Multilevel modelling of invalid voting rates in the Bulgarian local elections 2015 and 2019

Boyko Amarov

March 29, 2021

Abstract

High levels of invalid votes in democratic elections can erode the public's trust in the fairness of the electoral process and can thus undermine the elected government's legitimacy. While the rate of invalid voting in Bulgarian parliamentary elections (about 5%) are similar to the ones observed in other post-communist countries, recent local elections (2015, 2019) exhibit much higher invalid vote rates, reaching 15% on average in 2019 and up to 60% in some polling stations in the same year. The aim of this paper is to study the variation of invalid vote rates with socio-eocnomic and demographic factors. We use geo-referenced polling station level data from the the 2015 and 2019 local elections combined with city and municipality-level census data on ethnic composition, demographic characteristics, employment and poverty measures as well as neighbourhood-level data on school performance (4th and 7th grade) as a proxy measure of voters' education and socio-economic status. We use a Bayesian multi-level logistic regression model to account for section-level heterogeneity and for spatial autocorrelation of invalid voting rates between sections that are geographically close to each other. Within the model we test the hypothesis whether polling stations in highly contested municipalities have higher rates of invalid voting, which could indicate election fraud. Preliminary results indicate systematic variation of invalid vote rates with respect to the ethnic composition of cities and and the education proxy.