Telegram. He left the country too, after obtaining Saint Kitts and Nevis citizenship.

[17] For instance, a day before the first protest gathered over 5,000 participants its organizers were debating whether a threshold of 500 people would be surpassed.

[18] For instance, see Levada survey in 2011 (http://bit.ly/2nv4Nyb, p. 135) or VTSIOM study in 2011 (http://bit.ly/2on8h4Z)

of electoral fraud were quickly growing in numbers, documented both by independent observers and by regular voters. In a vast majority of cases, electoral fraud favored the incumbent party United Russia. Videos of ballot staffing and 'carousel' voting (i.e. same voters voting multiple times on different poll stations) started to circulate in the Web and in social media. Startling differences between exit polls and official results began to emerge; some exit polls reported 23.6% of the votes going to United Russia in Moscow, which was 20% lower relative to the official result.[19] Scholars later confirmed that the amount of fraud was sizable using statistical analysis. For instance, Enikolopov et al. (2013) showed that the presence of randomly assigned independent observers decreased United Russia's results by 11% on average (from 47% to 36%). Clear evidence of electoral fraud, together with the absence of its acknowledgment by the government, became a source of outrage for thousands of people and urged some of them to take to the streets.

On December 5, 2011, five to six thousand people appeared at a rally in the center of Moscow. 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 a precedent for the future, more massive ones. The next anti-fraud rallies were held on December 10 and 24, and had a record attendance both in Moscow (near 100,000 participants on both dates) and across the country (more than 100 cities participated). The subsequent waves of protests were less popular and involved fewer cities. Moscow and St. Petersburg, however, still hosted major rallies almost each month. The tipping point of the movement was reached on May 6, 2012, a few days before Vladimir Putin's inauguration as a President. Whereas all previous demonstrations were peaceful and non-violent, the rally on May 6 broke out in a number of serious clashes with the police forces. Within a few days, more than 30 activists were charged with allegedly inciting mass riots and using violence against the police, many then faced 3-4 years in prison. This trial, together with the absence of tangible achievements, marked the decline of protests as a form of political struggle in Russia.

## 2.4 VK and Protest Activity

In 2011, online social networks, including VK, became an important source of information in Russia, where traditional media was largely controlled by the state. Reports of electoral fraud, for example, were widely available online, often accompanied by pictures and YouTube videos. Most of the traditional media, however, did not cover the topic. Robertson (2015) shows that VK users were more likely to be aware of the activities of Golos, the most prominent electoral monitoring organization in Russia, as compared with non-VK users. Reuter and Szakonyi (2015) show that being a user of one of the online social networks was a strong predictor of respondent's being aware of the electoral fraud. Based on an online survey of protest participants, Dokuka (2014)

[19] Note that these exit poll results were later deleted from the corresponding polling agency's websites.

provides evidence that 67% of them learned about the upcoming protests from VK while another 22% obtained this information from other online social networks or online newspapers.

VK was also widely used for coordinating protest activities. VK users utilized open online protest communities for sharing information about protest demonstrations in their cities and coordinating organizational details. As most of the user profiles in VK, these communities were open, and anyone with an account in VK could see all the content posted in these communities.[20] According to our data, out of 133 cities that had protests, 87 had VK communities or events created with the purpose of organizing protest demonstrations after the December 2011 parliamentary elections. Most of these communities were created within the first several days after the parliamentary elections.[21]

# 3 Data

We use several sources of data. Our sample consists of 625 Russian cities with populations over 20,000 according to the 2010 Census. We exclude Moscow and Saint Petersburg from our sample as outliers.

To measure the penetration of VK across the cities, we collect information about the city of residence for all the users of VK with public accounts who joined VK before the autumn of 2011.[22] Based on this information, we compute the number of users in each city as of the end of summer of 2011, i.e. before the parliamentary elections were scheduled and before the electoral campaign began. More details about data collection are available in the Online Appendix.

We use hand-collected data on political protests that occurred between December 2011 and May 2012. When the protests began in December 2011, we started monitoring newspaper databases and online resources to record information about political protests in any Russian city mentioned in this context. The monitoring was repeated every week until the protests subsided in summer 2012. For each event, we recorded the number of protesters, as reported by three alternative sources: i) the police; ii) organizers of the protest; ii) a news source that wrote about the protest.[23] As a result of this monitoring, we have collected a comprehensive city-level database on political protests in Russia in 2011-2012. We aggregate this information to city-week level by constructing

[20]Note that these exit poll results were later deleted from the corresponding polling agency's websites.

[21]Protest communities were identified by searching for several standard keywords (e.g. "For Fair Elections") in the names of these communities, so it is possible that we underestimate the number of cities with online protest communities.

[22]Public accounts contain some basic information on the VK users, e.g. their home city, that is available to anyone on the Internet. The timing of account creation can be inferred from the account ID. Note that more that 90% of the accounts in VK are public.

[23]We have data on all the three estimates in 9.5% of the cases. Only one estimate is available in 64% of the cases. As a result, we primarily use the estimates reported by journalists in various news sources.

two variables: an indicator for the existence of a protest in a given city in a given week and the number of protesters, computed by taking the average number of protesters as reported by the police, organizers, and the news source.[24] If there were more than one protest event in a city during the same week, we take the number of protesters at the biggest event.

We use information on the city of origin of the students who studied at Saint Petersburg State University and other top Russian universities.[25] Since, unfortunately, administrative records on the admitted students are not available, this data is based on the information on the year of birth, university attended, and years of study provided in public accounts of the Odnoklassniki users. Note that as of 2014, when this data was collected, 80% of Russian adult population reported having an account in Odnoklassniki,[26] so the coverage of our data is reasonably large. More specifically, for each university in the sample, we calculate the number of students coming from each city in five-year cohorts. We mostly focus on three cohorts in our analysis: i) those who were born the same year as the VK founder or within two years from his birthday, either earlier or later; ii) those who were born from three to seven years earlier than the VK founder; iii) those who were born from three to seven years later than the VK founder.[27] Although using data from social media to measure the distribution of students across cities may introduce a measurement bias, the identifying assumption is that, while controlling for the number of Odnoklassniki users, this bias does not vary across cohorts in a way that is correlated with the outcomes of interest. Later on, we use various tests to provide evidence that this assumption holds.

Next, we use data on the number of Facebook users by city in 2011 and 2013. The data on Facebook penetration in 2011 was taken from Nikolai Belousov's blog.[28] The data on Facebook penetration in 2013 was collected manually for each city in the sample.[29]

We use three different sources of data for protests that occurred prior to the advent of social media. The data on protests in the late Soviet Union comes from Beissinger (2002). In the analysis, we first look at all the protests as a whole and then at the pro-democracy protests separately. The data on participants in the labor protests of 1997-2002 comes from Robertson (2011). Finally, we use

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

[25]In particular, we take all the universities from Moscow and Saint Petersburg in the top-100 Russian universities, as well as top 20 universities from other cities. To identify the elite schools, we use the university ranking of the RA Expert agency for 2014 (http://bit.ly/2ofLYgU).

[26]http://bit.ly/2nv9w2C

[27]Our results remain very similar if we use students' years of entrance to the university instead of their year of birth. We use the years of birth in the benchmark specifications, since these measures yield a stronger first stage in the IV regression (see below), which is likely to reflect the fact that age difference had an additional effect on the probability of joining VK. I.e., a 30-year-old graduate student was plausibly less likely to adopt VK than a 20-year-old undergraduate student, conditional on both studying at SPbSU with Durov.Our results are also robust to changes in the cohort definition and to inclusion of a larger number of cohorts in the model (see the robustness check section).

[28]http://bit.ly/2oWNTpg

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

information on the social protests of 2005 from the website of a communist organization,[30] though we understand that this data source is probably less reliable than the ones mentioned previously. For all three sources, we aggregate information at the city level and exploit two different measures: the maximum number of protesters in a city and an indicator for at least one protest in a city.

The data on electoral outcomes is coming from the Central Election Commission of the Russian Federation. We obtained the public opinion data from the MegaFOM opinion poll conducted by the Public Opinion Foundation (Fond Obschestvennogo Mneniya, or FOM) in October-November 2011.[31] This is a regionally representative survey of 56,900 respondents in 79 regions, of which 30,669 respondents come from 519 cities in our sample.

The city-level data on population, age, education, and ethnic composition comes from the Russian Censuses of 2002 and 2010. The data on the average wage and municipal budgets comes from the municipal statistics of Rosstat, the Russian Statistical Agency. Additional city characteristics (latitude, longitude, year of city foundation, and the location of administrative centers) come from the Big Russian Encyclopedia. Summary statistics for each variable employed in the analysis are presented in Table A1.

# 4 Theoretical Framework

In this section, we aim to construct a common theoretical framework to think about all the outcomes that we study. In particular, we seek to understand how better precision of information about the quality of government and/or about tactics of protests can eventually affect the number of people choosing the behavior of interest (voting or protest participation). To do that, we extend the model of Little (2015), who analyze how communication technology can be related to protest participation. Finally, we investigate if we should expect critical mass to matter for the relationship between VK penetration and protests.

## 4.1 Voting in Autocracy

We study how communication technology can affect the support of government in an autocracy. In contrast to the theories of voting in democracy, we consider what happens if a citizen decides about the expression of support of a single ruling party, rather than compare utilities from different candidates, as we think that this setup better matches the reality of quasi-authoritarian elections in Russia.

There is a continuum of risk-neutral citizens, $i \in [0,1]$. The nature draws a common prior belief

[30] http://trudoros.narod.ru/

[31] We are grateful to the president of FOM, Alexander Oslon, for generously sharing the data.

about the regime popularity, $\omega$, which is distributed as $N(0, 1/\alpha_0)$. Then the public signal about $\omega$ is drawn, $s_\omega \sim N(\mu_s, 1/\alpha_s)$. A cost of voting is drawn as $c_i \sim N(\mu_c, \sigma_c^2)$. Each citizen then maximizes:

$$u_v(v_i) = v_i[\bar{\omega} + \lambda_v V - c_i]$$

, where $v_i$ is a voting decision indicator ($= 1$ if $i$ votes for an autocrat, $= 0$ if not), $\lambda_v \geq 0$ is a taste-for-conformity parameter, and $V$ is the proportion of citizens who voted for an autocrat. In this version of the model, all citizens update their priors about the regime in the same direction, denoted by the fact that $\bar{\omega}$ does not vary with $i$. However, the cost $c_i$ is drawn at random for each individual, so their decisions to vote or not to vote will still differ.

Citizens' belief about the regime's popularity $\omega$ upon observing the public signal $s_\omega$ is:

$$\bar{\omega} = \mathbb{E}[\omega + s_\omega] = \frac{s_\omega \alpha_s}{\alpha_0 + \alpha_s}$$

Having updated their beliefs about the regime, citizens compare the estimated benefits of voting with their individual costs $c_i$. Note that citizens with $c_i < s_\omega \alpha_s/(\alpha_0 + \alpha_s)$ would vote no matter what others do, while citizens with $c_i > s_\omega \alpha_s/(\alpha_0 + \alpha_s) + \lambda_v$ are not going to vote no matter what. To identify the voting cut-off $\hat{c}$, we need to calculate the expected size of a protest given the public signal $s_\omega$:

$$\mathbb{E}[V|s_\omega] = Pr[c_i \leq \hat{c}(s_\omega)|s_\omega] = Pr\left[\frac{c_i - \mu_c}{\sigma_c} \leq \frac{\hat{c}(s_\omega) - \mu_c}{\sigma_c}\right] = \Phi\left[\frac{1}{\sigma_c}(\hat{c}(s_\omega) - \mu_c)\right] \tag{1}$$

Hence, the cut-off level $\hat{c}$ is determined by the following equation:[32]

$$\frac{s_\omega \alpha_s}{\alpha_0 + \alpha_s} + \lambda_v \Phi\left[\frac{1}{\sigma_c}(\hat{c}(s_\omega) - \mu_c)\right] = \hat{c}(s_\omega) \tag{2}$$

The comparative statics of the cost cut-off w.r.t. $s_\omega$ and $\alpha_s$ are as follows:

$$\frac{\partial \hat{c}}{\partial s_\omega} = \frac{\frac{\alpha_s}{\alpha_0 + \alpha_s}}{1 - \frac{\lambda_v}{\sigma_c}\phi(\frac{1}{\sigma_c}(\hat{c} - \mu_c))} \tag{3}$$

$$\frac{\partial \hat{c}}{\partial \alpha_s} = \frac{\frac{s_\omega \alpha_0}{(\alpha_0 + \alpha_s)^2}}{1 - \frac{\lambda_v}{\sigma_c}\phi(\frac{1}{\sigma_c}(\hat{c} - \mu_c))} \tag{4}$$

We assume that the taste-for-conformity parameter $\lambda_v$ is sufficiently small so that the denominator in both of the above fractions is positive. This is necessary for a meaningful equilibrium in which

[32] Note that in the case of $\lambda_v = 0$ the solution becomes a simple Bayesian updating.

a positive public signal about an autocrat increase the amounts of votes in favor of the regime, i.e. $\partial \hat{c}/\partial s_\omega > 0$. In this, social media increases support for an autocrat ($\partial \hat{c}/\partial \alpha_s > 0$) whenever public signal is favorable to the regime ($s_\omega > 0$) and decreases support ($\partial \hat{c}/\partial \alpha_s < 0$) whenever public signal is unfavorable ($s_\omega < 0$). Hence, we draw the following empirical prediction from this part of the model:

***Prediction 1.*** *Higher social media penetration, i.e. higher $\alpha_s$, leads to higher (lower) voting in favor of the ruling regime if the content of social media, i.e. public signal $s_\omega$, is, on average, positive (negative).*

## 4.2 Protests in Autocracy

There is a continuum of risk-neutral citizens, $i \in [0,1]$. First, nature draws common priors on regime popularity ($\omega$) and protest tactics ($\theta$). The common priors on $\omega$ and $\theta$ are distributed as $N(0,1/\alpha_0)$ and $N(0,1/\beta_0)$, respectively. Then the public signals are drawn, $s_\omega \sim N(\omega,1/\alpha_s)$ and $s_\theta \sim N(\theta,1/\beta_s)$. A random cost of protesting, which is separate from the costs of mismatching tactics, is drawn as $c_i \sim N(\mu_c,\sigma_c^2)$. Then each citizen maximizes:

$$u_p(p_i,t_i) = p_i[-\bar{\omega} + \lambda_p P - k(t_i-\theta)^2 - c_i]$$

, where $p_i$ is a protest decision ($=1$ if $i$ goes out to protest, $=0$ if not), $P$ is the proportion of citizens who turned out to protest, and $\lambda_p \geq 0$ is a social image parameter as in Enikolopov et al. (2017a),which could also bear other interpretations,.

Identical to the case of voting in the previous section, citizens update their beliefs about regime popularity $\omega$ upon observing the public signal $s_\omega$ as follows:

$$\bar{\omega} = \mathbb{E}[\omega + s_\omega] = \frac{s_\omega \alpha_s}{\alpha_0 + \alpha_s}$$

Similarly, upon observing signal $s_\theta$, citizens update their beliefs about the tactics of the upcoming protest:

$$\bar{\theta} = \mathbb{E}[\theta + s_\theta] = \frac{s_\theta \beta_s}{\beta_0 + \beta_s}$$

Since citizens would like to match the true $\theta$ as closely as possible, in optimum, they set $t_i = \bar{\theta}$. By definition, the expected level of discrepancy between $\bar{\theta}$ and $\theta$ is equal to the variance of $\bar{\theta}$, i.e., formally:

$$\mathbb{E}[k(\bar{\theta}-\theta)^2] = \frac{k}{\beta_0 + \beta_s}$$

Hence, each citizen decides on protest participation by weighing the following expected benefits and costs:

$$-\frac{s_\omega \alpha_s}{\alpha_0+\alpha_s}+\lambda_P \mathbb{E}[P|s_\omega, s_\theta] \lesseqgtr \frac{k}{\beta_0+\beta_s}+c_i \tag{5}$$

The only piece left to derive is the expected share of protesters. Again, we know for a fact that citizens with $c_i < -s_\omega \alpha_s/(\alpha_0+\alpha_s) - k/(\beta_0+\beta_s)$ are going to protest no matter what their compatriots do, while citizens with $c_i > \lambda_p - s_\omega \alpha_s/(\alpha_0+\alpha_s) - k/(\beta_0+\beta_s)$ will not participate no matter how many others do. We search for a cut-off level $\hat{c}$ that equalize the two sides of the equation (5) and will separate participants and non-participants. Similar to the previous section, such cut-off level $\hat{c}$ is determined by the following equation:

$$-\frac{s_\omega \alpha_s}{\alpha_0+\alpha_s}+\lambda_P \Phi\left[\frac{1}{\sigma_c}(\hat{c}(s_\omega, s_\theta)-\mu_c)\right]-\frac{k}{\beta_0+\beta_s}-\hat{c}(s_\omega, s_\theta)=0 \tag{6}$$

The comparative statics of the cut-off w.r.t. the public signal of regime strength ($s_\omega$) and social media (i.e. increased precision of public signals, $\alpha_s$ and $\beta_s$) are as follows:

$$\frac{\partial \hat{c}}{\partial s_\omega}=\frac{-\frac{\alpha_s}{\alpha_0+\alpha_s}}{1-\frac{\lambda_p}{\sigma_c}\phi(\frac{1}{\sigma_c}(\hat{c}-\mu_c))} \tag{7}$$

$$\frac{\partial \hat{c}}{\partial \beta_s}=\frac{\frac{k}{(\beta_0+\beta_s)^2}}{1-\frac{\lambda_p}{\sigma_c}\phi(\frac{1}{\sigma_c}(\hat{c}-\mu_c))} \tag{8}$$

$$\frac{\partial \hat{c}}{\partial \alpha_s}=\frac{-\frac{s_\omega \alpha_0}{(\alpha_0+\alpha_s)^2}}{1-\frac{\lambda_p}{\sigma_c}\phi(\frac{1}{\sigma_c}(\hat{c}-\mu_c))} \tag{9}$$

Assuming that $\lambda_p$ is sufficiently small so that the denominator of the above fractions is positive, we conclude that: (i) protest size decreases with a more favorable public signal about the regime (i.e. $\partial\hat{c}/\partial s_\omega < 0$), (ii) protest size increases when citizens receive a more precise signal about the protest tactics (i.e. $\partial\hat{c}/\partial \beta_s > 0$), and (iii) protest size increases when citizens receive a more precise signal about the regime conditional on the signal being negative (i.e. $\partial\hat{c}/\partial \alpha_s > 0$ if $s_\omega < 0$) and decreases with signal precision if the signal provides positive information about the regime (i.e. $\partial\hat{c}/\partial \alpha_s < 0$ if $s_\omega > 0$). We conclude by deriving the following empirical prediction from this analysis:

***Prediction 2.*** *Higher social media penetration (higher $\alpha_s$ and $\beta_s$) leads to higher protest participation against the ruling regime if the content of social media (public signal $s_\omega$) is, on average, negative. However, even when the content online is positive, social media could increase protest participation if the gains in coordination (higher $\beta_s$) are high enough.*

## 4.3 Social Media Penetration and the Critical Mass

Suppose all citizens fall into two categories - those who adopted social media and those who did not. The share of those who adopted social media is $m$. In this section, for illustrative purposes, we assume that the precision of the public signal about the regime, through any source of information, is the same for *all* citizens, including non-adopters. However, only adopters enjoy a higher precision of the tactics signal from social media, i.e. $\beta_s^a > \beta_s^n$ where $a$ denotes a social media adopter while $n$ indicates that a citizen did not adopt social media. We take the adoption decision as exogenous throughout this section.

Following the previous section, one can easily show that adopters and non-adopters would have different participation thresholds, $\hat{c}_a$ and $\hat{c}_n$, defined by the following pair of equations:

$$-\frac{s_\omega \alpha_s}{\alpha_0+\alpha_s}+\lambda_p P=\frac{k}{\beta_0+\beta_s^a}+\hat{c}^a \tag{10}$$

$$-\frac{s_\omega \alpha_s}{\alpha_0+\alpha_s}+\lambda_p P=\frac{k}{\beta_0+\beta_s^n}+\hat{c}^n \tag{11}$$

Note that now the share of protesters consists of two different types of participants - adopters and non-adopters:

$$P=mPr[c_i \leq \hat{c}^a|\bar{s}_\omega, \bar{s}_\theta^a]+(1-m)Pr[c_i \leq \hat{c}^n|\bar{s}_\omega, \bar{s}_\theta^n]$$

The goal of this section is to understand how protest participation changes with $m$ and whether, other things held constant, higher social media adoption could potentially trigger a protest after reaching a certain critical mass. To this end, we will study the first the cost thresholds, $\hat{c}^a$, and protest participation, $P$, to social media penetration. Subtracting equation (11) from equation (10), we get:

$$\hat{c}^a-\hat{c}^n=\frac{k(\beta_s^a-\beta_s^n)}{(\beta_0+\beta_s^a)(\beta_0+\beta_s^n)}=K>0$$

Note that $\hat{c}^a > \hat{c}^n$, meaning that the fraction of adopters who participate in protests is higher than the fraction of non-adopters who do. Expressing $\hat{c}^a$ in terms of $K$ and $\hat{c}^n$ and plugging in the result in (11), one gets:

$$-\frac{s_\omega \alpha_s}{\alpha_0+\alpha_s}+\lambda_p\left[m\Phi\left(\frac{1}{\sigma_c}(\hat{c}^n+K-\mu_c)\right)+(1-m)\Phi\left(\frac{1}{\sigma_c}(\hat{c}^n-\mu_c)\right)\right]=\frac{k}{\beta_0+\beta_s^n}+\hat{c}^n \tag{12}$$

For the ease of exposition, denote $\bar{c}^n=(\hat{c}^n+K-\mu_c)/\sigma_c$ and $\underline{c}^n=(\hat{c}^n-\mu_c)/\sigma_c$. Applying the implicit function theorem to equation (12), we derive the first derivative of the non-adopters participation, $\hat{c}^n$, w.r.t. social media penetration $m$:

$$\frac{\partial \hat{c}^n}{\partial m} = \frac{\lambda_p \left[\Phi(\bar{c}^n) - \Phi(\underline{c}_n)\right]}{1 - \frac{\lambda_p}{\sigma_c}\left[m\phi(\bar{c}^n) + (1-m)\phi(\underline{c}^n)\right]} > 0 \tag{13}$$

Hence, as the take-up of social media in population grows, non-adopters go out to protest with a higher probability. Note that, as a result, $P$ is also monotonically increasing with $m$:

$$\frac{\partial P}{\partial m} = \left(\Phi(\bar{c}^n) - \Phi(\underline{c}_n)\right) + \frac{\partial \hat{c}^n}{\partial m}\frac{1}{\sigma_c}\left(m\phi(\bar{c}^n) + (1-m)\phi(\underline{c}^n)\right) > 0$$

Assume now that, after citizens made their participation decisions, a protest gets organized only if the total share of citizens who would like to participate exceeds some threshold $P^*$.[33] Since the share of people willing to participate is monotonically increasing in $m$, there is a unique threshold of social media penetration $m^*$ such that, other parameters held equal, protests are organized in cities above this threshold and are not in cities below it. Hence, we conclude with the following empirical prediction:

***Prediction 3.*** *Higher rates of social media adoption (higher m) lead to higher protest participation (higher P). Moreover, if protests take place after a certain critical mass of potential participants is accumulated, we expect protests to occur only after social media penetration reaches a certain threshold.*

# 5 Identification Strategy

Our main hypothesis is that social media penetration (specifically, VK penetration) had a positive effect on political participation, whether it is protest participation, voting, or expressed support of the government. Thus, we estimate the following model:

$$Participation_i = \beta_0 + \beta_1 VKpenetration_i + \beta_2 \mathbf{X_i} + \varepsilon_\mathbf{i} \tag{14}$$

where $Participation_i$ is either a measure of protest activity, i.e. the logarithm of the number of protesters in city $i$ in the first weekend of the protests (December 10th and 11th) plus one or an indicator variable for the occurrence of at least one protest in city $i$ in the first weekend of the protests,[34] or support of the government, either through voting or expressed attitudes; $VKpenetration_i$ is the logarithm of the number of VK users in city $i$ in 2011; $X_i$ is a vector of control variables that includes a fifth-order polynomial of population, an indicator for being a regional or a subregional

[33] Such threshold behavior naturally arises if political protests are modeled in a more elaborate global game setting (e.g. as in Edmond, 2013).

[34] We focus on the first protests to avoid a possibility of dynamic effects within and across the cities.

(rayon) administrative center, average wage, number of city residents of different age cohorts, distance to Moscow and Saint Petersburg, an indicator for the presence of a university in a city, share of population with higher education in 2010 for each age cohort separately and the average share of population with higher education in 2002, ethnic fractionalization, and internet penetration. In some specifications, $X_i$ also includes the outcomes of the pre-2006 parliamentary elections to control for pre-existing political preferences of the population. Standard errors in all regressions are clustered at the regional level.

## 5.1 Identification Strategy

The OLS estimates of the equation (5) are likely to be biased, as the unobserved characteristics that make people more likely to become VK users can also make them more likely to participate in political activities. To address this issue, we use fluctuations in the origin of the students who have studied at SPbSU as a source of exogenous variation in VK penetration that does not have an independent effect on protest participation. In particular, we exploit the fact that the distribution of the home cities of the students who studied at SPbSU at the same time as the VK founder predicts the penetration of VK across cities in 2011, but the distribution of the home cities of the students who studied at SPbSU several years earlier and later does not. Specifically, we compute the number of students from each city in three five-year student cohorts (to match the Census definition of cohorts): (i) those who have studied at the same year as Durov, as well as one or two years earlier or later, (ii) those who studied from three to seven years earlier than Durov, and (iii) those who studied from three to seven years later than Durov.

The identifying assumption is that, conditional on population, education, and other observables, fluctuations of the student flows from different cities to Saint Petersburg State University in the 2000's are orthogonal to the unobserved determinants of protest participation.

Table A2 in the Online Appendix presents a full distribution of the SPbSU student cohorts by their home cities. Note that in all but one case the number of students is less than 40 students per home city, for all three cohorts.[35] Thus, the numbers are sufficiently small to allow for random fluctuations in the distribution of students across cities to happen.

Note that students were coming to study at Saint Petersburg State University from all over the country. These students arrived from 73 out of 79 Russian regions included in our study. Students in the Durov's cohort came from 237 different cities (more than a third of all Russian cities), while students from an older cohort came from 222 cities and students from a younger cohort came from 214 different cities.

[35] We also check that our results are robust to exclusion of cities with more than 10 students in the Durov's cohort.

## 5.2 Determinants of VK penetration

For our identification strategy to work, we first need to show that our instrument is relevant. Table 1 provides evidence on the determinants of VK penetration across Russian cities in 2011, and, in particular, on the effect of the number of Saint Petersburg State University students in different cohorts on VK adoption in their home cities. The results indicate that, once population controls are included, the five-year cohort of Pavel Durov, the VK founder, is positively and significantly (at 1% level) correlated with the subsequent VK penetration, in contrast to the younger and older cohorts, for which the corresponding coefficients are not significant. The coefficient for the number of SPbSU students in the Durov's cohort is stable across the specifications (2)-(7). In particular, it does not depend on the age and education distributions in a city, as we control for the number of people in each of the five-year age cohorts over 20 years of age, and for the education level in each of these cohorts. The magnitude of the effect implies that a 10% increase in the size of the VK founder's cohort coming from a given city leads to a 14% increase in the number of VK users in that city in 2011. The coefficient for the size of an older cohort is much smaller in magnitude, is not statistically significant across specifications (2)-(7). The coefficient for the size of a younger cohort is consistently negative significantly different from the effect of the Durov's cohort. These results are summarized in graphical form in Figure 1.

In addition, we provide evidence that the origin of the students in the Durov's cohort affects VK penetration in 2011 via its effect on early adoption of the network. We look at the determinants of VK penetration at the by-invitation-only stage, i.e. for the first 5,000 users (see Table A3). While the coefficient patterns for the number of SPbSU students are similar to those in Table 1, other controls, such as population, education by cohort, or ethnic fractionalization, become insignificant, consistent with our claim that initial VK penetration was largely idiosyncratic. The corresponding cohort coefficients, together with their confidence intervals, are shown graphically in Figure A2.

Overall, our results in Tables 1 and A3 suggest that the city differences in early VK penetration were, at least in part, generated by the year-to-year fluctuations in student flows from different cities and that these small initial differences in early adoption have had important long-term consequences for the later penetration of the social network. In the subsequent sections we employ additional tests to ensure that our results are not driven by other types of unobserved heterogeneity.

# 6 Empirical results

## 6.1 VK Penetration and Protest Participation

### 6.1.1 Reduced Form Estimation

We start by presenting the results of the reduced form estimation. Specifically, we look at how participation in rallies during the first weekend after the parliamentary elections is related to the number of the SPbSU students in different cohorts. Table 2 shows how the protest occurrence (columns (1)-(4)) and the size of these protests (columns (5)-(8)) are related to the number of the SPbSU students in different cohorts. We find that the size of the VK founder cohort has a positive and significant effect on both the incidence and the size of the protests, while the coefficients for other cohorts are much smaller and not statistically significant. Moreover, the sign of the coefficient for the older cohort is consistently negative across specifications. The difference between coefficients for different cohorts is statistically significant for the incidence of protests in all specifications. Figures 2A and 2B report these results graphically.

To assess the possible degree of omitted variable bias under the assumption that selection on the observables proportional to selection on the unobservables, we follow the approach of Oster (2016). In particular, we compute how important, in terms of explanatory power, should be unobservables relative to observables in order to fully explain the coefficient for the VK founder's cohort. We find that unobservables should be negatively correlated with observables and that their importance should be more than seven times higher to be able to explain the results for the size of the protests and three times higher to explain the results for the protest incidence. These results stand in sharp contrast to the standard assumption of equal selection, i.e. that unobservables are positively correlated with observables and are equally important (Antonji, Taber, and Elder, 2005).

Taken together, the results presented in Table 2 and Figures 2A-2B indicate that the SPbSU student cohort of the VK founder is positively and significantly associated with protest participation, in contrast to the older and younger SPbSU cohorts, and that these results are unlikely to be driven by omitted variable bias.

### 6.1.2 IV Results for Protest Participation

Reduced form results in Table 2 suggest that the SPbSU student cohort of the VK founder, through its impact on VK penetration, has affected protest activity in 2011. However, reduced form regressions do not allow us to quantify the magnitude of the effect of social media penetration on protests. In this section, we estimate equation (5) using the number of SPbSU students in the VK founder's cohort as an instrument for VK penetration in summer 2011, controlling for the numbers of SPbSU students in the older and younger cohorts.

First, we test the hypothesis that protests are more likely to occur if social media penetration is higher. The results in columns (1)-(4) of Panel A of Table 3 indicate that social media penetration had a quantitatively large and statistically significant effect on the incidence of protests. To be able to combine IV estimation with clustered standard errors and weak instrument tests, we use a linear probability model. The results indicate that VK penetration had a positive and statistically significant effect on the probability that a protest occurs in a city. A 10% increase in the number of VK users led to a 4.3-4.7 percentage points higher probability of a protest being organized.

One potentially important concern for our estimation is the weak instruments problem. Lack of a sufficiently strong first stage could lead to unreliable IV estimates and inference. The traditional Stock and Yogo (2005) thresholds for the F-statistic were derived for the case of homoscedastic errors, and thus cannot be applied to a model with clustered standard errors. For this reason, we use a recently developed methodology by Montiel Olea and Pflueger (2013) who derived a test for weak instruments similar to the one in Stock and Yogo (2005), but for the case of clustered standard errors. All of the corresponding Montiel Olea-Pflueger F-statistics for a test of 10% potential bias and a 5% significance level in our specifications take values above 270 and well exceed the required threshold level of 23. To be conservative, we conduct a weak instrument test using their methodology after each IV specification we employ. We also check that our instrument would not be considered weak if we used a more traditional Cragg-Donald F-statistic, as it exceeds 300 in all specifications, and is well above the Stock and Yogo (2005) thresholds under the assumption of homoscedasticity.

For comparison purposes, we show the OLS estimates for the same second-stage specifications in columns (5)-(8) of Panel A of Table 3. The coefficients are still highly significant, but are much smaller in magnitude than the corresponding IV estimates. Our explanation for the difference between OLS and IV is the negative selection bias. For example, if people with higher unobserved income are more likely to become VK users, but are less likely to participate in protests, this would lead to a downward bias in then the OLS estimates of the impact of VK penetration on protest participation.

Next we examine the effect of VK penetration on the number of protest participants. According to these estimates, a 10% increase in the number of VK users led to a 19% increase in the number of protesters. Although this effect appears to be large, it is important to have in mind that while VK users constituted a reasonably large share of city population (the average VK penetration in 2011 was 15 percent) while protest participants formed only a tiny fraction. Our data suggests that, for the cities with protests, only 0.4% of the city population participated in protests. As the average city population in our sample was 117 thousand (see Table A1), the aforementioned counterfactual of a 10% increase in VK penetration implies that an increase in the number of VK users by 1,000 leads to an increase in the number of protestors by approximately 50.

The results presented in Panel B of Table 3 assume a log-log relationship between the number of VK users and the number of protestors. To examine this association non-parametrically, we estimate a locally weighted regression between VK penetration and the number of protest participants. These results are presented in Figure 3. The downside of this approach is that it does not account for the endogeneity of VK penetration and does not take into account control variables. However, it provides some intuition on the functional form of the relationship. In particular, Figure 3 indicates that there is a threshold level of VK penetration, below which there is no relation between VK penetration and protests, and that the effect of VK penetration on protest participation is observed only after this tipping point, which is consistent with Prediction 3 of our model. The graph looks similar if we take both VK penetration and the number of protestors as a share of city population (see Figure A3). Thus, these results are consistent with the predictions of the threshold models of collective action (e.g., Granovetter, 1978; Lohmann, 1993, 1994). We confirm that the existence of a threshold level of VK penetration by estimating a nonlinear threshold model in which we allow the coefficient for the effect of VK penetration on protest activity to change at some point.[36] The results of this estimation indicate that indeed, there is a threshold level, below which there is no significant relationship between VK penetration and protest activity and above which there is a strong positive relationship. The threshold is approximately 24000, which corresponds to approximately 20% of the population.

## 6.2 VK Penetration, Voting, and Political Attitudes

Our theoretical framework suggests that as long as public signal about the quality of the regime is positive, there should be a positive effect of VK penetration on the support of the regime. Note that our content analysis of posts in VK before 2011 elections suggest that Putin, Medvedev, and the ruling party were mentioned much more often in blog posts than opposition candidates. According to standard content analysis measures most of these posts were neutral, and very few were negative, with the majority of posts being jokes and funny stories, sometimes even poems about the ruling candidates. Overall, our content analysis suggest that, at least on average, information in social media preceding the elections was positive.

We test whether an increase in VK penetration led to a decrease in electoral support for pro-governmental candidates in the elections that took place after the creation of VK. Table 4 presents the results of the estimation of equation (5) with electoral support for pro-government parties and candidates after 2006 as the outcome variables. In particular, we look at the share of votes received by the government party United Russia in the parliamentary elections of 2007, 2011, and 2016, as well as the share of votes received by Dmitry Medvedev in the presidential elections of 2008 and

[36] Note that this specification we still cannot account for the endogeneity of VK penetration, since we do not have an additional instrument for the threshold.

by Vladimir Putin in 2012. The results show that higher VK penetration consistently led to higher, not lower, electoral support for the government. This effect is not statistically significant for 2007 but is positive and significant for the remaining four elections.[37] Interestingly, OLS results show a statistically significant negative relationship between VK penetration and electoral support for pro-governmental candidates, which indicates that people, who are more likely to join VK are less likely to support the government, but this relationship is driven by endogenous self-selection.

One possible explanation for the positive causal effect of VK penetration on electoral support for pro-governmental candidates is that, on average, there was more pro-governmental than oppositional content in the network, so that higher VK penetration actually decreased the share of people who support the opposition. At the same time, reduction in the costs of collective action associated with higher VK penetration increased the probability that people supporting the opposition would go on protest, and the latter effect outweighed the former. An alternative explanation is that availability of VK increased political polarization, so that it increased both the number of pro-government supporters and the number of people who strongly opposed the government. It is also possible that the official electoral results were contaminated by electoral fraud and did not reflect the actual preferences of the population, although the results in Table 4 could be explained by electoral fraud only if higher VK penetration was associated with greater extent of electoral fraud, which does not sound plausible.

To address these potential alternative explanations, we complement our analysis of electoral outcomes with the analysis of the results of a large-scale opinion poll conducted right before the 2011 parliamentary elections. Respondents were asked about their support of the President Dmitry Medvedev, the Prime Minister Vladimir Putin, and of the government in general on a 6-point scale. They were also asked about their voting intentions in the upcoming parliamentary elections and about their readiness to participate in a hypothetical protest demonstrations.

The results of this analysis presented in Table 5 are consistent with the effect on voting outcomes identified in Table 4. Respondents in cities with higher VK penetration were more likely to give the highest support to the President Dmitry Medvedev, the Prime Minister Vladimir Putin, and the government in general. They were also more likely to report their intentions to vote for the pro-governmental party United Russia in the upcoming elections. We find no evidence of a polarizing effect of social media as there was no increase in the number of respondents with the lowest support for the President, Prime Minister, and the government as a whole.

Importantly, higher VK penetration led to a lower number of respondents who reported their readiness to participate in protests (the effect is significant at 10% level). Thus, right before the actual protests took place, the penetration of VK had a negative effect on the number of potential

[37] We also looked at voter turnout, but we did not find any significant impact of VK penetration on voter turnout after 2007.

participants of the protest, i.e. the information mechanism was working in the direction of reducing the probability of protests. These results are also consistent with the reduction in the cost of collection action being the main channel through which social media affects political protests, despite the fact that the information mechanism was working in the opposite direction.[38]

Overall, the results in Tables 4 and 5 suggest that the information mechanism is unlikely to be the main channel through which social media affected protest participation in the context of our study.

## 6.3 Identifying assumptions checks

### 6.3.1 Placebo Results for Earlier Protests

Table 6 presents results of the placebo regressions in which we estimate the same IV specifications as in columns (1)-(4) of Table 3, but with the measures of pre-VK protests as the dependent variables. Specifically, we look at the protests that occurred in the late Soviet Union in 1987-1992 (both total and pro-democracy as a separate category), labor protests in 1997-2002, and social protests in 2005. The results indicate that there is no significant association between VK penetration in 2011 and any of the placebo outcomes. Moreover, the relationship between VK penetration and protests in post-Soviet Russia is negative in all of the specifications. These results are consistent with the assumption that there is no time-invariant unobserved taste-for-protest heterogeneity that is driving our results. Although for the results in Panel B of Table 6 we cannot reject the hypothesis for the equality of the IV coefficients for the protests of December 2011 and the pre-VK protests because of the large standard errors, in Panel A, we can reject the hypothesis for the equality of the IV coefficients for the protests in December 2011 and similar coefficients for pro-democracy protests in 1987-1992 and the labor protests in 1997-2002.

### 6.3.2 Placebo Results for Earlier Electoral Outcomes

To further ensure that our results are not driven by unobserved heterogeneity, we also conduct a series of placebo tests for electoral outcomes. In particular, we replicate the results in Table 4 using various pre-2006 voting outcomes as the dependent variables. These voting outcomes capture pre-existing political preferences, and the results in Table 3 suggest that they are collectively important for predicting the protest activity of 2011. Table 7 summarizes the results of the placebo tests. Each cell in this table represents the coefficient for VK penetration in an IV regression similar to that in

[38] It is possible, however, that only the information about the electoral fraud that appeared after the elections mattered for protest participation, so that the direction of the information effect changed its sign in a matter of days. This is not fully consistent with the nature of the protest, as the protesters were making general political claims that were not limited to the issues of electoral fraud (Greene, 2014).

column (1) of Table 4, but with various voting outcomes as dependent variables. The specifics of each voting outcome are outlined in the title of each column, while the year of elections is being reported in the row name. Overall, we find that, out of the 34 corresponding regression coefficients, only two were significant at the 5% level and one at the 10% level. These numbers are very close to what could have been attributed to pure chance in multiple hypotheses testing, and they largely support our argument. To further ensure that our results are not driven by pre-existing political preferences, we include voting outcomes as controls for each set of results in the paper.

#### 6.3.3 Placebo Results for Other Universities

We use the distribution of home cities for three different cohorts of the SPbSU students to overcome the problem of unobserved heterogeneity between cities. Nevertheless, it is still theoretically possible that the cohort that studied during the same years as Durov happened to be an unusual cohort and that these people might, for some reason, have had a higher demand for education, a higher demand for social media, and a higher propensity to protest at the same time. To cope with this possibility, we collect data on other 64 Russian universities of comparable quality. Next, we replicate the results of the first-stage regression for the same cohorts in each of the 64 universities in our sample. We then compare the resulting coefficients with that of the corresponding SPbSU cohort. Figures 4A-C show the emerging empirical cumulative distribution functions of the coefficients for the Durov's cohort (Figure 4A), for the older cohort (Figure 4B), and for the younger cohort (Figure 4C).[39] We highlight other universities in Saint Petersburg as they could have experienced spillovers because of their proximity to SPbSU, i.e. their students could also have been more likely to join VK earlier.

Figure 4A indicates that the coefficient for the Durov's cohort at SPbSU lies at the top end of the distribution and, out of four universities with higher coefficients, two are located in Saint Petersburg. At the same time, the coefficients for the younger and older cohorts at SPbSU lie close to the medians of the corresponding distributions in Figures 4B and 4C. Thus, the results in Figures 4A-C indicate that, out of all the cohorts in SPbSU, only the Durov's cohort looks special for the prediction of VK penetration in 2011, relative to those in other Russian universities of similar quality. This is consistent with the idea that students from the corresponding cohort of in the Saint Petersburg State University played a special role in the subsequent penetration of the network.

#### 6.3.4 Student Data and Odnoklassniki

One potential concern with our approach is that we do not have administrative records on student cohorts and instead rely on the information from the profiles of Odnoklassniki users to infer the

[39] Figure A4 in the Online Appendix provides the corresponding graphs for the reduced form regressions.

number of students in each university at each point in time. As was noted in Section 3, these concerns are partially mitigated by the fact that 80% of adults in Russia had an account in Odnoklassniki at the time our data collection took place. This proportion was probably even higher for the younger cohorts, which further improves the representativeness of our data. Additionally, in order to correct for a possible measurement error bias due to the non-random variation in Odnoklassniki penetration, we control for the number of Odnoklassniki users in each city in all of our regressions.

Despite the aforementioned measures, a concern remains that people could be more likely to have an account in Odnoklassniki in cities with a higher VK penetration, and potentially more so in places with a greater number of SPbSU students in the Durov's cohort. To deal with this concern, we conduct two additional tests. First, we check whether the number of Odnoklassniki users is correlated with the number of VK users in a city at different stages of development of the VK network. The results in columns (1)-(3) of Table A4 in the Online Appendix indicate that early VK penetration (the number of users in a city among the first 5000, 50 000, or 100 000 users of the network) is negatively, though not significantly, related to the subsequent penetration of Odnoklassniki. This is consistent with the hypotheses that the initial diffusion of VK was not driven by general preferences for social media and that there might have been a substitution effect between different social networks. VK penetration in 2011 is, however, positively related to Odnoklassniki penetration at the time of the data collection in 2014, although this effect is not statistically significant (see column (4)), which suggests that in the long run penetration of different social networks was driven by the same fundamentals.

Second, we test whether Odnoklassniki penetration was related to student flows from Russian cities to Saint Petersburg State University. The results in columns (5)-(8) indicate that there is no such association, with the standard errors being substantially larger than the coefficients for the VK founder's cohort in all specifications. We conclude that the potential selection introduced by our data collection process is unlikely to bias our results.

#### 6.3.5 Protest Data and Media Reports

Another potential concern with our data collection is that the measures of protests, which were calculated based on media reports, could contain a measurement error that is correlated with VK penetration. It might have been the case that political protests were less likely to be covered by mass media, if they had not been discussed in social media in the first place. This concern is likely to be more relevant for smaller cities as the probability of non-reporting error should be substantially smaller for bigger cities. To address this issue, we examine how the magnitudes of our main IV coefficients vary depending on the size of the cities. Figure A5 in the Online Appendix shows that the IV coefficients for the effect of VK penetration on both the incidence of protests and the number of participants tend to increase when we restrict our sample to larger and larger cities. The coefficients remain largely significant despite a substantial increase in standard errors caused by a reduction in sample size. In particular, our baseline sample of 625 cities with population above 20,000 goes down to only 87 cities with population above 200,000, while the IV coefficients for the latter threshold remain significant at the 5% level and increases in magnitude . Thus, our results are unlikely to be driven by selective media reporting of protests in small cities.

## 6.4 Additional Evidence on Mechanisms

### 6.4.1 Protest Participation and Protest Online Communities

Before proceeding with the analysis, we provide suggestive evidence that VK was indeed used by protest participants to coordinate their activities. While we do not have full information on the content of VK during protests, our descriptive measures suggest that 87 out of 133 cities with protest activity had public VK communities directly related to the corresponding protest events. These communities were accessible to all VK users and were used for informing and coordinating offline protests. To provide evidence that availability of such communities was systematically related to offline protests, Table A5 shows that the number of VK users in protest communities was positively associated with offline protest participation (columns 1-4). In particular, a 10% increase in the number of people in protest communities was associated with a 1.2% increase in the number of protest participants. Similarly, a 10% increase in the number of people in VK protest communities was associated with a 3% increase in the probability of having a protest demonstration in a city (columns 5-8). Overall, these results provide suggestive evidence that there was at least one particular kind of online activity in VK (protest communities) that was directly associated with the spread of offline protests. These results, however, should be interpreted with caution since they do not have a causal interpretation and do not take into account the fact that protest communities represent only one of the channels through which VK could affect protest participation.

### 6.4.2 Fractionalization

To provide further evidence on the mechanisms behind the effect, we take advantage of the fact that Facebook was a close competitor of VK and was also used in protest activities. We look at the distribution of social media users between the two networks.[40] In particular, we compute a fractionalization index, i.e. the probability that two randomly picked users of online social media in a city belong to the same network. In the simplest case of non-overlapping audiences of the two networks, it can be computed as $fract_i = 1 - \sum_j s_{ij}{}^2$ , where $s_{ij}$ is a share of users in network j in city i among all the users of online social networks in city i. Since we do not have information

[40]In contrast to VK and Facebook, Odnoklassniki online social network was not actively employed in the protest movement (Reuter and Szakonyi, 2015), so we do not include it in the analysis.

on the overlap in the audiences of the two social networks we compute fractionalization using this simplified formula and check that our results are robust to a change in the fractionalization index that allows for partial overlap between users from different networks (up to a 70% overlap).

We examine how fractionalization of social media users between the two networks affected protest activity, conditional on the total number of social media users in any of the two networks. In particular, we estimate the following specification:[41]

$$\log(protesters)_i = \beta_0 + \beta_1 fract_i + \beta_2 \log(total\ users)_i + \beta_3 X_i + \varepsilon_{it} \qquad (15)$$

Information channel would imply a zero coefficient for fractionalization since the information critical of the government was available in both networks. Thus, the information effect depends on the total number of users in both networks and not on their sorting into the two networks. The mechanisms associated with a decrease in the costs of collective actions, however, would imply that the coefficient for fractionalization is negative, since both coordination and social pressure work primarily within a network (regardless of which one). Thus, the more divided the users are between the networks, the harder it is for the collective action channel to operate.

Table 8 displays the results of the estimation of equation (6). These results imply that, fractionalization is negatively associated with both protest participation and the incidence of the protests, which is consistent with the collective action channel. These effects, however, are significant only for the large cities, e.g. for a subsample of cities with population over 100,000. This is not surprising as the costs of collective action are presumably higher in large cities. The results presented in column (5) of Table 8 indicate that in larger cities a one standard deviation increase in the network fractionalization is associated with a 42% lower protest participation and a 9 percentage points lower probability of protests occurring.

Overall, these findings are consistent with the hypothesis that social media penetration affects protest participation primarily by reducing the costs of collective action. The finding that the effect of fractionalization is driven by the cities with populations above 100,000 suggests that in smaller cities other means of interpersonal communication may play a greater role in protest coordination.

### 6.4.3 Heterogeneity of the Results

Finally, our approach allows us to check whether there was any heterogeneity in our results with respect to the city characteristics. Table A6 reports our baseline IV results for various subsamples. We find that the effect comes mostly from the cities with higher incomes (columns 1-2), and with higher levels of interpersonal trust (columns 3-4). There is also evidence that the effect is observed mostly from the cities with more educated people, but this result is not statistically significant

[41] Note that we are forced to use OLS for this specification, as we do not have a good instrument for fractionalization.

(columns 5-6),

## 6.5 Implications and Robustness Checks

### 6.5.1 Policy Outcomes

If social media penetration affects protest participation, this, in turn, can influence policy outcomes. In the context of the Russian political protests of 2011-2012, protesters' demands were directed primarily at the national-level policies and appealed mainly to the federal government, which means that we do not necessarily expect to see any variation at the city level. Nevertheless, in an attempt to assess whether there were any differences in the local policies, we looked at the impact of VK penetration on municipal revenues and spending before and after the protests.[42]

Table 9 presents the corresponding results. In all of the specifications, we control for the 2008 values of the dependent variables, thus effectively focusing on the changes in the policy outcomes as opposed to their levels. Panel A tests how federal transfers to municipalities over different years depend on the level of VK penetration. We find that higher VK penetration does not translate into any significant changes in transfers before 2012, but it leads to a decrease in federal transfers in the years of 2012-2014. The magnitudes of these effects are fairly large, with a 10% increase in VK penetration leading to a 26% reduction in federal transfers in 2014. A potential explanation for this effect is that the national government punishes cities for allowing the protests to take place.

Panel B of Table 9 looks at a similar specification with the municipality tax revenue as an outcome variable. We find a negative effect of VK penetration on municipal tax revenues in 2012-2014, but it becomes at least weakly significant only in 2014. One potential explanation for this result is that consistent lack of transfers in previous years has reduced the tax collection capacity of the municipalities. Finally, Panel C of Table 9 checks whether a similar pattern holds for total municipal spending. Although the coefficients for VK penetration are consistently negative with relatively large magnitudes after 2011, the effect is not statistically significant. Thus, we do not find enough evidence that VK penetration led to lower municipal spending after 2011.

These results are consistent with the existing anecdotal evidence that federal and regional government often use municipal transfers as a political tool. A few months prior to the 2011 Parliamentary elections, several government officials were spotted arguing that their cities' municipal finance will be cut in case United Russia does not receive sufficient number of votes.[43] An independent

[42] Note, however, that the municipal data collection in Russia is not consistently implemented, which results in a large number of missing values.

[43] "Sarapul received 30 mln on roads and sidewalks this summer. Glazov received only 10 mln. We ourselves, Glazov residents, refused these extra 20 mln in the previous elections. (You) refused the good roads you could have been driving on. [...] Because United Russia oversees many various projects across the country. And they determine how to work with each city." – said the head of the presidential and government administration of Udmurtia Alexander Goriyanov on 5 Nov 2011 (http://bit.ly/2ofpOKa).

mayor of Yaroslavl, Yevgeny Urlashov, after winning the 2012 elections against a United Russia candidate, faced the lack of regional funding for teacher wages. In this context, it would not be entirely surprising if cities indeed received less transfers from the federal government as a result of political protests.

A serious limitation of these results, however, is that they do not separate between the effect of political protests caused by higher VK penetration from other channels through which VK penetration could affect policy outcomes. A potential way to identify the effect of political protests would be to use weather shocks as an instrument for protest participation (as in Madestam et al., 2013). Unfortunately, we were not able to find a specification with a sufficiently high predictive power in the first-stage regression.[44]

Overall, the results in Table 9 indicate that higher VK penetration led to lower federal transfers to municipal budgets starting from 2012, the first year after the onset of the protests, which suggests that the national government punished cities for allowing the protests to happen.

### 6.5.2 Additional Robustness Checks

We perform several additional robustness checks to ensure that our results are not driven by our choice of specification. We check that our results are robust to using 3-year or 7-year cohorts instead of 5-year cohorts. When 1-year cohorts are used, the results become noisier but still point in the same direction. Our results are also robust to including two older and two younger cohorts instead of one each. In our benchmark specification we chose to keep only one younger and one older cohorts, as our source of data for students is more complete for those cohorts. The results are also robust to using the years of study instead of the year of birth to compute different cohorts.

# 7 Conclusion

This paper provides evidence that social media penetration had a causal effect on both the occurrence and the size of the protest demonstrations in Russia in December 2011. At the same time, social media increased, not decreased, expressed support of the government. Additional evidence suggests that social media affects protest activity by reducing the costs of collective action, rather than by spreading information critical of the government. Thus, our results imply that social media can increase the ability of people to overcome collective action problems.

However, our results should be generalized with caution. First, Russian protests of 2011-2012 were unexpected and the government did not have time to prepare for them. If the threat of collective action is stable over time, governments might use various strategies to counteract social media

[44] Nevertheless, we are very grateful to wunderground.com for their help with obtaining the necessary weather data.

activism (King et al., 2013, 2014). Second, as our theoretical framework highlights, while social media is expected to lower the costs of coordination, information effects of social media could go either way, depending on whether the content of social media is, on average, positive to the government. Overwhelmingly critical content can influence political participation by diminishing support of the government and promoting protests at the same time.

We believe that our methodology can be used for studying the impact of social media penetration on other forms of collective action. For example, consumers, who would like to lower tariffs or discipline companies' misbehavior through boycotts, are also facing the same collective action problems. Similarly, collective action is important for fundraising campaigns of charitable or educational institutions, or for environmental activism. We expect social media to reduce the costs of collective action in all of these circumstances; at least as long as social norms imply that participation in collective action is desirable. More generally, our identification approach, which relies on the shocks to social distance from the inventor, is likely to be applicable to studying the impact of technology adoption in other settings, and can complement the methods of identification based on the physical distance (e.g., Dittmar, 2011; Cantoni and Yuchtman, 2014). In sum, our paper is the first step to finding the ways through which social media can change the opportunities for societies. More research is needed to understand whether similar results hold for other outcomes and in other contexts.

# References

ACEMOGLU, D., T. A. HASSAN, AND A. TAHOUN (2017): "The power of the street: Evidence from Egypt's Arab Spring," *The Review of Financial Studies, forthcoming*.

ADAY, S., H. FARRELL, M. LYNCH, J. SIDES, AND D. FREELON (2012): "Blogs and Bullets II: New Media and Conflict after the Arab Spring," *United States Institute of Peace*.

ADAY, S., H. FARRELL, M. LYNCH, J. SIDES, J. KELLY, AND E. ZUCKERMAN (2010): "Blogs and Bullets: New Media in Contentious Politics," *United States Institute of Peace*.

ADENA, M., R. ENIKOLOPOV, M. PETROVA, V. SANTAROSA, AND E. ZHURAVSKAYA (2015): "Radio and the Rise of Nazis in pre-War Germany," *Quarterly Journal of Economics*, 130, 1885–1939.

AIDT, T. AND R. FRANCK (2015): "Democratization Under the Threat of Revolution: Evidence from the Great Reform Act of 1832," *Econometrica*, 83, 505–547.

ANTONJI, J., C. TABER, AND T. ELDER (2005): "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools,," *Journal of Political Economy*, 113, 151–184.

BARBERÀ, S. AND M. JACKSON (2016): "A Model of Protests, Revolution, and Information," *Working Paper*.

BATTAGLINI, M. (2017): "Public Protests and Policy Making," *The Quarterly Journal of Economics*, 132, 485–549.

BEISSINGER, M. (2002): *Nationalist Mobilization and the Collapse of the Soviet State*, Cambridge University Press.

BHULLER, M., T. HAVNES, E. LEUVEN, AND M. MOGSTAD (2013): "Broadband Internet: An Information Superhighway to Sex Crime?" *Review of Economic Studies*, 80, 1237–1266.

BOND, R. M., C. J. FARISS, J. J. JONES, A. D. KRAMER, C. MARLOW, J. E. SETTLE, AND J. H. FOWLER (2012): "A 61-million-person Experiment in Social Influence and Political Mobilization," *Nature*, 489, 295–298.

CAMPANTE, F., R. DURANTE, AND F. SOBBRIO (2014): "Politics 2.0: The Multifaceted Effect of Broadband Internet on Political Participation," *Working Paper*.

CANTONI, D. AND N. YUCHTMAN (2014): "Medieval Universities, Legal Institutions, and the Commercial Revolution," *Quarterly Journal of Economics*, 129, 823–887.

CHIANG, C.-F. AND B. KNIGHT (2011): "Media Bias and Influence: Evidence from Newspaper Endorsements," *Review of Economic Studies*, 78, 795–820.

CZERNICH, N., O. FALCK, T. KRETSCHMER, AND L. WOESSMANN (2011): "Broadband Infrastructure and Economic Growth," *The Economic Journal*, 121, 505–532.

DELLAVIGNA, S., R. ENIKOLOPOV, V. MIRONOVA, M. PETROVA, AND E. ZHURAVSKAYA (2014): "Cross-Border Effects of Foreign Media: Serbian Radio and Nationalism In Croatia," *American Economic Journal: Applied Economics*, 6, 103–132.

DELLAVIGNA, S. AND E. KAPLAN (2007): "The Fox News Effect: Media Bias and Voting," *Quarterly Journal of Economics*, 122, 807–860.

DITTMAR, J. (2011): "Information Technology and Economic Change: The Impact of the Printing Press," *Quarterly Journal of Economics*, 126, 1133–1172.

DOKUKA, S. V. (2014): "Practices of Using On-line Social Networks,," *Sociological Studies*, 1, 137–145.

EDMOND, C. (2013): "Information Manipulation, Coordination, and Regime Change," *Review of Economic Studies*, 80, 1422–1458.

EISENSEE, T. AND D. STRÖMBERG (2007): "News Droughts, News Floods, and U.S. Disaster Relief," *The Quarterly Journal of Economics*, 122, 693–728.

ENIKOLOPOV, R., V. KOROVKIN, M. PETROVA, K. SONIN, AND A. A. ZAKHAROV (2013): "Field Experiment Estimate of Electoral Fraud in Russian Parliamentary Elections," *Proceedings of the National Academy of Sciences*, 110, 448–452.

ENIKOLOPOV, R., A. MAKARIN, M. PETROVA, AND L. POLISHCHUK (2017a): "Social Image, Networks, and Protest Participation," *Working Paper*.

ENIKOLOPOV, R., M. PETROVA, AND A. K. SONIN (2017b): "Social Media and Corruption," *American Economic Journal: Applied Economics, forthcoming*.

ENIKOLOPOV, R., M. PETROVA, AND E. ZHURAVSKAYA (2011): "Media and Political Persuasion: Evidence from Russia," *American Economic Review*, 101, 3253–3285.

ESFANDIARI, G. (2010): "The Twitter Devolution," *Foreign Policy*.

FALCK, O., R. GOLD, AND S. HEBLICH (2014): "E-Lections: Voting Behavior and the Internet," *American Economic Review*, 104, 2238–2265.

GAVAZZA, A., M. NARDOTTO, AND T. VALLETTI (2015): "Internet and Politics: Evidence from UK Local Elections and Local Government Policies," *Working Paper*.

GENTZKOW, M., N. PETEK, J. SHAPIRO, AND M. SINKINSON (2015a): "Do Newspapers Serve the State? Incumbent Party Influence on the US Press, 1869–1928," *Journal of the European Economic Association*, 13, 29–61.

GENTZKOW, M. AND J. SHAPIRO (2011): "Ideological Segregation Online and Offline," *Quarterly Journal of Economics*, 126, 1799–1839.

GENTZKOW, M., J. SHAPIRO, AND M. SINKINSON (2011): "The Effect of Newspaper Entry and Exit on Electoral Politics,," *American Economic Review*, 101, 2980–3018.

GENTZKOW, M., J. SHAPIRO, AND M. TADDY (2015b): "Measuring Polarization in High-dimensional Data: Method and Application to Congressional Speech," *Working Paper*.

GLADWELL, M. (2010): "Small change," *The New Yorker*, 4, 42–49.

GRANOVETTER, M. (1978): "Threshold Models of Collective Behavior," *American Journal of Sociology*, 1420–1443.

GREENE, S. (2014): *Moscow in Movement: Power and Opposition in Putin's Russia*, Stanford University Press.

HALBERSTAM, Y. AND B. KNIGHT (2016): "Homophily, group size, and the diffusion of political information in social networks: Evidence from Twitter," *Journal of Public Economics*, 143, 73 – 88.

HARDIN, R. (1982): *Collective Action*, Johns Hopkins University Press.

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

HENDEL, I., S. LACH, AND Y. SPIEGEL (2015): "Consumers' Activism: The Facebook boycott of Cottage Cheese," *Working Paper*.

JENSEN, R. (2007): "The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector," *The Quarterly Journal of Economics*, 122, 879–924.

KING, G., J. PAN, AND M. E. ROBERTS (2013): "How Censorship in China Allows Government Criticism but Silences Collective Expression," *American Political Science Review*, 107, 326–343.

——— (2014): "Reverse-Engineering Censorship In China: Randomized Experimentation And Participant Observation," *Science*, 345, 1–10.

KLIMEK, P., Y. YEGOROV, R. HANEL, AND S. THURNER (2012): "Statistical Detection of Systematic Election Irregularities," *Proceedings of National Academy of Sciences*, 109, 16469–16473.

KONONOV, N. (2012): *Kod Durova*, Mann, Ivanov and Ferber.

KURAN, T. (1991): "Now Out of Never: The Element of Surprise in the East European Revolution of 1989," *World Politics*, 44, 7–48.

LITTLE, A. (2015): "Communication Technology and Protest," *Journal of Politics*, 78, 152–166.

LOHMANN, S. (1993): "A Signaling Model of Informative and Manipulative Political Action," *American Political Science Review*, 87, 319–333.

——— (1994): "The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989–91," *World Politics*, 47, 42–101.

LYNCH, M. (2011): "After Egypt: The Limits and Promise of Online Challenges to the Authoritarian Arab State," *Perspectives on Politics*, 9, 301–310.

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, 1633–1685.

MANACORDA, M. AND A. TESEI (2016): "Liberation Technology: Mobile Phones and Political Mobilization in Africa," *Working Paper*.

MONTIEL OLEA, J. L. AND C. PFLUEGER (2013): "A Robust Test for Weak Instruments," *Journal of Business & Economic Statistics*, 31, 358–369.

MOROZOV, E. (2011): *The Net Delusion: How Not to Liberate the World*, Penguin UK.

OLSON, M. (1965): *The Logic of Collective Action*, Harvard University Press.

OSTER, E. (2016): "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*.

OSTROM, E. (1990): *Governing the Commons: The Evolution of Institutions for Collective Action*, Cambridge University Press.

PASSARELLI, F. AND G. TABELLINI (2017): "Emotions and Political Unrest," *Journal of Political Economy, forthcoming*.

PENDERGRAST, M. (2010): *Uncommon Grounds: The History of Coffee and How It Transformed Our World*, Basic Books.

PIERSKALLA, J. H. AND F. M. HOLLENBACH (2013): "Technology and collective action: The effect of cell phone coverage on political violence in Africa," *American Political Science Review*, 107, 207–224.

QIN, B. (2013): "Chinese Microblogs and Drug Quality," *Working Paper*.

QIN, B., D. STRÖMBERG, AND Y. WU (2017): "Why Does China Allow Freer Social Media? Protests versus Surveillance and Propaganda," *Journal of Economic Perspectives*, 31, 117–140.

REUTER, O. J. AND D. SZAKONYI (2015): "Online Social Media and Political Awareness in Authoritarian Regimes," *British Journal of Political Science*, 45, 29–51.

ROBERTSON, G. (2011): *The Politics of Protest in Hybrid Regimes*, Cambridge University Press.

——— (2015): "Political Orientation, Information and Perceptions of Election Fraud: Evidence from Russia," *British Journal of Political Science*, 1–20.

SHIRKY, C. (2008): *Here Comes Everybody: The Power of Organizing Without Organizations*, London, UK: AllenLane: PenguinGroup.

——— (2011): "The political power of social media: Technology, the public sphere, and political change," *Foreign affairs*, 28–41.

SNYDER, J. AND D. STRÖMBERG (2010): "Press Coverage and Political Accountability," *Journal of Political Economy*, 118, 355–408.

STEINERT-THRELKELD, Z., D. MOCANU, A. VESPIGNANI, AND J. FOWLER (2015): "Online Social Networks and Offline Protest," *EPJ Data Science*, 4.

STOCK, J. AND M. YOGO (2005): *Testing for Weak Instruments in Linear IV Regression*, Cambridge: Cambridge University Press.

STRÖMBERG, D. (2004): "Radio's Impact on Public Spending," *Quarterly Journal of Economics*, 119, 189–221.

TREISMAN, D. (2011): *The Return: Russia's Journey from Gorbachev to Medvedev*, New York: Free Press.

TUFEKCI, Z. AND C. WILSON (2012): "Social Media and the Decision to Participate in Political Protest: Observations from Tahrir Square," *Journal of Communication*, 62, 363–379.

YANAGIZAWA-DROTT, D. (2014): "Propaganda and Conflict: Theory and Evidence from the Rwandan Genocide," *Quarterly Journal of Economics*, 129, 1947–1994.

Figure 1. VK Penetration in 2011 and SPbSU student cohorts.

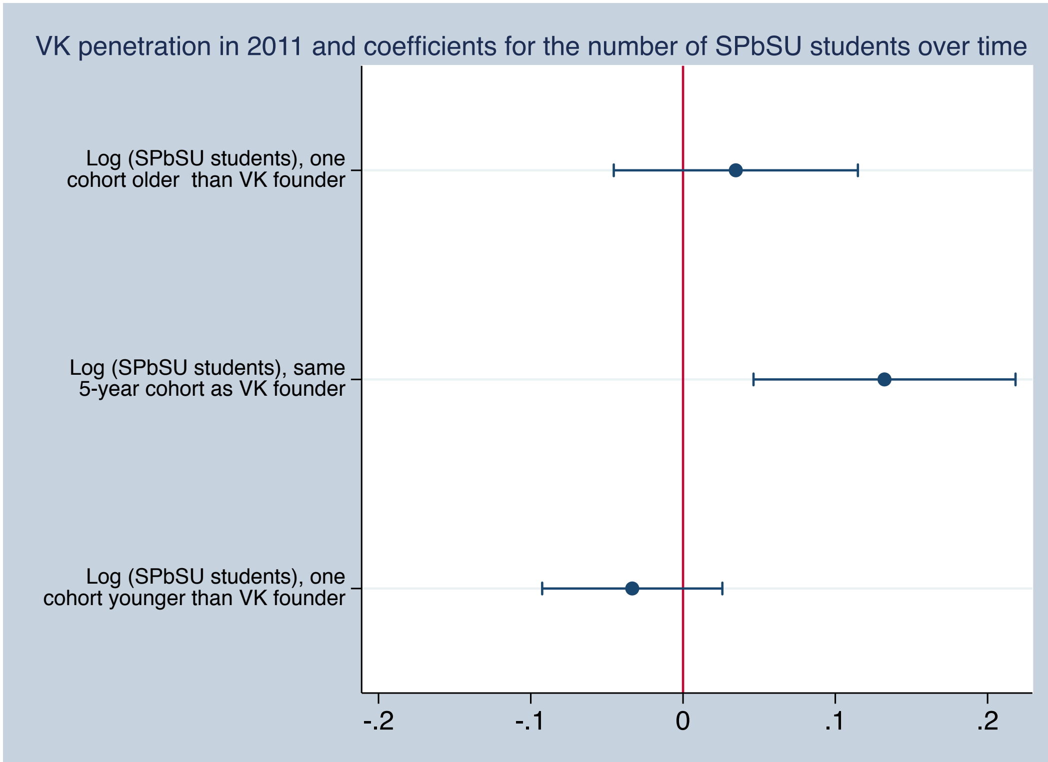
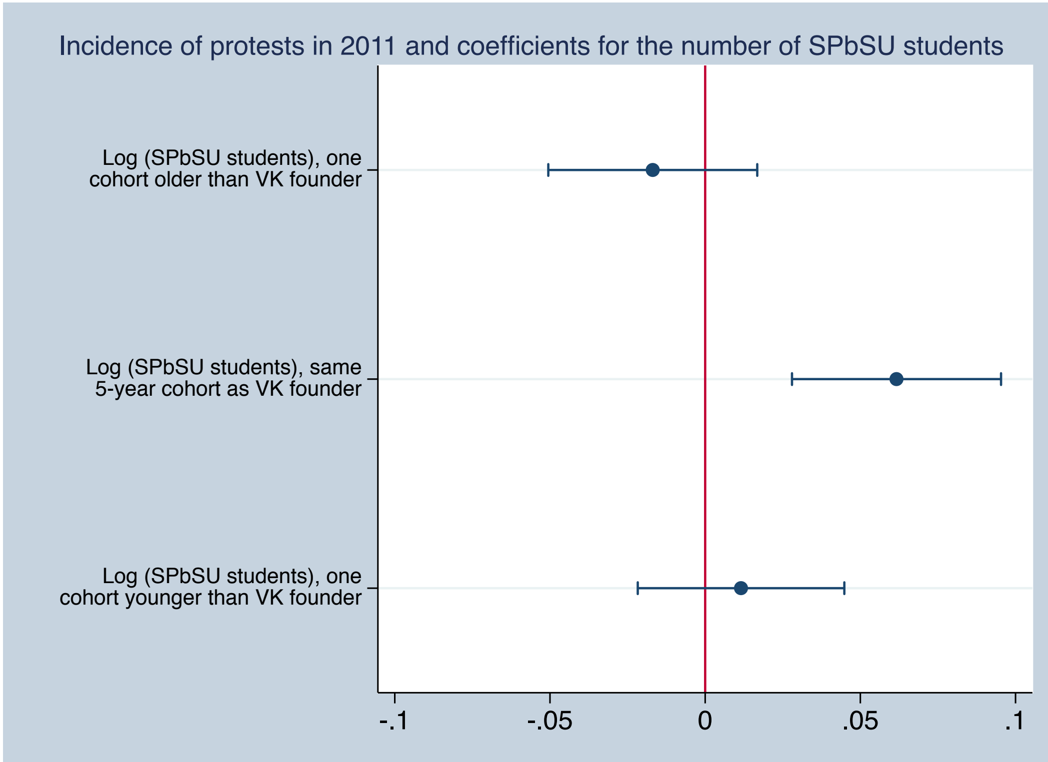
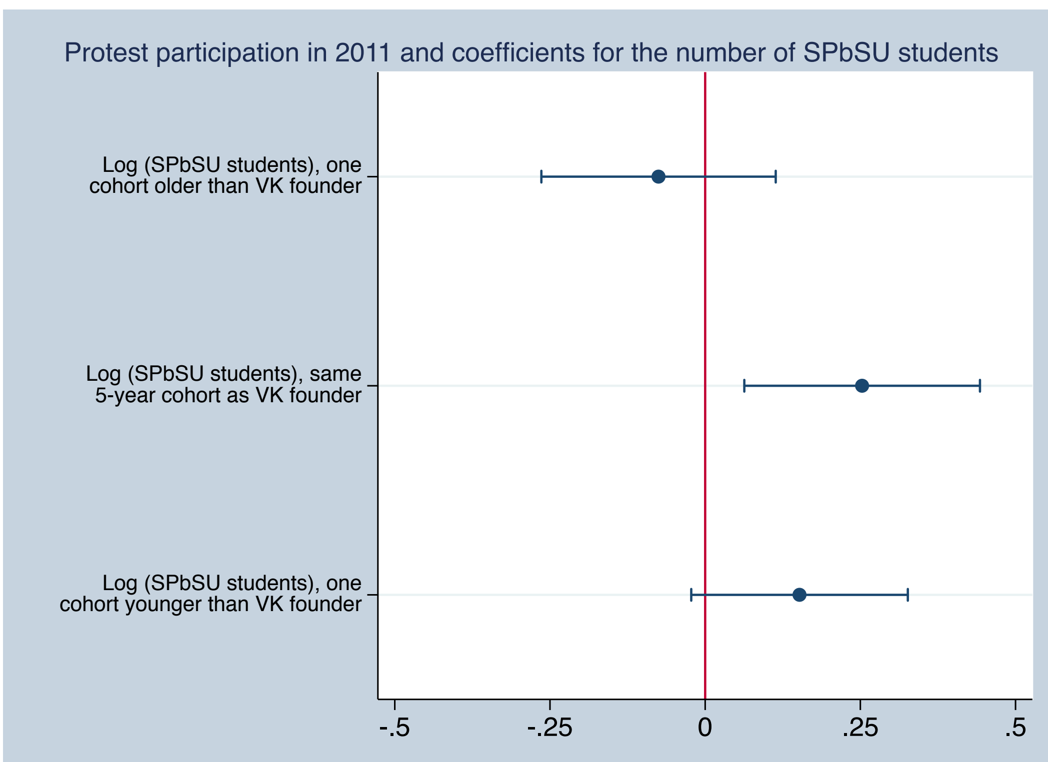


Figure 2. Protest acitivity and SPbSU student cohorts

A. SPbSU Cohorts from Different Cities and The Incidence of Protests




B. SPbSU Cohorts from Different Cities and Protest Participation



Figures 3. Nonparametric Relationship Between VK Penetration and Number of Protesters.

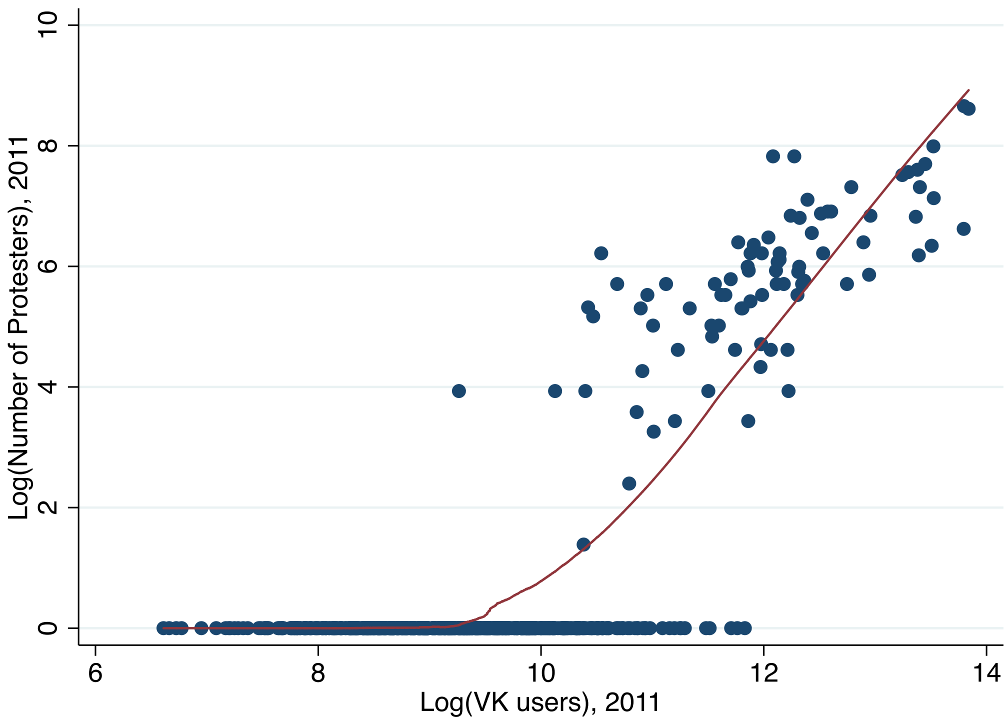
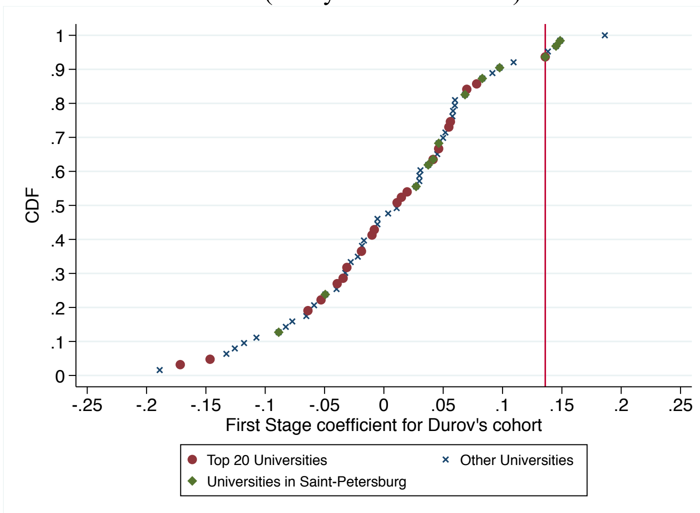
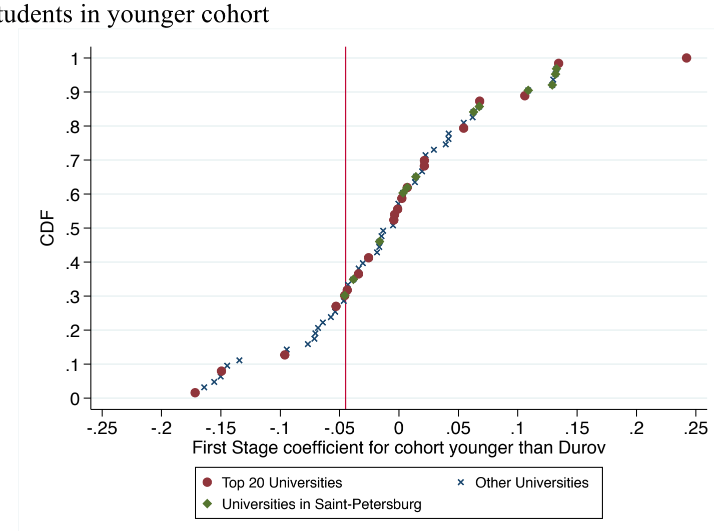


Figure 4. First Stage Coefficients for 65 Universities in Russia.

A: Distribution of students in Durov's cohort (+- 2 years from Durov).




B: Distribution of students in younger cohort




C: Distribution of students in older cohort

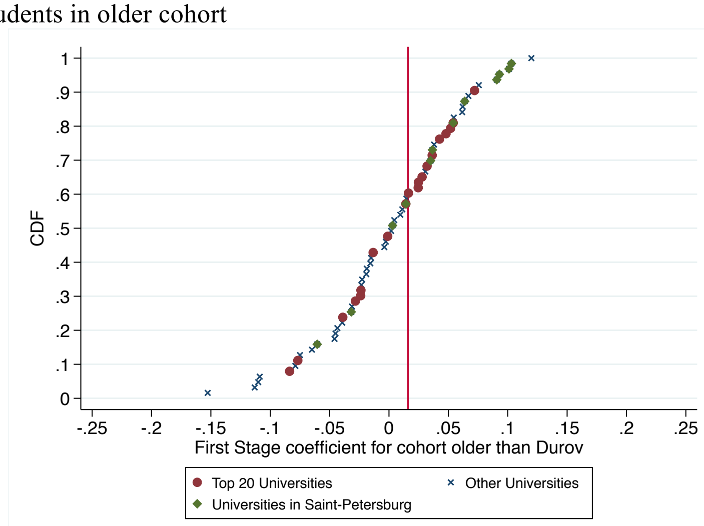


**Table 1. Determinants of VK penetration in 2011 (first stage regression).**

| | Log (number of VK users), Aug 2011 | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Log (SPbSU students), same 5-year cohort as VK founder | 0.5006*** | 0.1749*** | 0.1332*** | 0.1323** | 0.1452*** | 0.1385*** | 0.1404*** |
| | [0.1381] | [0.0442] | [0.0503] | [0.0517] | [0.0511] | [0.0497] | [0.0509] |
| Log (SPbSU students), one cohort younger than VK founder | 0.5612*** | -0.0323 | -0.0195 | -0.0333 | -0.0254 | -0.0364 | -0.0300 |
| | [0.1040] | [0.0522] | [0.0359] | [0.0355] | [0.0356] | [0.0379] | [0.0372] |
| Log (SPbSU students), one cohort older than VK founder | 0.3687** | 0.0945** | 0.0351 | 0.0347 | 0.0280 | 0.0224 | 0.0266 |
| | [0.1726] | [0.0448] | [0.0476] | [0.0482] | [0.0490] | [0.0461] | [0.0458] |
| Regional center | | 0.1992* | 0.2946** | 0.1860 | 0.1542 | 0.1864 | 0.1864 |
| | | [0.1115] | [0.1279] | [0.1393] | [0.1290] | [0.1310] | [0.1261] |
| Distance to Saint Petersburg, km | | | -0.0000 | -0.0001 | -0.0001 | -0.0001 | -0.0002 |
| | | | [0.0002] | [0.0002] | [0.0002] | [0.0002] | [0.0002] |
| Distance to Moscow, km | | | -0.0001 | -0.0000 | -0.0000 | -0.0000 | 0.0000 |
| | | | [0.0002] | [0.0002] | [0.0002] | [0.0002] | [0.0002] |
| Rayon center (county seat) | | | -0.0104 | -0.0200 | -0.0343 | -0.0358 | -0.0181 |
| | | | [0.0735] | [0.0683] | [0.0605] | [0.0678] | [0.0633] |
| Log (average wage), city-level, 2011 | | | 0.1604 | 0.1179 | 0.0526 | 0.0244 | 0.0501 |
| | | | [0.1493] | [0.1501] | [0.1547] | [0.1507] | [0.1445] |
| Presence of a university in a city, 2011 | | | | 0.1229 | 0.1609* | 0.1395 | 0.1480 |
| | | | | [0.0963] | [0.0937] | [0.0954] | [0.0948] |
| Internet penetration, region-level, 2011 | | | | 0.1958 | 0.1451 | 0.1665 | 0.1938 |
| | | | | [0.2254] | [0.2127] | [0.2382] | [0.2215] |
| Log (number of Odnoklassniki users), 2014 | | | | 0.0887 | 0.1099 | 0.1250 | 0.1408* |
| | | | | [0.0851] | [0.0786] | [0.0792] | [0.0790] |
| Ethnic fractionalization, 2010 | | | | 0.3894* | 0.4285* | 0.5763** | 0.3517* |
| | | | | [0.2205] | [0.2203] | [0.2277] | [0.2044] |
| Observations | 625 | 625 | 625 | 625 | 625 | 625 | 625 |
| R-squared | 0.4428 | 0.8614 | 0.9031 | 0.9063 | 0.9127 | 0.9105 | 0.9116 |
| Population controls | | Yes*** | Yes* | Yes** | Yes* | Yes* | Yes* |
| Age cohort controls | | | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** |
| Education controls | | | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** |
| Electoral controls, 1995 | | | | | Yes*** | | |
| Electoral controls, 1999 | | | | | | Yes** | |
| Electoral controls, 2003 | | | | | | | Yes* |
| p-value for equality of coefficients for three cohorts | 0.706 | 0.038** | 0.033** | 0.025** | 0.024** | 0.026** | 0.027** |
| p-value for equality of coefficients of Durov's and younger cohort | 0.762 | 0.011** | 0.011** | 0.008*** | 0.007*** | 0.007*** | 0.008*** |
| p-value for equality of coefficients of Durov's and older cohort | 0.583 | 0.279 | 0.229 | 0.231 | 0.159 | 0.145 | 0.156 |

*** 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. Logarithm of any variable is calculated with 1 added inside. When "Yes" is added to indicate 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. Education controls include the share of population with higher education overall according to 2002 Russian Census and separately in each of the age cohorts according to 2010 Russian Census, to account for both the levels and the change in education. 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 2. Student cohorts and protest participation in 2011. Reduced form estimation.**

| | Incidence of protests, dummy, Dec 2011 | | | | Log (number of protesters), Dec 2011 | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Log (SPbSU students), same 5-year cohort as VK founder | 0.062*** | 0.062*** | 0.064*** | 0.065*** | 0.253** | 0.259** | 0.263** | 0.274** |
| | [0.020] | [0.020] | [0.020] | [0.021] | [0.114] | [0.114] | [0.115] | [0.116] |
| Log (SPbSU students), one cohort younger than VK founder | 0.012 | 0.011 | 0.009 | 0.012 | 0.152 | 0.150 | 0.137 | 0.160 |
| | [0.020] | [0.020] | [0.020] | [0.020] | [0.105] | [0.105] | [0.105] | [0.106] |
| Log (SPbSU students), one cohort older than VK founder | -0.017 | -0.016 | -0.018 | -0.015 | -0.075 | -0.072 | -0.082 | -0.068 |
| | [0.020] | [0.020] | [0.020] | [0.020] | [0.113] | [0.113] | [0.112] | [0.113] |
| Regional center | -0.015 | -0.013 | -0.009 | -0.014 | 0.287 | 0.288 | 0.318 | 0.292 |
| | [0.099] | [0.097] | [0.096] | [0.098] | [0.488] | [0.480] | [0.480] | [0.487] |
| Distance to Saint Petersburg, km | 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, km | -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) | -0.001 | 0.001 | -0.007 | -0.011 | 0.003 | 0.005 | -0.029 | -0.051 |
| | [0.009] | [0.009] | [0.010] | [0.011] | [0.044] | [0.046] | [0.048] | [0.054] |
| Log (average wage), city-level, 2011 | 0.021 | 0.039 | 0.007 | -0.014 | 0.100 | 0.147 | 0.001 | -0.068 |
| | [0.034] | [0.037] | [0.036] | [0.034] | [0.176] | [0.190] | [0.193] | [0.184] |
| Presence of a university in a city, 2011 | 0.196** | 0.195** | 0.195** | 0.200** | 0.870** | 0.876** | 0.860** | 0.898** |
| | [0.098] | [0.098] | [0.097] | [0.097] | [0.423] | [0.423] | [0.422] | [0.426] |
| Internet penetration, region-level, 2011 | -0.013 | 0.005 | -0.003 | -0.007 | 0.138 | 0.181 | 0.175 | 0.149 |
| | [0.045] | [0.045] | [0.054] | [0.048] | [0.243] | [0.240] | [0.280] | [0.257] |
| Log (number of Odnoklassniki users), 2014 | 0.032* | 0.024 | 0.041* | 0.034* | 0.104 | 0.081 | 0.157 | 0.133 |
| | [0.017] | [0.019] | [0.021] | [0.019] | [0.109] | [0.120] | [0.123] | [0.119] |
| Ethnic fractionalization, 2010 | -0.089 | -0.081 | -0.071 | -0.067 | -0.580* | -0.516 | -0.468 | -0.506 |
| | [0.059] | [0.061] | [0.062] | [0.062] | [0.321] | [0.335] | [0.337] | [0.343] |
| Observations | 625 | 625 | 625 | 625 | 625 | 625 | 625 | 625 |
| R-squared | 0.776 | 0.780 | 0.781 | 0.781 | 0.823 | 0.826 | 0.828 | 0.826 |
| Population controls | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** |
| Age cohort controls | Yes** | Yes*** | Yes*** | Yes*** | Yes* | Yes** | Yes** | Yes** |
| Education 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-value for equality of coefficients for three cohorts | 0.078* | 0.071* | 0.058* | 0.069* | 0.271 | 0.271 | 0.250 | 0.247 |
| p-value for equality of coefficients of Durov's and younger cohort | 0.089* | 0.073* | 0.067* | 0.079* | 0.528 | 0.489 | 0.430 | 0.487 |
| p-value for equality of coefficients of Durov's and older cohort | 0.031** | 0.032** | 0.025** | 0.028** | 0.115 | 0.111 | 0.099* | 0.102 |

*** 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. Logarithm of any variable is calculated with 1 added inside. When "Yes" is added to indicate 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. Education controls include the share of population with higher education overall according to 2002 Russian Census and separately in each of the age cohorts according to 2010 Russian Census, to account for both the levels and the change in education. 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 3. VK penetration and protest participation in 2011.**

**Panel A. Probability of protests**

| | Incidence of protests, dummy, Dec 2011 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | IV | IV | IV | IV | OLS | OLS | OLS | OLS |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Log (number of VK users), Aug 2011 | 0.466** | 0.428*** | 0.459*** | 0.466*** | 0.060*** | 0.058*** | 0.055*** | 0.066*** |
| | [0.189] | [0.161] | [0.176] | [0.174] | [0.018] | [0.018] | [0.019] | [0.018] |
| Log (SPbSU students), one cohort younger than VK founder | 0.027 | 0.022 | 0.026 | 0.026 | 0.029 | 0.028 | 0.026 | 0.030 |
| | [0.024] | [0.023] | [0.025] | [0.024] | [0.021] | [0.021] | [0.021] | [0.020] |
| Log (SPbSU students), one cohort older than VK founder | -0.033 | -0.028 | -0.029 | -0.027 | 0.003 | 0.005 | 0.003 | 0.006 |
| | [0.031] | [0.029] | [0.028] | [0.029] | [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** |
| Education controls | Yes | Yes | Yes* | Yes | Yes | Yes | Yes | Yes |
| Other 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 | 625 | 625 | 625 | 625 | 625 | 625 | 625 | 625 |
| Effective F-stat (Montiel Olea and Pflueger 2013) | 278.8 | 276.0 | 276.0 | 276.0 | | | | |

**Panel B. Number of protesters**

| | Log (number of protesters), Dec 2011 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | IV | IV | IV | IV | OLS | OLS | OLS | OLS |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Log (number of VK users), Aug 2011 | 1.911** | 1.787** | 1.900** | 1.951** | 0.377*** | 0.360*** | 0.347*** | 0.399*** |
| | [0.924] | [0.809] | [0.872] | [0.866] | [0.098] | [0.102] | [0.105] | [0.101] |
| Log (SPbSU students), one cohort younger than VK founder | 0.216* | 0.196* | 0.207* | 0.218* | 0.221** | 0.217** | 0.209* | 0.231** |
| | [0.117] | [0.113] | [0.119] | [0.118] | [0.107] | [0.108] | [0.108] | [0.107] |
| Log (SPbSU students), one cohort older than VK founder | -0.141 | -0.123 | -0.124 | -0.120 | -0.004 | 0.004 | -0.002 | 0.009 |
| | [0.151] | [0.141] | [0.136] | [0.145] | [0.093] | [0.092] | [0.090] | [0.093] |
| Population controls | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** |
| Age cohort controls | Yes | Yes | Yes | Yes | Yes* | Yes* | Yes** | Yes** |
| Education controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Other 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 | 625 | 625 | 625 | 625 | 625 | 625 | 625 | 625 |
| Effective F-statistics (Olea Montiel and Pflueger 2013) | 278.8 | 276.0 | 276.0 | 276.0 | | | | |

*** 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. Logarithm of any variable is calculated with 1 added inside. When "Yes" is added to indicate 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. Education controls include the share of population with higher education overall according to 2002 Russian Census and separately in each of the age cohorts according to 2010 Russian Census, to account for both the levels and the change in education. 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 Petersburg, log (average wage), share of people with higher education in 2002, internet penetration in 2011, log (Odnoklassniki users in 2014).

**Table 4. VK penetration and Voting Outcomes.**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Voting share for United Russia, 2007 | | | | | | | |
| | IV | IV | IV | IV | OLS | OLS | OLS | OLS |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Log (number of VK users), Aug 2011 | 0.043 | 0.023 | 0.054 | 0.004 | -0.027* | -0.025** | -0.018 | -0.032*** |
| | [0.060] | [0.047] | [0.056] | [0.042] | [0.014] | [0.011] | [0.013] | [0.011] |
| Log (SPbSU students), one cohort younger than VK founder | -0.007 | -0.004 | -0.006 | -0.007 | -0.007 | -0.004 | -0.006 | -0.007 |
| | [0.009] | [0.007] | [0.008] | [0.006] | [0.008] | [0.007] | [0.007] | [0.006] |
| Log (SPbSU students), one cohort older than VK founder | 0.001 | 0.001 | -0.001 | -0.003 | 0.006 | 0.004 | 0.004 | -0.001 |
| | [0.008] | [0.007] | [0.008] | [0.006] | [0.007] | [0.006] | [0.007] | [0.005] |
| | Voting share for Medvedev, 2008 | | | | | | | |
| | IV | IV | IV | IV | OLS | OLS | OLS | OLS |
| Log (number of VK users), Aug 2011 | 0.153* | 0.132* | 0.165* | 0.113* | -0.007 | -0.009 | -0.004 | -0.013 |
| | [0.089] | [0.072] | [0.085] | [0.065] | [0.012] | [0.010] | [0.011] | [0.009] |
| Log (SPbSU students), one cohort younger than VK founder | -0.006 | -0.005 | -0.006 | -0.005 | -0.006 | -0.004 | -0.006 | -0.005 |
| | [0.010] | [0.008] | [0.010] | [0.007] | [0.008] | [0.006] | [0.007] | [0.006] |
| Log (SPbSU students), one cohort older than VK founder | -0.001 | -0.001 | -0.004 | -0.005 | 0.011 | 0.009 | 0.007 | 0.004 |
| | [0.010] | [0.009] | [0.010] | [0.008] | [0.007] | [0.006] | [0.007] | [0.005] |
| | Voting share for United Russia, 2011 | | | | | | | |
| | IV | IV | IV | IV | OLS | OLS | OLS | OLS |
| Log (number of VK users), Aug 2011 | 0.281* | 0.206* | 0.276* | 0.210 | -0.047*** | -0.043** | -0.034* | -0.052*** |
| | [0.169] | [0.118] | [0.154] | [0.130] | [0.017] | [0.016] | [0.017] | [0.014] |
| Log (SPbSU students), one cohort younger than VK founder | -0.004 | -0.001 | -0.002 | -0.002 | -0.004 | 0.001 | -0.003 | -0.001 |
| | [0.016] | [0.012] | [0.015] | [0.012] | [0.012] | [0.010] | [0.011] | [0.010] |
| Log (SPbSU students), one cohort older than VK founder | 0.000 | 0.005 | -0.001 | -0.005 | 0.023* | 0.023** | 0.020* | 0.015 |
| | [0.018] | [0.014] | [0.016] | [0.014] | [0.012] | [0.011] | [0.010] | [0.010] |
| | Voting Share for Putin, 2012 | | | | | | | |
| | IV | IV | IV | IV | OLS | OLS | OLS | OLS |
| Log (number of VK users), Aug 2011 | 0.155* | 0.129* | 0.153* | 0.110 | -0.015 | -0.014 | -0.011 | -0.021** |
| | [0.093] | [0.077] | [0.087] | [0.071] | [0.012] | [0.010] | [0.012] | [0.009] |
| Log (SPbSU students), one cohort younger than VK founder | 0.001 | 0.001 | 0.002 | 0.001 | 0.001 | 0.003 | 0.002 | 0.001 |
| | [0.010] | [0.008] | [0.009] | [0.007] | [0.008] | [0.007] | [0.007] | [0.006] |
| Log (SPbSU students), one cohort older than VK founder | 0.006 | 0.006 | 0.004 | 0.001 | 0.018** | 0.017** | 0.015** | 0.011* |
| | [0.011] | [0.010] | [0.010] | [0.009] | [0.007] | [0.007] | [0.007] | [0.006] |
| | Voting share for United Russia, 2016 | | | | | | | |
| | IV | IV | IV | IV | OLS | OLS | OLS | OLS |
| Log (number of VK users), Aug 2011 | 0.212* | 0.141 | 0.185* | 0.130* | 0.001 | 0.012 | 0.018 | -0.000 |
| | [0.116] | [0.088] | [0.095] | [0.077] | [0.017] | [0.016] | [0.018] | [0.012] |
| Log (SPbSU students), one cohort younger than VK founder | -0.002 | 0.002 | 0.001 | -0.001 | -0.002 | 0.003 | 0.000 | -0.001 |
| | [0.013] | [0.010] | [0.011] | [0.009] | [0.011] | [0.010] | [0.010] | [0.009] |
| Log (SPbSU students), one cohort older than VK founder | 0.011 | 0.015 | 0.009 | 0.006 | 0.025** | 0.024** | 0.020* | 0.015* |
| | [0.015] | [0.013] | [0.014] | [0.012] | [0.011] | [0.011] | [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 |
| Education controls | Yes | Yes | Yes | Yes | Yes*** | Yes*** | Yes*** | Yes*** |
| Other 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 | 625 | 625 | 625 | 625 | 625 | 625 | 625 | 625 |
| Effective F-statistics (Olea Montiel and Pflueger 2013) | 276.1 | 273.3 | 273.3 | 273.3 | | | | |

*** 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. Logarithm of any variable is calculated with 1 added inside. When "Yes" is added to indicate 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. Education controls include the share of population with higher education overall according to 2002 Russian Census and separately in each of the age cohorts according to 2010 Russian Census, to account for both the levels and the change in education. 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 Petersburg, log (average wage), share of people with higher education in 2002, internet penetration in 2011, log (Odnoklassniki users in 2014).

**Table 5. VK Penetration and Political Attitudes.**

| | How do you assess the work of president Dmitry Medvedev | | | | | |
|---|---|---|---|---|---|---|
| | Good and getting better | Good and remains the same | Good and getting worse | Bad, but getting better | Bad and remains the same | Bad and getting worse |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Log (number of VK users), Aug 2011 | 0.255** | -0.069 | -0.060 | -0.094 | -0.026 | 0.026 |
| | [0.127] | [0.130] | [0.062] | [0.059] | [0.076] | [0.061] |
| Log (SPbSU students), one cohort younger than VK founder | -0.013 | 0.010 | 0.001 | 0.013** | 0.003 | 0.005 |
| | [0.016] | [0.009] | [0.007] | [0.005] | [0.009] | [0.008] |
| Log (SPbSU students), one cohort older than VK founder | -0.016 | -0.017 | -0.001 | 0.006 | -0.011 | -0.006 |
| | [0.019] | [0.014] | [0.010] | [0.008] | [0.009] | [0.008] |
| | How do you assess the work of prime minister Vladimir Putin | | | | | |
| | Good and getting better | Good and remains the same | Good and getting worse | Bad, but getting better | Bad and remains the same | Bad and getting worse |
| Log (number of VK users), Aug 2011 | 0.205* | -0.072 | 0.004 | -0.061 | -0.068 | -0.016 |
| | [0.124] | [0.124] | [0.047] | [0.042] | [0.075] | [0.056] |
| Log (SPbSU students), one cohort younger than VK founder | -0.019 | 0.012 | -0.000 | 0.008** | 0.007 | 0.004 |
| | [0.016] | [0.009] | [0.006] | [0.003] | [0.009] | [0.007] |
| Log (SPbSU students), one cohort older than VK founder | -0.011 | -0.021 | -0.007 | 0.005 | -0.002 | -0.002 |
| | [0.018] | [0.016] | [0.007] | [0.006] | [0.011] | [0.007] |
| | How do you assess the work of the government | | | | | |
| | Good and getting better | Good and remains the same | Good and getting worse | Bad, but getting better | Bad and remains the same | Bad and getting worse |
| Log (number of VK users), Aug 2011 | 0.313** | 0.100 | -0.124* | -0.078 | -0.075 | -0.027 |
| | [0.133] | [0.129] | [0.074] | [0.079] | [0.104] | [0.091] |
| Log (SPbSU students), one cohort younger than VK founder | -0.017 | 0.015 | 0.004 | 0.013** | -0.001 | 0.001 |
| | [0.018] | [0.013] | [0.008] | [0.006] | [0.012] | [0.009] |
| Log (SPbSU students), one cohort older than VK founder | -0.019 | -0.026 | 0.007 | 0.006 | -0.014 | 0.001 |
| | [0.020] | [0.018] | [0.012] | [0.010] | [0.012] | [0.011] |
| | Which party are you planning to vote for in December elections | | | | | |
| | United Russia | Just Russia | LDPR | KPRF | Patriots of Russia | Yabloko |
| Log (number of VK users), Aug 2011 | 0.260* | 0.050 | -0.056 | -0.041 | -0.002 | -0.005 |
| | [0.155] | [0.055] | [0.055] | [0.067] | [0.009] | [0.013] |
| Log (SPbSU students), one cohort younger than VK founder | -0.006 | -0.000 | 0.006 | 0.003 | 0.001 | 0.002 |
| | [0.016] | [0.005] | [0.005] | [0.005] | [0.001] | [0.001] |
| Log (SPbSU students), one cohort older than VK founder | -0.043* | -0.004 | 0.005 | 0.002 | 0.000 | -0.002 |
| | [0.023] | [0.007] | [0.009] | [0.008] | [0.001] | [0.002] |
| | Do you personally admit or exclude a possibility to take part in any protests | | | | | |
| | Admit | Exclude | Difficult to answer | | | |
| Log (number of VK users), Aug 2011 | -0.278* | 0.101 | 0.186 | | | |
| | [0.164] | [0.184] | [0.146] | | | |
| Log (SPbSU students), one cohort younger than VK founder | -0.001 | -0.002 | 0.002 | | | |
| | [0.014] | [0.015] | [0.012] | | | |
| Log (SPbSU students), one cohort older than VK founder | 0.027 | -0.024 | -0.005 | | | |
| | [0.021] | [0.025] | [0.022] | | | |

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in brackets are adjusted by clusters within regions. Unit of observation is an individual respondent. Logarithm of any variable is calculated with 1 added inside. The table presents results of 27 separate IV regressions. In all regressions the number of observations is 31,728 and the effective F-statistics (Olea Montiel and Pflueger 2013) is 30,415. All regressions include the following city-level controls:5th polynomial of population, the number of people aged 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50 and older years, the share of population with higher education in each of the age cohorts separately, dummy for regional and county centers , distances to Moscow and St Petersburg, log (average wage), share of population with higher education in 2002, internet penetration in 2011, log (Odnoklassniki users in 2014).

**Table 6. VK Penetration and pre-VK Protests.**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **Panel A. Incidence of earlier protests** | | | | | | | | |
| | Incidence of protests, 1987-1992 | | | | Incidence of pro-democracy protests, 1987-1992 | | | |
| Log (number of VK users), Aug 2011 | 0.009 | 0.006 | -0.015 | 0.023 | -0.011 | -0.019 | -0.023 | 0.004 |
| | [0.281] | [0.271] | [0.263] | [0.273] | [0.194] | [0.186] | [0.189] | [0.192] |
| P-value for equality of coefficients with that in Table 3 | 0.182 | 0.182 | 0.139 | 0.176 | 0.094* | 0.086* | 0.077* | 0.086* |
| | Incidence of labor protests, 1997-2002 | | | | Incidence of social protests, 2005 | | | |
| Log (number of VK users), Aug 2011 | -0.056 | -0.053 | -0.022 | -0.018 | -0.070 | -0.058 | -0.170 | -0.035 |
| | [0.238] | [0.211] | [0.228] | [0.227] | [0.239] | [0.210] | [0.233] | [0.247] |
| P-value for equality of coefficients with that in Table 3 | 0.108 | 0.086* | 0.120 | 0.111 | 0.041** | 0.045** | 0.019** | 0.051* |
| **Panel B. Participation in earlier protests** | | | | | | | | |
| | Log (number of protesters), 1987-1992 | | | | Log (pro-democracy protesters), 1987-1992 | | | |
| Log (number of VK users), Aug 2011 | 0.533 | 0.410 | 0.281 | 0.477 | 0.144 | -0.010 | 0.017 | 0.136 |
| | [1.904] | [1.880] | [1.831] | [1.887] | [1.494] | [1.449] | [1.476] | [1.524] |
| P-value for equality of coefficients with that in Table 3 | 0.482 | 0.475 | 0.397 | 0.448 | 0.298 | 0.270 | 0.263 | 0.283 |
| | Log (participants in labor protests), 1997-2002 | | | | Log (participants in social protests), 2005 | | | |
| Log (number of VK users), Aug 2011 | -0.312 | -0.280 | -0.075 | -0.041 | -0.562 | -0.515 | -1.366 | -0.481 |
| | [1.625] | [1.426] | [1.552] | [1.549] | [1.850] | [1.644] | [1.774] | [1.884] |
| P-value for equality of coefficients with that in Table 3 | 0.268 | 0.238 | 0.306 | 0.295 | 0.194 | 0.188 | 0.080* | 0.201 |
| Population, Age cohorts, Education, and Other 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 | 625 | 625 | 625 | 625 | 625 | 625 | 625 | 625 |

*** 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. Logarithm of any variable is calculated with 1 added inside. "Yes" indicates inclusion of a corresponding group of controls. Significance level is NOT reported after each group of controls for the purpose of brevity. 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. Education controls include the share of population with higher education overall according to 2002 Russian Census and separately in each of the age cohorts according to 2010 Russian Census, to account for both the levels and the change in education. 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 Petersburg, log (average wage), share of people with higher education in 2002, internet penetration in 2011, log (Odnoklassniki users in 2014). P-values for equality of coefficients are calculated relative to a corresponding coefficient in columns (1)-(4) of Tables 4 and 5, using a 3sls framework.

**Table 7. VK Penetration and Pre-VK Voting Results.**

**Panel A. Parliamentary elections**

| | Dependent variable | | | | | |
|---|---|---|---|---|---|---|
| | Pro-government party vote share | Yabloko vote share | Communists vote share | LDPR vote share | Turnout | Against all share |
| Voting results in 1995, IV with SPbSU cohorts | -0.016 | -0.019 | 0.101 | 0.021 | 0.023 | -0.009 |
| | [0.028] | [0.019] | [0.072] | [0.052] | [0.038] | [0.008] |
| Voting results in 1999, IV with SPbSU cohorts | 0.059 | 0.000 | 0.047 | -0.009 | -0.085 | -0.002 |
| | [0.050] | [0.015] | [0.048] | [0.010] | [0.059] | [0.007] |
| Voting results in 2003 IV with SPbSU cohorts | -0.003 | -0.018 | -0.016 | -0.007 | -0.013 | -0.017 |
| | [0.003] | [0.011] | [0.023] | [0.024] | [0.041] | [0.012] |

**Panel B. Presidential elections**

| Year 1996, 1st round | Yeltsin vote share | Yavlinsky vote share | Zyuganov vote share | Lebedev vote share | Turnout | Against all share |
|---|---|---|---|---|---|---|
| Voting results, IV with SPbSU cohorts | -0.114 | 0.120 | 0.007 | -0.009 | 0.013 | -0.002 |
| | [0.080] | [0.089] | [0.016] | [0.041] | [0.025] | [0.003] |
| **Year 1996, 2nd round** | **Yeltsin vote share** | | **Zyuganov vote share** | | **Turnout** | **Against all share** |
| Voting results, IV with SPbSU cohorts | -0.108 | - | 0.124 | - | 0.008 | -0.008 |
| | [0.088] | - | [0.092] | - | [0.029] | [0.010] |
| **Year 2000** | **Putin vote share** | **Yavlinsky vote** | **Zyuganov vote** | **Tuleev vote share** | **Turnout** | **Against all share** |
| Voting results, IV with SPbSU cohorts | 0.129* | -0.027** | -0.056 | 0.004 | 0.001 | -0.012** |
| | [0.074] | [0.013] | [0.054] | [0.028] | [0.029] | [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 8. Fractionalization of Networks and Protest Participation.**

**Panel A. Network fractionalization and protest participation.**

| | Log (number of protesters), Dec 2011 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Whole sample | | | | Cities with more than 100 000 inhabitants | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Fractionalization of social media networks (Facebook+Vkontakte) | -0.894 | -0.997 | -1.021 | -0.868 | -4.797** | -4.789** | -4.468** | -5.562*** |
| | [0.744] | [0.782] | [0.746] | [0.745] | [2.140] | [2.248] | [2.184] | [2.071] |
| Log (number of users in both networks) | 1.896*** | 1.878*** | 1.819*** | 1.914*** | 1.233** | 1.299* | 1.478** | 1.460** |
| | [0.373] | [0.361] | [0.372] | [0.365] | [0.618] | [0.657] | [0.645] | [0.625] |
| Population, Age cohorts, Education, and Other 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 | 625 | 625 | 625 | 625 | 158 | 158 | 158 | 158 |
| R-squared | 0.838 | 0.841 | 0.842 | 0.841 | 0.821 | 0.836 | 0.838 | 0.840 |

**Panel B. Network fractionalization and the incidence of protest**

| | Incidence of protests, dummy, Dec 2011 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Whole sample | | | | Cities with more than 100 000 inhabitants | | | |
| Fractionalization of social media networks (Facebook+Vkontakte) | -0.135 | -0.149 | -0.158 | -0.122 | -0.983** | -0.898** | -0.914** | -1.113*** |
| | [0.143] | [0.149] | [0.143] | [0.144] | [0.435] | [0.424] | [0.426] | [0.414] |
| Log (number of users in both networks) | 0.265*** | 0.263*** | 0.251*** | 0.268*** | 0.072 | 0.106 | 0.103 | 0.129 |
| | [0.074] | [0.072] | [0.073] | [0.072] | [0.122] | [0.121] | [0.127] | [0.122] |
| Population, Age cohorts, Education, and Other 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 | 625 | 625 | 625 | 625 | 158 | 158 | 158 | 158 |
| R-squared | 0.783 | 0.787 | 0.786 | 0.786 | 0.768 | 0.789 | 0.786 | 0.799 |

*** 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. Logarithm of any variable is calculated with 1 added inside. When "Yes" is added to indicate 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. Education controls include the share of population with higher education overall according to 2002 Russian Census and separately in each of the age cohorts according to 2010 Russian Census, to account for both the levels and the change in education. 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 Petersburg, log (average wage), share of people with higher education in 2002, internet penetration in 2011, log (Odnoklassniki users in 2014).

**Table 9. VK Penetration and Municipal Budgets.**

| **Panel A. VK penetration and federal transfers to municipalities** | | | | | | |
|---|---|---|---|---|---|---|
| | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
| | Log (transfers to municipality) | | | | | |
| Log (number of VK users), Aug 2011 | -0.256 | -1.347 | -1.788 | -3.660** | -2.499* | -2.569** |
| | [0.867] | [1.103] | [1.143] | [1.742] | [1.350] | [1.275] |
| Population controls | Yes | Yes | Yes | Yes | Yes** | Yes |
| Age cohort controls | Yes | Yes | Yes | Yes | Yes** | Yes |
| Education controls | Yes** | Yes | Yes | Yes | Yes | Yes |
| Observations | 289 | 305 | 305 | 309 | 311 | 290 |
| Effective F-statistics (Olea Montiel and Pflueger 2013) | 136.8 | 138.8 | 140.5 | 139.2 | 134.2 | 119.2 |
| **Panel B. VK penetration and municipality's tax revenues** | | | | | | |
| | Log (municipal tax revenues) | | | | | |
| Log (number of VK users), Aug 2011 | -0.240 | -0.135 | 0.133 | -0.487 | -0.430 | -0.826* |
| | [0.287] | [0.356] | [0.317] | [0.378] | [0.363] | [0.474] |
| Population controls | Yes | Yes* | Yes | Yes | Yes | Yes |
| Age cohort controls | Yes | Yes* | Yes | Yes* | Yes* | Yes |
| Education controls | Yes*** | Yes* | Yes | Yes*** | Yes | Yes |
| Observations | 437 | 453 | 442 | 441 | 439 | 423 |
| Effective F-statistics (Olea Montiel and Pflueger 2013) | 197.8 | 202.1 | 197.1 | 195.9 | 192.1 | 183.5 |
| **Panel C. VK penetration and municipal spending** | | | | | | |
| | Log (municipal total spending) | | | | | |
| Log (number of VK users), Aug 2011 | -0.174 | 0.067 | -0.053 | -0.437 | -0.399 | -0.523 |
| | [0.296] | [0.288] | [0.256] | [0.350] | [0.381] | [0.398] |
| Population controls | Yes | Yes** | Yes | Yes | Yes | Yes |
| Age cohort controls | Yes** | Yes* | Yes** | Yes** | Yes | Yes |
| Education controls | Yes* | Yes | Yes | Yes*** | Yes | Yes |
| Observations | 377 | 388 | 407 | 398 | 417 | 396 |
| Effective F-statistics (Olea Montiel and Pflueger 2013) | 165.2 | 191.3 | 187.1 | 183.3 | 183.1 | 172.6 |

*** 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. IV estimates are reported. Logarithm of any variable is calculated with 1 added inside. When "Yes" is added to indicate 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. Education controls include the share of population with higher education overall according to 2002 Russian Census and separately in each of the age cohorts according to 2010 Russian Census, to account for both the levels and the change in education. Other controls include dummy for regional and county centers, distances to Moscow and St Petersburg, log (average wage), share of people with higher education in 2002, internet penetration in 2011, log (Odnoklassniki users in 2014). All specifications control for initial (2008) values of corresponding dependent variable.

**APPENDIX.**

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

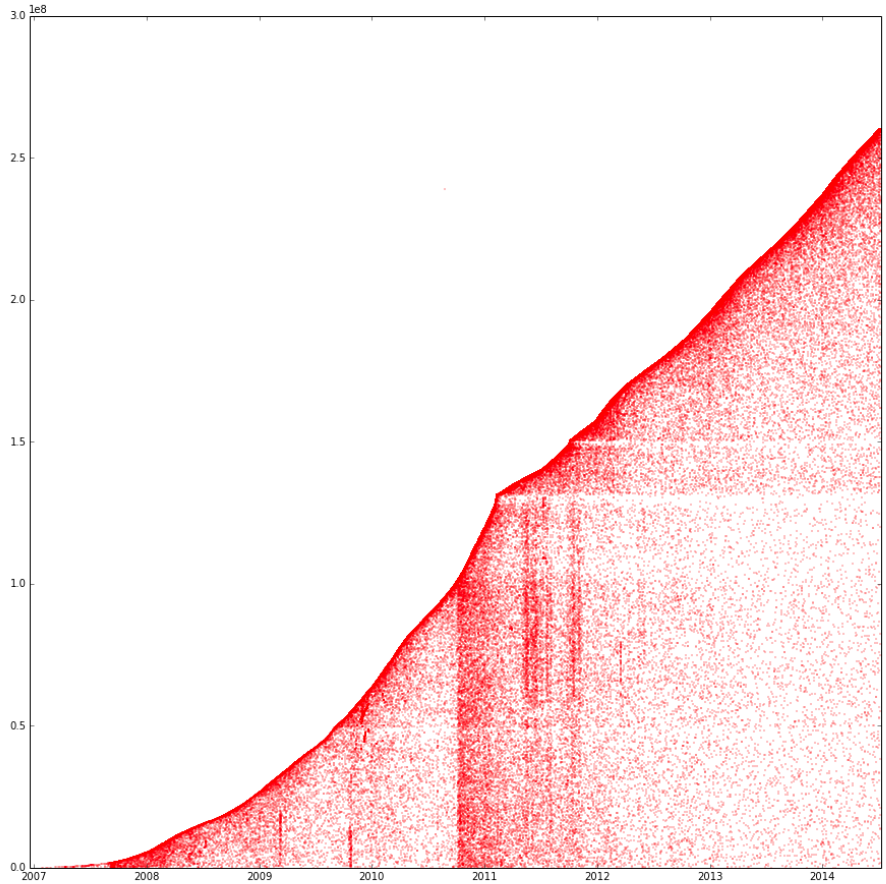
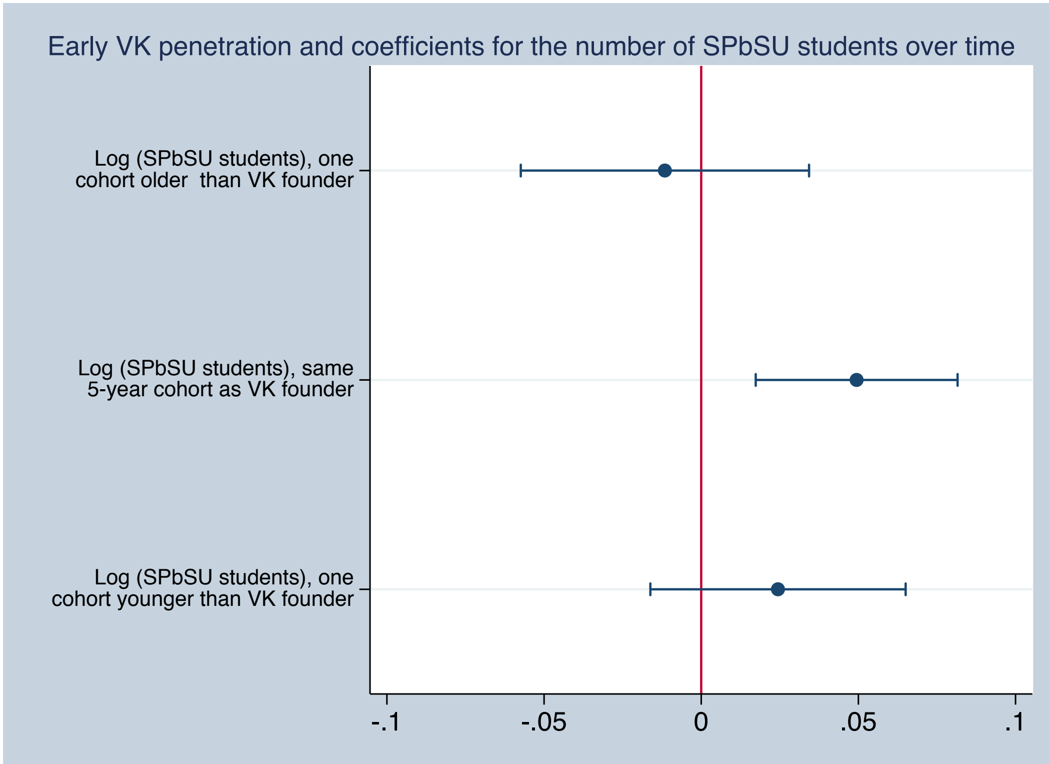


Figure A2. VK Penetration in November 2006 and SPbSU student cohorts.



Figures A3. Nonparametric Relationship between VK Penetration and Number of Protesters (in shares).

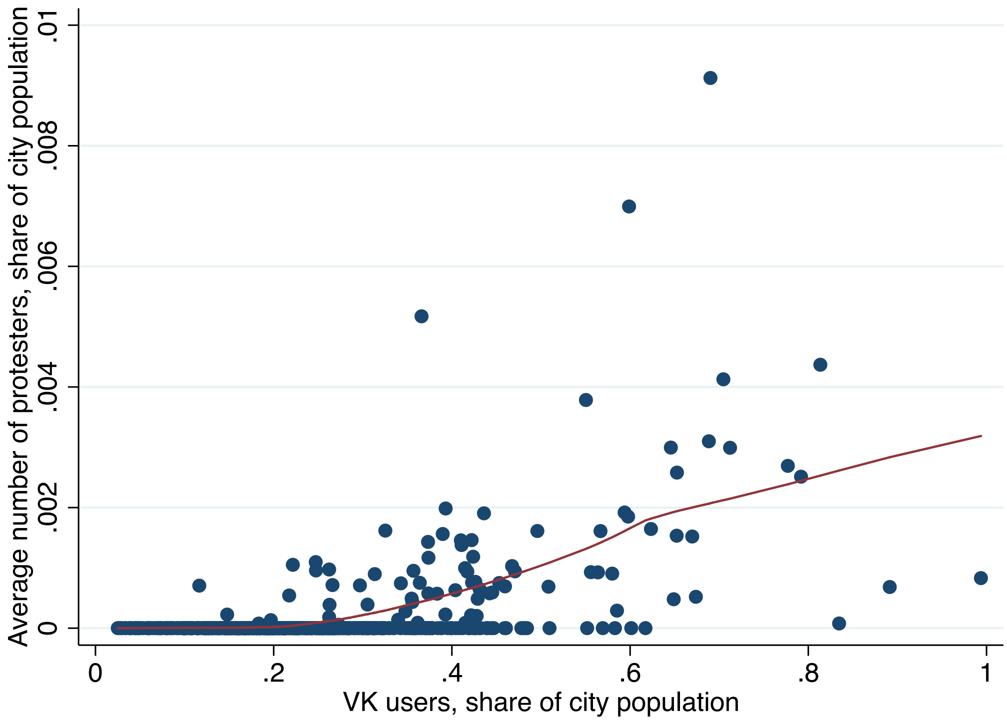
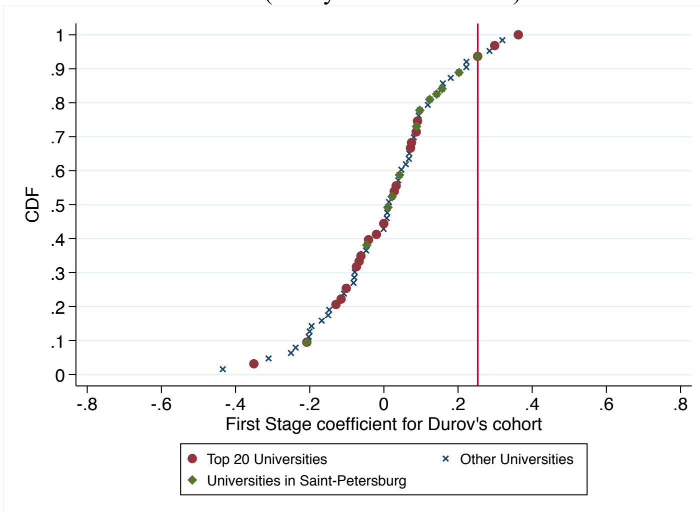
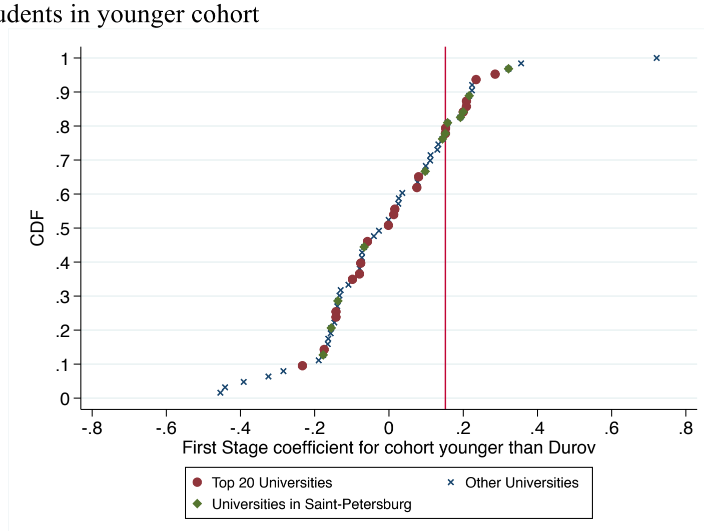


Figure A4. Reduced Form Coefficients for 65 Universities in Russia.

A: Distribution of students in Durov's cohort (+- 2 years from Durov).




B: Distribution of students in younger cohort




C: Distribution of students in older cohort

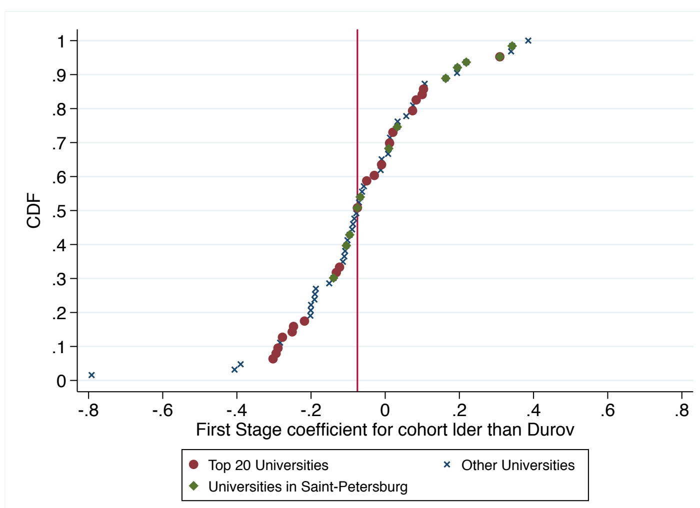
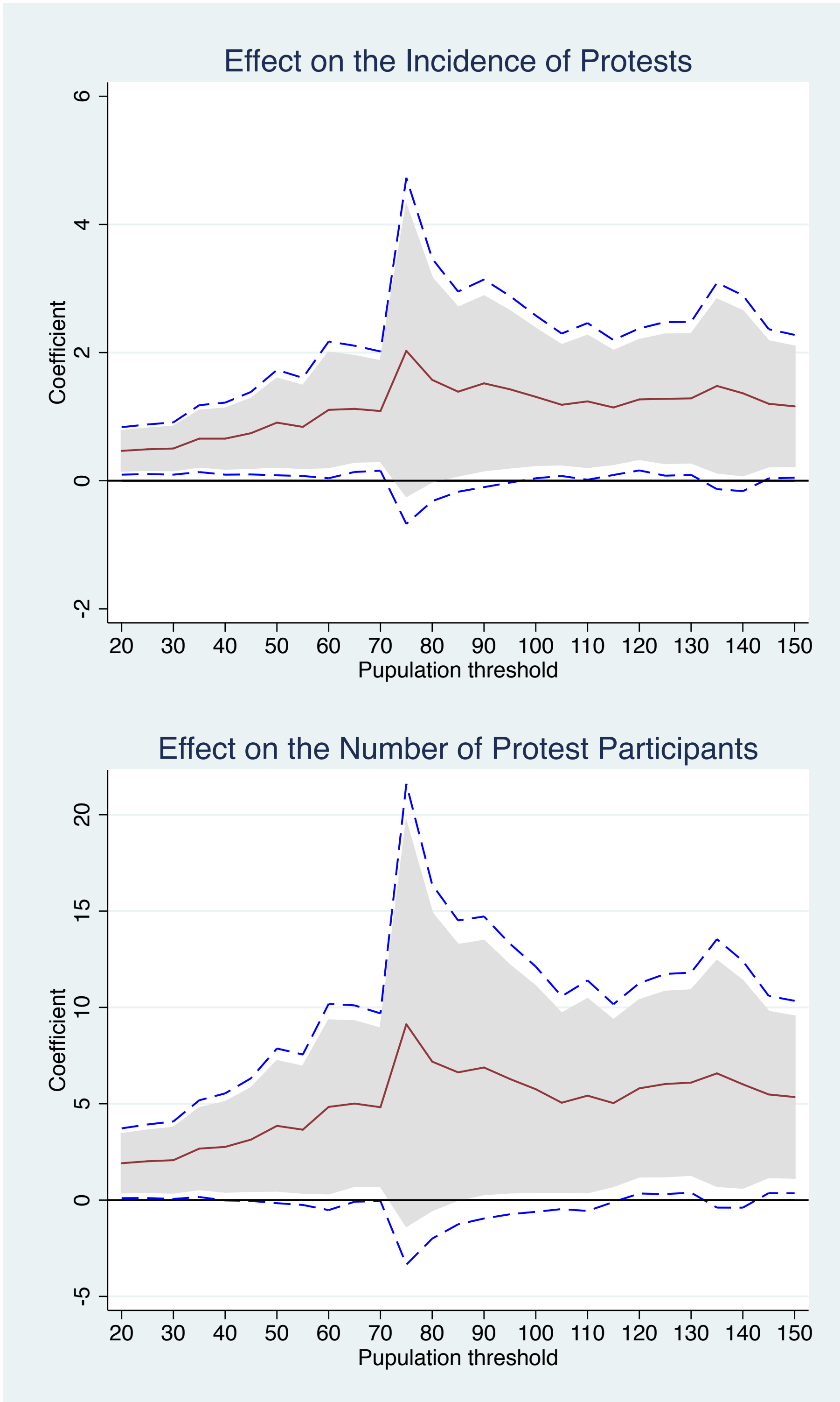


Figure A5. Magnitude of the Effect as a Function of Population Threshold.




Notes: The graphs show the magnitude of the coefficients for Log (VK users in 2011) on Log (protesters in December 2011) for specification reported in column 1 of Table 4 (upper panel) and a dummy for protest event in December 2011 for specification reported in column 1 of Table 5 (lower panel) as a function of the population threshold (in thousands). Grey areas show the 10% confidence intervals. Dashed lines show 5% confidence intervals.

**Table A1. Summary statistics.**

| | Observations | Mean | Standard deviation | Median | Min | Max |
|---|---|---|---|---|---|---|
| Incidence of protests, Dec 2011 | 625 | 0.13 | 0.34 | 0 | 0 | 1 |
| Incidence of protests, USSR, 1987-1992 | 625 | 0.22 | 0.41 | 0 | 0 | 1 |
| Incidence of pro-democratic protests, USSR, 1987-1992 | 625 | 0.18 | 0.38 | 0 | 0 | 1 |
| Incidence of anti-monetization protests, 2005 | 625 | 0.19 | 0.39 | 0 | 0 | 1 |
| Incidence of labor protests, 1997-2000 | 625 | 0.61 | 0.49 | 1 | 0 | 1 |
| Log (number of protest participants), Dec 2011 | 625 | 0.77 | 2.02 | 0 | 0 | 8.66 |
| Log (number of protest participants), USSR, 1987-1992 | 625 | 1.41 | 2.77 | 0 | 0 | 12.99 |
| Log (number of participants in pro-democratic protests), USSR, 1987-1992 | 625 | 1.38 | 3.08 | 0 | 0 | 13.93 |
| Log (number of participants in anti-monetization protests), 2005 | 625 | 1.28 | 2.7 | 0 | 0 | 9.21 |
| Log (number of participants in labor protests), 1997-2000 | 625 | 3.8 | 3.42 | 4.39 | 0 | 11.76 |
| Log (number of VK users), Aug 2011 | 625 | 9.54 | 1.33 | 9.31 | 6.61 | 13.84 |
| Log (number of early VK users), Nov 2006 | 625 | 0.08 | 0.3 | 0 | 0 | 3.5 |
| Log (number of VK users), 2013 | 625 | 10.13 | 1.27 | 9.84 | 7.65 | 14.3 |
| Log (number of Odnoklassniki users), 2014 | 625 | 10.72 | 1.12 | 10.45 | 7.94 | 14.36 |
| Log (number of Facebook users), 2013 | 625 | 6.9 | 2.06 | 6.76 | 0 | 12.3 |
| Log (SPbSU students, same 5-year cohort as VK founder) | 625 | 0.49 | 0.75 | 0 | 0 | 4.64 |
| Log (SPbSU students, one cohort younger than VK founder) | 625 | 0.4 | 0.63 | 0 | 0 | 2.77 |
| Log (SPbSU students, one cohort older than VK founder) | 625 | 0.44 | 0.7 | 0 | 0 | 3.53 |
| Internet penetration, region-level, 2011 | 625 | 0.27 | 0.17 | 0.22 | 0.01 | 0.63 |
| Population, in thousands, 2010 | 625 | 117.68 | 189.63 | 52.7 | 20 | 1393.5 |
| Regional center | 625 | 0.12 | 0.32 | 0 | 0 | 1 |
| Rayon center (county seat) | 625 | 0.79 | 0.41 | 1 | 0 | 1 |
| Distance to Saint Petersburg, km | 625 | 1481.62 | 839.41 | 1419 | 21.7 | 4646 |
| Distance to Moscow, km | 625 | 1152.76 | 875.97 | 1014 | 15.75 | 4174 |
| Log (average wage), 2011 | 625 | 9.89 | 0.35 | 9.83 | 9.08 | 11.19 |
| Log (number of people with age 20-24), 2010 | 625 | 8.46 | 1.07 | 8.19 | 6.79 | 11.83 |
| Log (number of people with age 25-29), 2010 | 625 | 8.53 | 1.02 | 8.28 | 6.83 | 11.84 |
| Log (number of people with age 30-34), 2010 | 625 | 8.47 | 1 | 8.21 | 6.8 | 11.69 |
| Log (number of people with age 35-39), 2010 | 625 | 8.41 | 0.99 | 8.16 | 6.84 | 11.59 |
| Log (number of people with age 40-44), 2010 | 625 | 8.27 | 0.99 | 8.03 | 6.78 | 11.41 |
| Log (number of people with age 45-49), 2010 | 625 | 8.42 | 0.97 | 8.21 | 6.79 | 11.52 |
| Log (number of people with age 50 and older), 2010 | 625 | 9.93 | 0.97 | 9.71 | 8.27 | 13.08 |
| % with higher education, 2002 | 625 | 0.15 | 0.06 | 0.13 | 0.05 | 0.45 |
| % with higher education among age 20-24, 2010 | 625 | 0.18 | 0.06 | 0.17 | 0.05 | 0.37 |
| % with higher education among age 25-29, 2010 | 625 | 0.34 | 0.1 | 0.33 | 0.11 | 0.67 |
| % with higher education among age 30-34, 2010 | 625 | 0.31 | 0.1 | 0.3 | 0.12 | 0.67 |
| % with higher education among age 35-39, 2010 | 625 | 0.28 | 0.08 | 0.26 | 0.13 | 0.58 |
| % with higher education among age 40-44, 2010 | 625 | 0.25 | 0.08 | 0.23 | 0.12 | 0.6 |
| % with higher education among age 45-49, 2010 | 625 | 0.23 | 0.08 | 0.21 | 0.09 | 0.62 |
| % with higher education among age 50-54, 2010 | 625 | 0.17 | 0.07 | 0.15 | 0.07 | 0.55 |
| Presence of a university in a city, 2011 | 625 | 0.15 | 0.35 | 0 | 0 | 1 |
| Ethnic fractionalization, 2010 | 625 | 0.2 | 0.17 | 0.14 | 0.01 | 0.85 |

**Table A2. 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 younger 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 A3. Determinants of Early VK Penetration.**

| | Log (number of early VK users), Nov 2006 | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Log (SPbSU students), same 5-year cohort as VK founder | 0.0925*** | 0.0510*** | 0.0517*** | 0.0495** | 0.0514*** | 0.0484** | 0.0509*** |
| | [0.0206] | [0.0190] | [0.0182] | [0.0193] | [0.0187] | [0.0194] | [0.0190] |
| Log (SPbSU students), one cohort younger than VK founder | 0.0756** | 0.0170 | 0.0194 | 0.0244 | 0.0278 | 0.0223 | 0.0265 |
| | [0.0378] | [0.0247] | [0.0228] | [0.0244] | [0.0250] | [0.0246] | [0.0245] |
| Log (SPbSU students), one cohort older than VK founder | 0.0343 | -0.0113 | -0.0071 | -0.0116 | -0.0143 | -0.0133 | -0.0144 |
| | [0.0324] | [0.0223] | [0.0276] | [0.0275] | [0.0271] | [0.0278] | [0.0276] |
| Regional center | | -0.0806 | -0.1068 | -0.1348 | -0.1208 | -0.1317 | -0.1430 |
| | | [0.0619] | [0.0706] | [0.0874] | [0.0885] | [0.0852] | [0.0866] |
| Distance to Saint Petersburg, km | | | 0.0001 | 0.0001 | 0.0001 | 0.0000 | 0.0000 |
| | | | [0.0000] | [0.0001] | [0.0001] | [0.0001] | [0.0001] |
| Distance to Moscow, km | | | -0.0000 | -0.0000 | -0.0000 | -0.0000 | -0.0000 |
| | | | [0.0000] | [0.0000] | [0.0000] | [0.0001] | [0.0000] |
| Rayon center (county seat) | | | -0.0132 | -0.0088 | -0.0148 | -0.0166 | -0.0083 |
| | | | [0.0131] | [0.0120] | [0.0141] | [0.0129] | [0.0139] |
| Log (average wage), city-level, 2011 | | | 0.0523 | 0.0365 | 0.0015 | 0.0167 | 0.0470 |
| | | | [0.0323] | [0.0315] | [0.0324] | [0.0303] | [0.0348] |
| Presence of a university in a city, 2011 | | | | 0.0951 | 0.1041 | 0.0976 | 0.0998 |
| | | | | [0.0631] | [0.0635] | [0.0615] | [0.0624] |
| Internet penetration, region-level, 2011 | | | | 0.0341 | 0.0222 | 0.0208 | 0.0289 |
| | | | | [0.0456] | [0.0437] | [0.0473] | [0.0499] |
| Log (number of Odnoklassniki users), 2014 | | | | -0.0194 | -0.0049 | -0.0082 | -0.0053 |
| | | | | [0.0200] | [0.0176] | [0.0195] | [0.0190] |
| Ethnic fractionalization, 2010 | | | | -0.0862 | -0.0712 | -0.0724 | -0.0630 |
| | | | | [0.0816] | [0.0856] | [0.0750] | [0.0758] |
| Observations | 625 | 625 | 625 | 625 | 625 | 625 | 625 |
| R-squared | 0.1805 | 0.5185 | 0.5333 | 0.5387 | 0.5470 | 0.5427 | 0.5452 |
| Population controls | | Yes | Yes | Yes | Yes | Yes | Yes |
| Age cohort controls | | | Yes** | Yes | Yes** | Yes | Yes |
| Education 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. Logarithm of any variable is calculated with 1 added inside. When "Yes" is added to indicate 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. Education controls include the share of population with higher education overall according to 2002 Russian Census and separately in each of the age cohorts according to 2010 Russian Census, to account for both the levels and the change in education. 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 A4. VK and penetration of Odnoklassniki**

| | Log (number of Odnoklassniki users), 2014 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Log (early VK users, from first 5000 users) | -0.042 | | | | | | | |
| | [0.059] | | | | | | | |
| Log (early VK users, from first 50000 users) | | -0.045 | | | | | | |
| | | [0.034] | | | | | | |
| Log (early VK users, from first 100000 users) | | | -0.024 | | | | | |
| | | | [0.037] | | | | | |
| Log (number of VK users), Aug 2011 | | | | 0.074 | | | | |
| | | | | [0.074] | | | | |
| Log (SPbSU students), same 5-year cohort as VK founder | | | | | 0.028 | 0.018 | 0.015 | 0.017 |
| | | | | | [0.048] | [0.045] | [0.043] | [0.045] |
| Log (SPbSU students), one cohort younger than VK founder | | | | | 0.084 | 0.072 | 0.073 | 0.068 |
| | | | | | [0.052] | [0.050] | [0.052] | [0.051] |
| Log (SPbSU students), one cohort older than VK founder | | | | | -0.049 | -0.032 | -0.034 | -0.028 |
| | | | | | [0.043] | [0.042] | [0.042] | [0.040] |
| Regional center | 0.259** | 0.269** | 0.267** | 0.252** | 0.261** | 0.228** | 0.256** | 0.254** |
| | [0.123] | [0.119] | [0.119] | [0.115] | [0.116] | [0.110] | [0.111] | [0.110] |
| Distance to Saint Petersburg, km | 0.001*** | 0.001*** | 0.001*** | 0.001*** | 0.001*** | 0.001*** | 0.001*** | 0.001*** |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| Distance to Moscow, km | -0.000*** | -0.000*** | -0.000*** | -0.000*** | -0.000*** | -0.001*** | -0.001*** | -0.001*** |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| Rayon center (county seat) | 0.053 | 0.051 | 0.053 | 0.054 | 0.046 | 0.071 | 0.073 | 0.067 |
| | [0.077] | [0.078] | [0.078] | [0.075] | [0.078] | [0.083] | [0.086] | [0.077] |
| Log (average wage), city-level, 2011 | 0.111 | 0.115 | 0.115 | 0.104 | 0.124 | 0.250** | 0.205** | 0.181* |
| | [0.106] | [0.105] | [0.106] | [0.105] | [0.104] | [0.097] | [0.100] | [0.104] |
| Presence of a university in a city, 2011 | 0.029 | 0.042 | 0.035 | 0.011 | 0.026 | -0.011 | -0.029 | -0.015 |
| | [0.097] | [0.098] | [0.097] | [0.097] | [0.095] | [0.088] | [0.082] | [0.087] |
| Internet penetration, region-level, 2011 | -0.479** | -0.469** | -0.467** | -0.492** | -0.471** | -0.354* | -0.287 | -0.354* |
| | [0.204] | [0.204] | [0.205] | [0.211] | [0.203] | [0.197] | [0.185] | [0.197] |
| Ethnic fractionalization, 2010 | -0.231 | -0.237 | -0.231 | -0.261 | -0.259 | -0.205 | -0.247 | -0.303** |
| | [0.166] | [0.167] | [0.167] | [0.162] | [0.168] | [0.162] | [0.163] | [0.151] |
| Population controls | Yes** | Yes** | Yes** | Yes** | Yes** | Yes** | Yes** | Yes** |
| Age cohort controls | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** |
| Education controls | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** |
| Observations | 625 | 625 | 625 | 625 | 625 | 625 | 625 | 625 |
| R-squared | 0.892 | 0.892 | 0.892 | 0.892 | 0.893 | 0.899 | 0.903 | 0.902 |

*** 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. Logarithm of any variable is calculated with 1 added inside. When "Yes" is added to indicate 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. Education controls include the share of population with higher education overall according to 2002 Russian Census and separately in each of the age cohorts according to 2010 Russian Census, to account for both the levels and the change in education. Results in columns (1)-(4) are robust to inclusion of electoral controls, but corresponding specifications are not shown to save space.

## Table A5. Online protest communities and protest participation. OLS estimates.

| | Log (number of protesters), Dec 2011 | | | | Incidence of protests, dummy, Dec 2011 | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Log (# of members in VK protest community in a city) | 0.121** | 0.120** | 0.119** | 0.121** | 0.030*** | 0.030*** | 0.030*** | 0.030*** |
| | [0.050] | [0.050] | [0.050] | [0.050] | [0.009] | [0.009] | [0.009] | [0.009] |
| Observations | 625 | 625 | 625 | 625 | 625 | 625 | 625 | 625 |
| R-squared | 0.824 | 0.827 | 0.829 | 0.826 | 0.783 | 0.786 | 0.787 | 0.786 |
| Population controls | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** |
| Age cohort controls | Yes | Yes | Yes* | Yes | Yes** | Yes** | Yes** | Yes** |
| Education 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. Logarithm of any variable is calculated with 1 added inside. 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. Education controls include the share of population with higher education overall according to 2002 Russian Census and separately in each of the age cohorts according to 2010 Russian Census, to account for both the levels and the change in education. 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), share of people with higher education in 2002, internet penetration in 2011, log (Odnoklassniki users in 2014).

**Table A6. VK penetration and protest participation. Heterogeneity of the effect. IV estimates.**

| | Log (number of protesters), Dec 2011 | | | | | |
|---|---|---|---|---|---|---|
| | Wage lower than median | Wage higher than median | Trust lower than median | Trust higher than median | Education lower than median | Education higher than median |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Log (number of VK users), Aug 2011 | 1.252 | 2.021** | 0.144 | 3.843** | 0.139 | 4.448 |
| | [1.602] | [0.995] | [1.830] | [1.558] | [0.210] | [2.789] |
| Log (SPbSU students), one cohort younger than VK founder | 0.031 | 0.314** | 0.152 | -0.076 | -0.032 | 0.117 |
| | [0.159] | [0.158] | [0.141] | [0.335] | [0.049] | [0.275] |
| Log (SPbSU students), one cohort older than VK founder | -0.094 | -0.210 | 0.213 | -0.675* | -0.040 | -0.348 |
| | [0.171] | [0.210] | [0.237] | [0.362] | [0.045] | [0.344] |
| Population controls | Yes** | Yes*** | Yes*** | Yes*** | Yes*** | Yes*** |
| Age cohort controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Education controls | Yes | Yes | Yes*** | Yes | Yes | Yes |
| Observations | 315 | 310 | 231 | 231 | 313 | 312 |
| Effective F-statistics (Olea Montiel and Pflueger 2013) | 141.3 | 113.2 | 125.4 | 104.0 | 42.46 | 197.6 |

*** 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. Logarithm of any variable is calculated with 1 added inside. Specification is the same as Table 4A, column (1) (only baseline controls included). When "Yes" is added to indicate 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. Education controls include the share of population with higher education overall according to 2002 Russian Census and separately in each of the age cohorts according to 2010 Russian Census, to account for both the levels and the change in education. Other controls include dummy for regional and county centers, distances to Moscow and St Petersburg, log (average wage), internet penetration in 2011, log (Odnoklassniki users in 2014).

Table A7. 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. The data was gathered manually from the open sources in the internet. Where possible, three participation 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 are treated equally. For protests with more than one estimate, an average number of participants 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 (see above). We identified 75 various demands in the dataset which we considered either anti-Soviet or pro-democratic. Examples of such demands are "Against Communist Party Privileges", "Decentralize Economic Administration", "Democratization of Political institutions", etc. A full list of anti-Soviet/pro-democratic demands is available upon request. For protests with more than one estimate of participation, an average number of participants 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 living in a given city, as of 2013. 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. By "valid" we mean “not blocked”. By "active" we mean that they were seen online at least once between June 21 and 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 5,000 VK users** | 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 were registered within less than a month in November 2006. |
| **Number of Odnoklassniki users in 2014** | Number of all registered Odnoklassniki users living in a given city, as of 2014. Manually collected data. |
| **Number of Facebook users in 2013** | Number of all registered Facebook users living in a given city, as of 2013. Manually collected data. |
| **Population in 2001, in thousands** | Collected from mojgorod.ru, which in turn stores data collected 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 2011** | Data gathered from Russian Federal State Statistics Service. |
| **Number of people with age xx-xx in 2010** | Data gathered from Russian Federal State Statistics Service. Based on Russian census in 2010. |
| **Presence of university** | 1 = city has at least one university; 0 = not. Data collected from Wikipedia. |
| **Percentage with higher education in 2010** | Percentage of adults with at least one university degree. Data gathered from Russian Federal State Statistics Service. Based on Russian census in 2010. |
| **Internet penetration in 2011, region-level** | Number of unique users in a region divided by its population according to the 2010 census. Data collected from liveinternet.com |
| **Number of SPbSU students, same 5-year cohort as VK founder** | Number of Odnoklassniki users who studied in Saint Petersburg State University in classes of 2004-2008, i.e. in the same age 5-year cohort together with Pavel Durov, former CEO of VK. Data manually collected from OK.ru. |
| **Number of SPbSU students, one cohort younger than VK founder** | Number of Odnoklassniki users who studied in Saint Petersburg State University in classes of 1999-2003, i.e. one 5-year cohort earlier than Pavel Durov, former CEO of VK. Data manually collected from OK.ru. |
| **Number of SPbSU students, one cohort older than VK founder** | Number of Odnoklassniki users who studied in Saint Petersburg State University in classes of 2009-2013, i.e. one 5-year cohort after Pavel Durov, former CEO of VK. Data manually collected from OK.ru. |