Routledge
Taylor & Francis Group

Journal of Elections, Public Opinion and Parties

ISSN: 1745-7289 (Print) 1745-7297 (Online) Journal homepage: http://www.tandfonline.com/loi/fbep20

The social desirability bias in autocrat's electoral ratings: evidence from the 2012 Russian presidential elections

Kirill Kalinin

To cite this article: Kirill Kalinin (2016): The social desirability bias in autocrat's electoral ratings: evidence from the 2012 Russian presidential elections, Journal of Elections, Public Opinion and Parties, DOI: 10.1080/17457289.2016.1150284

To link to this article: http://dx.doi.org/10.1080/17457289.2016.1150284

JOURNAL OF ELECTIONS, PUBLIC OPINION AND PARTIES, 2016
http://dx.doi.org/10.1080/17457289.2016.1150284

Routledge
Taylor & Francis Group

The social desirability bias in autocrat's electoral ratings: evidence from the 2012 Russian presidential elections

Kirill Kalinin

Department of Political Science, University of Michigan, Ann Arbor, MI, USA

ABSTRACT

In authoritarian regimes, election polls can be vastly polluted by measurement error, namely the social desirability bias, which can contribute to substantial inflation in the publicized estimates of an autocrat's electoral support and voter turnout, seeming to validate falsified election outcomes that match the inflated estimates. This study provides an in-depth analysis of the magnitude of social desirability bias in polling estimates released before and after the 2012 Russian presidential elections by focusing on the implications of Noelle-Neumann's "spiral of silence" theory. The empirical data analysis is based on list experiments from four data samples collected during the Russian presidential campaign. The estimated magnitude of the social desirability bias in Putin's electoral support is statistically significant and reaches approximately 15%. For voter turnout, however, I find social desirability bias of the same order as in Western democracies. My main conclusions are further validated by an alternative urns experiment conducted by one of the national pollsters. The detection of significant social desirability bias in the Russian presidential campaign raises the issue of survey research quality in authoritarian regimes and its effect on election outcomes.

Introduction

The 2012 Russian presidential elections were marked by the spread of massive protests associated with growing public awareness of election fraud allegations. However, numerous election forecasts produced by major national polling organizations suggested an overwhelming lead by Putin in the electoral ratings that matched the official elections results within the margin of survey error (Shpilkin 2011; Kalinin and Shpilkin 2012; Enikolopov et al. 2013; Kalinin and Mebane 2013). This paper addresses a key question: if election and survey results are roughly similar, and we know the former are artificially heightened, then what is pushing the survey results artificially upward? The observed "nonsensitivity" of polling estimates to election fraud could be

CONTACT Kirill Kalinin kirill.kalinin@gmail.com Department of Political Science, University of Michigan, Ann Arbor, MI, USA

Supplemental data for this article can be accessed at 10.1080/17457289.2016.1150284.

attributed to the inflation of electoral ratings caused by a social desirability bias.

The proposed theory builds on the notion of the "matching game", as well as Noelle-Neumann's seminal work on the "spiral of silence". It shows that when an individual feels unsafe about expressing ideas contrary to official policy or fears isolation, she will most likely hide her private preferences and display public preferences in favor of a candidate she thinks is most accepted by the general public. This observation relates especially well to the general climate of the 2012 Russian presidential elections, in which extensive media propaganda and the abuse of administrative resources by the incumbent made clear that Vladimir Putin was the most socially desirable candidate. The resulting inflation in forecasts of the election outcome contrasted against lower observed electoral figures could have permitted election administrators to organize election fraud up to the level of the discrepancy, effectively hiding the extent of election rigging.

The phenomenon of social desirability bias has been studied with respect to a wide range of issues in both autocracies (Geddes and Zaller 1989; Bischoping and Schuma 1992; Anderson 1994; Sieger 1990; Beltran and Valdivia 1999), as well as democracies (Belli, Traugott, and Beckmann 2001; Karp and Brockington 2005; Holbrook and Krosnick 2010b). According to Presser (1990) and to Traugott (2008), over-reporting of voter turnout has been found in every major validation study. In this paper, social desirability bias is understood as the difference in the probability of the socially desirable response and the honest response, where anonymity is strictly guaranteed. This becomes possible through the use of the list experiment or item count technique (ICT) (Miller 1984), which over the years has become a popular tool providing a reliable control for social desirability bias (Tsuchiya 2005; Chaudhuri and Christofides 2007; Corstange 2009; Imai 2011; Green and Kern 2012; Glynn 2013).

In contrast to previous work, this paper makes three novel contributions. First, it claims that the social desirability bias is a key factor explaining the close match between falsified official electoral results and election polling results on the incumbent's electoral support. Second, unlike the political science literature focusing on exploration of desirability bias with respect to turnout, this study (while not ignoring possible turnout bias) concentrates on the autocrat's electoral support. Third, within the context of my theory, the computed estimate of the social desirability bias can be treated as a proxy measure of election fraud. By and large, this study helps to bridge the gap between election forensics and survey methodology in autocracies and helps to validate a technique that has proved to be efficient in democracies. In a broader perspective, it adds leverage to our understanding of how the study of election polls can complement the field of election forensics.

The structure of this paper is as follows. The first section lays out Noelle-Neumann's "spiral of silence" theory and offers a set of key hypotheses for testing. The second section discusses the analytic strategy. The third section describes the political context of the 2012 Russian presidential election. The fourth and fifth sections provide description of the data and results of my empirical analysis. In the final part, I draw conclusions based on these findings and discuss prospects for future research.

Theory and hypotheses

Authoritarian regimes, with high cohesion or compliance of the state apparatus and control over the mass media, enable autocrats to effectively manage public sentiment towards the regime. When the regime dominates public political discourse, a "spiral of silence" is likely to emerge, where citizens conceal their private preferences and instead report socially desirable preferences (Noelle-Neumann 1984). Therefore, the autocrat is always concerned about the possibility of social desirability bias and tries to keep herself informed of the private preferences of constituencies by withholding and manipulating polling data (Otava 1988). This is especially true for election polls, which demonstrate respondent's compliance with the regime. Key questions, such as a voter's decision to vote for the autocrat, are almost always perceived as politically sensitive by the respondents, who are then more likely to falsify their answers. These questions are therefore most troubling for the autocrat as well. For instance, social desirability bias and low response rates have been attributed to mail surveys in the USSR, since Soviet citizens refrained from giving sincere responses even to officially approved questions (Kaplowitz and Shlapentokh 1982).

While the presence of a close match between inflated polling estimates and official electoral results guarantees the most desirable outcome for the autocrat, the occurrence of a discrepancy can trigger allegations of election fraud and mass protests. For instance, the autocrat's inflated support in pre-election polls followed by clean and fair election would most likely end up in a serious mismatch between the two figures, indicating autocrat's lower public support than otherwise expected by the citizenry, and thus potentially triggering collective action with relatively high costs for the autocrat. Another potential setback for the autocrat can be due to unbiased polling estimates creating an expectation of incumbent's relatively low public support, thus making election fraud easily detectable by the voters, consequentially raising the probability of mass protests. Indeed, the importance of matching outcomes vs. mismatching outcomes between the election results and election polls can by described be a simple "matching pennies" game, in which matched pennies result in a win of one player (the autocrat), and unmatched pennies contribute to a win of the other player (the voters). Unsurprisingly,

any information leading voters to conclude a "mismatched" outcome has occurred and the consequent possibility of protest activities incentivizes the autocrat to repress various information channels designed to provide the citizenry with information on the true popularity of the incumbent and untie her hands for voter fraud (Wintrobe 1998; Egorov, Guriev, and Sonin 2009). Therefore even if autocracies and democracies shared relatively similar levels of social desirability bias in polling estimates, the social desirability bias in autocracies can be much more consequential: matched outcomes in autocracies due to election fraud, and unmatched outcomes in democracies due to its absence.

Noelle-Neumann's seminal work on the "spiral of silence" helps us to explore the mechanism by which social desirability evolves in autocracies, such as Russia (Noelle-Neumann 1984). It implies that when one faction in society possesses total public visibility while the other has been completely marginalized, the individual will assess the political climate personally or through the media. This assessment impacts the public behavior of an individual and her willingness to reveal her private political preferences. In particular, the revelation of her private preferences is less likely when an individual feels unsafe about expressing ideas contrary to official policy, or has a fear of isolation. Under these circumstances, she will be inclined to falsify her political preferences in favor of those she thinks are acceptable to the public, or simply withdraw from discussion (Noelle-Neumann 1984, 5). For instance, the general setting of the Russian presidential elections 2012 with the incumbent's dominant position was likely to trigger sensitivity to questions related to his electoral support; supporters of political opposition, being in the minority, would feel pressured to remain silent or to falsify their preferences in favor of the incumbent.

This research will be primarily focused on testing the central implication of the "spiral of silence" theory: on inflation in the estimates of the incumbent's electoral support. In the social sciences, the individual's inclination to falsify her preferences is referred to as social desirability bias, and in political science as preference falsification (Kuran 1991). The concept implies that polled individuals may give dishonest answers to conform with societal norms and so as not to be embarrassed by their responses, thus contributing to an increase in response bias and measurement error. According to Tourangeau, Rips, and Rasinski (2000, 257) the "notion of sensitive questions presupposes that respondents believe there are defining desirable attitudes and behaviors, and that they are concerned enough about these norms to distort their answers to avoid presenting themselves in an unfavorable light". According to Kuran, preference falsification "is the act of misrepresenting one's genuine wants under perceived social pressures" (Kuran 1991, 37–57). Social desirability and self-censoring can substantially affect respondents' responses due to social sanctions and risks arising from the respondent's

decision to voice her support for the opposition or contentious opinions (Tourangeau, Rips, and Rasinski 2000). When there is almost no benefit in answering the questions truthfully, the individual would be more likely to subscribe to this strategy (Corstange 2009, 2–3).

With respect to voting and voter turnout, social desirability has been explored in the works of Streb et al. (2008), Belli, Traugott, and Beckmann (2001), Holbrook and Krosnick (2010a, 2010b), Comsa and Postelnicu (2012). In their cross-national study of turnout in several democracies, Karp and Brockington (2005) find that in national settings with higher levels of participation, the tendency to over-report turnout is greater than in settings with low participation levels. There are several studies focused on survey experiments in an authoritarian setting. Weghorst (2011) in his paper on support for opposition violence against an incumbent party shows the difference between the results from direct and indirect self-reporting. Gonzalez-Ocantos et al. (2011) provide evidence showing that after the 2008 Nicaraguan municipal elections, a direct question on vote buying, compared with responses from a survey experiment, clearly underestimated the proportion of those who reported this behavior.

How much social desirability bias would we expect to find in our study? In their research Karp and Brockington (2005) by comparing official voting records with survey responses on turnout from several democracies conclude that the social desirability bias in democracies not only positively correlates with the turnout rate, but also demonstrates high levels reaching 27% for Britain, 26% for Sweden and 40% for the US. Turning to incumbent support, since other mature democracies have nothing like the problem that the US has in terms of information and bandwagon effects with regard to incumbent's support, the US seems to be the most appropriate case for such comparison. For instance, if in the US presidential election there appears little bias, reaching 1.5%. For the House of Representatives and US Senate races it is slightly higher, reaching 4%, for gubernatorial races it is about 5% (Wright 1993, 295).

Which US election can be taken as a benchmark for our analysis? Since the Russian presidential election belongs to the low information environment where it is likely that many will not have heard of any candidate apart from the front-runner, as a benchmark the measure from similar setting needs to be taken. Even though at first glance the US presidential elections seem to provide the most plausible benchmark for comparison, their high salience and competitiveness make such comparison disputable. Instead, however, by taking into consideration the US gubernatorial and congressional elections, known for their low salience and strong incumbency effects, I thus find the most appropriate way of handling this problem. Presumably, if the inflation in turnout exceeds 30% or in Putin's votes exceeds 5%, one can argue that there is no doubt about the significance of the additional bias found in

Russia. I surmise that any such discrepancy is due in large part to "spiral of silence" motivations. Thus my hypothesis is as follows:

> H1. Due to social desirability effects, Russian presidential elections in 2012 were characterized by a substantial inflationary bias in the estimates for incumbent and voter turnout, exceeding the democratic benchmark of 5% for the incumbent's support and 30% for turnout.

Another question of interest is related to the variability of the social desirability bias across different demographic subgroups. Individuals displaying high levels of social desirability bias in contexts analogous to support for Putin are those who are relatively isolated, with weaker self-confidence, less interest in politics and fear disrespect or unpopularity. Among those most disposed to speaking out publicly are rather men than women, younger people than older ones, those belonging to a higher social strata than those from lower strata (Noelle-Neumann 1984, 24). However, when it comes to turnout, a number of studies have found that individuals most likely to over report voting have the same characteristics as those likely to vote: those who are highly educated, supportive of the regime, higher status individuals are most likely to falsify their preferences in a survey (Silver, Anderson, and Abramson 1986; Bernstein, Chadha, and Montjoy 2001). They are more likely trying to create a good impression on the interviewer and feel pressured to vote, leading to a greater desire to falsify their responses (Silver, Anderson, and Abramson 1986; Bernstein, Chadha, and Montjoy 2001). Those respondents for whom the norm of voting is most salient will be most likely to over report their preferences. Hence, here I formulate different hypotheses for Putin support than for turnout:

> H2.1. The social desirability bias in estimates for an autocrat's electoral support is expected to vary across different social groups: women, the elderly, persons with low education and the poor are expected to exhibit greater levels of misreporting compared to men, the young, those with higher education and the wealthy.
> H2.2. Over-reporting individuals with respect to voter turnout are expected to have the same demographic characteristics as the likely voters, i.e. those individuals who truthfully report having voted.

The next question is related to the temporal persistence of preference falsification throughout the electoral campaign. According to Noelle-Neumann (1984, 31), one might expect a recognizable pre-election tendency of people claiming they are going to vote for the incumbent, but a post-election tendency to claim support for the incumbent can be even more salient, with a greater proportion of people falsifying their preferences. This comes as no surprise, especially, when the autocrat has managed to manufacture much of her official electoral support, thus creating the illusion of enormous popularity contrasted with the chronic weakness of political opposition (Simpser 2013). If, however, the autocrat did worse in the elections than pre-election polls

reported, the magnitude of social desirability bias would most likely decrease and contribute to the growth of public opposition (Kuran 1991, 18). Since Putin's official results matched public expectations quite well, I expect the presence of temporal persistence of preference falsification to be without any significant changes between pre- and post-electoral periods.

H3. For pre-election and post-election polls, one can expect the presence of temporally persistent social desirability bias for both voter turnout and incumbent's electoral support.

Analytic strategy

Social desirability bias can be simply understood as the difference in the probability of the socially desirable response and the honest response, where anonymity is strictly guaranteed. While the socially desirable response is measured by a direct questioning technique, the honest response is measured by indirect techniques. My analytic strategy employs both types of techniques, making it possible to measure the magnitude of the social desirability bias.

Indirect questioning techniques are specifically designed to offer respondents an opportunity to answer truthfully without a fear of retribution. A list experiment or ICT, which is a type of indirect questioning technique, uses random assignment of respondents to treatment and control groups (Miller 1984). Both groups receive the same number of non-sensitive items, and the treatment group receives an extra item of a sensitive nature. The respondents are asked to provide the number of items they agree with. The estimated true proportion of respondents supporting the sensitive item is computed as the difference between the average number of statements reported by the treatment group and the average number of statements reported by the control group.

Many studies have shown that ICT provides a reliable control of the social desirability bias (Tsuchiya 2005; Chaudhuri and Christofides 2007; Corstange 2009; Imai 2011; Green and Kern 2012; Glynn 2013). Nonetheless, the technique has several established limitations. For instance, Kiewiet and Nickerson (2014) illustrate that ICT can provide extremely conservative estimates of high incidence behaviors; Glynn (2013) mentions its relative inefficiency and its failure to provide individual-level measures for the sensitive item. Moreover, the anonymity condition can be violated due to a ceiling effect (when all of the statements are chosen by a respondent) or a floor effect (when none of the statements is chosen), which can be partly overcome by increasing the length of the list (Tsuchiya, Kirai, and Ono 2007; Holbrook and Krosnick 2010a). While the core assumptions of the technique hold, list experiments can generate valid estimates of social desirability bias. However, violations of these assumptions can substantially undermine the validity of obtained results.

In the present study, I have conducted two list experiments designed to measure genuine electoral preferences for Putin and voter turnout:

> Here is the list of four/five assertions. Please, listen to them all and then tell me how many you agree with. Do not tell me which assertion you agree or disagree with, just give me the total estimate.
> • I usually read at least one newspaper or magazine a week;
> • I want to see Russia as a country with high-living standards;
> • I can recall the name of the head of Russian Constitutional Court;
> • ***I will vote [voted for] Vladimir Putin in the upcoming/most recent Presidential elections (March 4);***
> • I am satisfied with the level of my income.
> I agree with___ (number of assertions)

> Here is the list of four/five assertions. Please, listen to them all and then tell me how many you agree with. Do not tell me which assertion you agree or disagree with, just give me the total estimate.
> • My family has a car;
> • I can recall the name of the head of the Federation Council;
> • I watch TV daily;
> • ***I'll vote [I voted] in the upcoming/most recent Presidential elections (March 4);***
> • The level of pension in our country is quite high.
> I agree with___ (number of assertions)

Context

The Russian presidential campaign 2012 was given a jumpstart when then-President Medvedev proposed then-Prime Minister Vladimir Putin run for a third term in the fall 2011. This pre-arranged move ignited widespread public discontent and set the tone for both the upcoming parliamentary and presidential elections. Since numerous election fraud allegations at the parliamentary elections held in December 2011 provoked the rise of massive protests in Moscow and St Petersburg, the Kremlin urgently launched a series of reforms aimed to provide electoral transparency of the forthcoming presidential elections, such as installation of transparent ballot boxes and web cameras in every polling station across the country. Procedural transparency was overshadowed by the incumbent's excessive use of media and state resources. Some radio and television media outlets allocated 90% of programming to favorable stories about Putin (OSCE 2012). In addition to this, the incumbent's heavy reliance on the use of the state apparatus, namely local political machines, enabled him to secure an electoral victory by using a wide range of legal and illegal techniques (Kalinin and Mebane 2013). Numerous instances of electoral violations and election fraud, reported by election observers and election forensics specialists, provide enough evidence to conclude that elections were far from being clean and fair. For example, according to some of the studies, election fraud amounted to 5% and 10% of Putin's

electoral support and turnout, respectively (Kalinin and Shpilkin 2012); other studies provide the estimated proportion of precincts with election fraud, about 40% (Klimek et al. 2012). Surprisingly, however, all national pollsters demonstrated close proximity of the election forecasts (FOM: 58.7%, VCIOM: 59.9%, Levada-Center: 66.0%) to official statistics 63.6%.

Data

The fieldwork stage of the project was conducted over a period of two months between early February and early April 2012. Three surveys were conducted by Levada-Center (Omnibus/Courier Study, 17–20 and 24–27 February; March 2012), and one by the Demoscope (Russian Election Study, March–April 2012) with 1500–1600 respondents sampled for each survey (Colton et al. 2014). The major characteristics of the surveys can be found in Table A1 of the online Appendix. First, the sampling design employed by Levada-Center is typically utilized by the majority of national survey organizations. It is based on a proportional stratified sampling with the selection of households completed by the random route method and the selection of respondents by using quota sampling. In contrast, Demoscope's sample is based on multistage probability sampling with respondents selected by the nearest birthday method. All the surveys encompass both direct and indirect (ICT or list experiment) questions on Putin's electoral support and turnout, thus making an estimation of social desirability bias possible. The ICT experiment was preceded by direct question asked 5–10 minutes earlier as in Droitcour et al. (1991), because of concerns on the part of the pollsters about the unforeseen effects of list experiments on the estimates derived from the direct question. Moreover, the ICT experiment on turnout preceded the experiment on voting, being 5–6 questions apart from the latter.

Table 1 contains a summary of weighted demographic characteristics across four surveys. By and large, the collected data samples illustrate the absence of large differences across key demographic indicators, with the exception of wealthy/poor and rural/urban category for the post-electoral dataset. In order to simplify subsequent empirical analysis, I resort to data pooling, first by producing pre-electoral and post-electoral datasets, and second, by producing one single data set from all of the data.

In both cases, my data pooling has been justified by a statistical testing procedure equivalent to a likelihood ratio test for the pooled and unpooled data samples. All the missing values in the independent variables have been replaced with the medians or means.

In their methodological study, Blair and Imai (2012) list three identification assumptions for an ICT to be met: (a) randomization of the treatment, that is, randomization of respondents into the treatment and control groups; (b) no design effect, that is, the presence of the sensitive item in the list does not

Table 1. Demographic characteristics of respondents from four surveys, in percentages.

	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5
Gender					
Male	46.3	46.3	46.3	44.6	45.0
Female	53.7	53.7	53.7	55.4	55.0
Missing	0.0	0.0	0.0	0.0	0.0
Total	100.0	100.0	100.0	100.0	100.0
Age					
18–24	13.9	13.9	13.9	15.0	13.5
25–34	21.0	19.8	20.1	17.9	19.6
35–44	16.3	18.3	15.8	17.3	17.1
45–54	18.6	17.9	20.0	18.0	18.3
55–64	16.6	17.8	18.0	16.7	17.0
65–74	9.0	8.4	8.7	9.1	9.3
75 and older	4.5	3.9	3.5	6.0	5.2
Missing	0.0	0.0	0.0	0.0	0.0
Total	100.0	100.0	100.0	100.0	100.0
Education					
High school	70.7	70.7	70.7	67.6	69.25
College	29.3	29.3	29.3	31.6	30.61
Missing	0.0	0.0	0.0	0.8	0.14
Total	100.0	100.0	100.0	100.0	100.0
Well-being					
Wealthy	21.4	26.1	26.0	14.4	21.5
Poor	78.2	73.7	73.6	84.6	77.9
Missing	0.4	0.2	0.4	1.0	0.6
Total	100.0	100.0	100.0	100.0	100.0
Residence					
Urban	51.4	51.1	50.0	70.0	55.7
Rural	48.6	48.9	50.0	30.0	44.3
Missing	0.0	0.0	0.0	0.0	0.0
Total	100.0	100.0	100.0	100.0	100.0
Sample size	1601	1601	1633	1682	6517

Notes: Survey 1 – Levada-Center (17–20 February); Survey 2 – Levada-Center (24–27 February); Survey 3 – Levada-Center (March); Survey 4 – Demoscope (March–April) and Survey 5 –Pooled data.

affect the way the respondent thinks about the control items; (c) no liars, that is, respondents share their truthful answers when asked an ICT question. Any violation of these assumptions can potentially lead to the failure of the list experiment, and thus distort the estimates. The overall quality of designed experiments is assessed on the basis of whether all three assumptions were satisfied.

For determining the overall validity of the conducted ICT experiment, I test the presence of any apparent violations in the key ICT assumptions, such as randomization of treatment and the absence of liars. The randomization assumption is tested by regressing treatment assignment on covariates using both pre-electoral and post-electoral data sets. Almost all of the covariates are statistically insignificant, which is indicative of successful randomization of the treatment (Table A3). However, it has been impossible to achieve completely balanced treatment assignments. In both regressions, one of the covariates turns out to be statistically significant: "Sex" in the case of the pre-electoral data set and "Wealth" in our post-electoral data.

The assumption that the presence of floor and ceiling effects compromises anonymity has been also satisfied. According to Table A4, depicting the frequencies of item counts for voting and turnout; this problem has been minimized by the proper pre-selection of non-sensitive items, which reduced the frequency with which 0s and 5s appear in the data. Finally, since it is impossible to empirically test the design effect assumption in my experimental setting, I am forced to make strong assumption about the absence of design effect in my analysis.

Results

My initial findings with respect to the means for direct and indirect ICT self-reporting (ICT) as well as the difference between them, that is, social desirability bias (Δ) and associated bootstrapped standard errors are presented in Table 2. One of the potential challenges when using the ICT measures is the loss in efficiency of the estimator, that is, computed standard errors for the list experiment are four times larger than for the direct self-report. In spite of this fact, the presence of a strong statistically significant social desirability effect is observed in both pre-electoral and post-electoral settings. The difference in means estimator (DIM section of the table) shows the percentage of pre-electoral support for Putin among the participants reaching 47.2%, with s.e. 4.2%, while the direct self-report yields the estimate of 66.1% (1.3%). The difference between both figures gives us the estimate of social desirability bias of 18.9%, which is statistically significant at $\alpha = 0.05$ for a one-tailed *t*-test. The pre-electoral estimates for turnout show somewhat similar patterns: the indirect self-report yields an estimate of 50.7% (4.1%), and the direct

Table 2. Electoral support using direct and indirect self-report (%)[a].

		Pre-electoral period			Post-electoral period		
		Direct	Indirect	**Δ**	Direct	Indirect	**Δ**
DIM	Incumbent (All)	45.3	28.5	**16.8**	47.8	34.6	**13.2**
		(1.1)	(3.7)	**(3.7)**	(0.9)	(3.2)	**(3.2)**
	Incumbent (Voted)	66.1	47.2	**18.9**	68.8	52.1	**16.7**
		(1.3)	(4.2)	**(4.3)**	(1.0)	(3.4)	**(3.6)**
	Turnout (All)	67.1	50.7	**16.4**	75.9	55.3	**20.6**
		(1.0)	(4.1)	**(4.0)**	(0.8)	(3.7)	**(3.7)**
Modeled	Incumbent (All)	67.1	42.2	**24.9**	68.6	36.7	**31.9**
		(6.5)	(16.3)	**(16.2)**	(7.9)	(16.0)	**(11.5)**
	Turnout (All)	72.9	64.7	**8.2**	75.6	72.6	**3.0**
		(6.4)	(24.5)	**(19.5)**	(10.7)	(24.5)	**(15.3)**
	Sample size		3202			3315	

[a]DIM – estimates computed with standard difference in means estimator (weighted); Modeled – estimates computed using multivariate regression analysis with *list() package*. *Incumbent(All)* – percentage of Putin's supporters among all respondents; *Incumbent(Voted)* – percentage of Putin's supporters among those who intend to vote [voted]; *Turnout* – percentage of those who will vote [voted]. Monte Carlo standard errors are in parentheses. According to official election results, Putin received 63.6% of the popular vote, turnout reached 65.34%.

technique yields 67.1% (1.0%). The difference of 16.5% is statistically significant. In addition, using Blair and Imai's (2012) package on multivariate analysis of ICT, I also control for ceiling and floor effects (see "Modeled" section in the table). When using this control, even though the presence of social desirability seems to be prevalent for Putin's electoral support, the estimates for turnout seem to be more ambiguous. None of the social desirability estimates is statistically significant. Hence, my theoretical implications from *hypothesis 1* are well supported by the data: both the difference in means estimator and the modeled portion of my analysis demonstrate the presence of strong social desirability bias only in the incumbent's estimates, though failing to agree in the case of turnout. Basic findings illustrate that the magnitude of social desirability bias with respect to turnout is quite comparable with democracies while with respect to Putin's support observed inflation turns out to be higher compared to democratic benchmark defined earlier.

The comparison of the pre-electoral and post-electoral estimates shows the absence of statistically significant difference. Social desirability bias is persistent across both settings without any notable changes in the estimates. The statistically significant difference between pre-electoral 67.1% (1.0%) and post-electoral estimates 75.9% (0.8%) for voter turnout is the only exception to this. The observed higher levels of social desirability bias with regard to turnout reveals the growth in proportion of those who falsify their true preferences after incumbent's win, which is partly supportive of the "spiral of silence" theory. Thus, my hypothesis 3 seems to be also supported by the data.

Holbrook and Krosnick (2010a) argue that comparison between survey estimates and official estimates can be troublesome, because the numerous methodological problems associated with surveys – such as undercoverage, nonresponse – produce different kinds of survey biases in the estimates. However, it is still potentially useful to explore how well our survey estimates match the results from official statistics. According to official election results, Vladimir Putin received 63.6% of the popular vote with voter turnout reaching 65.3%. Both Putin's estimated support with the direct self-report and turnout for pre-election and post-election studies illustrate close proximity between the survey estimates and official statistics, a difference that is statistically indistinguishable from zero.

Along with the basic difference-of-proportions estimates, I also resort to multivariate analysis to test the hypotheses about the levels of bias across different subgroups.

This is done by regressing the answers to the direct questions on the set of socio-demographic covariates using the binary logit, as well as regressing the answers to the list experiment on assignment to the treatment condition, the set of socio-demographic covariates, and interactions between assignment and these covariates (Holbrook and Krosnick 2010a; Kiewiet 2015). I further implement the list experiment models using the Blair and Imai (2012)

maximum likelihood estimator, which enables me to obtain more efficient estimates of the social desirability bias. The set of predictors included in the model is as follows: *sex, age, education, a measure of subjective well-being of an individual* (also termed as *wealth* or *rich vs. poor* in the text) and type of residence (*rural/urban*).

Table 3 illustrates the estimated treatment effect (*Treatment*) in bold, that is, the magnitude of social desirability bias (Δ) after the inclusion of demographic regressors and interaction terms. The models 2 and 6 based on the list experiment show the estimated share of genuine support for the incumbent reaching 0.31(0.18) before the elections, and 0.42(0.18) after the elections; as far as turnout is concerned, the estimates for models 4 and 8 are 0.43(0.19) and −0.19(0.19), respectively.

Although the regression results exhibit the failure of post-electoral estimates regarding turnout to demonstrate statistical significance, in regard to support for Putin this analysis provides us with solid evidence about the presence of social desirability bias by controlling for socio-demographic imbalances between our two groups.

More importantly, Table 3 also displays socio-demographic profile of Putin's supporters. Specifically, when all other variables are held constant at their mean values for both pre- and post-electoral models, the probability of Putin's direct support increases if the respondent is a woman, younger and less educated. Putin's indirect support, however, shows quite different patterns. Individuals most likely to vote for Putin are older, better educated, wealthier and living in rural areas. Regarding turnout, the post-electoral surveys indicate that respondents most likely to vote are older, better educated, wealthier and live in rural areas; participation measured by indirect technique indicates that women, better educated and wealthy individuals are most likely to attend elections.

Since the ordinary least squares estimates can be inefficient when dealing with the list experiments, I also apply an alternative estimation procedure using Blair and Imai's (2012) linear and non-linear least square estimators as well as two models using maximum likelihood estimators. My findings from this auxiliary analysis demonstrate moderate consistency in the direction of the demographic effects between the direct and indirect questions (see Table A5 in the online Appendix).

In order to make the results more readily interpretable, in the next stage I resort to computing individual-level measures of social desirability bias by finding the differences in predicted probabilities for direct (the binary logit model) and indirect responses (maximum likelihood estimator, as suggested by Blair and Imai (2012)). All measures have been aggregated up to the level of social groups by computing corresponding means and standard errors. Figure 1(a) and 1(b) depicts the means as filled circles and 95% confidence intervals as solid lines for all three computed measures. The direction of social

Table 3. Effects of demographic variables on incumbent's electoral support and turnout among all respondents (weighted).

	Pre-electoral period (February 2012)				Post-electoral period (March–April 2012)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	1.363***	0.706***	−0.127	0.886***	1.999***	0.777***	−1.351***	1.220***
	(0.376)	(0.125)	(0.376)	(0.133)	(0.423)	(0.125)	(0.423)	(0.133)
Treatment		**0.314***		**0.432****		**0.424****		**−0.194**
		(0.176)		**(0.188)**		**(0.178)**		**(0.191)**
Sex	−0.520***	0.001	−0.105	0.024	−0.474***	0.096**	−0.170	0.139***
	(0.125)	(0.042)	(0.125)	(0.044)	(0.132)	(0.042)	(0.132)	(0.045)
Age	−0.008**	0.006***	0.019***	−0.001	0.002	0.007***	0.034***	0.001
	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)	(0.001)
Education	−0.105*	0.059***	−0.040	0.088***	−0.197***	0.089***	0.199***	0.062***
	(0.063)	(0.021)	(0.063)	(0.022)	(0.069)	(0.021)	(0.069)	(0.023)
Wealth	0.094	0.162**	0.101	0.186***	−0.042	0.090***	0.206***	0.084***
	(0.075)	(0.025)	(0.075)	(0.027)	(0.078)	(0.024)	(0.078)	(0.026)
Urban	−0.184	−0.089**	−0.164	−0.049	−0.355**	−0.034	−0.394**	−0.033
	(0.130)	(0.043)	(0.130)	(0.046)	(0.137)	(0.043)	(0.137)	(0.046)
Treatment:Sex		−0.052		0.005		−0.154***		−0.089
		(0.059)		(0.063)		(0.059)		(0.064)
Treatment:Age		0.002		0.003		0.002		0.005**
		(0.002)		(0.002)		(0.002)		(0.002)
Treatment:Education		0.020		0.006		−0.014		0.095***
		(0.030)		(0.031)		(0.029)		(0.032)
Treatment:Wealth		−0.040		−0.029		0.043		0.110***
		(0.036)		(0.038)		(0.034)		(0.036)
Treatment:Urban		0.050		0.040		−0.196		−0.180***
		(0.061)		(0.065)		(0.060)		(0.065)
N sample	2478	3130	2478	3134	2268	3260	2268	3260

Notes: Models: (1) and (5) – direct self-report for voting; (2) and (6) – indirect self-report for voting; (3) and (7) – direct self-report for turnout; (4) and (8) – indirect self-report for turnout.

*$p \leq .1$.

**$p \leq .05$.

**$p \leq .01$.

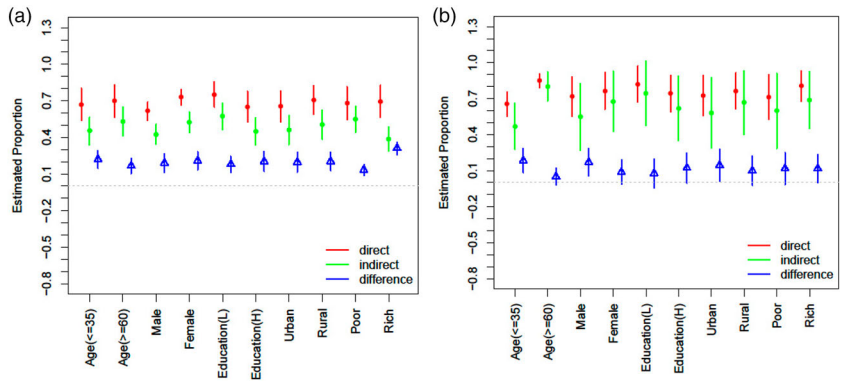

Figure 1. Estimated proportions for voting and turnout by subgroups.

desirability bias is marked as upper/lower triangle characters, depending on the sign of obtained bias. The estimates of bias for each subgroup can be found in the online Appendix (Table A6). Figure 1(a) illustrates that almost all of the estimated proportions are bounded between 0 and 1 (with the exception of the *rich* category), which is consistent with our expectations. Specifically, Putin's public support is in line with the previous work, indicating that he is more likely to be publicly supported by richer than poorer, rural than urban residents (Rose, Mishler, and Munro 2011).

Indeed, the social desirability bias is not evenly distributed across the sample population. In Figure 1(a), all social groups demonstrate high and statistically significant levels of social desirability bias: the rich, young and female demonstrate the highest values of social desirability bias while the poor, old, uneducated demonstrate the lowest.

The younger age group is likely to inflate their electoral support of Putin by 22%, and the old, by 17%. Men are likely to falsify incumbent support by about 19%, and women, by 21%. Those with higher education contribute to inflation by about 20%, and those with lower education, by about 18%. The poor seem to be least biased inflating their support for Putin by 13%, while the rich are the most biased, inflating it by 31%. A small distinction in the levels of the social desirability bias is observed between urban and rural residents. Since for Putin's electoral support, the likely voters – women and the rich – seem to demonstrate maximum (although not statistically significant) bias, hypothesis H2.1 is not well supported. The pattern of over-reporting for Putin's support is more like what would be expected for turnout, though not significantly so. Finally, Figure 1(b) representing turnout depicts that none of the distinctions appear significant apart from the distinction between young and old cohorts, male and female, urban and rural. Since most of the computed biases are statistically insignificant, we lack evidence to confirm or reject our

Table 4. External validity check of Putin's electoral support (rounded percentages)[a].

	Official results	Urns Experiment		ICT Experiment	
		Direct	Indirect	Direct	Indirect
Moscow	47	56	51	46	42
		(52; 60)	(49; 53)	(29; 62)	(25; 58)
St Petersburg	59	64	51	59	36
		(61; 67)	(48, 54)	(34; 85)	(11; 62)
Altaiskii region	57	50	44	63	57
		(48, 52)	(42, 46)	(33; 92)	(28; 87)
Sverdlovskii region	65	72	59	52	13
		(70; 74)	(57; 61)	(8; 96)	(−31; 57)

[a]Numbers in parentheses are 95% confidence intervals. Sample sizes for the urns experiment: Moscow (direct: 675, indirect: 2283), St Petersburg (direct: 1284, indirect: 927), Altajskij region (direct: 2299, indirect: 2208), Sverdlovskij region (direct: 2346, indirect: 1578). Sample sizes for the pooled data set: Moscow (direct: 346, indirect: 534), St Petersburg (direct: 146, indirect: 218), Altajskij region (direct: 146, indirect: 205) and Sverdlovskij region (direct: 68, indirect: 105).

hypothesis H2.2. In order to gain greater confidence in the result of the list experiment, I refer to an additional external validity check by utilizing an urns experiment, organized by one of the national pollsters (VCIOM) prior to the 2012 elections. The experiment is a street survey based on non-probability sampling design: on the first day a randomly chosen respondent is directly asked by the interviewer about her electoral preferences, while on the second day a randomly chosen respondent fills out a questionnaire by herself and drops it into an urn, thus ensuring anonymity of response. The geography of the experiment included four Russian regions (Moscow, St Petersburg, Sverdlovskij region and Altajskij region), each containing four primary sampling units. The declared response rate is in the range of 54–68%.

The results of this analysis are presented in Table 4. In both types of experiments, the indirect self-reporting always yields consistently lower estimates than for direct self-reporting. The estimates based on the ICT experiments seem to largely agree with the urns experiments: almost all of the confidence intervals overlap. According to the table, when both types of experiments are compared similar or greater differences between direct and indirect estimates for the urns experiment than for the list experiment are observed, making the original findings appear more conservative in light of this robustness test. Thus, the results from the urns experiment support my main findings from the list experiments.

Conclusion

A study of social desirability bias in authoritarian regimes helps to structure and reconcile many of the diverse arguments and contradictory interpretations about the role of survey research in the electoral politics in those regimes. Even though the majority of the previous research has been focused on understanding the sources of mis-prediction of electoral outcomes by pollsters,

oftentimes due to the inflationary or deflationary effects of the social desirability bias (Bischoping and Schuma 1992; Bodor 2012), this research has taken a different approach by building on the striking observation that the election polls in authoritarian regimes are accurate even though the validity of official results is seriously questioned by the presence of blatant election fraud. In this case, the observed close congruence between the figures of official electoral support and the polling estimates can contribute to well-grounded suspicions concerning the compromised quality of election polls.

Based on anecdotal evidence from election observers and extensive scholarly research on this topic, this paper argues that in the 2012 Russian presidential elections the estimates released by Russian national polling organizations contain a substantial degree of inflationary bias with respect to Putin's electoral support, which was persistent throughout the electoral campaign. An important finding here is that the similar level of inflation in turnout for Russia and the US suggests that the inflationary mechanism is similar in both countries. Most likely this mechanism has an alternative explanation: if regarding the candidate support respondents might fear repercussions for failing to mention the "right" candidate, in turnout there are no repercussions for suggesting a failure to vote. For those who falsely claim to have voted, the resulting need to name the candidate voted for gives rise to an "information effect" (Bartels 1996) in favor of the leading candidate. Given the similarity of turnout bias in the two countries, a similar information effect can be expected in Russia, which, in this paper, we have estimated on the basis of a suitable US benchmark at 5%. As discussed earlier in the text, regarding support for Putin, the presence of discrepancy between the expected information effect of 5% derived from the US benchmark and observed bias of 20% provided by my empirical data analysis cannot be plausibly explained by any other concept than the "spiral of silence". Therefore the estimated inflationary bias once compared with the US benchmark for incumbents provides us with an estimate of about 15%. This estimate is somewhat comparable with the level of election fraud by election administrators. Though being slightly higher than election forensics research suggests, this discrepancy can be explained by the presence of measurement error in the data. My basic findings are also supported by an alternative urns experiment conducted during the period of the study, reinforcing evidence of the vulnerability of electoral polling in authoritarian regimes.

My analysis demonstrates that the distribution of the social desirability bias is not even across the sample population, yet all social groups demonstrate high and statistically significant levels of social desirability bias in regard to Putin's support with the rich, the young and women demonstrating the highest values of social desirability bias. The social desirability estimates for turnout exhibit somewhat weaker findings: the rich, men and urban residents contribute most to inflated turnout estimates.

Hence, this paper by addressing theoretical implications from Noelle-Neumann's seminal work on the "spiral of silence" has strongly suggests that when an individual feels unsafe about expressing ideas contrary to official policy or having a fear of isolation, she will most likely hide her private preferences in favor of an incumbent she thinks is most accepted by the general public. My theory suggests that the social desirability bias, inflating the autocrat's support, can be viewed as a valuable resource for the autocratic regime, effectively hiding the vote-rigging needed to convincingly win the election. Once the "spiral of silence" is effectively installed, the general public can be deceived into accepting the election fraud necessary to match the inflationary gap. The observed inflation in election ratings permits election administrators to deliver the results predicted by polls, matching the outcome desired by the autocrat.

While this research reveals the presence of significant social desirability bias in the Russian public opinion polls, on a broader scale it poses the problem of the persistence of biased responses and survey errors in authoritarian regimes for a large number of politically sensitive questions.

Acknowledgement

The author is thankful to Walter Mebane for his useful advice and suggestions, the editor and three anonymous reviewers for their important insights. He would also like to express his deepest gratitude to Henry Hale (George Washington University), Timothy Colton (Harvard University), Polina Kozyreva (Higher School of Economics), Alexei Grazhdankin (Levada-Center, Russia), Valery Fedorov and Julia Baskakova (VCIOM, Russia) who helped me with the data collection for this project.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The fieldwork stage of this project was funded by Department of Political Science (University of Michigan).

Notes on contributor

Kirill Kalinin is a doctoral candidate of Political Science at the University of Michigan, Ann Arbor. Before joining the program he was a Fulbright visiting scholar and Carnegie fellow at the Center of Political Studies (Institute for Social Research, University of Michigan). He received his M.A. in Political Science from the European University in St Petersburg and a B.A. in Public Policy from the Volgograd Academy of Public Administration. His main research interests are in the politics of authoritarian regimes, Russian politics, political methodology, election forensics and survey methodology. He has been involved in multiple collaboration projects and methodological task forces for

major Russia's polling organizations, focusing on survey experiments and survey quality analysis. He is also one of the creators of the Election Forensics Toolkit website, a prototype designed to implement fraud detection diagnostics for the use of policy-makers, practitioners, and scholars.

References

Anderson, Leslie. 1994. "Neutrality and Bias in the 1990 Nicaraguan Preelection Polls: A Comment on Bischoping and Schuman." *American Journal of Political Science* 38 (2): 486–494.

Bartels, Larry M. 1996. "Uninformed Votes: Information Effects in Presidential Elections." *American Journal of Political Science* 40 (1): 194–230.

Belli, Robert F., Michael W. Traugott, and Matthew N. Beckmann. 2001. "What Leads to Voting Overreports? Contrasts of Overreporters to Validated Voters and Admitted Nonvoters in the American National Election Studies." *Journal of Official Statistics* 4 (17): 479–498.

Beltran, Ulises, and Marcos Valdivia. 1999. "Accuracy and Error in Electoral Forecasts: The Case of Mexico." *International Journal of Public Opinion Research* 11 (2): 115–134.

Bernstein, Robert, Anita Chadha, and Robert Montjoy. 2001. "Overreporting Voting: Why It Happens and Why It Matters." *The Public Opinion Quarterly* 65 (1): 22–44.

Bischoping, Katherine, and Howard Schuma. 1992. "Pens and Polls in Nicaragua: An Analysis of the 1990 Preelection Surveys." *American Journal of Political Science* 36: 331–350.

Blair, Graeme, and Kosuke Imai. 2012. "Statistical Analysis of List Experiments." *Political Analysis* 20 (1): 47–77.

Bodor, Tamas. 2012. "Hungary's 'Black Sunday' of Public Opinion Research: The Anatomy of a Failed Election Forecast." *International Journal of Public Opinion Research* 24 (4): 450–471.

Chaudhuri, Arijit, and Tasos C. Christofides. 2007. "Item Count Technique is Estimating the Proportion of People with Sensitive Feature." *Journal of Statistical Planning and Inference* 137: 589–593.

Colton, Timothy J., Henry E. Hale, Mikhail Kosolapov, Elena Tonis, and Anna Prudnikova. 2014. "Russian Post-Presidential Election Survey in 2012." Technical Report, Harvard University. UNF:5:QgdZMJsagudKfmUflZSmdA== Harvard Dataverse Network [Distributor] V1 [Version]. http://dx.doi.org/10.7910/DVN/24202.

Comsa, Mircea, and Camil Postelnicu. 2012. Measuring Social Desirability Effects on Self-Reported Turnout Using the Item Count Technique. *International Journal of Public Opinion Research*. doi:10.1093/ijpor/eds019.

Corstange, Daniel. 2009. "Sensitive Questions, Truthful Answers? Modeling the List Experiment Multivariately with LISTIT." *Political Analysis* 17: 45–63.

Droitcour, Judith, Rachel A. Caspar, Michael L. Hubbard, Teresa L. Parsley, Wedy Visscher, and Trena M. Ezzati. 1991. "The Item Count Technique as a Method of Indirect Questioning: A Review of Its Development and a Case Study Application." In *Measurement Errors in Surveys*, edited by Paul Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz, and Seymour Sudman, 185–210. New York: John Wiley & Sons.

Egorov, Georgy, Sergei Guriev, and Konstantin Sonin. 2009. "Why Resource-poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data." *American Political Science Review* 103 (4): 645–668.

Enikolopov, Ruben, Vasily Korovkin, Maria Petrova, Konstantin Sonin, and Alexei Zakharov. 2013. "Field Experiment Estimate of Electoral Fraud in Russian Parliamentary Elections." *Proceedings of the National Academy of Sciences of the United States of America* 110 (2): 448–452.

Geddes, Barbara, and John Zaller. 1989. "Sources of Popular Support for Authoritarian Regimes." *American Journal of Political Science* 33 (2): 319–347.

Glynn, Adam N. 2013. "What Can We Learn With Statistical Truth Serum? Design and Analysis of the List Experiment and Analysis of the List Experiment." *Public Opinion Quarterly* 77: 159–172.

Gonzalez-Ocantos, Ezequiel, Chad Kiewiet de Jonge, Carlos Melendez, Javier Osorio, and David Nickerson. 2011. "Vote Buying and Social Desirability Bias: Experimental Evidence from Nicaragua." *American Journal of Political Science* 56 (1): 202–217.

Green, Donald P., and Holger L. Kern. 2012. "Modeling Heterogeneous Treatment Effects in Survey Experiments with Bayesian Additive Regression Trees." *Public Opinion Quarterly* 76: 491–511.

Holbrook, Allyson, and Jon A. Krosnick. 2010a. "Social Desirability Bias in Voter Turnout Reports: Tests Using the Item Count Technique." *Public Opinion Quarterly* 74 (1): 36–67.

Holbrook, Allyson, and Jon A. Krosnick. 2010b. "Measuring Voter Turnout by Using the Randomized Response Technique: Evidence Calling into Question the Method's Validity." *Public Opinion Quarterly* 74 (2): 328–343.

Imai, Kosuke. 2011. "Multivariate Regression Analysis for the Item Count Technique." *Journal of the American Statistical Association* 106 (494): 407–416.

Kalinin, Kirill, and Walter Mebane. 2013. "Understanding Electoral Frauds through Evolution of Russian Federalism: the Emergence of Signaling Loyalty." Paper prepared for the Annual Meeting of Midwest Political Science Association, Chicago.

Kalinin, Kirill, and Sergey Shpilkin. 2012. "Kompleksnaya diagnostika falsifikatcii na rossiskih prezidentskih vyborah 2012." http://trv-science.ru/2012/03/27/shpilkin-kalinin/.

Kaplowitz, Stan, and Vladimir Shlapentokh. 1982. "Possible Falsification of Survey Data: An Analysis of a Mail Survey in the Soviet Union." *The Public Opinion Quarterly* 46 (1): 1–23.

Karp, Jeffrey A., and David Brockington. 2005. "Social Desirability and Response Validity: A Comparative Analysis of Overreporting Voter Turnout in Five Countries." *The Journal of Politics* 67 (3): 825–840.

Kiewiet, Chad P. 2015. "Who Lies About Electoral Gifts? Experimental Evidence from Latin America." *Public Opinion Quarterly*. doi:10.1093/poq/nfv024.

Kiewiet, Chad P., and David W. Nickerson. 2014. "Artificial Inflation or Deflation? Assessing the Item Count Technique in Comparative Surveys." *Political Behavior* 36: 659–682.

Klimek, Peter, Yuri Yegorov, Rudolf Hanel, and Stefan Thurner. 2012. "Statistical Detection of Systematic Election Irregularities." *Proceedings of the National Academy of Sciences of the United States of America* 109 (41): 16469–16473.

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

Miller, Judith. 1984. "A New Survey Technique for Studying Deviant Behavior." PhD thesis, The George Washington University.

Noelle-Neumann, Elisabeth. 1984. *The Spiral of Silence. A Theory of Public Opinion – Our Social Skin*. Chicago, IL: University of Chicago Press.

OSCE. 2012. "Election Observation Mission Final Report: Russian Presidential Election 4 March 2012." Technical report, OSCE.

Otava, Jiri. 1988. "Public Opinion Research in Czechoslovakia." *Social Research* 55: 247–260.
Presser, Stanley. 1990. "Can Changes in Context Reduce Vote Overreporting in Surveys?" *Public Opinion Quarterly* 54: 586–593.
Rose, Richard, William Mishler, and Neil Munro. 2011. *Popular Support for an Undemocratic Regime: The Changing Views of Russians*. Cambridge: Cambridge University Press.
Shpilkin, Sergei. 2011. "Statistika issledovala vybory." http://www.gazeta.ru/science/2011/12/10a3922390.shtml.
Sieger, Karin. 1990. "Opinion Research in East Germany: A Challenge to Professional Standards." *International Journal of Public Opinion Research* 2 (4): 323–344.
Silver, Brian D., Barbara A. Anderson, and Paul R. Abramson. 1986. "Who Overreports Voting?" *American Political Science Review* 80 (2): 613–624.
Simpser, Alberto. 2013. *Why Governments and Parties Manipulate Elections*. New York: Cambridge University Press.
Streb, Matthew J., Barbara Burrel, Brian Frederick, and Michael A. Genovese. 2008. "Social Desirability Effects and Support for a Female American President." *Public Opinion Quarterly* 72: 76–89.
Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2000. *The Psychology of Survey Response*. New York: Cambridge University Press.
Traugott, Michael W. 2008. *37 Validation Studies*. 408–417. London: Sage. doi:10.4135/9781848607910.
Tsuchiya, Takashiro. 2005. "Domain Estimators of the Item Count Technique." *Survey Methodology* 1 (31): 41–51.
Tsuchiya, Takashiro, Yoko Kirai, and Shigeru Ono. 2007. "A Study of the Properties of the Item Count Technique." *Public Opinion Quarterly* 71: 253–272.
Weghorst, Kieith. May 2011. "Political Attitudes and Response Bias in Semi-Democratic Regimes: A Survey Experiment Comparing the List Experiment and Randomized Response in Tanzania." Working Paper.
Wintrobe, Ronald. 1998. *The Political Economy of Dictatorship*. Cambridge: Cambridge University Press.
Wright, Gerald C. 1993. "Errors in Measuring Vote Choice in the National Election Studies, 1952–88." *American Journal of Political Science* 37: 291–316.