Russian Election Reform and the Effect of Social Conformity on Voting and the Party System: 2007 and 2008

Coleman, Stephen

Metropolitan State University

30 January 2009

Russian Election Reform and the Effect of Social Conformity on Voting and the Party System: 2007 and 2008

Stephen Coleman*

Abstract

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 the 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 [1,2,6] reported 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 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.

Introduction

The analysis is based on a predictive mathematical model that captures the effect of social conformity on people's collective social choices [1,2]. The model is derived from several principles of conformity [2,3,4]. 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. Conformity is greatly affected by nonconscious cognitive processes. People's conformity with a social norm can occur without their conscious awareness, and they may attribute their behavior to other reasons.

Applied to voting behavior, the model of conformity makes the following

* E-mail to steve.coleman@metrostate.edu. Or see www.popculardelusions.org.

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 [5]. 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. Prospective voters can get this information through informal social channels, opinion polls (frequently reported in Russia), and by their information as to how others voted in previous 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, 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 by rational-choice voting models, for example. 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 [1,2,6] and which also holds strongly in national elections of many other countries, including the United States, Germany, Japan, and Ukraine, among others [1,2].

The effect of social conformity on voting has received little attention in voting research, which has been dominated by rational-choice theory [7]. 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 behavior [8]. 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. But this line of research has 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. Supporting the conformity model more directly is evidence of a "bandwagon" effect, when voters increasingly shift their vote to the party most likely to win. 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 [9]. Evidence on political communication further

supports that interaction between people affects their voting choice [10], while other research shows a strong and consistent relationship between voting participation and conformity with other social norms [11,12].

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 seats were assigned in proportion. Previously, half the seats were assigned by proportional representation and half by winner-take-all in electoral districts, but President Putin abolished the system of electoral districts. 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

Application of the conformity model to an election first requires a measure for conformity, but there is no standard measure at hand. If there are only two alternatives, it is easy to see which is favored by the majority. When more than two alternative choices are involved, or the number can change, it is not obvious how to measure conformity. The model builds on the principle of conformity that when deciding which of several actions to take, people may take into account the relative frequencies or proportions of other people who are making these choices. This assumes that people have direct knowledge of others' likely behavior or, at least, a naïve, implicit statistical knowledge about the relative frequencies or probabilities of behavior and social categories in their society. Psychological research shows, in fact, that people acquire such statistical knowledge automatically, without conscious awareness or intent [13]. So if people are using such statistical information when making voting decisions, one must have a measure for conformity that captures this cognitive representation.

Although one might imagine a variety of possible conformity measures based on probability information, the Shannon entropy measure [14] of statistical information theory is already well known and has clear links to both predictability of events and cognitive decision making. The human brain responds to changes in event probability in relation to the entropy of the situation [15] and has a cognitive limit of about 3 (bits) on the entropy scale [16] when it has to discriminate between several alternatives of different probabilities. The entropy measure incorporates the idea that more unexpected events carry more information. 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 $i = 1, 2, \ldots k$, and $\sum p_i = 1$ is $H(S) = -\sum p_i \log p_i$ with the logarithm to base 2. As differences between the probabilities increase, entropy decreases; when the probabilities tend toward equality, entropy increases. Let p_i represent the probability of people in an election voting for the *i*-th party. If voters shift their vote to one or two large parties from smaller parties so that p_i increase for the larger parties, entropy will decrease. Therefore if people are becoming more conformist, aligning themselves with larger political party voting groups, the entropy of the collective voting decision is decreasing. So entropy is an inverse measure of conformity. *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 $p = 1$.

Entropy or nonconformity in the turnout choice is identified as *H(T)*. One can calculate this retrospectively using the fractional turnout rate *t,*

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

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, conformity expresses itself as high turnout above 50% and closer to 100%. If, however, the norm is to abstain from voting, which sometimes happens, conformity implies a very low turnout below 50%. One can make a similar calculation for the entropy in the party vote choice *H(P)* using the actual voting share p_i received for each party in the election.

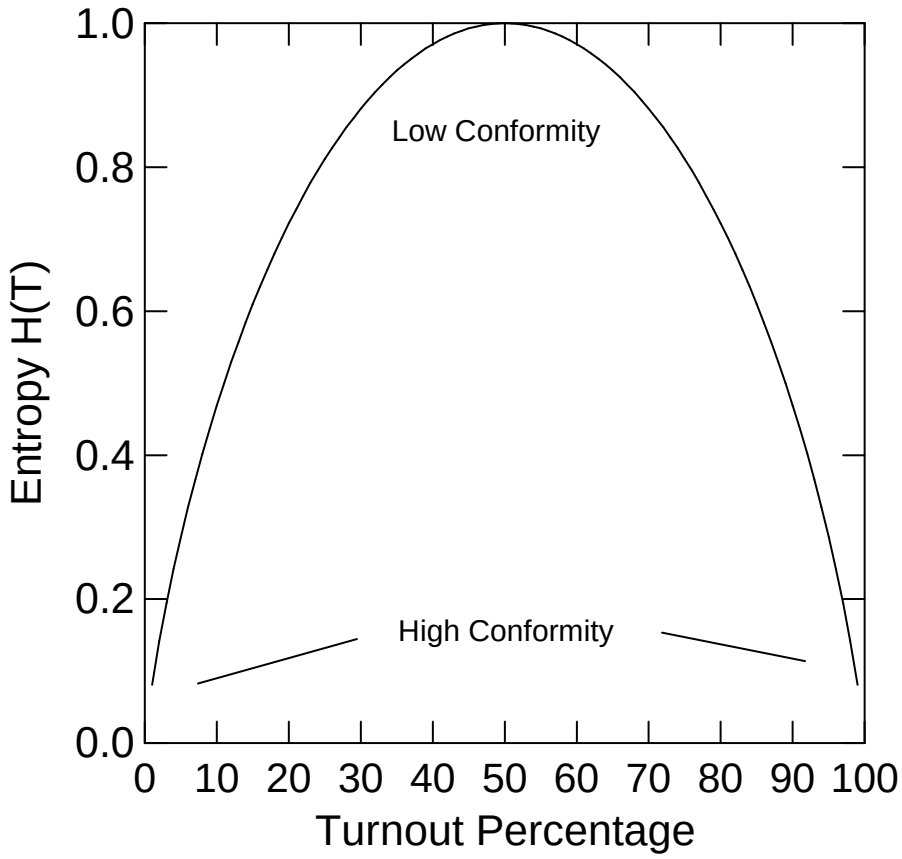
The principle that people are fairly consistent in their degree of conformity from one choice situation to another implies a positive relationship between the two voting entropy measures in an election, that is, between voter turnout and the distribution of votes across parties. One can express the idea 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 approximately $t = 0.50$. Electoral units of analysis, whether measured over time or cross-sectionally,

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.

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; (2) a maximum of *H(P)* 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 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 [17]. Neither rational choice theory nor any other voting theory would make such a general prediction.

Although one can start with the assumption that *k* should equal the number of parties on the ballot, it is usually the situation that the number of parties on the ballot 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 number of parties will bump against people'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 the parties on the ballot from memory. 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.

Figure 1. Turnout, entropy, and social conformity.



Testing and Estimating the Models

Duma election of 2007. The election data is analyzed for 85 regions. 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 1a. 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 1a. 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 model, if it fits, can be estimated as $t_{max} = -a_1/(2\, a_2)$. For this model, a bootstrap procedure [18] 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 heteroscadacity. 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 1b.

Table 1b. Descriptive statistics for the 2008 Presidential election.

	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 2) 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, was 2.04 in 2007 and 1.40 in 2008. The estimate of log *k* = 2.04 (95% CI is 1.99 -2.08) 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. Sell also Figures 2 and 3.

Table 2. 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: k = 2^log k; standard errors are corrected for heteroscadacity; p < 0.00001

The parabolic regression models (Table 3, Figures 2 and 3) have virtually the same strength of fit as the first models when measured by R^2—0.88 for 2007 and 0.74 for 2008. The estimated coefficients a_0 are not statistically different from zero, failing to reject a difference at p< .05; so the parabolic curves approximately intersect the origin as predicted by the hypothesis. 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.

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.

Table 3. Regression coefficients 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 heteroscadacity; $p < 0.00001$

Figure 2. Predicted vs Observed Models: 2007

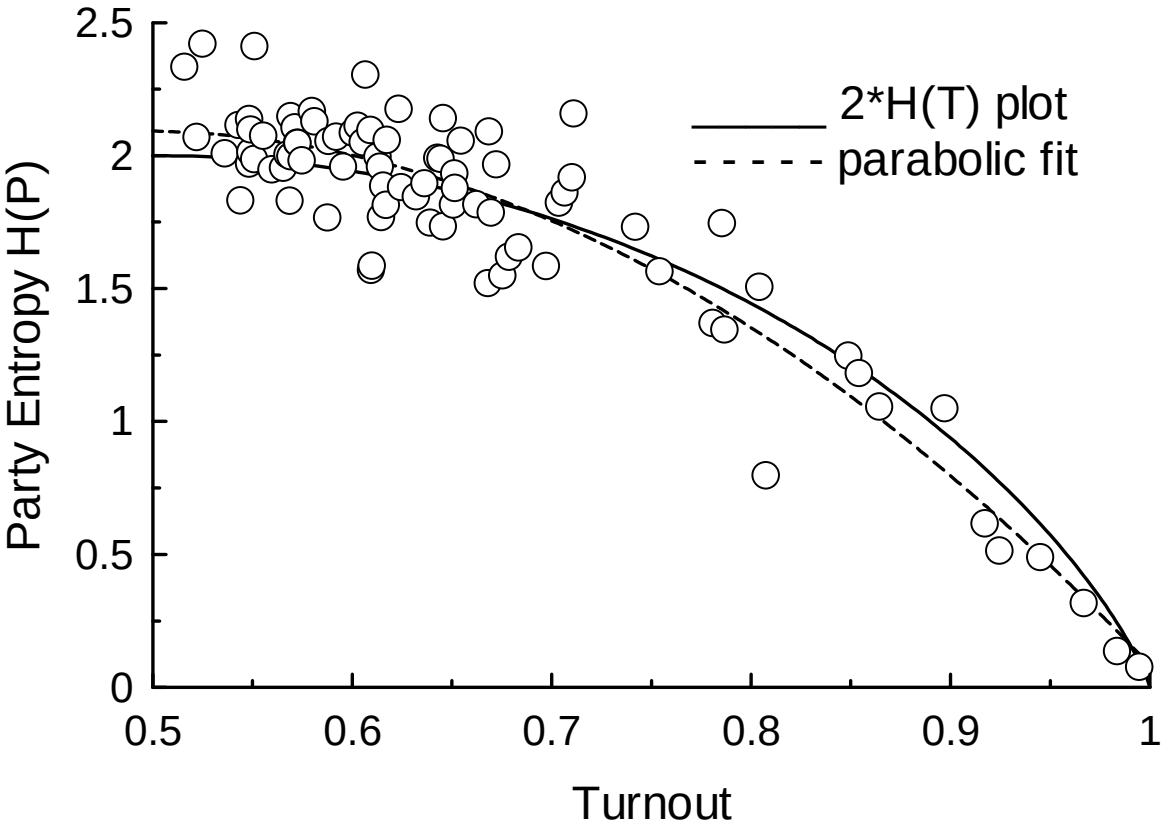
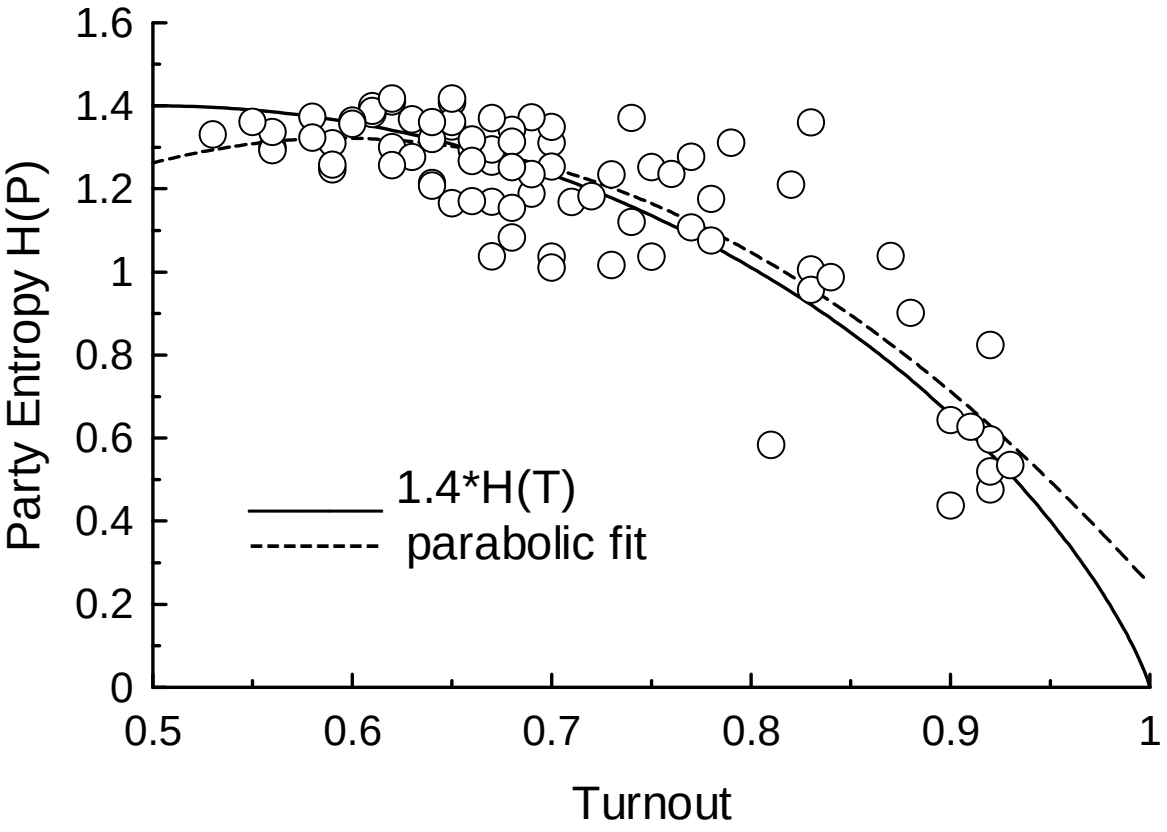


Figure 3. Predicted vs Observed Models: 2008



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 6 regions that are outliers on the turnout scale having turnouts over 90% (Dagestan, Kabardino-Balkaria, Ingushetia, Karachay-Cherkssia, 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 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:

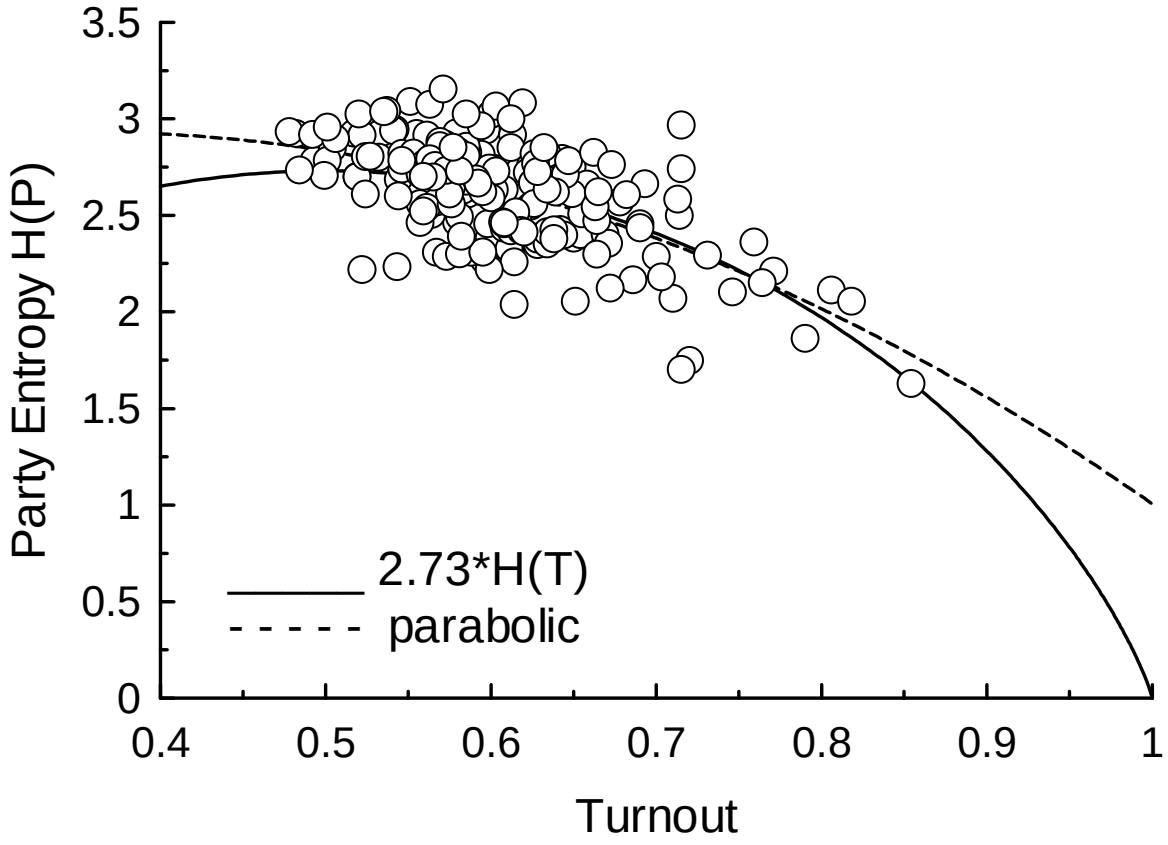
Q17-t. How many political parties does Russia need now? (Percentage in agreement)

	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 also increased so that the effect of conformity is now more pronounced. The party system has evolved, however. The fragmentation of the party system has substantially decreased, as 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) [2]; and the regression of the other 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 psychological limit of about 3 bits, which is rare and not likely to be sustainable in a society. These extreme entropy levels were characteristic of 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 Observed Models: 1999



Were the 2007 and 2008 elections less fair or less democratic than earlier elections? 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 [18]. Other research argues against this, however [19]. Fewer concerns were raised about the presidential election. One might ask whether this research can address the question of fairness. 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. 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 patterns seen in earlier elections; they are not unusual or suspect. 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. The heightened conformity now in evidence in Russia, however, suggests 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. Heightened social pressure might be perceived by observers as undue influence on voting behavior, but that does not necessarily imply fraud.

References

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

2. Coleman, S. *Popular Delusions: How Social Conformity Molds Society and Politics.* Youngstown, NY: Cambria Press, 2007.

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

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

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

6. Бородин, А.Д., [Borodin, A.D.] Согласованность коллективных действий в поведении российских избирателей [The Effect of Social Conformity on the Behavior of Russian Voters], ЭКОНОМИЧЕСКИЙ ЖУРНАЛ ВШЭ [Economic Journal HSE]. 2005. Tom 9. № 1: 74-81.

7. Mueller, D.C. *Public Choice II*. Cambridge, UK: Cambridge University Press, 1989.

8. Green, D.P. and Shapiro, I. *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science.* New Haven, CT: Yale University Press, 1994.

9. MacAllister, I.R, Johnston, J., Pattie, C., Tunstall, H., Dorling, D.F.L., and Rossiter, D.J. Class Dealignment and the Neighborhood Effect: Miller Revisited. *British Journal of Political Science*. 2001. Vol. 31, P. 41-60.

10. Huckfeldt, R. and Sprague, J. *Citizens, Politics, and Social Communication: Information and Influence in an Election Campaign.* NY: Cambridge University Press, 1995.

11. Knack, S. Civic Norms, Social Sanctions, and Voter Turnout. *Rationality and Society.* 1992. Vol. 4. P. 133-156.

12. Knack, S. and Kropf, M.E. For Shame! The Effect of Community Cooperative Context and the Probability of Voting. *Political Psychology.* 1998. Vol. 19. P. 585-599.

13. Hasher, L. and Zacks, R. Automatic Processing of Fundamental Information: The Case of Frequency of Occurrence. *American Psychologist.* 1984. Vol. 39. P. 1372-1388.

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

15. Norwich, K. *Information, Sensation and Perception*. San Diego, CA: Academic Press, 1993.

16. Miller, G. The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. *Psychological Review*. 1956. Vol. 63. P. 81-97.

17. Coleman, S. Testing Theories with Qualitative and Quantitative Predictions. *European Political Science*. 2007. Vol. 6. P. 124-133.

18. "Russian Duma elections 'not held on a level playing field', say parliamentary observers." Council of Europe. 3.12.2007.

19. Rose, R. and Mischler, W. 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.

25.03.2009