

Russian Election Reform and the Effect of Social Conformity on Voting and the Party System: 2007 and 2008¹

S. Coleman²

In 2007 Russian voters elected representatives to the State Duma under new electoral procedures that President V. Putin had instituted. A presidential election followed in 2008, won by D. Medvedev, leading to Putin's new role as prime minister. To many observers, the reforms and election campaigns resulted in a party system manipulated to the advantage of the government, although Putin's reported goal was to reduce the number of political parties. Earlier research found that social conformity exerted a strong, persistent, and predictable influence on voting in national elections from 1991 to 2003. This analysis examines how the effect of social conformity on Russian voters and the party system might have changed from earlier elections as a result of the electoral reforms and campaign practices. Specific questions addressed are how well the political party system now aligns with the interests of voters, and whether this type of analysis can speak to fairness of the elections.

Key words: elections, Duma, procedures, conformity

JEL classification: D7, P16

Introduction

The analysis is based on a predictive mathematical model that captures the effect of social conformity on people's collective social choices (Coleman, 2004, 2007a; Бородин, 2005). The model is derived from several principles of conformity in social psychology (Cialdini, 1990, 1993; Coleman, 2007a). First is the fact that people often align their behavior with the behavior of others or act in accord with a widely held social norm. This can depend, however, on whether people's attention is brought to bear on the specific norm or behavior involved. Second, when more people are doing something, then others are increasingly likely to do the same, although this process is self-limiting as not everyone will come into conformity. Third is the principle that people are fairly consistent in their degree of collective conformity from one situation to another. The closer the connection is between situations, the more likely people are to act consistently. Conformity is greatly affected by nonconscious cognitive processes.³ This type of cognition involves the prefrontal cortex, a different brain system than is used for conscious reasoning but one that fits with Kahneman's (2003) explanation of the effects seen in prospect theory and framing that limit rational decision making.

¹ Acknowledgments: My thanks to Rein Taagepera and Alexander Borodin for their constructive suggestions for this research.

² Metropolitan State University, 1450 Energy Park Drive, Suite 100A, St. Paul, Minnesota USA, steve.coleman@metrostate.edu; www.populardelusions.org

People's conformity with a social norm can occur without their conscious awareness, and they may attribute their behavior to other reasons. Individuals may correctly see that social conformity affects others' behavior while holding the illusion that their own behavior is not affected by social influence (Pronin, Berger, and Moluki, 2007). People will conform in private situations as well as in public and change their attitudes and preferences to conform with others.

Many economists ignore the psychological foundations of conformity and try to explain it by rational choice models. But whether conformity can be analyzed within the scope of rational choice is questionable. The rational choice assumption is that if people conform with a social norm, it must be because they see a benefit for conforming or a possible loss or punishment for not conforming. For example, Fehr and Gaechter (1998) attribute conformity with social norms, such as fairness, to an expectation of reciprocity. Bikhchandani, Hirshleifer, and Welch (1998) develop a conformity model to explain herding in economic decisions, including consumer marketing and such effects as stock market bubbles or crashes. Their model assumes that people's decisions are conscious and rational but based on observational information. They assert that when people are trying to decide which of several alternatives is the best decision to take, and they use information about others' choices as a guide to what's best, then a cascade of like decisions will result. Along the same lines, Durlauf and Young (2001) offer an expanded discussion of models for merging individual behavior and group conformity from a rational choice economics perspective. Behavioral economists, by contrast, have demonstrated that social influence and collective norms often dominate self-interest and change the order of individual preferences, violating assumptions of rational choice models (Heinrich, et. al., 2004). In his extensive analysis of whether conformity with social norms fits the rational choice model, Elster (1989) concludes that social norms cannot be reduced to rationality or any other optimizing individual decision.

Applied to voting behavior, the model of conformity makes the following argument. The social norm that citizens should vote exerts a strong influence on behavior, and research indicates this is the primary reason that people vote (Blais, 2000). The great attention given to a national election heightens people's awareness that they should vote and strengthens the effect of the voting norm. People are influenced toward voting by how often they expect others to vote and for whom (Gerber and Rogers, 2009). Prospective voters can get this information through informal social channels, opinion polls (frequently reported in Russia), and by their knowledge of how others voted in previous

³ See Coleman (2007a) for a review of the research.

elections. The effect of conformity on the decision to vote then spills over to a consistent degree among many people in their choice of political parties, affecting the distribution of votes among the parties or candidates (for a single office). As more people decide to vote out of a desire to conform, following the majority norm, they also are more likely to vote for the political parties that most other people are voting for. These principles lead to a mathematical model of the relationship between voter turnout and the distribution of votes among parties when influenced by social conformity.

The applicability of the model becomes a test for the effect of conformity on voting behavior and its degree of influence. Some voters may vote for reasons other than social conformity, as predicted, for example, by rational choice voting models; the model does not apply to this group. But if social conformity has a strong effect on a significant proportion of voters, it will have a unique and predictable impact on the election result—it leaves a signature or fingerprint of its effect that one can detect through quantitative analysis. It is this model that was demonstrated previously for Russian elections (Coleman, 2004, 2007a; Бородин, 2005) and which also holds strongly in national elections of many countries, including the United States, Germany, Japan, India, and Ukraine, among others (Coleman, 2004, 2007a).

The effect of social conformity on voting has received little attention in voting research, which has been dominated by rational choice theory (Mueller, 1989). Following an economic analogy, this theory assumes that people vote because the potential gain for themselves outweighs the costs of voting; party choice follows the same self-serving logic. This theory has been sharply criticized, however, for failing to explain voting participation (Green and Shapiro, 1994). Other explanations for voting behavior examine sociological aspects but not specifically conformity. Nevertheless, it is well known from the earliest behavioral studies of voting that people often vote the same as their families, friends, community, labor union, or ethnic group. Political scientists have not sorted out whether in these situations people vote like others because of a desire to conform with them or because they share a common self-interest with others in their affiliated group. Psychological research (Cohen, 2003), however, demonstrates that political beliefs are dominated more strongly by the group influence of a person's political party than by a person's ideology or objective analysis of policy issues. Supporting the conformity model more directly is the well-known "bandwagon" effect, when voters shift their vote to the party most likely to win (Mehrabian, 1998). Also, neighborhood effects that cause people to change their vote to be in line with the majority in their neighborhood have been reported frequently in Britain (MacAllister et al., 2001). Evidence on political communication further supports that

interaction between people affects their voting choice (Huckfeldt and Sprague, 1995). Prior to this line of conformity research, there has been no direct research on the question of a spillover of conformity from the voting participation choice to the political party choice. But there is substantial evidence of a strong and consistent relationship between voting participation and conformity with other important social norms that is indicative of a general spillover effect involving the voting norm (Knack, 1992; Knack and Kropf, 1998).

This analysis concerns the Duma election of 2007 and the presidential election of 2008. The Duma election was held on December 2, 2007. This was the first election in which all voting was for party lists and all seats were assigned by proportional representation. Previously, half the seats were assigned proportionally to parties and half by winner-take-all in electoral districts, but President Putin abolished the system of electoral districts. This eliminated independent candidates. Minimum thresholds for party membership also were established to further reduce opportunities for minor parties. Eleven parties competed in 2007 after a few others had been ruled ineligible. The threshold for a party to gain a seat was 7%, which meant that only four parties won representation in parliament. United Russia, which Putin supported, gained 63% of the vote. The presidential election was held on March 28, 2008. D. Medvedev, with the backing of several political parties, won with 71% of the vote; the remainder divided among the other three candidates.

Elaboration of the Conformity Model

The model builds on the principle of conformity that when deciding between several actions, people often take into account the expected frequency or probability of others' behavior. This assumes that people have direct knowledge of others' likely behavior or, at least, a naïve, implicit statistical knowledge of the relative frequencies or probabilities of behavior. Psychological research shows, in fact, that people acquire such statistical knowledge automatically, without conscious awareness or intent (Hasher and Zacks, 1984). Not having direct knowledge of what these individual probability estimates are, one can use the results of the election to get a retrospective estimate of the probability of an arbitrary citizen voting, or the probabilities of a randomly chosen voter voting for any party. If voters are fairly accurate in their individual estimates, at least on average, then the retrospective estimates should be sufficient to test the model. There is little direct evidence, however, as to the accuracy of people's expectations about the future outcome of elections. Opinion polls ask prospective voters about their voting intentions, not their predictions. But prediction markets, such as

the Iowa Electronic Market, show that collectively people can accurately predict the outcome of an election and estimate probabilities (Wolfors and Zitzewitz, 2004).

If people are using such statistical information when making voting decisions, one must have a measure for conformity that captures this cognitive representation. There is no standard measure of conformity at hand, however. If there are only two alternatives, it is easy to see which is favored by the majority and by how much. But when more than two alternative choices are involved, and the number can change, it is not obvious how to measure conformity. Although one might imagine a variety of possible conformity measures based on probability information, the Shannon entropy measure (Shannon and Weaver, 1969; Theil, 1967; Khinchin, 1957) of statistical information theory (communication theory) is already well known and has clear links to both predictability of events and cognitive decision making. In short, the model interprets behavior as a form of communication in society and applies methods of communication theory to its analysis. Entropy is a measure of unpredictability, and greater unpredictability in social behavior equates with less social conformity. The human brain responds to changes in event probability in relation to the entropy of the situation (Norwich, 1993) and has a cognitive limit of about 3 (bits) on the entropy scale (Miller, 1956) when it has to discriminate between several alternatives of different probabilities. The entropy measure incorporates the idea that more unexpected events carry more information, but the relationship is a logarithmic function of event probability.

The entropy measure $H(S)$ for a set S of k alternatives of probability p_i where $i = 1, 2, \dots, k$, and $\sum p_i = 1$ is $H(S) = -\sum p_i \log_2 p_i$ with the logarithm to base 2; units are bits.⁴ As differences between the probabilities increase, entropy decreases; when the probabilities tend toward equality, entropy increases. $H(S)$ is never less than zero and attains its maximum when all alternatives are equally probable or $p_i = 1/k$; the maximum is $\log k$. (For $k = 2$, maximum $H(S) = 1$; for $k = 3$, maximum is 1.58; for $k = 4$, maximum is 2, etc.). $H(S) = 0$ at $p = 0$ or $p = 1$.

Consider how one can apply entropy to voting as a measure of conformity.⁵ Let p_i represent the probability that a randomly chosen voter will vote for the i -th party. If voters shift their vote to one

⁴ Note that entropy is a set function, not the function of a single variable, and it is an intensive quantity (like temperature), meaning that it does not depend on the size or population of the social system. In contrast to subjective expected utility, entropy is measured on a ratio scale. One can also interpret entropy as average information, that is, an average of terms $-\log p_i$, where each term represents the statistical information of an event of probability p_i . This entropy measure does not have the same meaning as the entropy of statistical physics or thermodynamics, which is an extensive measure of disorder or randomness.

⁵ For other examples of the application of entropy to voting, see Darcy and Aigner (1980), Cohen and Warwick (1983), and Gill (2004).

or two large parties from smaller parties so that p_i increase for the larger parties, entropy decreases. Therefore if people are becoming more conformist, aligning themselves with larger political party voting groups, the entropy of the collective voting decision is decreasing. In other words, entropy is an inverse measure of conformity in voting or other collective social decisions. Using the actual voting share p_i received for each party in the election, one can calculate retrospectively the entropy $H(P)$ of the choice among political parties.

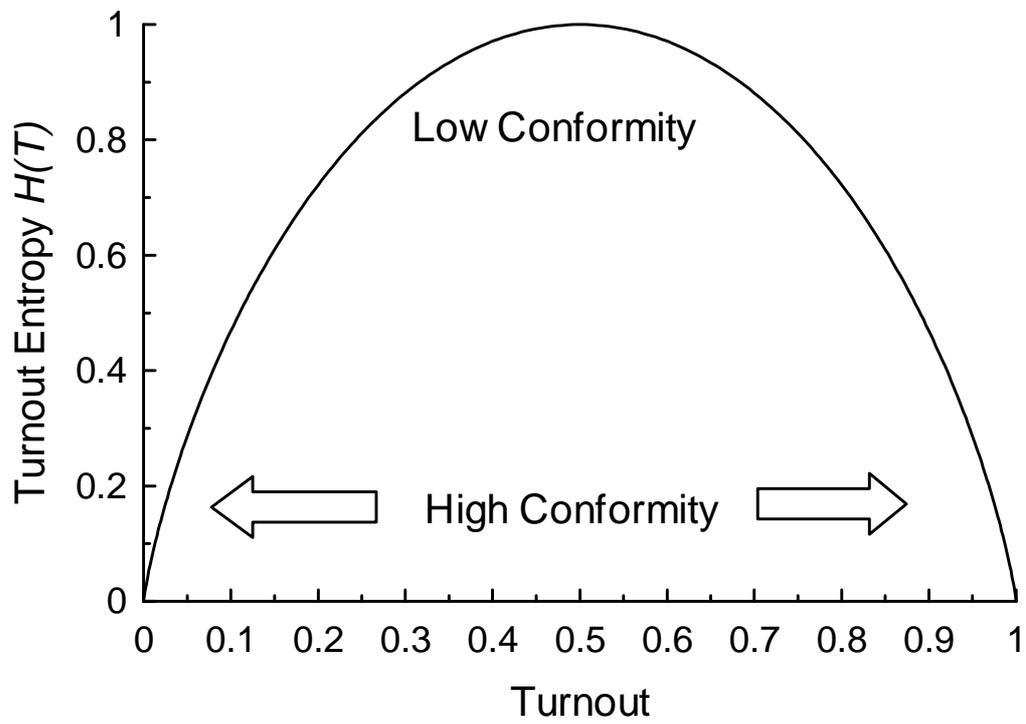
One can also apply the entropy measure to the choice of whether to vote or not. For the two-choice situation—to vote or abstain—entropy or nonconformity in the turnout choice is identified as $H(T)$. One can calculate this retrospectively using the fractional turnout rate t and abstention rate $1 - t$,

$$H(T) = -t \log_2 t - (1-t) \log_2 (1-t).$$

This is shown in Figure 1 where one sees a nearly parabolic relationship between $H(T)$ and turnout, having a maximum $H(T) = 1$ at $t = 0.50$ when the probabilities of voting and abstaining are equal. If the prevalent social norm is that people should vote, high conformity expresses itself as high turnout above 50% and closer to 100%. If, however, the norm is to abstain from voting, which sometimes happens, high conformity implies a very low turnout below 50%.

The principle that people are fairly consistent in their degree of conformity from one choice situation to another implies a positive relationship between the two voting entropy measures in an election, that is, between voter turnout and the distribution of votes across parties. Both entropy measurements are mappings of the same underlying state of conformity. One can express the idea of consistency mathematically and make a more definitive hypothesis as to an isomorphic relationship between the two measures. Let us put both entropy measures on the same scale by taking the number of choices into account for each measure. $H(P)/\log k$ and $H(T)/\log 2$ both have range $[0,1]$. So full consistency or agreement between the two measures along the scale implies $H(P)/\log k = H(T)/\log 2$, which simplifies to $H(P) = \log k H(T)$. An immediate prediction from this equation is that $H(P)$ must have the same relationship to turnout as the parabolic curve in Figure 1, but multiplied by $\log k$. The maximum of the $H(P)$ curve is predicted to be at $t = 0.50$. Electoral units of analysis, whether measured over time or in cross-section, should be approximately on such a curve when their party entropy is plotted against turnout—a relationship that can be tested and estimated with regression analysis.

Figure 1. Turnout, Turnout Entropy, and Conformity



Applicability of the model in an election is indicated by: (1) a nearly parabolic shape to the relationship of $H(P)$ to t , as in Figure 1, with approximately $H(P) = 0$ at $t = 0$ and $t = 1$; (2) a maximum of the parabolic fit near 50% turnout; (3) a ratio of $H(P)$ to $H(T)$ approximately equal to $\log k$, when voters perceive k parties competing; and (4) a strong fit of the model to election data, indicating a strong consistency in voting conformity across the two voting choices. These also test the ratio-scale measurement property of entropy. The predictions are a definitive test for a strong presence of social conformity in voting because they follow directly and uniquely from a conformity theory and are very unlikely to be true if other explanations of voting hold (Coleman, 2007b). Neither rational choice theory nor any other voting theory has made such specific predictions that would apply widely to elections. Previous statistical research on Russian elections has followed the rational choice model, relating party vote share to factors such as regional economic voting (Treisman, 2001), economic dissatisfaction (Richter, 2006), and voters' candidate or policy preferences (Schofield and Zakharov, forthcoming). These are typical of rational choice models for voting in that they are retrospective explanations of the results of particular elections, not general models, and lack quantitative specificity. The only predictions in regression models are the signs and statistical significance of the coefficients.⁶ In any case, they do not speak to conformist voting behavior in Russia, which is not tested in the models. Note that unlike the rational choice models, the conformity model does not try to explain the level of voter turnout or why a particular party has won the election—typical goals of voting research. Here the assertion is only that given the turnout level, there will be a certain division of the vote among parties at that turnout level, which can be predicted from a conformity model. The level of turnout reflects the degree of compliance with the social norm that good citizens vote in elections. Regional variability in the degree of turnout conformity in a country allows statistical analysis of its relationship with party voting.

Although one can start with the assumption that k should equal the number of parties on the ballot, it is often the situation that this number does not accurately reflect what voters perceive as their choices. So the analysis estimates the number of parties rather than simply counting the number on the ballot. Often voters are ignorant of minor parties on the ballot and, more generally, not all party choices may register distinctly in voters' cognitive decision making. The number that a person can consider is constrained by a person's short-term or working memory capacity, especially when there is little time to make the choice and reflect on information about all the parties. And an excessive

⁶ See Coleman (2007b) for more discussion of the weakness of regression models for developing predictive social science

number of parties will bump against a person’s cognitive entropy limit on information processing, causing errors in classification. One might get an estimate of the psychological number of parties, for example, by asking voters just after they voted to name quickly from memory the parties on the ballot. Time and information constraints on the voting decision also make it more likely that a voter’s decision making will fall back on a simple heuristic such as the relative sizes of the expected vote for the parties.

Testing and Estimating the Models

Duma election of 2007. The election data is analyzed for 85 regions.⁷ The data includes voting for 11 parties plus a very small percentage of votes that were invalid (average 1% per region); this was counted as an additional voting category in the analysis, much as earlier elections allowed voters to vote “against all.” There were some votes for each party in each region. Basic statistics are in Table 1. The analysis first tested an OLS regression model of $H(P) = \log k H(T)$, which gives an estimate of both the strength of the model and an estimate of k as approximately the average number of parties that voters perceived as choices in the election.

Table 1

Descriptive Statistics for the 2007 Duma Election

	Turnout	$H(P)$	$H(T)$	$H(P)/H(T)$
N of cases	85	85	85	85
Mean	0.66	1.77	0.87	1.99
Median	0.63	1.93	0.96	2.03
Std Dev	0.12	0.49	0.20	0.25

The test of the model $H(P) = \log k H(T)$ has a potential shortcoming in that it offers little opportunity to examine alternative hypotheses. For example, if the true relationship is not parabolic or the maximum is not at turnout 50%, the analysis would not detect that directly. So to complement the first analysis, and provide a way to check for alternatives, $H(P)$ also was regressed on a parabolic (quadratic) model of turnout, that is, $H(P) = a_0 + a_1 t + a_2 t^2$. The location of the maximum of this

theory.

⁷ Election data in Russian is available from the Central Election Commission of the Russian Federation, <http://www.cikrf.ru>. Or see Alexander Kireev’s English website, Electoral Geography 2.0, at <http://www.electoralgeography.com/new/en>.

model, if it fits, can be estimated as $t_{\max} = -a_1/(2 a_2)$. For this model, a bootstrap procedure (Efron and Tibshirani, 1998) was used to estimate a 95% confidence interval (CI) for the location of the maximum, which would be very difficult calculate analytically. The regression model was re-estimated 1,000 times, randomly choosing 85 data points with replacement from the original data set. The maxima were calculated and ranked from least to greatest and the 25th and 975th values define the confidence interval. Standard errors of the coefficients were corrected to compensate for heteroscedasticity. Analysis was done with Systat.

Presidential election of 2008. This analysis was based on 83 regions, having deleted two that represented voters outside the Russian borders. Four parties were included in the calculation of the party entropy, and analysis followed the same procedures as with the 2007 election data. Basic statistics are in Table 2.

Table 2

Descriptive Stistics for the 2008 Presidential Eection

	Turnout	$H(P)$	$H(T)$	$H(P)/H(T)$
N of cases*	83	83	83	83
Mean	0.70	1.17	0.84	1.42
Median	0.68	1.26	0.90	1.40
Std Dev	0.10	0.24	0.17	0.19

*Statistics for voters outside Russia were excluded.

Results

The first model tested, the theoretically derived model, (Table 3) fits very well in both elections with R^2 equal to 0.86 in 2007 and 0.72 in 2008. The estimate of $\log k$, an approximate average number of political parties perceived by the voters, is 2.04 in 2007 and 1.40 in 2008. A post hoc test of the model for 2007 rejects the hypothesis that $\log k \neq 2$ at $p = .10$. The estimate of $\log k = 2.04$ (95% CI is 1.99 -2.08) in 2007 is virtually the same as found by taking the average of $H(P)/H(T)$ over the cases; the average ratio is 1.99 (95% CI is 1.94 to 2.05). Since $\log 4 = 2$ corresponds exactly to a four-choice situation, one can infer that the 2007 election was to most voters, on average, a four-party contest. In other words, the parliamentary election looks statistically more like a four-party contest than an 11-party contest from a conformity perspective. This corresponds well to the actual results of the election wherein, because of the 7% threshold, only four parties won seats and together amassed 92% of the total vote. For 2008, the estimate of $\log k = 1.40$ is below 1.58, the predicted value

for three parties. One can convert this to a number of parties by raising 2 to the 1.40 power; $2^{1.40} = 2.6$ suggesting the 2008 election was perceived as a two-choice or three-choice contest by most voters. See also Figures 2 and 3.

Table 3

Regression coefficient for the equation $H(P) = \log k H(T)$

Election	$\log k$ (std error)	k parties	R^2
2007	2.04 (0.02)	4.1	0.86
2008	1.40 (0.01)	2.6	0.72

Notes: standard errors are corrected for heteroscedasticity; $p < 0.0001$

The empirical parabolic regression models (Table 4, Figures 2 and 3) have virtually the same strength of fit as the theoretical models when measured by R^2 —0.88 for 2007 and 0.74 for 2008. As Figure 1 shows, the ideal regression curve would intersect the x-axis at $t = 0$ and $t = 1$ and be symmetrical in a reflection around the $t = 0.5$ axis. In fact, the estimated coefficients a_0 for 2007 and 2008 are not statistically different from zero, failing to reject a difference at $p < .05$; so the parabolic curves approximately intersect the origin as predicted.⁸ In 2007 the fitted curve equals 0.09 at $t = 1$, close to the predicted zero (Figure 2); but the difference at $t = 1$ is greater for 2008. One can make a simple test for an intersect at $t = 1$ by reflecting the turnout data around the $t = 0.5$ axes and redoing the quadratic regression; that is, let $t' = 1 - t$ and regress on t' . For 2007 a post hoc test on the reflected data regression rejects the hypothesis that $a_0 \neq 0$ at $p = .34$; but for 2008 it fails to reject inequality at $p = .02$.

The parabolic maxima in 2007 and 2008 are estimated at turnouts 49.2% (36% - 56% CI) and 59.5% (52% - 63% CI), respectively. Neither regression is sufficiently different from the predicted model to cause us to consider an alternative theory to the conformity model.⁹

⁸ A post hoc test that $a_0 \neq 0$ is rejected at $p = .75$ in 2007 and at $p = .10$ in 2008.

⁹ Alternatively, nonlinear regression estimates the maxima in 2007 and 2008 at $t = 0.49$ and $t = 0.55$, respectively. Corresponding Wald 95% CI are [0.41 - 0.57] and [0.48 - 0.63], both including $t = 0.50$.

Table 4

Regression Cefficients for the Model $H(P) = a_0 + a_1 t + a_2 t^2$

Election	a_0 (std err)	a_1 (std err)	a_2 (std err)	R^2	parabolic maximum	bootstrap 95% confidence interval for parabolic maximum	
2007	0.22 (0.68)	7.64 (1.85)	-7.77 (1.20)	0.88	0.492	0.36	0.56
2008	-1.00 (0.55)	7.81 (1.55)	-6.56 (1.06)	0.74	0.595	0.52	0.63

Note: standard errors of the coefficients are corrected for heteroscedacity; model p < 0.00001

Figure 2. Predicted vs Empirical Models: 2007

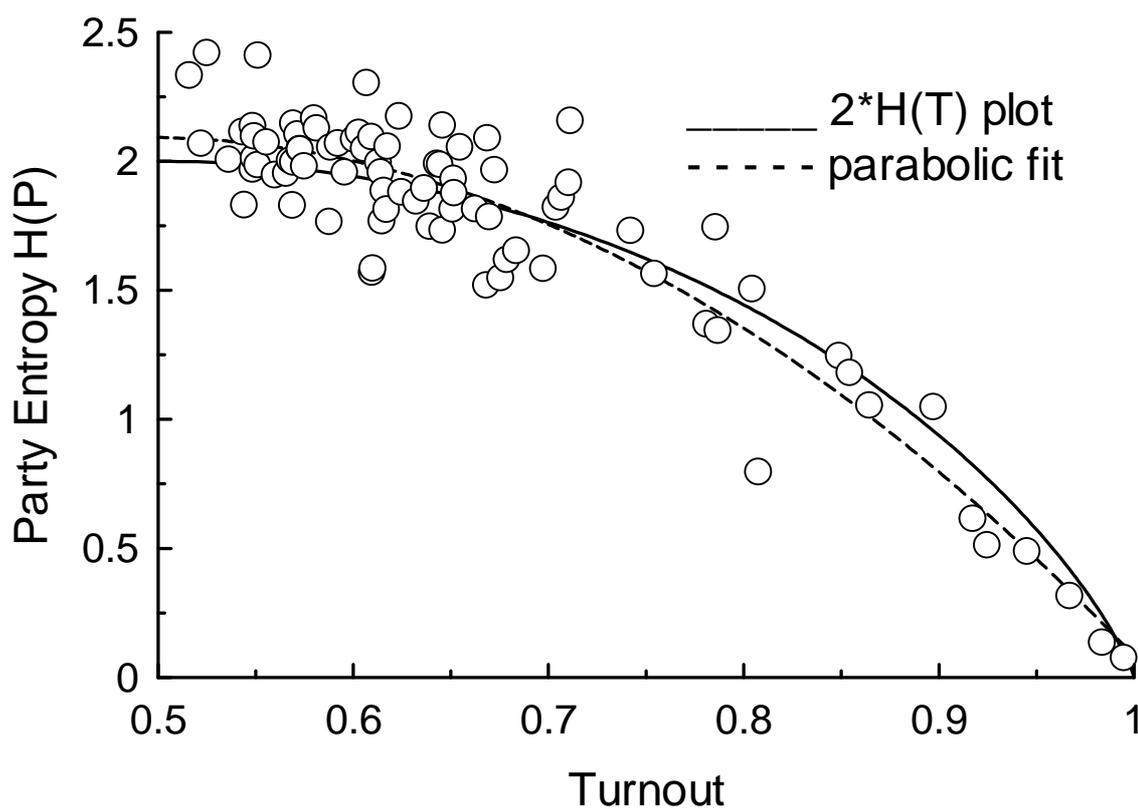
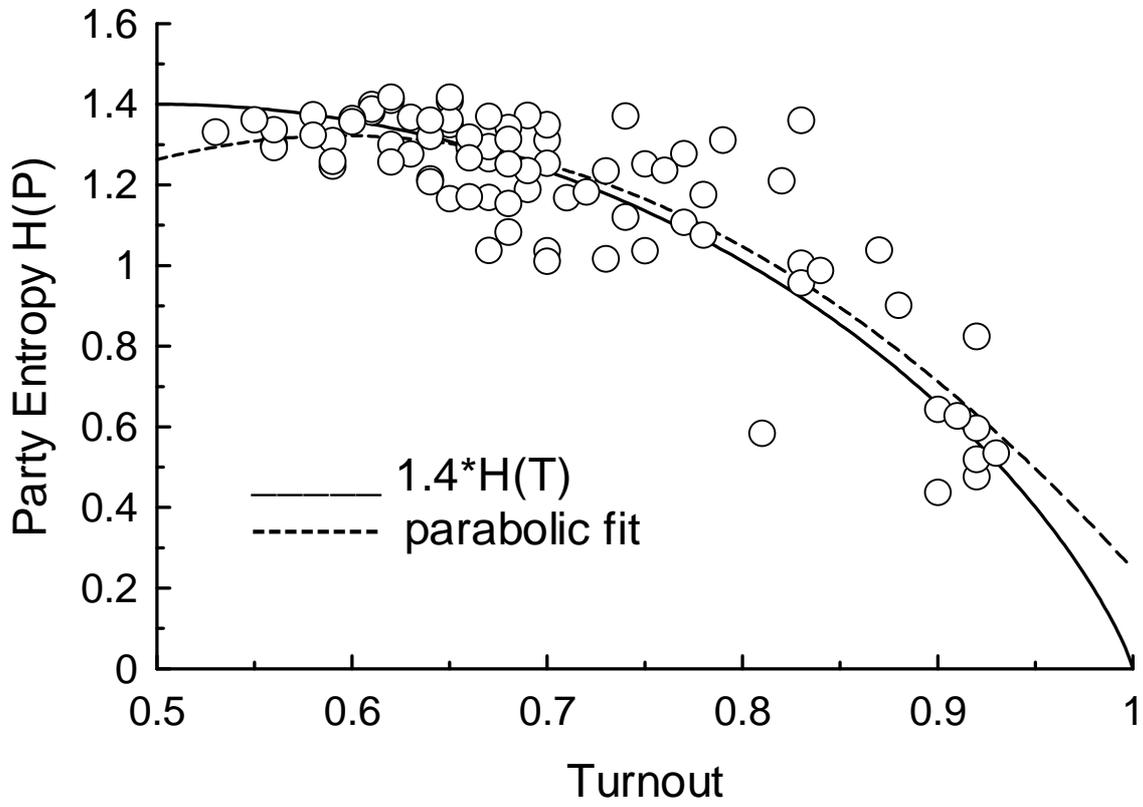


Figure 3. Predicted vs Empirical Models: 2008



After the 2007 election many observers questioned the election procedures in several areas that had the highest turnouts and where almost everyone apparently voted for the same party. In fact, there are six regions that are outliers on the turnout scale having turnouts over 90% (Dagestan, Kabardino-Balkaria, Ingushetia, Karachay-Cherkessia, Mordovia, and Chechnya.) These same regions also had very high turnouts in 2008. One can characterize them as being on the geographical fringes of Russia and having relatively large minority ethnic populations. To check on the influence of these cases, which fit the predicted model well, the analysis for 2007 was redone excluding these regions. The result was again that the model $H(P) = 2.04 H(T)$ fits the data very well, just as it did with those regions included. R^2 is reduced to 65%, however, as must be anticipated when the range of turnout is limited. The 2008 models were also checked excluding regions with turnout over 90%. Results were very similar to the earlier model but again with R^2 reduced. The model estimate is $H(P) = 1.40 H(T)$ at $R^2 = 52\%$. Parabolic models also echo the previous results. So one can infer that although the highest turnout regions have a substantial influence on the strength of fit of the regression models, and are at the extreme of Russian voting participation, they are not exceptions to the predicted relationship between turnout and party entropy nor do they significantly influence model predictions.

Unexplained variation and prediction errors in the analysis can arise when some voters vote for reasons unrelated to conformity. Another source of error in the models is when people who are prone to conform mistakenly estimate the relative frequencies of other voters' choices. Any irregularities in voting procedures will also affect the goodness of fit, although isolated local voting problems will have less impact on the results here, which use regional data. Errors in the calculation of voter turnout because of ballot over-counting, or inaccurate estimates of the voting age population and the number of eligible voters (the electorate) will further bias the results and, specifically, the estimate of the location of the parabolic maximum.

Discussion

All the evidence points to the conclusion that both elections fit the theoretically predicted relationship between party entropy and turnout that is diagnostic of the effect of social conformity on voting. Furthermore, from the viewpoint of predictive modeling, there is no appreciable difference between the empirical parabolic model and the more constrained, theoretically derived model $H(P) = \log k H(T)$. Thus one can reject alternatives to the theoretical model that might involve a large shift in location of the maximum of the curve or have a distinctly different shape. The correspondence of the

two models also attests to the strength of the observed relationship between party entropy and turnout and, thus, the strength of conformity spillover from the turnout choice to the party choice. The only parameter allowed to vary in the more constrained model for the Duma election was the number of parties, which was estimated to be 4—a number in agreement with the final allocation of seats in the Duma and the great majority of votes. The presidential election was midway between a two-party election and a three-party election. Considering both elections, one can say that most voters made decisions as if the number of parties were between 2 and 4. The importance of social conformity in the two election outcomes is especially apparent when one considers that the model applies equally to each, although the elections were quite different as to type—parliamentary or presidential—number of choices, personalities, campaigns, issues, parties, and dates. Clearly, these political factors had a minor impact if any.

The estimated number of parties agrees with the number people would prefer, as reported in opinion polls, though not all voters prefer the same number of choices. One can compare the estimated number of parties with surveys about how people think of the party system. In 2007, for example, a large majority (68%) saw a need for only one to three large parties. This is from a survey by Levada Center in Moscow that asked the following question:

Table 5

Q17-t. How many political parties does Russia need now? (Percentage in agreement)

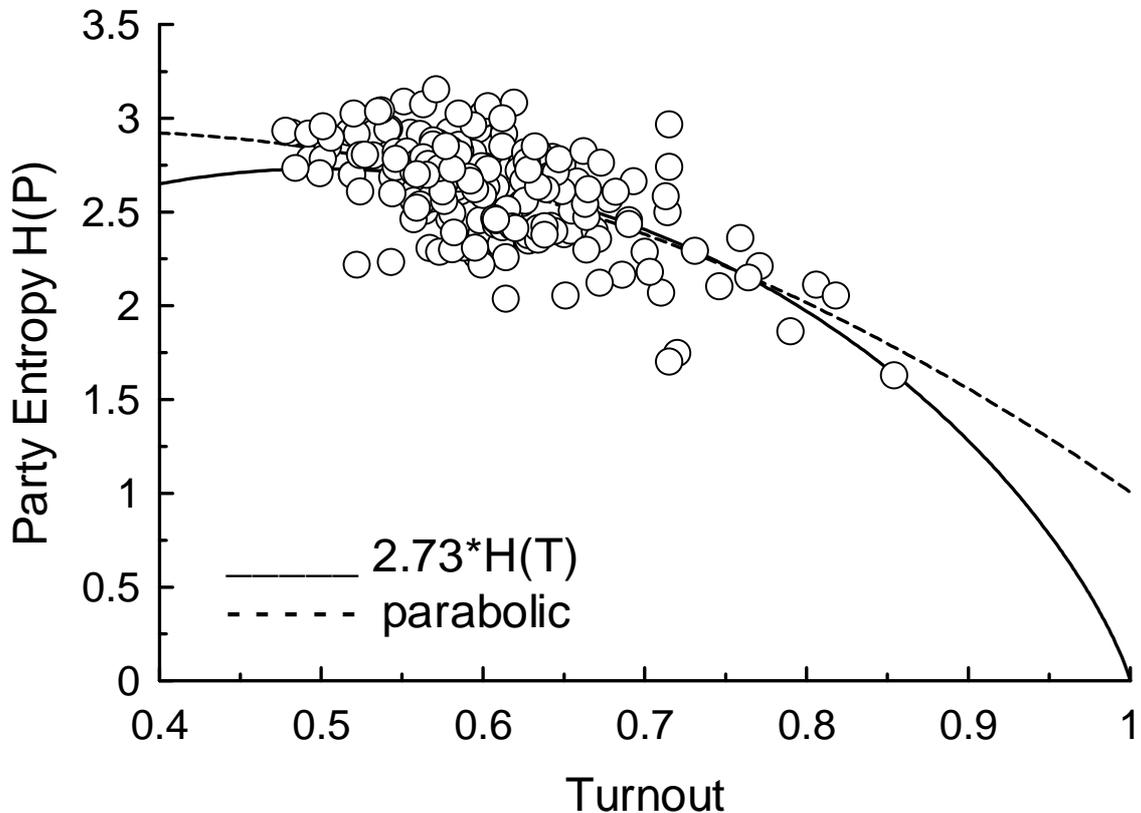
	Apr 04	Sept 04	Oct 05	July 06	Apr 07	Oct 07
One strong ruling party	34	34	38	32	30	28
2 or 3 big parties	41	44	39	42	46	40
Many small parties	8	6	4	5	7	9
No need for any parties	7	6	7	7	6	7
Don't know	11	9	12	14	13	17

Source: Levada Center, nationwide surveys, 2004-2007.

The results of the analysis are in line with models estimated for elections in the 1990s, but goodness of fit of the models has increased so that the effect of conformity is now more pronounced. The party system also has evolved, so that the fragmentation of the party system has substantially

decreased; the smallest parties now get smaller shares of the vote. This has led to a large decrease in average party entropy from the 1990s. Some regions have also increased their turnout substantially over earlier elections, as described above. To see this, compare Figures 2 and 3 with Figure 4, which shows the corresponding regional data for the Duma election in 1999 (to party lists). As in 2007, there were up to 11 party choices in the 1999 election, but in 1999 the number depends on the region. For the parabolic regression model, $R^2 = 0.36$ ($N = 222$ with two outliers removed, $p < .0001$) (Coleman, 2007a); and a regression estimate of the theoretical model yields $H(P) = 2.73 H(T)$ ($R^2 = 0.29$, estimated $k = 6.6$ parties, $p < .0001$). Observe that in 1999 many regions had party entropy levels near the individual psychological limit of about 3 bits. Previous research (Coleman, 2007a) showed that this same limit seems to apply, at least approximately, to collective voting decisions. Historical election results from many countries show that this degree of party entropy is rare and not likely to be sustainable in a society, but such extreme entropy levels were characteristic of Russian parliamentary elections in the 1990s. In sum, the party system has become a more normal party system over the last decade, with a number of major parties in closer agreement with voters' preferences and perceptions, and at a more normal party entropy level. It appears that Putin succeeded in reducing the number of political parties, practically speaking.

Figure 4. Predicted vs Empirical Models: 1999



One might ask whether this research can address the question of fairness. Were the 2007 and 2008 elections corrupted by fraudulent ballot counting? Were they less fair and democratic than earlier elections? If there was fraud, did it significantly alter the results of the election? Because of President Putin's changes to the party system, the apparent suppression of some parties, government control of the media, and reports of voting irregularities, many observers have questioned the fairness of the 2007 election ("Russian Duma elections . . . ", 2007). Other researchers argue against this, however (Rose and Mishler, 2009). By contrast, few concerns were raised about fairness in the presidential election as Putin was the clear favorite. Analytic studies point to several issues that suggest ballot manipulation on a large scale. One argument for this is the strong correlation between voter turnout and the vote share received by United Russia in 2007. A more complex analysis (Mygakov, Ordeshook, and Shakink, 2007) shows statistical anomalies in the distribution of reported turnout and votes at the local level in areas with high turnout. Accepting the possibility of widespread voting fraud, however, one must still question whether the results would have been significantly

different if there had been no fraud. This is a difficult issue to resolve. If turnout was artificially inflated for the benefit of United Russia, it is likely that an actual increase in turnout would have produced about the same results. In high turnout areas the great majority favored United Russia, so one might consider “imaginary” voters as proxies for nonvoters. Analysis of nonvoters in American elections offers equivocal findings on what might happen if everyone voted, but generally no significant changes in election results would be anticipated (Citrin, Schickler, and Sides, 2003).

From the perspective of this analysis, the latest elections do not appear to have been manipulated in a way that lessened the impact of social conformity or replaced its force by other motivations to vote. This analysis is not strictly comparable with other studies that focus only on the winning party or candidate, however. Here the analysis concerns the entire distribution of the vote across the parties. But the strong correlation of turnout with United Russia’s vote share in 2007 is completely consistent with a typical bandwagon effect predicted by a conformity model or the cascade model of Bikhchandani, Hirshleifer, and Welch (1998). As more people vote, they are increasingly likely to vote for the party expected to get the most votes. So a simple analysis of correlation is not sufficient evidence as to whether conformity or fraud is the better explanation. It is utterly inconceivable, moreover, that ballots could have been so carefully rigged as to produce the results seen here, such as the parabolic relation between party entropy and turnout with a maximum near the predicted value of 50% turnout in both 2007 and 2008 elections—a result predicted and observed in many other countries. Undoubtedly some fraud occurred, but it was not sufficient either to overturn the effect of conformity or to produce the effect.

Although voter fraud has been especially suspect in the highest turnout regions, they fit the same predicted model as other regions. In fact, the strength of the models, and for some regions the very high levels of turnout, now show a considerably stronger effect of conformity in 2007 and 2008 than in earlier elections. But the 2007 and 2008 elections continue the general patterns seen in earlier elections; 2007 does not stand out as an exception. Furthermore, the 2007 and 2008 elections show very similar results in the strength of fit of the models to each other, so there would be little reason to see 2007 as less fair than 2008. Indeed the exceptional degree of fit of the models in 2007 and 2008 would weigh against any other substantial voting factors in these elections. Conformity might be perceived by observers as undue influence on voting behavior or public officials, but that does not necessarily imply fraud. The strong spillover effect and, in some regions, heightened conformity suggest that one can expect to see increasing social pressure in voting as well as in other dimensions of

social behavior, especially in regions with the highest turnout levels where that degree of social pressure that sometimes can turn violent..

References

Бородин, А.Д. [Borodin A.D.] (2005). Согласованность коллективных действий в поведении российских избирателей [The Effect of Social Conformity on the Behavior of Russian Voters] // ЭКОНОМИЧЕСКИЙ ЖУРНАЛ ВШЭ [Economic Journal HSE]. Т. 3. № 1. С. 74-81.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1998). Learning from the Behavior of Others // Journal of Economic Perspectives. Vol., 12. P. 151-170.

Blais, A. (2000). To Voter or Not to Vote: The Merits and Limits of Rational Choice Theory. Pittsburgh, PA: University of Pittsburgh Press.

Cialdini, R.B. (1993). Influence: Science and Practice. 3rd ed. N.Y.: Harper Collins.

Cialdini, R.B., Reno, R.R., and Kallgren, C.A. (1990). A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places // Journal of Personality and Social Psychology. Vol. 58. P. 1015-1026.

Citrin, J., Schickler, E., and Sides, S. (2003). What if Everyone Voted? Simulating the Impact of Increased Turnout in Senate Elections // American Journal of Political Science. Vol. 47. P. 75-90.

Cohen, G. L. (2003). Party Over Policy: The Dominating Impact of Group Influence on Political Beliefs // Journal of Personality and Social Psychology. Vol. 85. P. 808-822.

Cohen, L. and Warwick, P. (1983). Political Cohesion in a Fragile Mosaic: The Yugoslav Experience. Boulder, CO: Westview Press.

Coleman, S. (2004). The Effect of Social Conformity on Collective Voting Behavior // Political Analysis. Vol. 12. P. 76-96.

Coleman, S. (2007a). Popular Delusions: How Social Conformity Molds Society and Politics. Youngstown, NY: Cambria Press.

Coleman, S. (2007b). Testing Theories with Qualitative and Quantitative Predictions // European Political Science. Vol. 6. P. 124-133.

Darcy, R. and Aigner, H. (1980). The Uses of Entropy in the Multivariate Analysis of Categorical Variables // American Journal of Political Science. Vol. 24. P. 155-174.

Durlauf, S. N. and Young, H. P. (2001). Social Dynamics. Cambridge, MA: MIT Press.

Efron, B. and Tibshirani, R. J. (1998). An Introduction to the Bootstrap. Baton Rouge, FL: Chapman and Hall.

Elster, J. (1989). The Cement of Society: A Study of Social Order. New York: Cambridge University Press.

- Fehr, E. and Gaechter, S.** (1998). Reciprocity and Economics: The Economic Implications of Homo Reciprocans // *European Economic Review*. Vol. 42. P. 845-859.
- Gerber, A., and Rogers, T.** (2009). Descriptive Social Norms and Motivation to Vote: Everybody's Voting and so Should You // *Journal of Politics*. Vol. 71: 178-191.
- Gill, J.** (2004). An Entropy Measure of the Uncertainty in Vote Choice // *Electoral Studies*. Vol. 24. P. 371-392.
- Green, D.P., and Shapiro, I.** (1994). *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science*. New Haven, CT: Yale University Press.
- Hasher, L., and Zacks, R.** (1984). Automatic Processing of Fundamental Information: The Case of Frequency of Occurrence // *American Psychologist*. Vol. 39. P. 1372-1388.
- Heinrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., and Gintis, H., Eds.** (2004). *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford: Oxford University Press.
- Huckfeldt, R., and Sprague, J.** (1995). *Citizens, Politics, and Social Communication: Information and Influence in an Election Campaign*. NY: Cambridge University Press.
- Kahneman, D.** (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics // *The American Economic Review*. Vol. 93. P. 1149-1475.
- Khinchin, A. I.** (1957) *Mathematical Foundations of Information Theory*. New York: Dover.
- Knack, S.** (1992). Civic Norms, Social Sanctions, and Voter Turnout. // *Rationality and Society*. Vol. 4. P. 133-156.
- Knack, S., and Kropf, M.E.** (1998). For Shame! The Effect of Community Cooperative Context and the Probability of Voting. *Political Psychology* // Vol. 19. P. 585-599.
- MacAllister, I.R, Johnston, J., Pattie, C., Tunstall, H., Dorling, D.F.L., and Rossiter, D.J.** (2001). Class Dealignment and the Neighborhood Effect: Miller Revisited // *British Journal of Political Science*. Vol. 31, P. 41-60.
- Mehrabian, A.** (1998). Effects of Poll Reports on Voters Preferences // *Journal of Applied Social Psychology*. Vol. 28. P. 2119-2130.
- Mueller, D.C.** (1989). *Public Choice II*. Cambridge, UK: Cambridge University Press.
- Myagkov, M., Ordeshook, P., Shakin, D.** (2009). *The Forensics of Election Fraud: Russia and Ukraine*. Cambridge: Cambridge University Press.
- Miller, G.** (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information // *Psychological Review*. Vol. 63. P. 81-97.
- Norwich, K.** (1993). *Information, Sensation and Perception*. San Diego, CA: Academic Press.

Pronin, E., Berger, J., and Moluki, S. (2007). Alone in a Crowd of Sheep: Asymmetric Perceptions of Conformity and Their Roots in an Introspection Illusion // Journal of Personality and Social Psychology. Vol. 92. P. 585-595.

Richter, K. (2006). Wage Arrears and Economic Voting in Russia // American Political Science Review. Vol. 100. P. 133-145.

Rose R., and Mishler W. (2009). How do Electors Respond to an Unfair Election? The Experience of Russians. CSPP Studies in Public Policy, Paper SPP46. Center for the Study of Public Policy, University of Aberdeen.

“Russian Duma elections ‘not held on a level playing field’, say parliamentary observers.” (3.12.2007). Council of Europe. <https://wcd.coe.int/ViewDoc.jsp?id=1221469&Site=DC>.

Schofield, N. and Zakharov, A. V. (forthcoming). A Stochastic Model of 2007 Russian Duma Election // Public Choice.

Shannon C.E., and Weaver W. (1969). The Mathematical Theory of Communication. Urbana, IL: University of Illinois Press.

Theil, H. (1967). Economics and Information Theory. Amsterdam: North-Holland.

Treisman, D. S. (2001). After the Deluge: Regional Crises and Political Consolidation. Ann Arbor, MI: University of Michigan Press.

Wolfers, J. and Zitzewitz, E. (2004). Prediction Markets // Journal of Economic Perspectives. Vol. 18. P. 107-126.

Submitted July 27, 2009