# Reevaluating the Middle Class Protest Paradigm: A Case-Control Study of Democratic Protest Coalitions in Russia

Bryn Rosenfeld

University of Southern California

# A Supplementary Appendix

## A.1 About the Surveys

### The Levada Center Protest Surveys

The Levada Center surveyed protesters at four demonstrations in Moscow from late 2011 through early 2013. A respected independent polling organization, the Levada Center used international best-practices in protest survey methodology to ensure representativeness in the achieved sample (Walgrave and Verhulst 2011). Stationary protests in Russia are generally cordoned off by special forces and riot police, creating a defined perimeter for the protest event. During the post-election demonstrations, police established entry points and required protesters to pass through metal detectors. Interviewers took advantage of these procedures to randomly sample protesters at a fixed skip interval. When this was not possible, for example, at moving demonstrations and marches, interviewers moved systematically though the columns of participants, selecting every n-*th* respondent at an interval set by fieldwork supervisors. Specifically, interviewers were instructed, upon completing an interview, to look at the row of marchers in front of them and count, typically, the 5-*th* respondent to their right/left, zigzagging through the column as they moved through rows of marchers. If the selected respondent refused to participate, interviewers were instructed to select

Bryn Rosenfeld, Assistant Professor, Department of Political Science, University of Southern California, 3518 Trousdale Parkway, Von Kleinsmid Center (VKC), Los Angeles, CA 90089. Phone: (213) 740–6974. Email: brosenfe@usc.edu

a new respondent at the same fixed skip interval. If the selected respondent was with a group, interviewers were instructed to ensure some distance from the group before beginning so that the respondent's answers were not overheard by others. In this way, interviewers worked their way from one end of the crowd to the other randomly selecting respondents without the use of quotas. Because protesters generally assembled by ideology, political party, or identity, this procedure ensured that all groups were represented. Interviews ranged in length from 5-15 minutes.[1] The dates and sample sizes of the protest surveys are as follows: December 24, 2011 (n=791), February 4, 2012 (n=1,344), September 15, 2012 (n=805), January 13, 2013 (n=325).

In each case, the surveys were agreed upon in advance with the protest organizers, who shared details about timing, entry points, location, and route that were taken into account in designing the sampling plan. All interviewers also carried a signed and stamped letter verifying their status from the director of the Levada Center, a well-known and respected independent outfit.

Experienced professional supervisors were on-site during the protests, observing the interviewing process and ensuring that interviewers followed the protocol. They were also available throughout the course of the interviewing to answer enumerators' questions.

Levada systematically gathered information about response rates and refusals to participate. For the first survey wave, interviewers were asked to record the gender and estimated age (in four categories) of all individuals who declined to take part. In each of the subsequent waves, interviewers

[1]During these brief interviews, respondents were not, for example, asked to assess either national economic conditions or their own personal economic well-being (except in terms of household consumption). Including standard questions on these topics, using wording consistent with the representative population survey, would have allowed for a more detailed analysis of economic grievances as a driver of protest. More generally, in choosing between nationally representative surveys that ask about specific protest events and surveys conducted at the the event itself, researchers face tradeoffs involving statistical power, the threat of omitted variables, and post-treatment bias. Thus, the choice of data must be motivated by the goals of the research with these caveats in mind.

were asked to report the total number of refusals along with the total number of interviews they had completed. In addition, interviewers were asked to answer several questions about themselves (gender, age, how they voted in the parliamentary election), which Levada used to ensure that interviewer characteristics did not unduly influence response rates.

Levada reports the following deviations from the sampling plan; their assessment and mine is that these issues did not significantly affect the survey results:

1. Not all interviewers were able to complete the number of interviews requested of them by their supervisors.

2. In addition to the respondents chosen at random according to the protocol, protest participants occasionally volunteered themselves as respondents.

Regarding the second point on self-selection into the survey sample, fieldwork supervisors provided guidance to interviewers about how to deal with such situations both in the interviewer instructions and orally. The instructions state that when approached by protest participants wishing to take part in the survey, enumerators should explain that the goal of the survey is "not only to know the opinions of individual protest participants but the opinions of participants as a whole. To that end, [I] must select respondents for the survey according to a protocol that has been strictly defined for me and which [I] do not have the right to alter."

Following the first protest survey, a question was added for interviewers at the end of the instrument about the random selection of respondents. The total share of non-randomly selected "volunteer" respondents is 4.6% across the three subsequent surveys. The main analyses in this paper focus on the last two survey waves in which this share was 4.2%. Excluding these respondents does not substantively change the results.

Again, the Levada Center is a very well-respected and professional research organization. The

instructions to interviewers were quite detailed and every effort was taken to document nonresponse, avoid deviations from the sampling plan, and ensure that interviewer characteristics did not bias the achieved sample.

**FOM GeoRating**

The Foundation for Public Opinion's (FOM) GeoRating surveys are both nationally and regionally representative of Russia's population. A minimum of 500 interviews are conducted in each region, including both the city of Moscow and Moscow oblast. Approximately, 93% of Russia's population lives in the 68 regions in which the survey was conducted. A three stage stratified random sample (at the level of the administrative region, the locale, and the household) was used. In total, 1954 locals were sampled, including 676 cities, 315 small towns, and 963 villages. Certain areas of the North Caucasus were excluded from the sample due to ongoing instability. The interview method was face-to-face in respondents' homes. Face-to-face surveys are the norm across the former-Soviet states where telephone penetration is generally insufficient to achieve a representative sample. The survey was conducted in November 2009. Its sample size is n=33,997.

When possible, I cross-checked the FOM data against the most recent Russian census data from 2010. The FOM sample closely mirrors the census data on gender, age, and education. The census data do not, however, include information about occupation, only broader measures of employment status (i.e. economically active vs. various categories of economic inactivity). The census data are archived at: `http://www.gks.ru/free_doc/new_site/perepis2010/croc/perepis_itogi1612.htm` and `http://moscow.gks.ru/wps/wcm/connect/rosstat_ts/moscow/ru/census_and_researching/census/national_census_2010/score_2010/score_2010_default`.

**FOM Monitoring Surveys**

The surveys referenced in Tables 3 and 4 were conducted by the Foundation for Public Opinion (FOM) in February, April, May, June, September, and October of 2009, among a nationally representative sample of the population, age 18 and over. Each of the surveys used the same multistage stratified random sampling design. The total sample size in each wave was 2,000 respondents. Additional Information can be found at `http://fom.ru` and in the following publications:

1. FOM. 2009. "Working Russians and the labor market." The Foundation for Public Opinion. Dominanty No. 23 `http://bd.fom.ru/pdf/d23rabros.pdf`.

2. FOM. 2009. "The Labor Market: The Position of Workers in the Labor Market." The Foundation for Public Opinion. Dominanty No. 38 `http://bd.fom.ru/pdf/d38trud.pdf`.

3. FOM. 2009. "The Labor Market: The Position of Workers in the Labor Market." The Foundation for Public Opinion. Dominanty No. 41 `http://bd.fom.ru/pdf/d41trud.pdf`.

**Russian Academy of Sciences Study**

The Institute of Sociology of the Russian Academy of Sciences surveyed a nationally representative sample of the population, age 18 and over in February 2014, using a multistage stratified random sampling design with quotas applied at two stages: at the regional level by urbanity and in the selection of respondents by gender, age, and settlement type. The total sample size was 1600 respondents. The same questionnaire was also administered to an oversample of 300 respondents with middle class educational and occupational characteristics. Additional information can be found in Russian Academy of Sciences (2014).

**Levada Center Survey of the Urban Middle Class**

The Levada Center's survey of the Russian middle class was conducted in 14 large and medium-sized cities among professionals, managers, entrepreneurs and business owners with at least some higher education. The monthly income cutoff was 60,000 rubles per family member in Moscow, 50,000 rubles in Saint Petersburg, and 35,000 rubles in all other cities. Gender and age quotas were applied in the selection of respondents. Fieldwork was conducted May 5-19, 2011 in Ekaterinburg, Kazan, Khabarovsk, Krasnodar, Krasnoiarsk, Moscow, Nizhny Novgorod, Novosibirsk, Omsk, Perm, Rostov, Saint Petersburg, Samara, Voronezh.

### A.1.1 Question Wordings

[`consumption`] To which of the following groups of the population do you belong?

1. We don't have enough money even for food;
2. We have enough money for food, but not for clothing;
3. We have enough money for food and clothing, but the purchase of more expense things, like a TV or refrigerator, is problematic;
4. We can buy some expensive things, like a refrigerator or TV, but we cannot buy a car;
5. We can buy a car, but we cannot say that we have no financial constraints;
6. We do not have to refuse ourselves anything.

[`motivation`] What brought you to this protest? (maximum of three responses allowed)

1. A desire to change the policies of those in power, to begin reforms;
2. Anger at the actions of those in power, the repression of protest participants and leader;
3. The goal of regime change, Putin's resignation and that of his team;
4. I felt deceived: the reforms promised by those in power [that is, the reforms begun under Medvedev] turned out to be fiction;
5. Solidarity with the position of one of the parties or leaders who organized the protest;
6. A desire to show my support for the ideas of the protest movement;
7. A sense of moral obligation, I couldn't do otherwise;
8. I want to live well;

9. My friends or acquaintances came to the protest and I came with them;
10. I just came to have a look;
11. Difficult to say.

[`resolve`] How do you think the issues that you consider most important for Russia can be resolved? (maximum of three responses allowed)

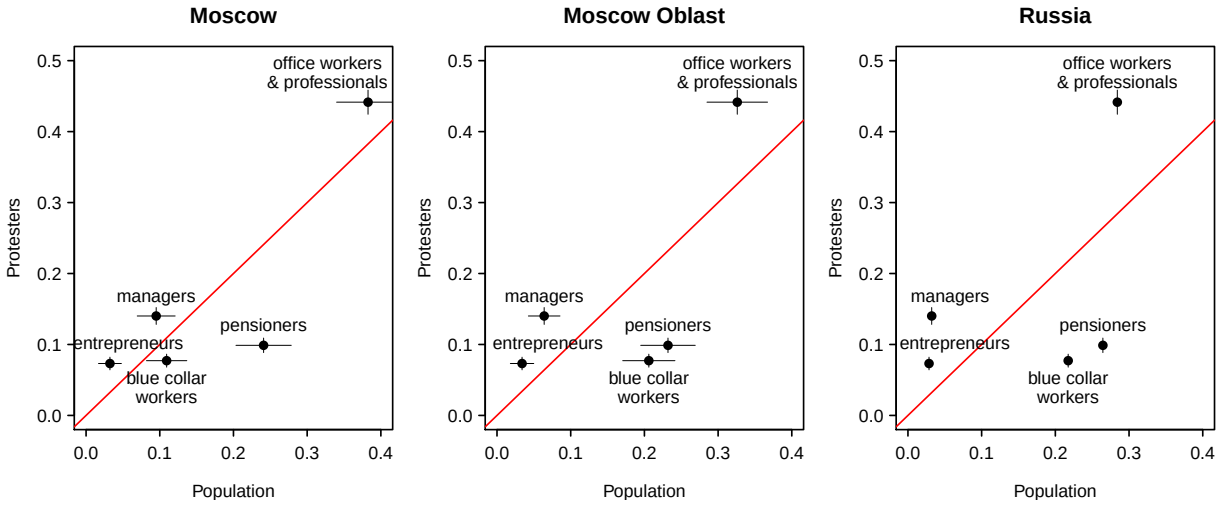
1. Fight for power in elections within the framework of existing laws;
2. Focus on informing citizens and recruiting new supporters;
3. Form a civic body to monitor the actions of those in power;
4. Develop independent media;
5. Participate in elections, voting for the parties that offer a real platform;
6. Form a broad coalition of civic groups, representing various interests;
7. Search for avenues of cooperation at various levels of power with different civic groups (unions, environmental groups, voluntary organizations);
8. Utilize the courts more broadly, filing suits against those in power who have broken the law;
9. Continue to attend peaceful protests and demonstrations;
10. Organize radical acts of protest with the goal of regime change;
11. Difficult to say.

## A.2 Additional Descriptive Tables and Figures

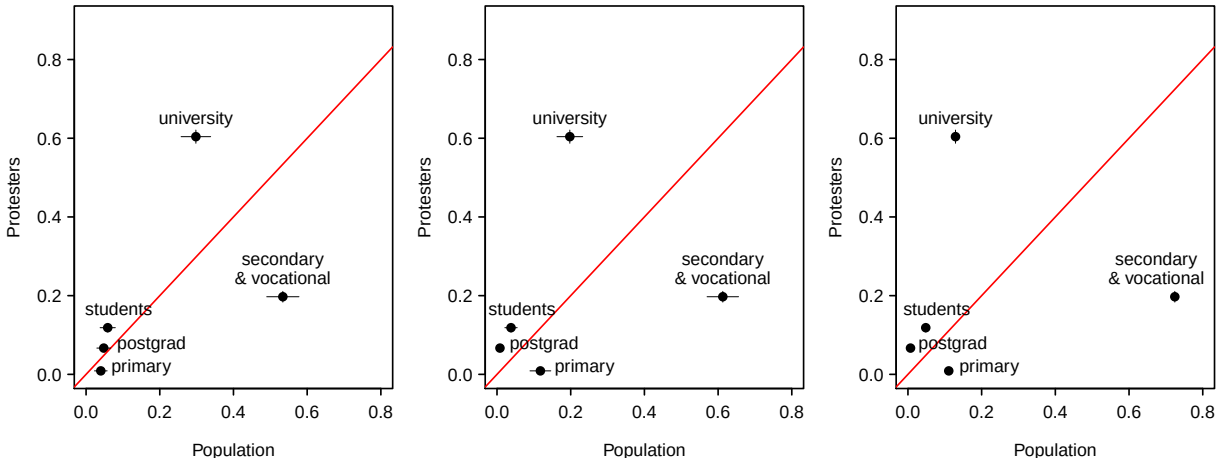
| | Moscow | Moscow Region | Russia |
|---|---|---|---|
| *Overall* | | | |
| Middle class | 2.08 | 2.99 | 5.64 |
| Non-middle class | 0.50 | 0.44 | 0.39 |
| State | 0.65 | 0.69 | 0.68 |
| Private | 1.12 | 1.10 | 1.10 |
| *Class by sector* | | | |
| State middle class | 1.18 | 1.67 | 2.55 |
| Non-state middle class | 2.13 | 3.09 | 6.85 |
| State non-middle class | 0.19 | 0.16 | 0.14 |
| Non-state non-middle class | 0.76 | 0.67 | 0.60 |
| *By occupation* | | | |
| Office workers & professionals | 1.15 | 1.35 | 1.55 |
| Managers | 1.47 | 2.19 | 4.35 |
| Entrepreneurs | 2.26 | 2.15 | 2.56 |
| Blue-collar workers | 0.71 | 0.37 | 0.36 |
| Pensioners | 0.41 | 0.43 | 0.37 |
| *By education* | | | |
| Post-graduate | 1.39 | 8.36 | 9.61 |
| University degree | 2.03 | 3.06 | 4.68 |
| University students | 2.02 | 3.12 | 2.45 |
| Secondary & vocational | 0.37 | 0.32 | 0.27 |
| Primary | 0.22 | 0.08 | 0.08 |

Source: The Levada Center protest surveys and FOM GeoRating

Table A1: **Index of Representation**. This index was constructed by dividing the share of protesters comprised by each demographic group by its share of the population for three referent populations: Moscow, Moscow Region, and Russia. An index of 1.00 means that a demographic group's representation in the protests was proportional to its share of the population. An index of greater than 1.00 means that it was overrepresented, while an index of less than 1.00 means that it was underrepresented.




(a) Occupation




(b) Education

Figure A1: **Detailed Demographic Comparison of Protesters and the Population.** The top panel of this figure compares the occupational characteristics of protesters with three referent populations from which protesters were recruited (Moscow, the Moscow region, and Russia). It shows that middle class professions (office workers and professionals, managers, and entrepreneurs) were overrepresented among the protesters, whereas blue-collar workers were underrepresented. The bottom panel shows the comparison in terms of educational attainment. While the university-educated were substantially overrepresented, those with secondary and vocational training were underrepresented. As in Figure 1, the location on the y-axis of each plotted point gives the share of protesters with a given profile, while position on the x-axis gives that same profile's share of the population. The 45 degree red line thus indicates proportional representation. The vertical bars are 95% confidence intervals based on the sample size of the pooled protest dataset. Horizontal bars indicate 95% confidence intervals based on the sample size of the population data. Data sources: The Levada Center and the Foundation for Public Opinion (FOM).

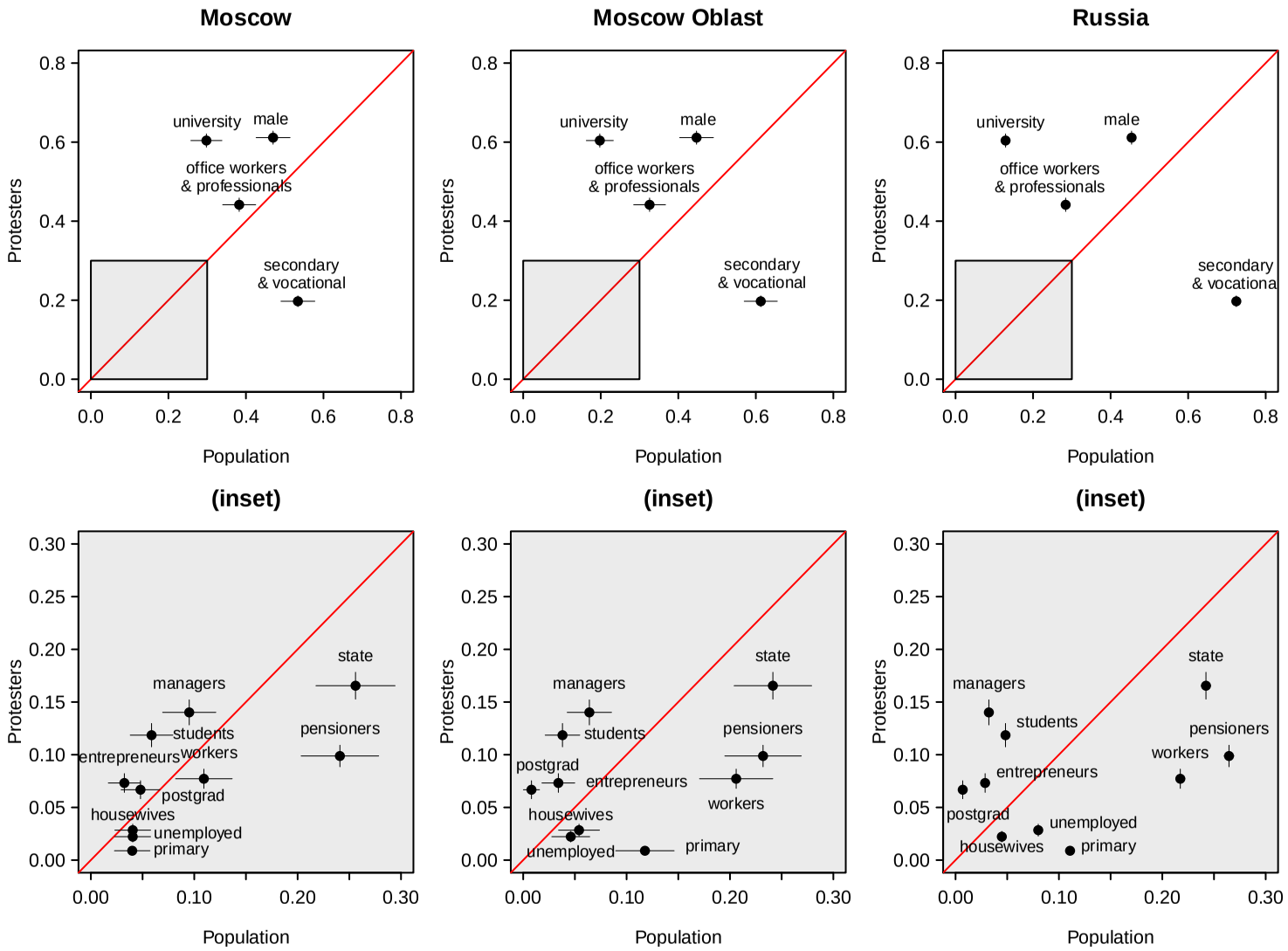



Figure A2: **Pooled Demographic Comparison of Protesters and the Population**. This figure compares the demographic composition of the protesters with three referent populations from which protesters were recruited. The regions shaded grey in the first row of plots are magnified in the second row of plots. The x-axis gives the share of the *population* with a given demographic characteristic, while the y-axis gives the share of *protesters* with that same characteristic. Points that lie above the 45 degree red line thus indicate groups that were overrepresented among the protesters, points below the line groups that were underrepresented. The vertical bars are 95% confidence intervals for the pooled protest data. Horizontal bars give 95% confidence intervals for the population data. Data sources: The Levada Center and the Foundation for Public Opinion (FOM).

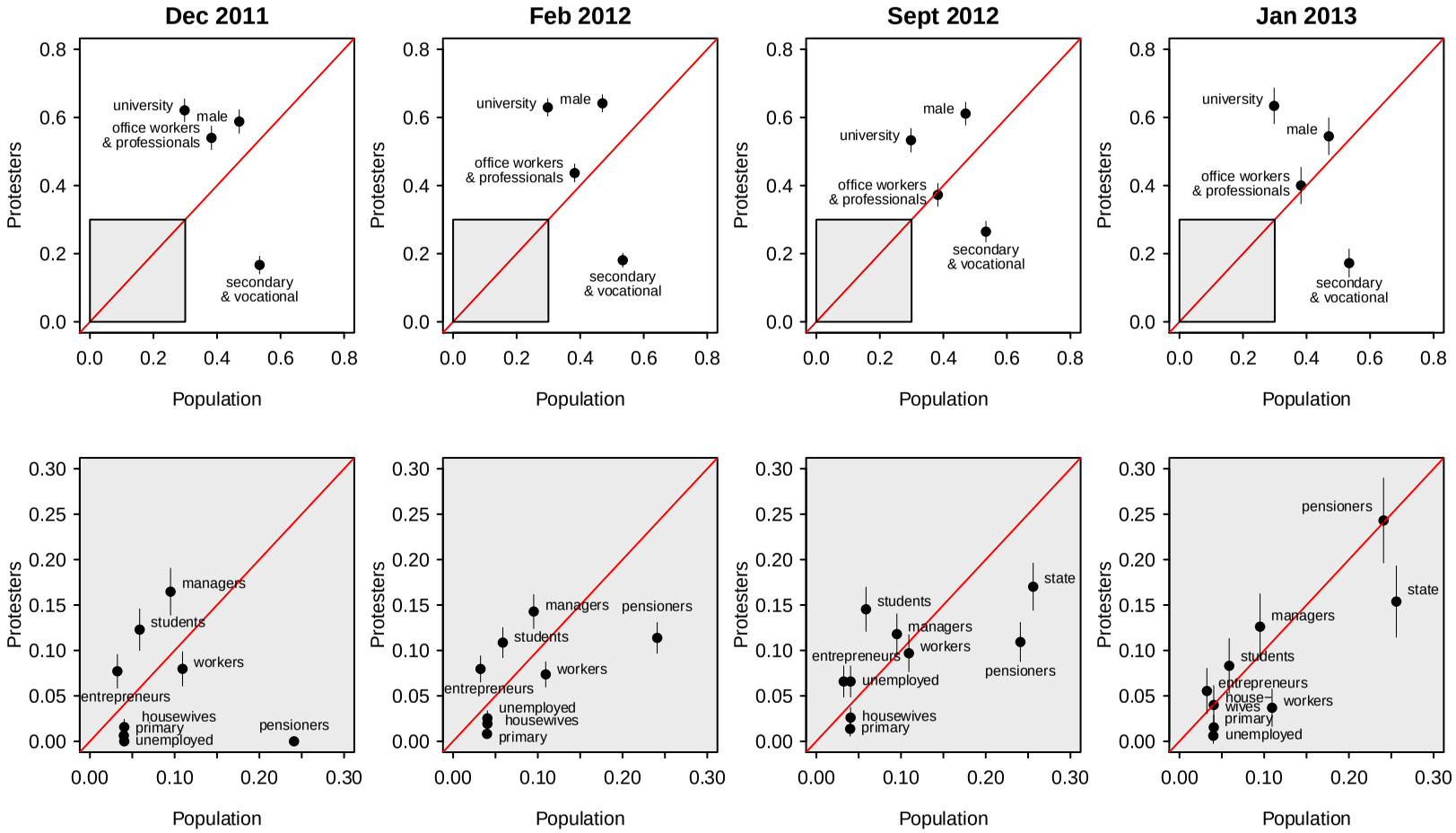



Figure A3: **Demographic Comparison of Protesters and the Population by Survey Wave**. This figure compares the demographic composition of the protests with the population of Moscow city for each of the four survey waves. Again, the regions shaded grey in the first row of plots are magnified in the second row of plots. The x-axis gives the share of the *population* with a given demographic characteristic, while the y-axis gives the share of *protesters* with that same characteristic. Points that lie above the 45 degree red line thus indicate groups that were overrepresented among the protesters, points below the line groups that were underrepresented. The vertical bars are 95% confidence intervals for each protest survey. Horizontal bars give 95% confidence intervals for the population data. Data sources: The Levada Center and the Foundation for Public Opinion (FOM).

### A.3 The Case-Control with Contaminated Controls Model

Let $Z_i$ denote true protest participation and $\eta_i = 1$ indicate inclusion in a sample that consists of $n_1$ randomly sampled protesters and $n_u$ observations randomly sampled from the population without regard to protest participation.[2] Further, let $\pi$ denote the population prevalence of protest participation (a parameter to be estimated), $X_i$ a vector of relevant covariates observed in both samples, $\alpha$ the intercept parameter and $\beta$ a vector of parameters associated with each of the J covariates ($\beta_j$, for $j = 1, ..., J$). Under case-control sampling with contaminated controls, we do not fully observe true protest participation, $Z_i$. In particular, we do not observe $Z_i$ for $n_u$ observations sampled from the population.

Based on the sampling protocol, the probability of protest participation given the covariates and inclusion in the sample, $P(Z_i = 1|X_i, \eta_i = 1)$, can be expressed as:

$$\psi_i = P(Z_i = 1|X_i, \eta_i = 1) = \frac{exp(ln(\frac{n_1}{\pi n_u} + 1) + \alpha + X_i^\top \beta)}{1 + exp(ln(\frac{n_1}{\pi n_u} + 1) + \alpha + X_i^\top \beta)} \quad (1)$$

Equation 1 implies the assumption that protest participation, which is only partially observed, follows a logistic distribution with an intercept shift to reflect the sampling protocol.

Next, letting $S_i$ denote a binary stratum indicator, which takes the value $S_i = 1$ for all observations for which the choice (protest participation) is known and $S_i = 0$ for all observations for which the choice is unknown, we can write the probability of observing protest participation, conditional on actual participation and selection into the sample, as:

$$\theta_i = P(S_i = 1|Z_i = 1) = \frac{n_1}{n_1 + \pi n_u} \quad (2)$$

Combining equations 1 and 2, we describe the process of drawing observations from the stratum in which protest participation is known versus the stratum in which protest participation is not

[2]The notation in this section follows: Rota, Christopher T. et al. 2013. "A Re-Evaluation of a Case-Control Model with Contaminated Controls for Resource Selection Studies." *Journal of Animal Ecology* 82(6):1165-73.

known as a Bernoulli trial.[3] This gives us the following likelihood function for the case-control with contaminated controls design:

$$L(\alpha, \beta, \pi | S_i, X_i, n_1, n_u) = \prod_{i=1}^{n} (\psi_i \times \theta_i)^{S_i} (1 - (\psi_i \times \theta_i))^{1-S_i} \tag{3}$$

Finally, for the estimation of this model, I use a Bayesian framework:

$$\begin{aligned} f(\alpha, \beta, \pi | S_i, X_i, n_1, n_u) &\propto L(\alpha, \beta, \pi | S_i, X_i, n_1, n_u) \\ &\times f(\alpha) f(\beta_1) \times ... \times f(\beta_j) g(\pi) \end{aligned} \tag{4}$$

## A.4 Estimation of the Model

I fit a Bayesian implementation of the case-control model with contaminated controls in `JAGS` via the `rjags` package for R.[4]

[3] It is assumed that the population is large enough that there is no overlap between the two samples.

[4] I specified independent normal priors, $\mathcal{N}(0, 10)$, for the intercept parameter and regression coefficients and a uniform prior, $\mathcal{U}(0, 1)$, for the prevalence parameter $\pi$. For each version of the model, I ran four chains from over-dispersed starting values for 22,500 iterations. I discarded the first 12,500 iterations as burn-in and kept every 10th iteration thereafter for a total of 4,000 draws from the marginal posterior distribution of model parameters. Visual inspection of the trace and autocorrelation plots as well as the Gelman-Rubin diagnostic suggest proper convergence.

## A.5 Additional Case-Control Results

| | Est | 95% CI |
|---|---|---|
| Middle class vs. Non-middle class | 2.31 | [1.91, 2.83] |
| State vs. Non-state | 0.62 | [0.51, 0.74] |
| State middle class vs. Non-state middle class | 0.74 | [0.56, 0.98] |

Source: The Levada Center and the Foundation for Public Opinion (FOM)

Table A2: **Relative Risk Estimates (Case-Control Model with Contaminated Controls, Moscow Sample)**. This table reports the fractional increase in the risk of protest participation for several groups relative to the given baseline. The results are based on a case control with contaminated controls design, using a random sample of protesters and a random sample of the population of Moscow and Moscow region. The same results are presented in Figure 2.

| | Est | 95% CI |
|---|---|---|
| Middle class vs. Non-middle class | 7.66 | [6.55, 9.01] |
| State vs. Non-state | 0.42 | [0.36, 0.50] |
| State middle class vs. Non-state middle class | 0.51 | [0.41, 0.63] |

Source: The Levada Center and the Foundation for Public Opinion (FOM)

Table A3: **Relative Risk Estimates (Case-Control Model with Contaminated Controls, Russia Sample)**. This table reports the results of an alternative case-control with contaminated controls specification, based on a random sample of Russia's population as a whole and the random sample of protesters used throughout the analysis in the main text. Like the preceding results, these findings indicate that state workers, in general, and the state middle class, in particular, were much less likely than their private sector counterparts to protest. In fact, the relative risk estimates for being a state employee (vs. non-state employee, 42%) and a middle class state-employee (vs. a middle class non-state employee, 51%) in rows 2 and 3, respectively, have confidence intervals that overlap the equivalent estimates in Table A2. The main difference, when we compare protesters with Russia's population as a whole, is that the effect sizes for all of the key variables, and especially the middle class variable, are larger.

### A.6 Robustness Checks: Population Prevalence of Protest Constrained

A great deal of controversy surrounds reports on the size of protest participation. Police estimates are widely criticized as being substantial underreports, a useful propaganda tool in the hands of the state. Meanwhile, opposition leaders have every incentive to exaggerate the number of participants. Since there is substantial uncertainty about total protest participation, in the preceding analysis, I allowed this figure to be estimated as a parameter in the model. I now test the sensitivity of my findings by setting bounds on the extent of protest participation. To do so, I incorporate evidence on the size of protest participation drawn from various media sources.[5] For the lower bound, I take Russian MVD figures and multiply them by an estimate of the share of protesters participating in multiple demonstrations. Based on these figures, I constrain $\pi$ to 0.3% of the population of Moscow and Moscow oblast for the lower bound estimates. For the upper bound, I take opposition estimates and make no assumption about repeat protest participation. This yields an upper bound estimate of $\pi$ equal to 6% of the population of Moscow and Moscow oblast. The 95% confidence interval of the model-based estimate of $\pi$ in the forgoing section lies just above this upper bound.

The results in Table A4 allow us to substantively evaluate the predicted probability (or absolute risk) of protest as a function of covariates across the likely range of protest prevalence. The results in Table A5 do the same for relative risk. The first thing to notice is that while absolute risk is sensitive to the value of $\pi$, relative risk remains fairly stable. Comparing the estimates of relative risk at the lower and upper bounds of $\pi$, we see that they have confidence intervals which are overlapping. These results imply that the main conclusion from the preceding analysis is robust across the likely range of $\pi$. A middle class composed of state workers is systematically less likely

[5]These sources are documented in Appendix Table A6.

| | Lower Bound $\pi$=0.003 | Upper Bound $\pi$=0.055 |
|---|---|---|
| All | 0.0038 | 0.0600 |
| | [0.0032, 0.0046] | [0.0523, 0.0707] |
| Middle class | 0.0062 | 0.0996 |
| | [0.0052, 0.0076] | [0.0850, 0.1163] |
| Non-middle class | 0.0021 | 0.0400 |
| | [0.0018, 0.0026] | [0.0328, 0.0549] |
| State employees | 0.0023 | 0.0390 |
| | [0.0017, 0.0030] | 0.0302, 0.0533 |
| Non-state employees | 0.0043 | 0.0671 |
| | [0.0036, 0.0052] | [0.0584, 0.0775] |
| State middle class - Non-state middle class† | −0.0021 | −0.0292 |
| | [−0.0040, −0.0003] | [−0.0535, −0.0044] |

†First difference in predicted probability. Data source: The Levada Center and the Foundation for Public Opinion (FOM)

Table A4: **Predicted Probabilities of Protest Participation – Bounded Estimates (Case-Control Model with Contaminated Controls, Moscow Sample)**. The table reports predicted probabilities of protest participation, using the case-control design with contaminated controls and incorporating knowledge about $\pi$. Since the population prevalence of protest is widely debated, with opposition politicians and the police offering sharply divergent estimates, the table gives results at both the lower and upper bound of reported protest participation. This analysis thus probes the sensitivity of the main results by setting bounds on the fraction of the population that actually engaged in protest, denoted in Equation 4 as $\pi$. Whereas the model in Equation 4 is parametrically identified without prior knowledge of $\pi$, with prior knowledge of $\pi$ the model is overidentified. Hence, the value of $\pi$ given in Table 2 is a model-based estimate, while here $\pi$ has been constrained to a minimum value of .3% and a maximum value of 5.5% of the population of Moscow.

to join opposition protests than a middle class that is independent of the state.

As a final robustness check, Table A7 gives an alternative set of results for a range of protest prevalence estimates using the prior-correction model for rare events in King and Zeng (2001). While the assumptions for this model are not fully met in the present data (since selection did not depend entirely on $Y$ and, indeed, $Y$ was not even measured in the population sample), as $Pr(Y = 1)$ becomes very small, the case-control design with contaminated controls increasingly resembles the classic case-control. Unsurprisingly then, the results are similar.

| | Lower Bound $\pi$=0.003 | Upper Bound $\pi$=0.055 |
|---|---|---|
| Middle class vs. Non-middle class | 2.89 | 2.46 |
| | [2.31, 3.61] | [1.88, 3.11] |
| State vs. Non-state | 0.53 | 0.59 |
| | [0.41, 0.70] | [0.47, 0.73] |
| State middle class vs. Non-state middle class | 0.68 | 0.73 |
| | [0.49, 0.96] | [0.55, 0.95] |
| State middle class vs. Non-middle class | 2.14 | 1.88 |
| | [1.56, 2.90] | [1.43, 2.51] |
| Non-state middle class vs. Non-middle class | 3.13 | 2.63 |
| | [2.44, 3.98] | [1.98, 3.40] |

Data source: The Levada Center and the Foundation for Public Opinion (FOM)

Table A5: **Bounded Relative Risk Estimates (Case-Control Model with Contaminated Controls, Moscow Sample)**. This table reports the estimated relative risk of protest participation, using the case-control design with contaminated controls and incorporating knowledge about $\pi$. As in Table A4, this analysis probes the sensitivity of the main results given in Figure 2 by setting bounds on the fraction of the population that actually engaged in protest. Since the population prevalence of protest is widely debated, with opposition politicians and the police offering sharply divergent estimates, the table gives results at both the lower and upper bound of reported protest participation.

| Date | MVD estimate | Opposition max estimate | Share unique | Notes |
|---|---|---|---|---|
| 5-Dec-11 | 2,000 | 10,000 | 1.00 | Chistye Prudy |
| 10-Dec-11 | 25,000 | 150,000 | 0.95 | Bolotnaya |
| 17-Dec-11 | 1,500 | 5,000 | 0.20 | Bolotnaya |
| 24-Dec-11 | 28,000 | 120,000 | 0.37 | "For Fair Elections;" Sakharov |
| 4-Feb-12 | 38,000 | 160,000 | 0.20 | "For Fair Elections;" Kaluzhskaya to Bolotnaya |
| 26-Feb-12 | 10,000 | 30,000 | 0.20 | Garden Ring |
| 5-Mar-12 | 14,000 | 25,000 | 0.20 | Pushkin Square |
| 10-Mar-12 | 10,000 | 25,000 | 0.20 | Novy Arbat |
| 6-May-12 | 8,000 | 120,000 | 0.20 | Yakimanka and Bolotnaya |
| 12-Jun-12 | 15,000 | 200,000 | 0.20 | Pushkin Square |
| 15-Sep-12 | 14,000 | 100,000 | 0.10 | "March of Millions;" Pushkin Sq. to Pr. Sakharov |
| 13-Jan-13 | 10,000 | 50,000 | 0.05 | "March Against Scoundrels;" anti-Magnitsky |
| 6-May-13 | 7,000 | 30,000 | 0.05 | Bolotnaya |

Table A6: **Protest Attendance**. This table gives contrasting estimates of protest attendance, based on police and opposition figures cited in media reports. The share unique column gives a rough, and necessarily somewhat arbitrary, estimate of the fraction of protesters who were first time participants. The police figures in the first column were then multiplied by the share of unique protesters to calculate the lower bound estimate of total protest participation (0.3% of the population of Moscow and Moscow region). The upper bound estimate (6%) was derived from the maximum figures given by the opposition without regard to uniqueness.

| | Lower Bound $\pi$=0.003 | Upper Bound $\pi$=0.055 |
|---|---|---|
| Middle class vs. Non-middle class | 2.84 | 2.53 |
| | [2.25, 3.50] | [2.07, 3.02] |
| State vs. Non-state | 0.53 | 0.54 |
| | [0.41, 0.70] | [0.44, 0.68] |
| State middle class vs. Non-state middle class | 0.68 | 0.72 |
| | [0.50, 0.97] | [0.56, 0.97] |
| State middle class vs. Non-middle class | 2.11 | 1.97 |
| | [1.55, 2.82] | [1.50, 2.55] |
| Non-state middle class vs. Non-middle class | 3.07 | 2.72 |
| | [2.40, 3.88] | [2.19, 3.30] |

Source: The Levada Center and the Foundation for Public Opinion (FOM)

Table A7: **Bounded Relative Risk Estimates (Bias-Corrected Logistic Regression for Choice-Based Sampling, Moscow Sample)**. This analysis probes the sensitivity of the main results to model specification. The table reports the estimated relative risk of protest participation, using the approach to logistic regression in rare events data recommended by King and Zeng (2001). Specifically, this alternative specification uses finite-sample bias correction for rare events and prior correction for choice-based sampling. The former addresses the problem that rare events data amplify finite sample bias, leading to the systematic underestimation of $Y_i = 1$. The latter is used to correct ordinary logistic regression estimates for intentional selection on the dependent variable—that is, for a sampling scheme which oversamples cases in which $Y_i = 1$ (i.e. protesters). Prior correction requires prior information about the fraction of ones in the population, $\pi$. Since the population prevalence of protest, $\pi$, is widely debated, with opposition politicians and the police offering sharply divergent estimates, the table gives results at both the lower and upper bound of estimates of protest participation. Relative risk is calculated using simulation: taking a random draw of M=1,000 $\beta$ from $P(\beta)$, computing relative risk, and averaging over the simulations. In contrast to the case-control design with contaminated controls used in the main text, the assumption implied by this model is that the "control" sample was drawn from those observations for which $Y_i = 0$ (i.e. non-protesters). Clearly, this assumption may not be warranted, making the case-control design employed here less suitable to the present data than the case-control with contaminated controls design discussed in the main text. Nevertheless, the results are very similar: the estimates all share overlapping confidence intervals with those in Table A5. This is not surprising given that, as Lancaster and Imbens (1996, 156) observe, as $\pi$ approaches 0, the sampling increasingly approximates pure choice-based sampling in which one subsample is drawn from those observations with $Y_i = 0$ and the other from those with $Y_i = 1$.

## A.7 Robustness Checks: Controlling for Cohort

| | *Avg. Age* | *Birth Year* | *Age in 1991* |
|---|---|---|---|
| Protesters | | | |
| State middle class | 41 | 1971 | 20 |
| Private sector middle class | 35 | 1977 | 14 |
| Democrats | 42 | 1970 | 21 |
| Communists | 46 | 1966 | 25 |
| Nationalists | 32 | 1980 | 11 |
| Population of Moscow and Moscow Oblast | | | |
| State middle class | 40 | 1972 | 19 |
| Private sector middle class | 36 | 1976 | 15 |
| Democrats | 43 | 1969 | 22 |
| Communists | 57 | 1955 | 36 |
| Nationalists | 33 | 1979 | 12 |

Source: Levada Center and Foundation for Public Opinion (FOM)

Table A8: **Descriptive Statistics on Age/Cohort**. This table displays the average age of various groups of protesters and the population.

Table A9: Case-Control with Contaminated Controls Model Predicting Protest Participation Robustness Check Including Birth Cohort

| | *Dependent variable:* protest participation | |
|---|---|---|
| | Moscow | |
| | Estimate (median) | CI [2.5%, 97.5%] |
| Middle class | 1.23 | [0.92, 1.56] |
| State employment | −1.45 | [-1.95, -0.96] |
| State x Middle class | 0.96 | [0.32, 1.65] |
| Male | 0.60 | [0.36, 0.84] |
| Age | −0.63 | [-1.16, -0.12] |
| $\text{Age}^2$ | 0.02 | [-0.02, 0.07] |
| Democrats | 9.59 | [4.19, 23.43] |
| Communists | 1.71 | [1.29, 2.17] |
| Nationalists | 1.08 | [0.65, 1.57] |
| Post-Cold War birth cohort | −0.79 | [-1.31, -0.28] |
| Constant | −1.01 | [-2.55, 0.57] |
| $\pi$ | 0.10 | [0.06, 0.15] |

*Note:* The table reports point estimates for the model parameters and 95% Bayesian credibility intervals. Data source: The Levada Center and the Foundation for Public Opinion (FOM)

Table A10: Case-Control with Contaminated Controls Model Predicting Protest Participation Robustness Check Including Age as a Categorical Variable

| | *Dependent variable:* protest participation | |
|---|---|---|
| | Moscow | |
| | Estimate (median) | CI [2.5%, 97.5%] |
| Middle class | 1.27 | [0.96, 1.61] |
| State employment | −1.36 | [-1.87, -0.86] |
| State x Middle class | 0.87 | [0.21, 1.54] |
| Male | 0.62 | [0.37, 0.87] |
| Age 26-35 | −0.30 | [-0.68, 0.07] |
| Age 36-45 | −0.25 | [-0.63, 0.15] |
| Age 46-55 | −0.75 | [-1.16, -0.35] |
| Age 56-65 | −0.46 | [-0.89, -0.005] |
| Over 65 | −1.17 | [-1.66, -0.70] |
| Democrats | 9.58 | [3.95, 23.63] |
| Communists | 1.70 | [1.28, 2.16] |
| Nationalists | 1.10 | [0.65, 1.65] |
| Constant | −3.02 | [-3.63, -2.51] |
| $\pi$ | 0.11 | [0.06, 0.15] |

*Note:* The table reports point estimates for the model parameters and 95% Bayesian credibility intervals. Data source: The Levada Center and the Foundation for Public Opinion (FOM)

Table A11: Democratic Ideology and Protest Participation
Robustness Check Including Birth Cohort and Categorical Age

| | *Dependent variable:* democrats | |
|---|---|---|
| | pooled | |
| | Cohort | Categorical Age |
| | (1) | (2) |
| Middle Class | 0.691*** | 0.709*** |
| | (0.133) | (0.139) |
| State Employment | 0.079 | 0.138** |
| | (0.049) | (0.063) |
| Middle Class x state employment | −0.496*** | −0.535*** |
| | (0.043) | (0.036) |
| Post-Cold War birth cohort | 0.241*** | |
| | (0.074) | |
| Age 26-35 | | 0.054 |
| | | (0.224) |
| Age 36-45 | | 0.302* |
| | | (0.179) |
| Age 46-55 | | 0.312 |
| | | (0.220) |
| Age 56-65 | | 0.801*** |
| | | (0.185) |
| Over 65 | | 1.079*** |
| | | (0.111) |
| Consumption-2 | 0.574 | 0.559 |
| | (0.559) | (0.589) |
| Consumption-3 | 0.480 | 0.456 |
| | (0.405) | (0.435) |
| Consumption-4 | 0.751* | 0.736 |
| | (0.452) | (0.494) |
| Consumption-5 | 0.916** | 0.908** |
| | (0.402) | (0.455) |
| Consumption-6 | 0.613** | 0.590* |
| | (0.280) | (0.315) |
| Male | −0.261*** | −0.262*** |
| | (0.047) | (0.055) |
| Age | 0.259 | |
| | (0.265) | |
| Age squared | 0.002 | |
| | (0.025) | |
| Constant | −2.874*** | −1.974*** |
| | (0.046) | (0.492) |
| Observations | 1,130 | 1,130 |

*Note:* *p<0.1; **p<0.05; ***p<0.01. Logit coefficients from regressions predicting ideological affinity with democrats. Standard errors are clustered by survey wave. The reference category for consumption is "we don't even have enough money for food." The main coefficients are virtually identical to those of Model 3 in Table A12 below. Data Source: The Levada Center.

## A.8 Robustness Checks: The Democratic Protest Coalition

Table A12 reports the coefficients from the model discussed in the main text and a series of robustness tests. The first four columns are based on the pooled data, while columns 5 and 6 report the results for waves three and four respectively. Column 1 is the base model and consists of just the two key independent variables and their interaction. Consistent with prior research, the middle class variable is positive and significant, implying that middle class protesters are significantly more likely than non-middle class protesters to identify with democratic forces. Substantively, a middle class protester is about 50 percent more likely than a non-middle class protester to be a democrat. The predicted probability for middle class protesters is about 37% versus only 25% for the non-middle class. Meanwhile, also as expected, the coefficient on state employment and the interaction of state employment and middle class are negative, though they do not reach conventional levels of statistical significance.

Column 2 adds the basic demographic controls, addresses some of the apparent confounding in the first model, and considerably clarifies the relationship between state employment and the ideology of protesters: Protesters who are middle class state employees share about the same low probability of demonstrating in support of democracy as non-middle class protesters. The effect of state employment for non-middle class protesters is essentially zero. Regardless of employment sector, working class protesters were unlikely to identify as democrats. For the middle class, by contrast, state patronage in the form of public-sector employment has a sizable negative effect on democracy support. The fact that the interaction between employment sector and middle class remains significant, even after the inclusion of the proxy for income, suggests that differences between private and public-sector employees cannot be attributed to a sectoral wage gap.

The remaining columns 4 through 6 in Table A12 report additional robustness checks. In col-

Table A12: Democratic Ideology and Protest Participation

| | *Dependent variable:* democrats | | | | | |
|---|---|---|---|---|---|---|
| | pooled | | | | wave 3 | wave 4 |
| | No Controls | Controls | Full Model | Occ/Edu | By wave | By wave |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Middle Class | 0.499*** | 0.775*** | 0.699*** | | 0.832*** | 0.366 |
| | (0.189) | (0.133) | (0.133) | | (0.198) | (0.304) |
| State Employment | −0.111 | 0.098*** | 0.083 | −0.334*** | 0.149 | 0.018 |
| | (0.076) | (0.029) | (0.058) | (0.018) | (0.533) | (0.909) |
| Middle Class x state employment | −0.153 | −0.546*** | −0.500*** | | −0.543 | −0.343 |
| | (0.099) | (0.050) | (0.055) | | (0.588) | (0.990) |
| Office worker/professional | | | | 0.626*** | | |
| | | | | (0.091) | | |
| Manager | | | | 0.241** | | |
| | | | | (0.099) | | |
| Entrepreneur/business owner | | | | 0.945*** | | |
| | | | | (0.283) | | |
| Student | | | | 0.491*** | | |
| | | | | (0.104) | | |
| Housewife | | | | 0.943** | | |
| | | | | (0.394) | | |
| Graduate degree | | | | 0.533 | | |
| | | | | (0.324) | | |
| Higher education | | | | 0.387*** | | |
| | | | | (0.131) | | |
| Consumption-2 | | | 0.567 | 0.595 | 1.216 | |
| | | | (0.560) | (0.518) | (1.116) | |
| Consumption-3 | | | 0.475 | 0.467 | 0.930 | 0.143 |
| | | | (0.402) | (0.378) | (1.083) | (0.576) |
| Consumption-4 | | | 0.744* | 0.720* | 1.241 | 0.205 |
| | | | (0.452) | (0.429) | (1.079) | (0.553) |
| Consumption-5 | | | 0.904** | 0.868** | 1.357 | 0.464 |
| | | | (0.405) | (0.383) | (1.089) | (0.609) |
| Consumption-6 | | | 0.609** | 0.505** | 0.935 | 0.616 |
| | | | (0.277) | (0.237) | (1.193) | (1.133) |
| Male | | −0.256*** | −0.265*** | −0.229*** | −0.195 | −0.293 |
| | | (0.028) | (0.049) | (0.006) | (0.164) | (0.243) |
| Age | | 0.112 | 0.063 | −0.077 | −0.092 | 0.626 |
| | | (0.207) | (0.192) | (0.097) | (0.308) | (0.458) |
| Age squared | | 0.011 | 0.017 | 0.030*** | 0.031 | −0.043 |
| | | (0.021) | (0.019) | (0.005) | (0.036) | (0.049) |
| Constant | −1.034*** | −1.694*** | −2.250*** | −2.208*** | −2.538** | −2.612** |
| | (0.223) | (0.290) | (0.170) | (0.182) | (1.202) | (1.109) |
| Observations | 1,130 | 1,130 | 1,130 | 1,130 | 805 | 325 |
| Log Likelihood | -695.869 | -681.942 | -679.195 | -674.715 | -467.462 | -207.876 |

*Note:* *p<0.1; **p<0.05; ***p<0.01. Logit coefficients from regressions predicting ideological affinity with democrats. Standard errors are clustered by survey wave in the pooled models (Columns 1-4). The reference category for consumption is "we don't even have enough money for food" (in all but column 6, where that category is combined with "we have enough money for food, but not enough for clothing," due to the small number of observations). For occupation, it is unemployed, pensioners, and blue-collar workers, while for education, it is less than a university education.
Data Source: The Levada Center.

umn 4, I disaggregate the middle class into its occupational and educational components, showing that each operates similarly on the dependent variable in terms of direction and significance. This justifies their inclusion in a single measure of middle case, as in the previous models. The coefficients in model 4 can be interpreted as the effect of each occupational and educational category in comparison with the occupational or educational profile of the non-middle class.[6] The positive and statistically significant coefficients[7] for office workers and professionals, managers, entrepreneurs and business owners, students, and having higher education indicate that these protesters were more likely than those with non-middle class occupations or lower levels of education to be part of the pro-democracy coalition. Meanwhile, the large, negative coefficient on state employment implies that, *ceteris paribus*, state workers who did join the protests were much less likely than those who were not employed by the state to be democrats. The results in column 4 are thus consistent with the previous models.

Last, in columns 5 and 6, I disaggregate the pooled data and rerun the full model separately for the third and fourth protest waves, respectively.[8] Although several of the coefficients in these models based on fewer observations fail to attain conventional levels of statistical significance, the results are substantially the same. Even at the end of the mobilizational cycle in January 2013 (wave 4), with the number of protesters dwindling and a sample size of fewer than 400 respondents, the basic patterns I have described are evident in the data and the coefficients are surprisingly

[6] In particular, the reference category for occupation pools all non-middle class occupations (pensioners, the unemployed, and blue-collar workers). Similarly, for education, the reference category is schooling consistent with a non-middle class career as a blue-collar or routine laborer (specialized post-secondary or vocational training, complete and incomplete secondary, and primary).

[7] Only having a graduate degree, which is a numerically smaller group and less precisely estimated, is statistically insignificant.

[8] Again, unfortunately, because state employment was not measured in the first two waves, they are excluded from the analysis.

stable.

Finally, Table A13 shows that the main results are substantively unchanged if an alternative dependent variable combining liberals and democrats is used. Table A14 highlights how models that consider only class status or occupational and educational measures fail to capture the important interactive effects of sector of employment.

Table A13: Liberal and Democratic Ideology and Protest Participation

| | *Dependent variable:* liberals or democrats | | | | | |
|---|---|---|---|---|---|---|
| | pooled | | | | wave 3 | wave 4 |
| | No Controls | Controls | Full Model | Occ/Edu | By wave | By wave |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Middle Class | 0.620** | 0.831*** | 0.706*** | | 0.864*** | 0.207 |
| | (0.290) | (0.172) | (0.171) | | (0.180) | (0.293) |
| State Employment | 0.026 | 0.198 | 0.147 | −0.202*** | 0.245 | −0.034 |
| | (0.176) | (0.067) | (0.095) | (0.058) | (0.466) | (0.859) |
| Middle class x state employment | −0.128 | −0.436** | −0.336* | | −0.475 | 0.203 |
| | (0.319) | (0.182) | (0.174) | | (0.520) | (0.936) |
| Housewife | | | | 0.659 | | |
| | | | | (0.610) | | |
| Office worker/professional | | | | 0.777*** | | |
| | | | | (0.041) | | |
| Manager | | | | 0.298*** | | |
| | | | | (0.032) | | |
| Entrepreneur/business owner | | | | 0.827*** | | |
| | | | | (0.092) | | |
| Student | | | | 0.716** | | |
| | | | | (0.283) | | |
| Graduate degree | | | | 0.298 | | |
| | | | | (0.404) | | |
| Higher education | | | | 0.431*** | | |
| | | | | (0.113) | | |
| Consumption-2 | | | 0.779 | 0.790 | 1.438 | |
| | | | (0.542) | (0.547) | (1.109) | |
| Consumption-3 | | | 1.031*** | 1.008*** | 1.342 | 0.749 |
| | | | (0.277) | (0.266) | (1.076) | (0.561) |
| Consumption-4 | | | 1.433*** | 1.383*** | 1.824* | 0.753 |
| | | | (0.364) | (0.354) | (1.072) | (0.537) |
| Consumption-5 | | | 1.663*** | 1.644*** | 1.951* | 1.327** |
| | | | (0.244) | (0.205) | (1.082) | (0.604) |
| Consumption-6 | | | 2.222*** | 2.128*** | 2.583** | 1.449 |
| | | | (0.364) | (0.312) | (1.161) | (1.273) |
| Male | | −0.289*** | −0.326*** | −0.297*** | −0.217 | −0.406* |
| | | (0.060) | (0.079) | (0.033) | (0.154) | (0.238) |
| Age | | −0.062 | −0.146*** | −0.114 | −0.050 | 0.175 |
| | | (0.053) | (0.024) | (0.108) | (0.289) | (0.438) |
| Age squared | | 0.024*** | 0.036*** | 0.033*** | 0.014 | 0.004 |
| | | (0.001) | (0.004) | (0.0004) | (0.034) | (0.048) |
| Constant | −0.519 | −0.653*** | −1.778*** | −2.219*** | −2.426** | −1.347 |
| | (0.326) | (0.035) | (0.426) | (0.032) | (1.180) | (1.029) |
| Observations | 1,130 | 1,130 | 1,130 | 1,130 | 805 | 325 |
| Log Likelihood | -766.532 | -756.300 | -744.623 | -739.611 | -518.008 | -214.059 |

*Note:* *p<0.1; **p<0.05; ***p<0.01. Logit coefficients from regressions predicting ideological affinity with democrats/liberals. Standard errors are clustered by survey wave in the pooled models (Columns 1-4). The reference category for consumption is "we don't even have enough money for food" (in all but column 6, where that category is combined with "we have enough money for food, but not enough for clothing," due to the small number of observations). For occupation, it is unemployed, pensioners, and blue-collar workers, while for education, it is less than a university education. Data Source: The Levada Center.

Table A14: Democratic Ideology and Protest Participation: The Effect of Class Alone

| | *Dependent variable:* democrats | | | |
|---|---|---|---|---|
| | pooled | | | |
| | No Controls | Controls | Add Income | Occ/Edu |
| | (1) | (2) | (3) | (4) |
| Middle class | 0.559*** | 0.654*** | 0.509*** | |
| | (0.095) | (0.065) | (0.073) | |
| Office worker/professional | | | | 0.318*** |
| | | | | (0.108) |
| Manager | | | | 0.254 |
| | | | | (0.225) |
| Entrepreneur/business owner | | | | 0.588** |
| | | | | (0.264) |
| Student | | | | 0.505*** |
| | | | | (0.021) |
| Housewife | | | | 0.584* |
| | | | | (0.316) |
| Graduate degree | | | | 0.662*** |
| | | | | (0.204) |
| Higher education | | | | 0.528*** |
| | | | | (0.151) |
| Consumption-2 | | | −0.001 | −0.054 |
| | | | (0.318) | (0.313) |
| Consumption-3 | | | 0.073 | −0.010 |
| | | | (0.187) | (0.178) |
| Consumption-4 | | | 0.539** | 0.455** |
| | | | (0.230) | (0.187) |
| Consumption-5 | | | 0.883*** | 0.764*** |
| | | | (0.210) | (0.181) |
| Consumption-6 | | | 0.829*** | 0.688*** |
| | | | (0.183) | (0.165) |
| Male | | −0.303*** | −0.346*** | −0.314*** |
| | | (0.038) | (0.054) | (0.060) |
| Age | | 0.528** | 0.395* | 0.421** |
| | | (0.215) | (0.214) | (0.174) |
| Age squared | | −0.039* | −0.022 | −0.026 |
| | | (0.023) | (0.023) | (0.018) |
| Constant | −1.094*** | −2.358*** | −2.514*** | −2.834*** |
| | (0.099) | (0.403) | (0.188) | (0.191) |
| Observations | 3,244 | 3,240 | 3,234 | 3,229 |
| Log Likelihood | -2030.194 | -1988.076 | -1953.562 | -1943.153 |

*Note:* *p<0.1; **p<0.05; ***p<0.01. Logit coefficients from a regression predicting ideological affinity with democrats. Standard errors are clustered by survey wave. The reference category for consumption is "we don't even have enough money for food." For occupation, it is unemployed, pensioners, and blue-collar workers, while for education, it is less than a university education. Data Source: The Levada Center.

## A.9 Robustness Checks: Selection

If democrats are less likely to work for the state, then the association between public-sector employment, democratic values, and protest participation may be due to sorting, rather than to selective incentives. To test this alternative empirically, I exploit the fact that selection effects should vary with the nature of the Russian regime. While under Yeltsin the state was, at least ostensibly, democratizing, during the Putin years, the state was quite obviously backsliding into autocracy. Medvedev's presidency presents a more ambiguous case, though one associated with the ascendance of liberal factions within the Kremlin and heightened democratic rhetoric. As such, if those who joined the state sector under Yeltsin, and perhaps also Medvedev, were more likely to protest than those who joined the state sector under Putin, then selection is likely to be a factor. If they are not, then we can more confidently ascribe differences in patterns of protest participation among public and private sector workers to differences in political preferences and values that result from their employment status.

A simple strategy for empirically testing this argument is to divide state sector protest participants into cohorts by the regime in power when they joined the public sector. More specifically, if individuals engage in self-sorting, choosing state sector careers under regimes that are compatible with their political views, those who joined Yeltsin's state should be more pro-democratic and more likely to protest than those who joined Putin's.[9] Similarly, it is possible that those who joined during Medvedev's presidency—when a more liberal faction within the Kremlin appeared ascendant and democracy and the rule of law received renewed rhetorical emphasis—would be more likely to protest than those who joined earlier (e.g., during the illiberal turn in Putin's second term). This case, however, is more ambiguous. Finally, while selection makes less clear predictions for those

[9]This analysis assumes that protest is correlated, though not synonymous, with democracy support.

who joined the public sector under communism, the mere fact that this group remained employed through both the Yeltsin and Putin years, and myriad changes in Russia's regime type, may suggest that they are ideologically flexible and pragmatic. If so, an account based on selection would expect this group to be least likely to protest.

Because the surveys lack a question about how long respondents have worked in their current job or sector, I assume that public-sector employees joined in their 20's and have remained in the public sector for the duration of their careers. In fact, occupational segmentation by sector and actual patterns of labor market participation in Russia support this stylization.[10] In Table A15, I estimate the same index of representation as in the preceding analysis for four cohorts of state employees under two different assumptions. The results in the first row of the table assume that respondents joined the public sector by the time they were 25, thus placing state employees who were younger than 30 at the time of the survey in the Medvedev cohort, 30-37 in the Putin cohort, 38-46 in the Yeltsin cohort, and 47 and over in the communist cohort. The second row tests the sensitivity of these results to the assumption that employment begins in the mid-20s, by adjusting it downward to age 22, immediately following university.

| | *joined state employment under...* | | | |
|---|---|---|---|---|
| | Communism | Yeltsin | Putin | Medvedev |
| Became state employees by age 25 | 0.44 | 0.87 | 0.79 | 0.78 |
| Became state employees by age 22 | 0.51 | 0.96 | 0.77 | 0.65 |

Data source: The Levada Center protest surveys and FOM

Table A15: **Index of Representation by Cohort for State Employees**. This table examines protest participation among state employees by cohort and shows that those state employees who joined under Yeltsin were more likely to protest against the regime than those who joined in any other period. The index was constructed by dividing the share of protesters in each demographic group by its share of the population, in this table the population of Moscow.

[10] According to the EBRD LiTS I survey, which includes a detailed employment history going back to 1989, 90% of those employed in the state sector in Russia in 2006, at the time of the survey, had been state sector employees for their entire career.

The evidence for selection in Table A15 is mixed. On the one hand, it shows that those who joined the public sector during the Yeltsin administration, between 1991 and 1999, were more likely to protest than those who joined after Putin came to power, beginning in 2000. This evidence accords with the argument that the Yeltsin era's liberal rhetoric attracted liberals who were more likely to protest an increasingly illiberal regime. That said, those who joined the state under Yeltsin were nevertheless underrepresented—suggesting that selective incentives still operated among this group, as predicted. Selection also cannot easily explain why mobilization was generally low among those who joined the public sector during the Soviet period, but increased markedly among those at or near retirement age with a pension to fall back on (see the right side of Figure A4).[11]

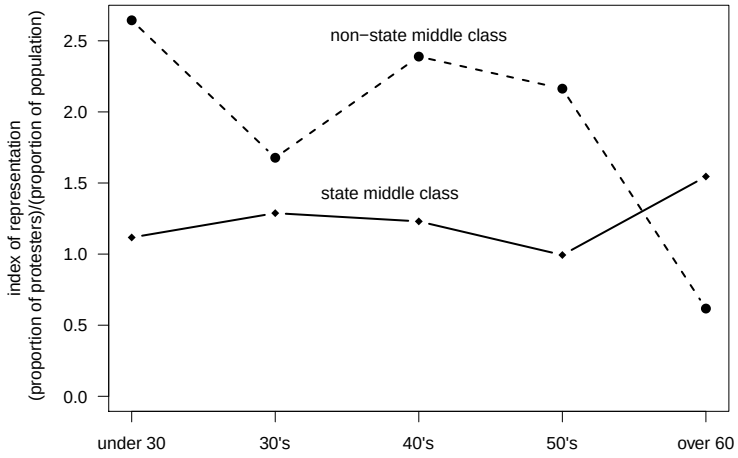



Figure A4: **Exit Options and Protest Participation (Middle Class).** This figure compares middle class mobilization by sector of employment at different points in the lifecycle. The y-axis plots an index of representation, with higher values indicating greater overrepresentation. On the x-axis are age cohorts. Note that this analysis holds constant educational and occupational characteristics (which do not vary within the middle class) along with the labor market (Moscow) which may also be associated with available alternatives. The results are consistent with the claim that lifecycle variation in available alternatives to state employment affects the power of state selective incentives and, thus, rates of mobilization. Data sources: The Levada Center and the Foundation for Public Opinion (FOM).

By contrast, an account based on selective incentives provides a straightforward explanation for

[11] This group has certainly been politically active before, as in the 2005 protests against the monetization of benefits. However, explanations which focus only on generational cohort effects cannot explain the observed differences across sectors.

these patters. Those state employees who had a pension to fall back on were most likely to protest, while those with the fewest exit options in the youngest cohort were least likely to jeopardize their jobs by taking to the streets. In sum, while we cannot exclude the possibility that selection may have played a role in higher protest participation among Yeltsin-era state employees, variation in the power of selective incentives arguably better accounts overall for the distinctive patterns in protest turnout observed in these data.

## A.10 Calculation of Counterfactual Protest Participation

Let $s_1$ be the rate of protest participation for the public sector and $s_0$ the rate of protest participation for the private sector.

The predicted probabilities of protest participation based on the model in Table 2 are:[12]

$$s_1 = 0.0698$$

$$s_0 = 0.1122$$

The number of public sector workers in Moscow and Moscow oblast, $n_{state}$, is calculated:

$$\begin{aligned} n_{state} &= N_{Moscow/MoscowOblast} \times share_{public-sector} \\ &= 18,500,000 \times 0.25 \\ &= 4,597,082 \end{aligned}$$

The number of additional protesters who would have taken to the streets if state workers protested at the rate of everyone else is thus:

$$\begin{aligned} \textit{Additional Protesters} &= s_0 \times n_{state} - s_1 \times n_{state} \\ &= 194,842 \end{aligned}$$

[12]These predicted probabilities were calculated by averaging over the empirical distribution of all other covariates in the population sample, using the observed-value approach in Hanmer and Kalkan (2013).

Similarly, let $m_1 s_1$ be the rate of protest participation for the public sector middle class and $m_1 s_0$ be the rate of protest participation for the private sector middle class.

Again, based on the model in Table 2:

$$m_1 s_1 = 0.1336$$

$$m_1 s_0 = 0.1814$$

The size of the public sector middle class in Moscow and Moscow oblast, $n_{state,mc}$, is computed as follows.

$$\begin{aligned} n_{state,mc} &= N_{Moscow/MoscowOblast} \times share_{public-sector,mc} \\ &= 18,500,000 \times 0.10 \\ &= 1,879,779 \end{aligned}$$

The number of additional protesters who would have taken to the streets if the state-sector middle class protested at the rate of the non-state middle class is thus:

$$\begin{aligned} \textit{Additional Protesters} &= m_1 s_0 \times n_{state,mc} - m_1 s_1 \times n_{state,mc} \\ &= 89,756 \end{aligned}$$