Прикладная эконометрика, 2022, т. 67, с. 74–96.
Applied Econometrics, 2022, v. 67, pp. 74–96.
DOI: 10.22394/1993-7601-2022-67-74-96

***E. A. Podkolzina*, L. E. Kuletskaya, O. A. Demidova**[1]

Spatial modelling of voting preferences: The "Mystery" of the Republic of Tatarstan

We argue neighbors play a crucial role in voting behavior for the main candidate in Russia. Moreover, the official status of the region and connectedness with the ruling party matter. The neighborhood effects we explain with the idea that voters base on public choices and illustrate it on the example of Privolzhskiy federal district regions with an emphasis on Tatarstan and its effect on voting on the municipal level. The Republic of Tatarstan is an interesting case also because it is the republic in Russia that has reference to sovereignty in its constitution and at the same time is loyal to the Kremlin. This paper presents a detailed spatial analysis of voters' responses at the municipal level covering Russian presidential elections in 2018 year using the example of the Republic of Tatarstan and its surrounding regions. The preferred 2-step OLS specification with instruments shows that Tatarstan had a strong positive effect on neighboring regions in terms of voting for the main candidate, while surrounding regions voted differently and negatively affected each other. Municipalities with better economic conditions had a negative impact on the share of votes for the main candidate and positive for the opposite.

Keywords: spatial autocorrelation; electoral preferences; local economic conditions; GMM; 2-step OLS; Russian 2018 presidential elections.

JEL classification: C21; C31; R5.

1. Introduction

Much empirical research provides the deep analysis of the political behavior of voters with spatial modeling of electoral choice and the determination of factors that affect the political preferences of voters. While most of these studies are focused on general economic or political factors in the country that affect the election results (and in this case, spatial econometric methods become only a necessary tool), the rest of the research is devoted to describing and predicting the impact of membership in social groups, mobility across regions on voting results or mutual influence of voters on the propensity to join certain parties and vote for specific candidates (Kuletskaya, 2021). The main position in such studies is the hypothesis of positive spatial autocorrelation: for almost all socio-economic and political actions spatially close units (territories) are more likely to behave similarly than spatially distant units (Darmofal, 2006; Huckfeldt,

[1] **Podkolzina, Elena** — HSE University, Moscow.
Kuletskaya, Lada — HSE University, Moscow; lada.kuletskaya@gmail.com.
Demidova, Olga — HSE University, Moscow; demidova@hse.ru.

1986); there is a spatial clustering of similar behaviors and preferences between voters living in neighboring territories.

Despite the large number of studies devoted to this topic, very few of them analyze Russian voting data. Basically, the subject of the study is the elections in the United States (Burnett, Lacombe, 2012; Dow, 2001; Kim et al., 2003; Lacombe, Shaughnessy, 2007) or Britain (Arzheimer, Evans, 2012; Cutts et al., 2014; Jensen et al., 2013). Based on these papers authors made the conclusions about:

1) positive spatial autocorrelation (Burnett, Lacombe, 2012; Kim et al., 2003; Wuhs, McLaughlin, 2019);

2) the importance of taking into account the geographical distance between the party and the electors (Arzheimer, Evans, 2012; Górecki, Marsh, 2012);

3) the importance of estimating contextual effects (Burbank, 1997; Eulau, Rothenberg, 1986; Pattie, Johnston, 2000).

Nevertheless, few studies of elections in Russia emphasize the authoritarian structure of the country and the importance of considering the regional factor as "some regional leadership use their tight political control to produce strong electoral support in federal elections" (Moraski, Reisinger, 2010, p. 1), what implies different approach comparing with Western democracies. Thus, we can't generalize implications from other papers (considering USA or Britain) and should provide new research.

In this paper, we continue the paper published by (Podkolzina et al., 2020) where the authors proved the significant and positive spatial autocorrelation in 2018 Russian presidential election by using global Moran's *I* index and divided all regions into local clusters and outliers by using local Moran's *I*. It was achieved that the Republic of Tatarstan had both local clusters and local outliers of territorial election commissions (TECs)[2] and, therefore, this republic is especially different, what makes its analysis interesting. The Republic of Tatarstan is an interesting case also because it is the republic in Russia that has reference to sovereignty in its constitution and at the same time is loyal to the Kremlin.

As was achieved by Moraski and Reisinger (2010), electoral deference over time occur in the form of geographic clusters with regions — leaders influencing its neighbors ("other regions witnessed the behavior of deferential leaders, perceived the likely benefits of such action, and changed their behavior accordingly" (Moraski, Reisinger, 2010, p. 2)). This paper reflects a deep spatial analysis of Tatarstan's influence on its neighboring regions in terms of voting in the Russian 2018 elections. Considering possible general direction of influence from Republic of Tatarstan, based on papers (Moraski, Reisinger, 2010; Reisinger, Moraski, 2009) we argue that Tatarstan increases votes for the main candidate in the bordering municipalities of other regions due to its economic development and historical loyalty to the Kremlin. In addition, discussing why Tatarstan's neighbors may copy behavior we based on statement published by Moraski and Reisinger (2010): "That is, given the emergence of a highly popular president actively pursuing reforms that directly influence regional interests, regions took a cue from Kabardino-Balkaria and Tatarstan that one potentially effective method for earning an audience with the Kremlin entailed showing the Kremlin that your region would be deferential when federal elections rolled around" (Moraski, Reisinger, 2010, p. 19).

[2] The grid of TECs roughly corresponds to the grid of municipalities and urban districts. In some cases, several TECs can form large municipalities.

In our research we develop ideas of (Borodin, 2005; Coleman, 2004, 2007, 2018; Aleskerov et al., 2005) that voters in their choices are led by public opinion and continue the conclusions (Moraski, Reisinger, 2010) of spatial influence of Privolzhskiy federal district regions on each other with an emphasis on Tatarstan (as deferential leader) and its effect on voting on municipal level. We include economic and social control variables as it was done in (Abrams, Butkiewicz, 1995; Boudreau et al., 2019; Cutts et al., 2014; Elinder, 2010; Forbes et al., 2020; Górecki, Marsh, 2012; Kim et al., 2003) to estimate its effect on the number of votes for the candidates in the Russian 2018 elections. Thus, we supplemented the basic idea of the influence of neighbors on each other with the help of socio-economic factors and it distinguishes our paper from the already mentioned studies on this topic.

We gathered voting results, social and economic factors, a neighborhood matrix for all the municipalities of Tatarstan and its surrounding regions in order to estimate mutual influence of municipalities using the 2-step OLS method. Included spatial lags help to decompose both effects from Tatarstan on its neighbors and neighbors on Tatarstan. We showed that spatial effects play a crucial role in voting, especially for the main candidate. Spatial variables explain up to half of all variance in voting. Moreover, the impact of Tatarstan on other bordering regions explains a bigger share than the inverse effect of bordering regions on Tatarstan and this influence increases the share of votes for the main candidate. We explain such impact by high historical loyalty of Tatarstan to the current political party in power and its economic development, which its neighbors are keen to replicate and therefore also vote for the main candidate.

Thereby, we firstly check if the previous findings that voting results exhibit spatial autocorrelation (were only obtained for a special subset of countries) can be expanded to other countries (autocracies, countries with different political structure). Secondly, we check if Tatarstan is a republic that influence its neighbors in terms of voting and presents a “hot spot” as was supposed in (Moraski, Reisinger, 2010). And thirdly, we use the municipal (not regional) data level as we believe that such a micro-level can significantly expand the practical application of spatial modeling and allow us to better identify the correct directions of possible neighborhood effects.

This paper is organized as follows. Literature review in Section 2 provides papers related to the research problem with an accent on Russian elections. Section 3 discusses details about the data sample and descriptive statistics. Section 4 contains chosen estimation method; all the findings are described in the Section 5 and tested in Section 6, and its summary is presented in Conclusion.

2. Literature review

In the literature there are many papers related to voting analysis considering spatial approach. Generally, the purpose of the theoretical spatial approach is to predict and explain voter preferences in particular locations surrounded by other electorate’s territories with a specific candidate’s policy position, as voters and candidates can be represented by points in an issue space (Black, 1948; Davis et al., 1970; Downs, 1957; Enelow, Hinich, 1982, 1989). Theoretical papers are usually characterized by a limited amount of voters who maximize their self-interest; the candidate reflects a voter’s fixed ideal point and closest to him received the vote (Poole, Rosenthal, 1984). One of the common examples of applying theoretical methods can be found in (Enelow, Hinich, 1989), where the authors presented a general probabilistic voting theory where two candidates compete and maximize expected votes. Their aim is to show the conditions for achieving equilibrium in the model and what are the factors that break this equilibrium. It was stated that the electorate makes a decision according to preferences of each subgroup taking into account the information level, subgroup size and preference intensity (Clinton, 2006; Enelow, Hinich, 1989; Erikson, Romero, 1990).

However, empirical researchers usually face a range of problems while estimating voter choice and defining key factors affecting this choice. Firstly, voter should face policy positions measured with errors and undesirable future events that can occur with a particular candidate in power. Thus, even if a voter is rational and informed about all policy positions of the candidates there is no strong way to predict the voter's final choice. Secondly, if a candidate has spatial mobility across regions, he can change his preferences due to the "social conformity" factor (Coleman, 2018).

Coleman (2004, 2007, 2018) emphasizes that commonly people in their choices are affected by the preferences of their affiliated group: family, relatives, friends, colleagues. However, it has not yet been found if this effect is connected with an intuitive wish to be like others or with sharing a common, rational self-interest with others. It has been noticed that people always compare their behavior with others, correcting it with widely spread behavioral patterns or accepting common opinion (Gerber, Rogers, 2004). In (Coleman, 2018) the author found evidence that social conformity existed in a large proportion of voters during Russian State Duma elections. The effect of conformity appeared in the presidential elections to the same degree. This idea seemed to be strongly correlated with the general spatial approach that voters living in neighboring localities influence each other. Russian elections also were studied by Borodin (2005) who proved the significant role of social conformity in the elections of 1995, 1999 and 2003. What is more, over time, social conformity plays a greater role in voting, however sometimes "voter–conformists" can be mistaken in forming an opinion about what the widely accepted political preferences are.

There are also a lot of research in behavioral politics and sociology regarding how social environment affects the behavior of voters (contextual effects) (Burbank, 1997; Cox, 1968; Durlauf, 2004; Ethington, McDaniel, 2007; Huckfeldt, Sprague, 1991; Johnson et al., 2002; Pattie, Johnston, 2000; Stoetzer, 2017). For instance, Huckfeldt and Sprague (1991) found that people suffering from peer pressure tend to vote the same way as their friends and colleagues even if they have different candidate preferences. Nieuwbeerta and Flap (2000) analyzed how social networks influenced voters' behavior on the example of The Netherland elections in 1998. It was found that personal social network composition affects political choices; such characteristics as religion, education and social status had significant effects on the respondent's voting behavior (please refer to the Section 3 to find more broad literature review and explanation of nature of contextual effects).

Some researchers investigate contextual effects analyzing not voters' social characteristics, but the economic and social environment during elections. (Cutts et al., 2014) examined the results from "spillover effects" in the 2010 British elections. In particular, authors considered the "spillover effects" of financing political campaigns in one locality on the voting results in neighboring localities. In other words, it was assumed that the more money a political party spends on campaigning in a certain locality, the more votes it will get from residents not only of that locality, but also from neighboring ones. The authors suppose that voters in neighboring localities listen to the same media, move freely between localities both in their daily lives and during active political campaigns, and, as a result, are aware of political activities taking place in neighboring localities. The similar results on the 2010 British elections were found by Jensen et al. (2013) who assumed that media resources are closely connected with social status (employed in senior position / student) as variables of social status can be a proxy for media. The authors concluded that such factors as the incumbent, senior position at work, and marital status positively affected votes for the Conservative party, while, in contrast, variables such as being a student, solo parent and having a low income had significant negative effect on support for the Conservative party. The main approach in this paper is the spatial dependences among regions, considering the spatial Durbin model as the best model to explain this issue.

In line with (Cutts et al., 2014), there is a range of articles considering not only spatial specification, but the relationship between economic indicators. Kim et al. (2003) tested the hypothesis that votes for the party currently in power are positively associated with income growth and negatively associated with an increase in the unemployment rate, using spatial error model as a main empirical approach. The advantages of spatial error model (SEM) were also highlighted by Lacombe and Shaughnessy (2007), who found the superiority of SEM models compared with OLS for making conclusions related to the significance of regressors. Such factors as economic growth and unemployment were also considered by Elinder (2010), who analyzed Swedish general elections at two levels: regional and municipal. On the municipal level, unemployment did not affect votes as much as at the regional level, while growth had no significant effect on government support on both levels. Hence, the economic and social environment undoubtedly played an important role influencing voters (Abrams, Butkiewicz, 1995).

There are few papers considering Russian elections. Moraski and Reisinger (2010) explored spatial features of Russian political development and loyalty for the current political party in power across regions. It was shown that republics and southern regions support the Kremlin more than other regions, while northwest Russia voted less for the current political party in power. The number of regions with high votes for the Kremlin increased over time and by 2004 there were seven regions with constantly high votes for the Kremlin, including Kabardino-Balkaria Republic and the Republic of Tatarstan. This distribution can be connected with more federal subsidies and budget money for loyal regions. Tatarstan was also called as deference leader in the 2000 presidential elections, although it presented their own president in 1991 elections and boycotted the 1993 Duma elections until republics were recognized as having greater sovereignty in comparison with other regions and until Tatarstan signed a power-sharing treaty with the federal government that was included in the 1993 Constitution. After such changes in the political status of Tatarstan, it entered in the range of regions with the highest number of votes for the current political party in power. What is more, Moraski and Reisinger (2010) also found positive spatial autocorrelation in voting across Russian regions in Duma election after 2003 year with some republics as “hot spots” in terms of spreading regional authoritarianism and promoting voting for a particular party.

In our research we develop idea of (Coleman, 2018) that voters in their choices are led by public opinion and continue conclusions found by Moraski and Reisinger (2010) of the spatial influence across regions with an emphasis on Tatarstan and its effect on voting on municipal level. This paper extends the existing literature on this topic in the following ways.

As it was outlined, previous findings about spatial autocorrelation in voting were obtained mostly only for a special subset of countries (USA, Britain) with different political structure and political regime, so to eliminate doubt the generalizability of these results to other countries with different political regime (autocracies), we investigate neighborhood effects using the example of recent Russian elections, that provides significant insights about the features of voting in Russia.

We use the municipal (not regional) data level as we believe that such a micro-level can significantly expand the practical application of spatial modeling and allow us to better identify the correct directions of possible neighborhood effects.

We check if Tatarstan is a republic that influence its neighbors in terms of voting and presents a "hot spot" as was supposed in (Moraski, Reisinger, 2010). This particular case helps in understanding features of voting in Russia and provides evidence of expanding such analysis to whole Russian regions and check the hypothesis of influence strong Russian region — leader on its neighbors.

3. Peculiarities of Russian regions

Regions in Russia are political units with large differences not only in size, population and other characteristics, but in status and federal significance. Unlike other regions (oblasts or krais), republics are described in the Russian Constitution as states, but this designation does not mean that they have high state sovereignty. The Constitution of the Republic of Tatarstan provides that sovereignty is an "inherent quality condition" of the Republic, interpreting it as the possession of full state power outside the jurisdiction and powers of Russia. The main law of the Republic is the Constitution of Tatarstan, adopted on November 30, 1992. In case of a conflict between Federal law and a normative legal act of the Republic of Tatarstan, the normative legal act of the Republic of Tatarstan should be applied. Besides Tatarstan, in our research we consider its neighboring regions: the republics of Udmurtia, Bashkortostan, Chuvashia, Mariy El and Orenburg, Samara, Ulyanovsk, and Kirov regions (oblasts).

Focusing on voting results across regions, Moraski and Reisinger (2010) outlined that it is difficult to determine what regions showed strong support to the current regime in the Kremlin in the 1991, 1996, and 2000 presidential elections, while we face with dramatic jump in the number of deferential regions between the 2000 and 2004 presidential elections. However, some tendencies can be highlighted in the number of regions showing deference to the Kremlin over time since 2000. According to (Moraski, Reisinger, 2010) the deference leaders in the 2000 and 2004 presidential elections were Dagestan, Ingushetia, Kabardino-Balkaria, and Tatarstan. Although Tatarstan was not in top-7 regions with highest number of votes for Putin, it reflects a strong commitment to stable and high voting results for this candidate.

Table 1 reflects voting results for Putin (or Medvedev in 2008) for presidential elections, taking into consideration Tatarstan and its surrounding regions. It can be seen in the 2012 and 2018 elections that Tatarstan always provides a high number of votes for the Kremlin, while others do not have such a tendency.

Table A1 in Appendix A, where voting in 2004, 2008, 2012, 2018 elections is presented in the context of the entire country, shows that Tatarstan was one of the top positions in voting for V. Putin / D. Medvedev among other regions.

In (Podkolzina et al., 2020) the authors showed that spatial dependence explains voting behavior across regions for voting results in the 2018 Russian presidential elections. The authors calculated global and local Moran's I and Geary indices (Anselin, 1995) to measure the degree of similarity of voting results of neighboring territories, and the Getis–Ord index to assess whether a particular territory can be attributed either to a cluster of high values of voting results or to a cluster of low values. The authors found that the shares of votes for all candidates in neighboring localities are similar across the country, with clusterization of mostly high values for all candidates. As for local indices that allow to divide all regions into clusters (where voters vote in a similar way) or outliers (where voters vote differently), the authors found that for the main candidate Tatarstan

Table 1. Voting results for V. Putin / D. Medvedev in 2004, 2008, 2012 and 2018 presidential elections in the Republic of Tatarstan and surrounding regions

Region	Share of votes for			
	V. Putin, 2004	D. Medvedev, 2008	V. Putin, 2012	V. Putin, 2018
Republic of Bashkortostan	0.918	0.880	0.753	0.777
Mariy El Republic	0.673	0.772	0.600	0.740
Republic of Tatarstan	0.826	0.792	0.827	0.821
Udmurt Republic	0.759	0.705	0.658	0.762
Chuvash Republic	0.671	0.665	0.623	0.773
Kirov oblast	0.655	0.763	0.579	0.704
Orenburg oblast	0.588	0.608	0.569	0.730
Samara oblast	0.633	0.641	0.586	0.758
Ulyanovsk oblast	0.659	0.669	0.582	0.743

has territorial electoral commissions (TECs) that can be included in the group with similar percentages of votes, and TECs with different percentages of votes with neighbors. In this research, we investigate the impact of Tatarstan on voting in neighboring regions as we found an ambiguous direction of spatial dependence in TECs.

According to the literature review and descriptive statistics of voting results we suppose that Tatarstan positively affects the propensity of neighboring regions to vote for the main candidate. We expect a positive influence as Tatarstan is loyal to the current political party in power and is a successful region in economic terms (Table 2). Its neighbors, therefore, are keen to replicate the model and vote for the main candidate.

This suggestion is strongly in line with the previous empirical works regarding contextual and spatial effects in voting (Boudreau et al., 2019; Burbank, 1997; Eulau, Rothenberg, 1986; Górecki, Marsh, 2012; Jessee, 2009; Johnson et al., 2002; Johnston, 2004; Pattie, Johnston, 2000; Wuhs, McLaughlin, 2019). As was mentioned by Burbank (1997) the main approach of contextual effects is that individuals' political views and actions can be represented by the individual characteristics and the attributes of other people in the overall social environment. The social environment of a person consists of friends, family, neighbors, colleagues and institutions (churches, clubs, schools, businesses, etc.). Thus, the information that people encounter is biased according to the social composition of the local environment. People are much more likely to change their votes in a certain direction if those with whom they discuss political issues support that direction, especially if they are family members (with whom they most often discuss politics) (Burbank, 1997). In addition, as Wuhs and McLaughlin (2019) argued, neighboring counties may exhibit similar voting patterns because neighboring districts can be influenced by the same unobservable and non-obvious effect or event. What is more, voters interact more often with voters from neighboring districts and political campaign raking place in some district affect its neighbors.

In this paper, we base on mentioned achievements and postulates in academic literature and argue that electorate copy the political preferences of their neighboring social environment.

We also assume that the stability of the region and its economic success matter for electorate. As Tatarstan is the most successful in economic terms region among its neighbors, it's more rational for current political party in power to translate the agitation and stimulate to vote for the main candidate through this region. Tatarstan neighbors, in turn, are more susceptible to such influence

Table 2. Tatarstan's and its neighbors' places in 2014–2018 among all Russian regions according to GRP per capita / Consumer Price Index (sorted by 2018)

#	Region	2014	2015	2016	2017	2018
15	Republic of Tatarstan	490.80	549.43	567.95	624.71	728.39
24	Orenburg oblast	424.19	451.05	450.70	488.12	583.68
28	Samara oblast	368.87	406.29	410.21	448.94	509.46
42	Udmurt Republic	322.72	371.09	386.96	403.41	464.33
44	Republic of Bashkortostan	336.41	351.52	359.13	373.37	448.37
63	Ulyanovsk oblast	239.57	262.61	284.24	296.30	301.08
66	Mariy El Republic	234.16	280.90	263.15	278.65	289.78
67	Kirov oblast	211.20	235.76	244.35	259.46	279.89
71	Chuvash Republic	217.73	230.80	240.45	255.12	278.85

from a stronger region and thus copy the behavior of Tatarstan's voters. In addition, Tatarstan shows a strong historical loyalty to the Kremlin and therefore increases the share of votes for the main candidate. Thus, our hypothesis is that Tatarstan positively affects the propensity of neighboring regions to vote for the main candidate.

4. Data and methodology

4.1. The data

For this research we collected data from various sources described below. A full description of all variables used in the final model is presented in Section 4.2.

1. Neighborhood matrix of municipalities. In order to build a neighborhood matrix for municipalities in the Republic of Tatarstan and eight surrounding regions we used open data for QGis 3.12 website from the data catalog site https://mydata.biz/ru/catalog/databases/borders_ru. The matrix itself was built manually, based on the visual map of Russian municipalities presented in the website.

2. Data on voting at the municipal level. The initial data on voting was taken from the website of the Central Election Commission (http://www.cikrf.ru, http://www.vybory.izbirkom.ru) on the results of the Russian presidential elections in 2018 at the level of TECs. These data contain information about the percentage of voters who voted for the candidates and the turnout at the polling stations. We converted data from the TEC level to the municipal level based on the available data on the number of ballots for each candidate and the total number of voters.

3. Social and economic factors of municipalities. The explanatory control variables were taken from the open state database "Indicators of municipalities" of the Federal state statistics service (https://rosstat.gov.ru/free_doc/new_site/bd_munst/munst.htm). The variables were collected for 2017 (i.e. the year preceding the election and forming the opinion of voters). When collecting the necessary variables, we were guided by the principle of visible economic indicators for voters. We believe that in their choice, voters are primarily guided by easily visible and tangible results of the current situation in the region/municipality (for example, the number of houses commissioned, sports and cultural events held, developed infrastructure, etc.), rather than by any General economic or demographic indicators.

In this paper we consider voting for two candidates: Vladimir Putin (who is defined as main candidate) and Ksenia Sobchak (who is defined as opposition candidate). Sobchak was chosen as she represented the "non-system" opposition in contrast with other candidates. Other candidates belong to the "system" opposition, participating in elections from time to time with the same political statements. In 2018, Sobchak participated in the election for the first time, and does not belong to any of the current political parties and is not dependent on the political establishment.

4.2. Methodology

To consider the interdependencies in voting across neighboring territories, we used the spatial autoregression model (SAR) as a benchmark model:

$$Y = X\beta + \rho WY + \varepsilon, \tag{1}$$

where Y is a dependent variable (voting results for a particular candidate), X is a matrix of explanatory variables, W is a boundary weighting matrix $N \times N$, where N is a number of municipalities (elements of matrix are $w_{ii} = 0$, $w_{ij} = 1/N_{ij}$, N_{ij} is the number of municipalities that have common border with municipality i, $i \neq j$), WY is a spatial lag (in our case: the average share of votes for a candidate in neighboring municipalities).

We proceed from the premise that residents of neighboring municipalities are affected by each other in voting decisions, therefore, the spatial lag WY is endogenous (the proportion of voters for the nominated candidate in a given municipality and in neighboring municipalities is correlated). That is why it is incorrect to evaluate model (1) using OLS as the corresponding coefficients' estimates will be biased.

To solve this problem, we use the generalized method of moments (Verbeek, 2008) and instrumental variables for WY. The most popular approach for selecting instruments was introduced by Kelejian and Prucha (1998) where columns of the WX, W^2X matrices are proposed as the instruments for the endogenous variable WY.

The technical explanation for choosing these tools is the following transformation of the model (1):

$Y = (1 - \rho W)^{-1}(X\beta + \varepsilon)$, we suppose that $|\rho| < 1$,

$$Y = (1 + \rho W + \rho^2 W^2 + \rho^3 W^3 + \ldots + \rho^n W^n + \ldots)(X\beta + \varepsilon),$$

i.e. $$WY = (WX + \rho W^2 X + \rho^2 W^3 X + \ldots + \rho^n W^{n+1} X + \ldots)\beta + (1 - \rho W)^{-1}\varepsilon. \tag{2}$$

According to formula (2) we can assume that WX, W^2X are correlated with WY, but since the factors forming the matrix X are exogenous, they do not correlate with errors. That's why WX and W^2X are relevant instruments.

Thus, for the average share of votes in i^{th} municipality neighboring, we offer average values of explanatory factors in neighboring municipalities (WX) and second-order neighboring municipalities (W^2X) as the instruments. We assume that these indicators are not correlated with the shocks of the i^{th} municipality. These assumptions seem plausible enough, but we also provide tests on the relevance and validity of the instruments.

As the dependent variables, we choose Y_{base} the share of votes for the main candidate, Y_{opp} the share of votes for the opposition candidate. The explanatory variables *X* for both dependent variables are the same, we discuss them in more detail below.

For each of the dependent variables we calculated spatial lags WY_{base} and WY_{opp}.

To identify the possible difference in the voting of the electorate living in Tatarstan and outside of it, we modified the model (1). Instead of a single *WY* spatial lag that reflects the overall influence of neighboring municipalities, we introduced several spatial lags that reflect the specifics of the influence of central municipalities on the bordering municipalities of Tatarstan and its surrounding regions. The final models for the main and opposition candidates, considering the introduced several spatial lags are represented by models (3) and (4), respectively:

$$Y_{base} = X\beta + \rho_1 WY_{base} D_{T_T} + \rho_2 WY_{base} D_{T_{NT}} + \rho_3 WY_{base} D_{NT_T} + \varepsilon, \tag{3}$$

$$Y_{opp} = X\beta + \rho_1 WY_{opp} D_{T_T} + \rho_2 WY_{opp} D_{T_{NT}} + \rho_3 WY_{opp} D_{NT_T} + \varepsilon, \tag{4}$$

where D_{T_T} is a dummy variable = 1, if the municipality is *located* in Tatarstan and does not have border(s) with municipalities of other regions;
$D_{T_{NT}}$ is a dummy variable = 1, if the municipality is *located* in Tatarstan and has a border(s) with municipalities of other regions;
D_{NT_T} is a dummy variable = 1, if the municipality is *not located* in Tatarstan and has a border(s) with municipalities of Tatarstan.

The different spatial lags included in the models estimate influence from central or bordering municipalities of the regions on each other in voting for the two candidates. Estimated coefficients ρ_1, ρ_2, ρ_3 reflect the level and significance of this influence.

The spatial lags in models (3), (4) are also endogenous, thus we used instrumental variables $Z_1, Z_2, ..., Z_l$ $(l \geq 1)$ that are columns of matrices WX and W^2X multiplied by $D_{T_T}, D_{T_{NT}}, D_{NT_T}$ (we used the approach of (Kelejian, Prucha, 1998)).

The estimates of parameters $\beta, \rho_1, \rho_2, \rho_3$ of models (3), (4) were calculated with the generalized method of moments (Baum et al., 2007; Stillman, 2007; Verbeek, 2008). Moment identities were based on orthogonality of $Z_1, Z_2, ..., Z_l$ and ε, as well as $X_1, X_2, ..., X_k$ and ε (where *k* is a number of columns of matrix *X*).

In order to take into account the influence of municipalities from neighboring regions of Tatarstan on each other, for the main candidate we also evaluate model (5), which is essentially model (3) in which the spatial lag $WY_{base} D_{T_T}$ was changed to $WY_{base} D_{NT_{NT}}$:

$$Y_{base} = X\beta + \rho_1 WY_{base} D_{NT_{NT}} + \rho_2 WY_{base} D_{T_{NT}} + \rho_3 WY_{base} D_{NT_T} + \varepsilon, \tag{5}$$

where $D_{NT_{NT}}$ is a dummy variable = 1, if the municipality is *not located* in Tatarstan and does not have border(s) with municipalities of Tatarstan.

For the opposition candidate, model (5) cannot be estimated as for this candidate the spatial lags are strongly correlated with each other (stronger than for the main candidate). In our work, we do not include all four spatial lags $(WY_{base} D_{NT_{NT}}, WY_{base} D_{T_T}, WY_{base} D_{T_{NT}}, WY_{base} D_{NT_T})$ at the same time, due to the problem of strong multicollinearity, since spatial lags are closely interrelated.

As usual, when using instruments, we need to check that they are relevant and valid. The relevance of instruments means that they are related to the variables for which they are used. If there is only 1 endogenous variable (denoted as $\widetilde{X}$), the estimated model should be identified as

$$\widetilde{X} = \beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + ... + \beta_l Z_l + \gamma_1 X_1 + \gamma_2 X_2 + ... + \gamma_k X_k + \varepsilon, \quad (6)$$

where l is the number of columns of matrix Z (instruments), k is the number of columns of matrix X (regressors), with $l \geq k$.

The following hypothesis is tested:

$$H_0: \beta_1 = ... = \beta_l = 0, \qquad H_1: \beta_1^2 + \beta_2^2 + ... + \beta_l^2 > 0,$$

with F-statistics. If H_0 is rejected, this is equivalent to the rank of the vector $\Pi = (\beta_1, \beta_2, ..., \beta_l)$ being equal to 1, so the instruments $Z_1, Z_2, ..., Z_l$ are relevant.

If there are several endogenous variables (in our case these variables are spatial lags $WYD_{T_T}, WYD_{T_{NT}}, WYD_{NT_T}$, where Y is equal to Y_{base} or Y_{opp} (depending on the candidate selected for consideration), the relevance test of the instruments $Z_1, Z_2, ..., Z_l$ reduces to check that the system of equations

$$\begin{aligned} WYD_{T_T} &= \beta_{01} + \beta_{11} Z_1 + \beta_{21} Z_2 + ... + \beta_{l1} Z_l + \gamma_{11} X_1 + \gamma_{21} X_2 + ... + \gamma_{k1} X_k + \varepsilon_1, \\ WYD_{T_{NT}} &= \beta_{02} + \beta_{12} Z_1 + \beta_{22} Z_2 + ... + \beta_{l2} Z_l + \gamma_{12} X_1 + \gamma_{22} X_2 + ... + \gamma_{k2} X_k + \varepsilon_2, \quad (7) \\ WYD_{NT_T} &= \beta_{03} + \beta_{13} Z_1 + \beta_{23} Z_2 + ... + \beta_{l3} Z_l + \gamma_{13} X_1 + \gamma_{23} X_2 + ... + \gamma_{k3} X_k + \varepsilon_3 \end{aligned}$$

has complete row rank which is 3, i.e. the matrix

$$\Pi = \begin{pmatrix} \beta_{11} & ... & \beta_{l1} \\ \beta_{12} & ... & \beta_{l2} \\ \beta_{13} & ... & \beta_{l3} \end{pmatrix}$$

has the rank 3.

To check this condition, *rk* statistics (Kleibergen, Paap, 2006) were applied. For details see (Baum et al., 2007).

The validity of $Z_1, Z_2, ..., Z_l$ refers to their orthogonality to ε, which can be tested with the Hansen's *J*-statistic (Baum et al., 2007; Verbeek 2008).

As for social and economic factors of municipalities (matrix X), we collected variables that can reflect electorate preferences in voting. We assume that, in addition to the opinion of the closest social environment, voters are influenced by how the municipality/district looks like. That is why it is important to include visible and tangible indicators for the population, on which voters can easily rely in their choice. Our suggestions and hypotheses are presented below. It should be noted that since the main candidate is also a current candidate, the voters assessing their readiness to vote base on what the candidate has done previously, how well he has followed the economic situation in the country and in the region of the potential voter, thus, we collected data for 2017 as it precedes 2018 presidential elections and forms voters' choices.

The following list of socio-economic indicators and *republic* dummy were included as explanatory variables in models (3), (4) and (5).

$X_1 = road\_length$ is the ratio of the length of local public roads to the area of the municipality. The length of roads may be a proxy for everyday mobility and reflect the availability of moving to neighboring regions (and for example to Tatarstan) or to the capital of Russia (and receive new information from different sources). That's why is it is unclear what effect will prevail and what sign the coefficient will have before this variable in the final model.

$X_2 = fam\_subsid$ is the ratio of the number of families receiving subsidies for housing and utilities to the population. This indicator characterizes the willingness of the authorities to spend money on the social well-being of citizens and will directly affect the willingness to vote for the main candidate, since at the time of the election he was also the current President. We assume that the impact of this variable cannot be clearly identified in advance: on the one hand, it is reasonable to assume that the more families receive subsidies, the more voters should support the main candidate. On the other hand, more subsides mean more households in bad economic conditions and the subsidies may be insufficient, and obtaining a subsidy may involve a number of bureaucratic problems, which may negatively affect voters' support for the main candidate.

$X_3 = goods\_services$ is the ratio of the goods and services produced in the municipality (in 10000 RUB) to the municipal population. According to our previous research, the more goods produced in the municipality, the fewer votes the main candidate received. We suppose this can be explained that such products are supplied by micro-enterprises, which are not supported by the main candidate's program.

$X_4 = houses$ is the ratio of the number of houses built in the municipality to the municipal population. We assume that a large number of new residential buildings is an indicator of the region's development (people want to live there) and has a positive effect on votes for the main candidate. In addition, the development of the city is directly an indicator of the development of the municipality and district for voters.

$X_5 = noncommerc\_fonds$ is the ratio of the full accounting value of fixed assets for non-commercial organizations (divided by 10) to the municipal population. We assume that a larger number of non-commercial organizations' core funds may reflect greater support from the government, thereby increasing the number of votes for the main candidate. In addition, since the activities of non-commercial organizations are often associated with the organization of sports, cultural and educational events, the larger number of non-commercial organizations' core funds will stimulate the growth of such events, which, in turn, will show voters the development of the municipality.

$X_6 = invest$ is the ratio of the investments in fixed capital at the expense of the budget of the municipality (divided by 10) to the municipal population. Investments in fixed assets most often represent the cost of new construction, expansion, and the reconstruction and modernization of facilities, the purchase of machinery, equipment, vehicles, etc. Spending on fixed assets has a positive effect on voting for the main candidate.

$X_7 = expenses$ is the ratio of local budget expenditure (divided by 10) to the municipal population. Similarly to the investment variable, the higher the local budget expenditures, the more support residents feel for their region and municipality, and the more votes they will give for the main candidate.

$X_8 = republics$ is a dummy variable = 1 if the municipality belongs to a republic (including Republic of Tatarstan) and 0 otherwise. In our research, we suppose that generally republics have a significant and positive influence on voting for the main candidate as republics have certain constitutional powers, right to set their own state languages, as well as have capitals and have been discussed as being more loyal to authoritarian regional politics and the Kremlin (Moraski, Reisinger, 2010; Stepan, 2000). This approach is also consistent with our previous work when we found that municipalities of Tatarstan neighboring Orenburg oblast (that is not a republic) voted less for the main candidate than residents of other municipalities of Tatarstan.

Table 3. Descriptive statistics of voting for V. Putin and K. Sobchak at 2018 Russian presidential elections

Locality	Variable	N	Mean	Std.Dev.	Min	Max	Median
Tatarstan (all)	putin_share	45	0.892	0.059	0.734	0.973	0.911
Not Tatarstan (all)	putin_share	281	0.771	0.049	0.652	0.923	0.771
Republics	putin_share	180	0.817	0.064	0.701	0.973	0.802
Oblasts	putin_share	146	0.751	0.048	0.652	0.923	0.747
Tatarstan (borders)	putin_share	27	0.896	0.054	0.767	0.973	0.898
Not Tatarstan (borders with Tatarstan)	putin_share	44	0.768	0.045	0.679	0.892	0.765
Not Tatarstan (centered municipalities)	putin_share	237	0.772	0.051	0.652	0.923	0.772
Tatarstan (all)	sobchak_share	45	0.007	0.004	0.002	0.019	0.006
Not Tatarstan (all)	sobchak_share	281	0.009	0.004	0.003	0.029	0.009
Republics	sobchak_share	180	0.009	0.004	0.002	0.023	0.009
Oblasts	sobchak_share	146	0.009	0.004	0.004	0.029	0.008
Tatarstan (borders)	sobchak_share	27	0.006	0.003	0.002	0.014	0.006
Not Tatarstan (borders with Tatarstan)	sobchak_share	44	0.008	0.003	0.004	0.013	0.008
Not Tatarstan (centered municipalities)	sobchak_share	237	0.009	0.004	0.003	0.029	0.009

Note. In this table and in all following tables we use such designations: N — number of observations; Std.Dev. — standard deviation; Min — minimum value; Max — maximum value.

Some of the collected socio-economic indicators are strongly connected, so in order to avoid the problem of multicollinearity, we initially included all factors in the model, and then tested the hypothesis of the joint insignificance of coefficients for a group of variables. If the hypothesis was not rejected, we deleted the corresponding variables. For the main specification, we had to exclude the variable *nocommerc_fonds*. However, Table B1 in Appendix B presents the results for different sets of control variables, showing the stability of signs before the coefficients and the direction of the influence of the regressors.

Table 4 reflects the descriptive statistics, median- and mean-comparison tests for all regressors divided by municipalities that belongs to the Tatarstan and that belongs to other regions. We firstly added to the Table 4 mean-comparison test to check if the means for regressors differ for mentioned types of municipalities. We found that variables *goods_services*, *houses*, *nocommerc_fonds* and *expenses* have significant differences for municipalities from Tatarstan and from other regions (Tatarstan has greater means). Secondly, we added to the Table 4 median-comparison test to check if the medians for regressors differ for mentioned types of municipalities. It was found that Tatarstan has greater medians for all variables, and the differences in medians are significant for all regressors. It is a confirmation that Tatarstan is more successful region in economic terms and is a better channel to political campaign.

Table 4. Difference (Diff.) in means and medians for regressors (Tatarstan and neighboring regions)

Variable	Mean	Diff. mean	Median	Diff. median (Pearson χ^2)	Diff. median (continuity corrected: Pearson χ^2)	Min	Max
road_length		–0.001		7.451**	6.599**		
– Tatarstan	0.005		0.003			0.001	0.060
– other regions	0.006		0.002			0.000	0.083
goods_services		198.69**		18.794***	17.428***		
– Tatarstan	420.46		215.773			35.70	3272.02
– other regions	221.77		100.012			3.105	4592.63
fam_subsid		0.003		11.369***	10.31***		
– Tatarstan	0.029		0.024			0.004	0.110
– other regions	0.025		0.013			0.000	1.837
houses		0.239**		9.307*	8.35*		
– Tatarstan	0.656		0.442			0.191	4.207
– other regions	0.417		0.332			0.008	5.195
nocommerc_fonds		17.59**		18.794***	17.428***		
– Tatarstan	73.37		67.424			28.78	315.67
– other regions	55.78		44.615			0.235	211.90
invest		0.217		7.45**	6.59**		
– Tatarstan	0.759		0.381			0.020	6.284
– other regions	0.543		0.209			0.000	14.063
expenses		6.827***		21.682***	20.21***		
– Tatarstan	30.38		29.441			16.75	62.05
– other regions	23.55		22.188			1.530	50.60

Note. For Tatarstan there are 45 observations (municipalities) and 281 for other regions.
* — p-value < 0.1; ** — p-value < 0.05; *** — p-value < 0.01.

5. Empirical results

Table 5 shows the results for the main specification with two sets of spatial lags for the main candidate. Table 5 also reflects tests for checking instruments for relevance and validity requirements. The instruments for almost all specifications are relevant and valid.

According to Table 5, spatial variables have significant influence on votes for the main candidate. The central municipalities of Tatarstan voted homogeneously for the ruling party and the surrounding border municipalities act in a similar way (the coefficients' estimate at WYD_{T_T}, $WYD_{T_{NT}}$ are positive and significant at 1% level). The border municipalities of Tatarstan also voted homogeneously for the ruling party and the border municipalities of other regions reproduce this pattern.

Table 5. Results for the main specification

Variable	Specification		
	(1) putin_share	(2) putin_share	(1) sobchak_share
WYD_{T_T}	0.1040***		–0.0184
	(0.0165)		(0.1975)
$WYD_{T_{NT}}$	0.1241***	0.0729**	–0.3227***
	(0.0126)	(0.0242)	(0.0790)
WYD_{NT_T}	–0.0005	–0.0463*	–0.0874
	(0.0085)	(0.0235)	(0.0525)
$WYD_{NT_{NT}}$		–0.0444*	
		(0.0232)	
road_length	–0.5897**	–0.5781**	0.1151**
	(0.2393)	(0.213)	(0.0360)
goods_services	–0.1444*	–0.0965	0.0169*
	(0.0649)	(0.0367)	(0.0056)
fam_subsid	–0.0481***	–0.0426***	0.0006
	(0.0138)	(0.0104)	(0.0004)
houses	0.0102	0.0162**	0.0011*
	(0.0053)	(0.0060)	(0.0005)
invest	0.0199	0.0145	–0.0002
	(0.0207)	(0.0222)	(0.0014)
expenses	–0.0017	0.0015	–0.0007**
	(0.0033)	(0.0036)	(0.0002)
republics	0.0392***	0.0454***	–0.0003
	(0.0053)	(0.0053)	(0.0004)
const	0.7592***	0.7822***	0.0096***
	(0.0103)	(0.0214)	(0.0007)
N	326	326	326
R^2 centred	0.5123	0.4710	0.2277
R^2 uncentred	0.9966	0.9963	0.8747
Underidentification test (Kleibergen–Paap *rk* LM statistic)	66.350***	58.245**	80.814***
Hansen *J*-statistic (overidentification test of all instruments)	40.254	49.298*	40.418

Note. Specification (1) reflects formulas (3) and (4) for main and opposition candidate respectively, specification (2) reflects formula (5) for main specification. Robust standard errors — in parentheses.
* — *p*-value < 0.05; ** — *p*-value <0.01; *** — *p*-value <0.001.

The results for municipalities in other regions (not Tatarstan) are not significant (the coefficient before WYD_{NT_T} variable is not significant), that's why we can conclude that the effect of spatial lags does not work so much for them. A negative sign of the coefficient's estimate before $WYD_{NT_{NT}}$ variable shows that voters in central municipalities of other regions (not Tatarstan) voted differently from the border municipalities, which are more influenced by Tatarstan. The influence comes not from the center of Tatarstan, but from its border regions.

For the opposition candidate, spatial variables have less impact — we found no significant spatial influence on voting results from voters located in the central municipalities of Tatarstan or in municipalities of other regions that have borders with Tatarstan. Voters from municipalities of Tatarstan that have borders with other regions voted differently from voters from neighboring municipalities of other regions and have a negative spatial effect on the number of votes for Sobchak. Therefore, we received the logical result: the more voters voted for the main candidate, the less votes opposition candidate received.

The length of roads and goods produced in the municipality have a significant negative impact on the voting for the main candidate (in contrast to the opposition candidate, where the signs and significance of coefficients before these variables are directly opposite), as we expected. Subsidies for housing and utilities affect significantly and negatively for the support for the main candidate, the effect of bad life conditions and large bureaucracy for processing small payments dominates here. The coefficient's estimate before dummy *republics* is significant and positive, thus, republics support the main candidate more than other regions.

Moreover, taking into consideration coefficient of determination, denoted as R^2, calculated for both candidates for all specifications, we estimated contribution of each of the spatial variables $(WYD_{NT_{NT}}, WYD_{T_T}, WYD_{T_{NT}}, WYD_{NT_T})$ to the variability of the dependent variable Y.

To do this, we used the factorial decomposition of R^2:

$$R^2 = \sum_{i=1}^{n} \hat{\beta}_i \frac{\widehat{\operatorname{cov}}(X_i, Y)}{\widehat{\operatorname{var}}(Y)}. \quad (8)$$

R^2 refers to the share of the explained variance of the dependent variable through included in the model regressors. To estimate the impact of a certain explanatory variable one may find the ratio of $\hat{\beta}_j \frac{\widehat{\operatorname{cov}}(X_j, Y)}{\widehat{\operatorname{var}}(Y)}$ to R^2.

In our estimation, we used R^2 that was calculated after using OLS models defined in (9):

$$Y = X\beta + \mu_1 \widehat{WYD_{T_T}} + \mu_2 \widehat{WYD_{T_{NT}}} + \mu_3 \widehat{WYD_{NT_T}} + \varepsilon \text{ (for main and opposition candidates),}$$

$$Y = X\beta + \mu_1 \widehat{WYD_{NT_{NT}}} + \rho_2 \widehat{WYD_{T_{NT}}} + \rho_3 \widehat{WYD_{NT_T}} + \varepsilon \text{ (for main candidate only),}$$

where

$$\widehat{WYD_{T_T}} = \hat{\beta}_{01} + \hat{\beta}_{11} Z_1 + \hat{\beta}_{21} Z_2 + \ldots + \hat{\beta}_{l1} Z_l + \hat{\gamma}_{11} X_1 + \hat{\gamma}_{21} X_2 + \ldots + \hat{\gamma}_{k1} X_k,$$

$$\widehat{WYD_{T_{NT}}} = \hat{\beta}_{02} + \hat{\beta}_{12} Z_1 + \hat{\beta}_{22} Z_2 + \ldots + \hat{\beta}_{l2} Z_l + \hat{\gamma}_{12} X_1 + \hat{\gamma}_{22} X_2 + \ldots + \hat{\gamma}_{k2} X_k,$$

$$\widehat{WYD_{NT_T}} = \hat{\beta}_{03} + \hat{\beta}_{13} Z_1 + \hat{\beta}_{23} Z_2 + \ldots + \hat{\beta}_{l3} Z_l + \hat{\gamma}_{13} X_1 + \hat{\gamma}_{23} X_2 + \ldots + \hat{\gamma}_{k3} X_k,$$

$$\widehat{WYD_{NT_{NT}}} = \hat{\beta}_{04} + \hat{\beta}_{14} Z_1 + \hat{\beta}_{24} Z_2 + \ldots + \hat{\beta}_{l4} Z_l + \hat{\gamma}_{14} X_1 + \hat{\gamma}_{24} X_2 + \ldots + \hat{\gamma}_{k4} X_k. \quad (9)$$

Table 6 represents what parts of Y variance can be explained by the spatial variables for the specifications of the models from Table 6. We see that spatial effects play a crucial role in voting, especially for the main candidate. Table 6 shows that for 1st specification for main candidate, spatial variables explain more than half of all Y variance, while for 2nd specification this value is lower but still high. For opposition candidate spatial lags do not play such a key role, for this candidate X factors are more important in Y explanation. The effect of Tatarstan on other bordering regions explains a bigger share than the inverse effect of bordering regions on Tatarstan. If we look at specification (1) (or (2)) for the main candidate, Tatarstan's impact explains 42.5% (or 32%) variation in votes for Putin in bordering regions, and all other regions explain only 0.1% (or 3%).

Table 6. Contribution of spatial lags to the variance of the dependent variable

Variable	Specification		
	(1) putin_share	(2) putin_share	(1) sobchak_share
WYD_{T_T}	23.5%	—	0.5%
$WYD_{T_{NT}}$	42.5%	32%	14.8%
WYD_{NT_T}	0.1%	3%	1.2%
$WYD_{NT_{NT}}$	—	9%	—
Aggregated spatial lags	66.1%	44%	16.5%

Note. Specification (1) reflects formulas (3) and (4) for main and opposition candidate respectively, specification (2) reflects formula (5) for main specification.

The bigger effect of Tatarstan on other bordering regions than inverse effect once again highlights the greater influence of this region on its neighbors. Thus, we confirmed that Tatarstan can be a channel of influence on neighboring regions.

All in all, we found that Tatarstan has a positive effect on the propensity of neighboring regions to vote for the main candidate, while for the opposition candidate we did not find such an effect. Surrounding republics support the main candidate too, however generally neighboring regions vote differently and have an ambiguous effect on voting for the main candidate.

As for control variables, we confirm our hypothesizes mentioned in Section 3 and outline significant aspects influencing voting results. These results are in line and continue (Podkolzina et al., 2020), where authors constructed spatial lag and spatial error models for municipalities of Tatarstan and found that for both models parameters for spatial effects factors (rho, lambda) are significant at the 5% significance level for the main candidate (for Sobchak we observed the significance of only the spatial error model).

6. Robustness check

We also found research (Enikolopov et al., 2013; Skovoroda, Lankina, 2017) assuming possible inaccurate data on turnout / the share of votes in some regions and, therefore, our initial data may not fully reflect the preferences of voters. Manipulations with votes could be realized using different channels — for example, pressure to vote in certain way and throw-ins. Pressure could influence behavior of the neighbors' and in this case, it is better not to exclude manipulated results, but throw-ins could bias results. We do not have information on the type of manipulation and prefer to control for such municipalities using dummies for excessive turnout[3]. We realized two additional ways to check our results for robustness.

1. For each region we calculated the significance of the difference in turnout in a single municipality and in the other municipalities in this region (on average). If the difference is significant and positive, then we assume that there could be an inflation and create a dummy variable with 1, in all other cases this dummy variable contains 0 (this dummy is called *dummy_1*).

2. We calculated the significance of the difference in turnout in a single municipality and in the other municipalities in our dataset (on average). If the difference is significant and positive, then we assume that there could be an inflation and create a dummy variable with 1, in all other cases this dummy variable contains 0 (this dummy is called *dummy_2*).

The initial results for the main candidate and the results with two new dummies are presented in Table 7.

Table 7. Robustness check of initial results (only spatial lags are presented) for putin_share

	Specification					
	Initial		With *dummy_1*		With *dummy_2*	
	(1)	(2)	(1)	(2)	(1)	(2)
WYD_{T_T}	0.1040***		0.0931***		0.076***	
	(0.0165)		(0.0127)		(0.013)	
$WYD_{T_{NT}}$	0.1241***	0.0729**	0.1173***	0.0836***	0.0760***	0.0606**
	(0.0126)	(0.0242)	(0.0084)	(0.0193)	(0.0133)	(0.0193)
WYD_{NT_T}	–0.0005	–0.0463*	–0.0027	–0.3333	0.0058	–0.0116
	(0.0085)	(0.0235)	(0.0079)	(0.2241)	(0.0079)	(0.0215)
$WYD_{NT_{NT}}$		–0.0444*		–0.0288		–0.0174
		(0.0232)		(0.0208)		(0.0204)
Underidentification test (Kleibergen–Paap *rk* LM statistic)	66.350***	58.245**	69.694***	49.187**	69.872***	55.050**
Hansen *J*-statistic (overidentification test of all instruments)	40.254	49.298*	46.095	46.690*	42.794	49.219*

Note. Specification (1) reflects formulas (3) and (4) for main and opposition candidate respectively, specification (2) reflects formula (5) for main specification. Standard errors in parentheses.
* — *p*-value < 0.05; ** — *p*-value <0.01; *** — *p*-value <0.001.

Based on Table 7, we can conclude that our initial results and main findings did not change after robustness check and including additional dummies can help with identifying possible incorrect official data. Thus, we confirm the conclusions made in previous section.

[3] We also looked at the research made by Sergey Shpilkin but found his method of data correction unacceptable in our case: using this approach to data correction strictly limits the number of votes in municipality from 80–90 to 30–40% which is deeply skeptical to accept (taking into account small number of observations — only 326 municipalities in total). Thus, we found our approach with dummies smoother and more suitable in this case.

7. Conclusion

Using electoral choices to support main or opposition candidates in the Russian 2018 presidential elections, we outline the importance of considering neighborhood effects while analyzing social or economic processes on the regional level. Focusing on Russia's regions at the municipal level, we confirmed the presence of spatial autocorrelation in Russian voting data and state that Tatarstan is a region — leader influenced its neighbors. Tatarstan provided a particularly interesting case of municipalities voting similarly and strongly in support for the main candidate. This tendency was previously highlighted by Moraski and Reisinger (2010) for earlier elections and was continued by this research. Our analysis also shows that strong support for the Kremlin was noticed in the republics surrounding Tatarstan. As for Tatarstan's neighbors, in general they voted differently and had a negative impact on the number of votes for the main candidate.

In this section we also want to discuss about reasons of high impact from Tatarstan on its neighbors. Although it can be still stated as an open question, we explain the impact from Tatarstan by several reasons. Firstly, from the literature review we found that there a lot of research confirming the sharing the same voting preferences of people from the same social environment and from the neighboring territories. Secondly, we estimated the economic success of Tatarstan in comparison with its neighbors and suggested that the strength of Tatarstan can be a factor of influencing on the neighbors. We confirmed thus suggestion by calculated contribution of spatial lags to variance of dependent variable.

Turning to the question of social and economic factors that affected the elections results, we found that the variables of the length of roads, locally produced goods and services, number of families with subsides had a significant negative impact on the voting for the main candidate in contrast to the opposition candidate, where the signs and significance of the coefficients before these variables are directly opposite.

Our future work on this topic will consider new tendencies of spatial influence on voting in Russian regions, using more advanced modeling techniques.

References

Abrams B. A., Butkiewicz J. L. (1995). The influence of state-level economic conditions on the 1992 U.S. presidential election. *Public Choice*, 85 (1–2), 1–10. DOI: 10.1007/BF01047898.

Aleskerov F., Borodin A., Kaspe S., Marshakov V., Salmin A. (2005). Analysis of electoral preferences in Russia in 1993–2003: Polarization index dynamics. *HSE Economic Journal*, 9 (2), 173–184 (in Russian).

Arzheimer K., Evans J. (2012). Geolocation and voting: Candidate-voter distance effects on party choice in the 2010 UK general election in England. *Political Geography*, 31 (5), 301–310. DOI: 10.1016/j.polgeo.2012.04.006.

Anselin L. (1995). Local indicators of spatial association — LISA. *Geographical Analysis*, 27 (2), 93–115. DOI: 10.1111/j.1538-4632.1995.tb00338.x.

Baum C., Schaffer M., Stillman S. (2007). Enhanced routines for instrumental variables/generalized method of moments estimation and testing. *The Stata Journal*, 7 (4), 465–506. DOI: 10.1177/1536867X0800700402.

Black D. (1948). On the rationale of group decision-making. *Journal of Political Economy*, 56 (1), 23–34. DOI: 10.1016/S0140-6736(02)71855-2.

Borodin A. (2005). Social conformity in the behavior of Russian voters. *HSE Economic Journal*, 9 (1), 74–81 (in Russian).

Boudreau C., Elmendorf C. S., MacKenzie S. A. (2019). Racial or spatial voting? The effects of candidate ethnicity and ethnic group endorsements in local elections. *American Journal of Political Science,* 63 (1), 5–20. DOI: 10.1111/ajps.12401.

Burbank M. J. (1997). Explaining contextual effects on vote choice. *Political Behavior*, 19 (2), 113–132. DOI: 10.1023/A:1024806024732.

Burnett J. W., Lacombe D. (2012). Accounting for spatial autocorrelation in the 2004 presidential popular vote: A reassessment of the evidence. *Review of Regional Studies*, 42 (1), 75–89. DOI: 10.1177/1091142106295768.

Clinton J. D. (2006). Representation in congress: Constituents and roll calls in the 106th house. *The Journal of Politics*, 68 (2), 397–409. DOI: 10.1111/j.1468-2508.2006.00415.x.

Coleman S. (2004). The effect of social conformity on collective voting behavior. *Political analysis,* 12 (1), 76–96. DOI: 10.1093/pan/mpg015.

Coleman S. (2007). *Popular delusions: How social conformity molds society and politics*. Cambria Press.

Coleman S. (2018). Voting and conformity: Russia, 1993–2016. *Mathematical Social Sciences*, 94, 87–95. DOI: 10.1016/j.mathsocsci.2017.10.005.

Cox K. R. (1968). Suburbia and voting behavior in the London metropolitan area. *Annals of the Association of American Geographers*, 58 (1), 111–127. DOI: 10.1111/j.1467-8306.1968.tb01639.x.

Cutts D., Webber D., Widdop P., Johnston R., Pattie C. (2014). With a little help from my neighbours: A spatial analysis of the impact of local campaigns at the 2010 British general election. *Electoral Studies,* 34, 216–231. DOI: 10.1016/j.electstud.2013.12.001.

Darmofal D. (2006). Spatial econometrics and political science. Society for Political Methodology Working Paper Archive.

Davis O. A., Hinich M. J., Ordeshook P. C. (1970). An expository development of a mathematical model of the electoral process. *The American Political Science Review*, 64 (2), 426–448. DOI: 10.2307/1953842.

Dow J. K. (2001). A comparative spatial analysis of majoritarian and proportional elections. *Electoral Studies*, 20 (1), 109–125. DOI: 10.1016/S0261-3794(99)00041-4.

Downs A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy,* 65 (2), 135–150. DOI: 10.1111/j.1467-6435.1986.tb01254.x.

Durlauf S. N. (2004). Neighborhood effects. In: *Handbook of regional and urban economics*, 2173–2242. Elsevier. DOI: 10.1016/B978-008044910-4.00480-6.

Elinder M. (2010). Local economies and general elections: The influence of municipal and regional economic conditions on voting in Sweden 1985–2002. *European Journal of Political Economy*, 26 (2), 279–292. DOI: 10.1016/j.ejpoleco.2010.01.003.

Enelow J. M., Hinich M. J. (1982). Ideology, issues, and the spatial theory of elections. *American Political Science Association*, 76 (3), 493–501. DOI: 10.2307/1963727.

Enelow J. M., Hinich M. J. (1989). A general probabilistic spatial theory of elections. *Public Choice,* 61 (2), 101–113. DOI: 10.1007/bf00115657.

Enikolopov R., Korovkin V., Petrova M., Sonin K., Zakharov A. (2013). Field experiment estimate of electoral fraud in Russian parliamentary elections. *Proceedings of the National Academy of Sciences*, 110 (2), 448–452. DOI: 10.1073/pnas.1206770110.

Erikson R. S., Romero D. W. (1990). Candidate equilibrium and the behavioral model of the vote. *The American Political Science Review*, 84 (4), 1103–1126. DOI: 10.2307/1963255.

Ethington P. J., McDaniel J. A. (2007). Political places and institutional spaces: The intersection of political science and political geography. *Annual Review of Political Science*, 10, 127–142. DOI: 10.1146/annurev.polisci.10.080505.100522.

Eulau H., Rothenberg L. (1986). Life space and social networks as political contexts. *Political Behavior*, 8 (2), 130–157. DOI: 10.1007/BF00987180.

Forbes J., Cook D., Hyndman R. J. (2020). Spatial modelling of the two-party preferred vote in Australian federal elections: 2001–2016. *Australian and New Zealand Journal of Statistics*, 62 (2), 168–185. DOI: 10.1111/anzs.12292.

Gerber A. S., Rogers T. (2004). Descriptive social norms and motivation to vote: Everybody's voting and so should you. *The Journal of Politics*, 71 (1), 178–191. DOI: 10.1017/S0022381608090117.

Górecki M. A., Marsh M. (2012). Not just "friends and neighbors": Canvassing, geographic proximity and voter choice. *European Journal of Political Research*, 51 (5), 563–582. DOI: 10.1111/j.1475-6765.2011.02008.x.

Huckfeldt R. (1986). *Politics in context: Assimilation and conflict in urban neighborhoods*. New York: Agathon Press.

Huckfeldt R., Sprague J. (1991). Discussant effects on vote choice: Intimacy, structure, and interdependence. *The Journal of Politics*, 53 (1), 122–158. DOI: 10.2307/2131724.

Jensen C. D., Lacombe D. J., McIntyre S. G. (2013). A Bayesian spatial econometric analysis of the 2010 UK General Election. *Papers in Regional Science*, 92 (3), 651–666. DOI: 10.1111/j.1435-5957.2012.00415.x.

Jessee S. A. (2009). Spatial voting in the 2004 presidential election. *American Political Science Review*, 103 (1), 59–81. DOI: 10.1017/S000305540909008X.

Johnson M., Shively W. P., Stein R. M. (2002). Contextual data and the study of elections and voting behavior: Connecting individuals to environments. *Electoral Studies*, 21 (2), 219–233. DOI: 10.1016/S0261-3794(01)00019-1.

Johnston R. (2004). Scale, factor analyses, and neighborhood effects. *Geographical Analysis*, 36 (4), 350–368. DOI: 10.1016/S0261-3794(01)00019-1.

Kelejian H. H., Prucha I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics*, 17 (1), 99–121. DOI: 10.1177/1536867x1301300201.

Kim J., Elliott E., Wang D. (2003). A spatial analysis of county-level outcomes in US presidential elections: 1988–2000. *Electoral Studies*, 22 (4), 741–761. DOI: 10.1016/S0261-3794(02)00008-2.

Kleibergen F., Paap R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133 (1), 97–126. DOI: 10.1016/j.jeconom.2005.02.011.

Kuletskaya L. E. (2021). Spatial modeling of voter choice: The survey of theoretical and empirical approach. *Spatial Economics*, 17 (2), 127–164 (in Russian). DOI: 10.14530/se.2021.2.127-164.

Lacombe D. J., Shaughnessy T. M. (2007). Accounting for spatial error correlation in the 2004 presidential popular vote. *Public Finance Review*, 35 (4), 480–499. DOI: 10.1177/1091142106295768.

Moraski B., Reisinger W. M. (2010). Spatial contagion in regional machine strength: Evidence from voting in Russia's federal elections. In: *APSA 2010 Annual Meeting Paper*. https://ssrn.com/abstract=1643282.

Nieuwbeerta P., Flap H. (2000). Crosscutting social circles and political choice. Effects of personal network composition on voting behavior in The Netherlands. *Social Networks*, 22 (4), 313–335. DOI: 10.1016/S0378-8733(00)00029-0.

Pattie C., Johnston R. (2000). "People who talk together vote together": An exploration of contextual effects in Great Britain. *Annals of the Association of American Geographers*, 90 (1), 41–66. DOI: 10.1111/0004-5608.00183.

Podkolzina E. A., Demidova O. A., Kuletskaya L. E. (2020). Spatial modeling of voting preferences in Russian Federation. *Spatial Economics*, 16 (2), 70–100 (in Russian). DOI: 10.14530/se.2020.2.070-100.

Poole K. T., Rosenthal H. (1984). U. S. Presidential elections 1968–80: A spatial analysis. *American Journal of Political Science*, 28 (2), 282–312. DOI: 10.2307/2110874.

Reisinger W. M., Moraski B. J. (2009). Regional voting in Russia's federal elections and changing regional deference to the Kremlin. In: *Annual National Conference of the Midwest Political Science Association, 67th (Chicago, Illinois)*. DOI: 10.17077/g02k-n8zq.

Skovoroda R., Lankina T. (2017). Fabricating votes for Putin: New tests of fraud and electoral manipulations from Russia. *Post-Soviet Affairs*, 33 (2), 100–123. DOI: 10.1080/1060586X.2016.1207988.

Stepan A. (2000). Russian federalism in comparative perspective. *Post-Soviet Affairs*, 16 (2), 133–176. DOI: 10.1080/1060586X.2000.10641484.

Stoetzer L. F. (2017). A matter of representation: Spatial voting and inconsistent policy preferences. *British Journal of Political Science*, 49 (3), 941–956. DOI: 10.1017/S0007123417000102.

Verbeek M. (2008). *A guide to modern econometrics*. John Wiley&Sons, Chichester.

Wuhs S., McLaughlin E. (2019). Explaining Germany's electoral geography: Evidence from the eastern states. *German Politics and Society*, 37 (1), 1–23. DOI: 10.3167/gps.2018.370101.

Received 22.06.2022; accepted 22.08.2022.

Appendix A

Table A1. Tatarstan's place among Russian regions according to Putin's (or Medvedev's in 2008) share of votes

Year	Rating
2004	11
2008	11
2012	8
2018	11

Source: Central Election Commission (http://www.cikrf.ru, http://www.vybory.izbirkom.ru).

Appendix B

Table B1. Results for different sets of control variables, specification (1)

	putin_share	sobchak_share	putin_share	sobchak_share
WYD_{T_T}	0.1155***	–0.0317	0.1158***	–0.0776
	(0.0173)	(0.1982)	(0.0172)	(0.1975)
$WYD_{T_{NT}}$	0.1443***	–0.3486***	0.1445***	–0.3974***
	(0.0123)	(0.0751)	(0.0119)	(0.0773)
WYD_{NT_T}	–0.0032	–0.0756	–0.0016	–0.0866
	(0.0092)	(0.0526)	(0.0093)	(0.0523)
roads_length	–0.4140	0.1131**	–0.5125*	0.1292***
	(0.2333)	(0.0349)	(0.2425)	(0.0356)
goods_services	–0.1576*	0.0169*	–0.1512*	0.0169*
	(0.0656)	(0.0068)	(0.0663)	(0.0068)
fam_subsid	–0.0646***	0.0009	–0.0645***	0.0013*
	(0.0174)	(0.0005)	(0.0168)	(0.0005)
houses	0.0163**	0.0011*	0.0154**	0.0011*
	(0.0053)	(0.0005)	(0.0054)	(0.0005)
nocommerc_fonds	0.0014	–0.0001	0.0016	–0.0001
	(0.0010)	(0.0001)	(0.0095)	(0.0001)
invest			0.0312	–0.0014
			(0.0206)	(0.0013)
expenses	0.0016	–0.0007**		
	(0.0037)	(0.0002)		
const	0.7599***	0.0097***	0.7622***	0.0082***
	(0.0107)	(0.0006)	(0.0076)	(0.0004)
N	326	326	326	326
R^2 centered	0.4425	0.2526	0.4456	0.2379
R^2 uncentered	0.9961	0.8760	0.9962	0.8736
Underidentification test (Kleibergen–Paap *rk* LM statistic)	76.537***	42.131	73.716***	37.731
Hansen *J*-statistic (overidentification test of all instruments)	40.061	43.075	36.937	43.986

Note. Specification (1) reflects formulas (3) and (4) for main and opposition candidate respectively.
Robust standard errors — in parentheses.
* — *p*-value < 0.05; ** — *p*-value < 0.01; *** — *p*-value < 0.001.