Oskari Harjunen, Tuukka Saarimaa
and Janne Tukiainen

Love Thy (Elected) Neighbor? Residential Segregation, Political Representation and Local Public Goods

Aboa Centre for Economics
Discussion paper No. 138
Turku
September 2021 (previous draft January 2021)

The Aboa Centre for Economics is a joint initiative of the economics departments of the University of Turku and Åbo Akademi University.

ISSN 1796-3133

Printed in Uniprint
Turku
September 2021 (previous draft January 2021)

Oskari Harjunen, Tuukka Saarimaa and Janne Tukiainen

Love Thy (Elected) Neighbor? Residential Segregation, Political Representation and Local Public Goods

Aboa Centre for Economics
Discussion paper No. 138
September 2021 (previous draft January 2021)

ABSTRACT

We study geographic political representation and geographic distribution of local public goods *within* local jurisdictions using geo-coded data on politicians, the electorate and elementary schools. Descriptive analysis reveals that poorer neighborhoods are under-represented and that local politicians have a strong support base in their home neighbourhoods. Based on randomized election outcomes due to personal vote count ties, geographic representation has a causal effect on school closures. The probability of closure is cut in half when a candidate living close to the school is randomly elected. High-income residents react to closures by moving away from the neighborhood, thus reinforcing segregation.

JEL Classification: C21, D72, H75, R23

Keywords: Geographic representation, random elections, residential segregation, school closure

Contact information

Oskari Harjunen
VATT Institute for Economic Research
Arkadiankatu 7, Helsinki, FI-00101
Email: oskari.harjunen (at) vatt.fi

Tuukka Saarimaa (corresponding author)
Aalto University School of Business and School of Engineering and Helsinki Graduate School of Economics
Ekonominaukio 1, Espoo, FI-02150
Email: tuukka.saarimaa (at) aalto.fi

Janne Tukiainen
University of Turku, and VATT Institute for Economic Research
Rehtorinpellonkatu 3, FI-20014 University of Turku
Email: janne.tukiainen (at) utu.fi.

Acknowledgements

We thank Toke Aidt, Benoit Crutzen, Paolo Fornaro, Essi Eerola, Brett Fischer, Olle Folke, Elias Oikarinen, Markku Siikanen, Stephane Wolton and seminar participants at Rotterdam, Cambridge & Essex, LSE, FEAAM 2020, Helsinki and EMUEA 2021 for useful comments.

1 Introduction

There is a growing body of evidence showing the harmful effects of residential segregation. For example, a number of recent papers have documented that childhood neighborhood context has a causal effect on various adulthood outcomes (see e.g. Damm and Dustmann 2014, Chetty et al. 2016, Chyn 2018, Chyn and Haggag 2019, Nakamura et al. 2021, Deutscher 2020 and Laliberté 2021) and intergenerational mobility (Chetty and Hendren 2018 and Chetty et al. 2020). By residing in a particular neighborhood, people choose not only their neighbors and peers, but also the local public goods they are able to consume. Thus, the political decisions of where to locate public goods are crucially important in shaping the contextual environment of residents in different neighborhoods. If public goods are not provided equally to all neighborhoods, we should ask why this is so, and in particular, how residential segregation and local political forces are intertwined in shaping this inequality.

This paper concentrates on the political underpinnings of the provision of local public goods across neighborhoods. More precisely, we study the link between geographic political representation and geographic distribution of public goods *within* local jurisdictions using Finnish data spanning three municipal council terms from 2005 to 2017. The analysis is facilitated by detailed geo-coded data on the residential location of all municipal election candidates, the electorate and the location of local public goods. The electoral context for our analysis is an at-large open-list proportional representation (PR) system.[1] As opposed to a ward system, when the elections are held at-large voters can vote for any candidate in the municipality, and thus, there is no guarantee that all neighborhoods are equally represented or that politicians have incentives to cater to local voters.

We start our analysis by describing the geographic representativeness of municipal councils by comparing the residential distribution of local politicians and the electorate across small neighborhoods. This descriptive analysis reveals that neighborhoods with poorer and less-educated electorates are under-represented relative to their share of municipal population, both at the extensive and intensive margins. We also document that candidates get systematically more votes from the polling districts they themselves live in compared to other polling districts and that the additional amount of votes increases

[1] According to the Database of Political Institutions (Scartascini et al. 2018), 94 out of 147 democracies employ PR, and about fourth of these use open-list procedures.

with the length of the candidates' residential spell in the district and with incumbency status.[2] This means that local politicians have a strong local support base, which should incentive them to cater to voters in their neighborhoods even though the elections are held at-large.

In the second part of the analysis, we ask whether the unequal geographic representation that we document translates into unequal geographic distribution of local public goods. We focus our analysis on closures of elementary schools as they have a well-defined geographic location and are a prime example of a local service where proximity is an important factor as young children need to attend the school on a daily basis. Moreover, the number of elementary schools in Finland diminished by a third during the time period of our analysis providing us with spatial and temporal variation in local service availability.[3] To make sure we can interpret our results as causal, we use election outcomes decided by a lottery, which take place when there is a tie in personal votes within a party for the last seat of the party list. We find that randomly increasing the representation of a local school in the municipal council, which makes decisions on school closures, decreases the probability of school closure during the election term. The effect is large as in our lottery sample the probability of closing down a school during the election term roughly halves from about 20% to 10% when a candidate close to the school is randomly elected (as opposed to a candidate from the same party near another school). We also show that the effect does not depend on whether the candidates themselves have school-aged children suggesting that political motives, rather than their personal needs, are at play.

Finally, we examine what happens to the neighborhoods when a local school is closed. Using difference-in-differences (DID) methods, we find evidence of neighborhood resorting. High-income residents of the neighborhoods experiencing a closure of a relatively large school "vote with their feet" by moving away from these neighborhoods, thus reinforcing residential segregation.

We contribute to several strands of literature both in economics and political science. First, we examine political representation and its effects from a new angle of geographic

[2]Municipalities are divided into polling districts. The polling districts are only used to allocate voters to polling stations and for vote counting purposes.

[3]Similar waves of school closures have taken place, for example, across the US (Brummet 2014), Denmark (Beuchert et al. 2018), the Netherlands (De Haan et al. 2016) and Sweden (Taghizadeh 2020).

representation at a very local level.[4] This is important not only because of the possible effects on the distribution of public goods and funds more generally, but also because interacting with one's neighbors may affect politicians' beliefs and preferences. To the extent that politicians are exposed to systematically more affluent and educated neighbors than the electorate in general as we document, their perceptions about public opinion may be biased (e.g. Broockman and Skovron 2018 and Enos 2017). Descriptive evidence on the within-jurisdiction geographic representation is still very limited as most candidate-level datasets analyzed in the literature lack the geographic detail of our data.

Related to this, our results are consistent with the idea of friends-and-neighbors voting whereby politicians have a core support base in their neighborhood and long-standing neighborhood ties are important for cultivating personal votes. We document this phenomenon in elections held at-large where voters can cast votes to individual candidates to express their preferences for locals due to open-lists (see also Carey and Shugart 1995, Shugart et al. 2005, Saarimaa and Tukiainen 2016 and Campbell et al. 2019).

Second, our results contribute to the literature on distributive politics. Several studies show convincingly that representation in a legislative body matters for the geographic distribution of centralized spending (e.g., Ansolabehere et al. 2002, Knight 2008, Dragu and Rodden 2011, Brollo and Nannicini 2012, Hodler and Raschky 2014 and Fiva and Halse 2016). However, as these studies focus on national or regional level spending, we still have an incomplete understanding of how distributive politics operate within local jurisdictions. This is a major gap in our knowledge as many important tasks have been delegated to local governments worldwide.[5] For example in Finland, municipalities are responsible for elementary schooling, primary health care, land use and zoning policies, public transportation and other such policies that influence the daily lives of the citizens. To our knowledge, we are the first to identify a causal link running from geographic representation to policy outcomes within local jurisdictions.[6] Taken as a whole, our

[4]A large body of empirical work has established a causal link between legislative representation of various groups and policies preferred by those groups (see e.g. Pande 2003 on minorities, Chattopadhyay and Duflo 2004 on gender or Hyytinen et al. 2018a on occupational background.)

[5]Cox (2009) highlights this mismatch between theoretical (e.g., Dixit and Londregan 1996) and empirical work on distributive politics. Specifically, district level studies are not informative about how resources are distributed across different groups of voters.

[6]The contemporaneously written study by Folke et al. (2021) shows using Swedish data that given the overall level of representation in a neighborhood, those neighborhoods with more representation from

findings suggest that the models of distributive politics are relevant also at the very local level.

Furthermore, prior research on geographic representation and spending mostly concerns ward-based systems where politicians have obvious electoral incentives to cater to their local ward. It is less clear that these incentives exist in at-large PR systems (e.g., Carey and Shugart 1995, Trebbi et al. 2008, Trounstine 2010, Abott and Magazinnik 2020 and Fiva et al. 2020). Our results indicate that also in at-large election systems with personal votes, such as the Finnish open-list PR system, geographic representation and public good provision can be linked.

Finally, our results speak to the literature on residential sorting and neighborhood effects. Recently this literature has focused on how a neighborhood's private amenities respond endogenously to its socio-economic makeup and how this reinforces residential sorting (e.g., Diamond 2016, Couture and Handbury 2020, Couture et al. 2020, Almagro and Domınguez-Iino 2020 and Su 2021). We highlight how residential sorting together with the local political system leads to inequality in representation and in publicly provided amenities across neighborhoods and how this may also reinforce residential sorting. Moreover, the link between residential segregation and the provision and quality of public goods across neighborhoods may be an important mechanism behind neighborhood effects, as suggested recently by Fogli and Guerrieri (2019) and Laliberté (2021).

The paper proceeds as follows. In the next section, we introduce the Finnish local election system, the school system and other municipal tasks and finances. In Section 3, we describe the data. In Section 4, we analyze the geographic representation, the effects of representation on school closures, and the effects of school closures on residential segregation. Finally, we conclude and discuss policy implications and interesting avenues for future research.

the majority bloc have fewer building permits for multifamily homes approved, and fewer proposals to close schools made. Beach et al. (2019) and Carozzi and Repetto (2019) are also closely related as they analyze local level close elections in the US and Spain, respectively. However, neither of these studies use information on the residential location of the politicians themselves, and thus, do not link geographic representation and policy outcomes as we do.

2 Institutional background

2.1 Municipal elections

Finland has a two-tier system of government consisting of the central government and municipalities. Municipal councils are the main seat of power in the municipal decision-making and the councils make the decisions on school closures. Municipal elections are held simultaneously in all municipalities and the council term lasts for four years. Our data span three council terms: 2005-2008 (elections held in October 2004), 2009-2012 (elections held in October 2008) and 2013-2017 (elections held in October 2012).[7] During our analysis period, the council size was a step function of population and varied between 13 and 85, the median being 27. At the same time, median municipality size was 7262 and ranged from 773 to 603,968. Each municipality has only one electoral district and no geographic quotas are in place.

The seat allocation is based on proportional representation (PR), using the open-list D'Hondt election rule. The elections are held at-large so that voters can vote for any candidate in the municipality. In the elections, each voter casts a single vote to a single individual candidate and they cannot vote for a party without specifying a candidate. Moreover, the list order in the ballot is alphabetical so parties cannot use list order to signal their preferences. These rules mean that voters (as opposed to parties) decide which candidates are eventually elected from a given list, because the number of votes that a candidate gets determines the candidate's rank on her party's list. The total number of votes over the candidates of a given party list determines the votes for each party and this determines how many seats each party gets.

An important feature of this election system is that in many cases, there is an exact tie in the number of votes at the margin where the last available seat (or seats) for a given party list is allocated. For example, it is possible that a party gets k seats in the council and that the k^{th} and $(k+1)^{th}$ ranked candidates of the party receive exactly the same number of votes. The Finnish law dictates that in this case, the winner of the marginal seat has to be decided randomly. Typically, the seat is allocated by drawing a ticket (name) from a hat. We make use of these randomly decided election outcomes in

[7] From 2017 onward the start of council term was moved forward from January to June. The 2017 elections were held in April 2017.

our empirical analysis.[8]

2.2 Elementary school system

Municipalities are responsible for elementary education, which consists of a nine-year compulsory school starting in the year the child turns seven. Almost all elementary schools are public schools and they are free of charge. In most municipalities, school intake is catchment area based so that each address in a municipality is assigned to a catchment area of at least one school and children living within the catchment area of a particular school are guaranteed a place in that school. With this institutional setup, most children attend the school closest to where they live.

The number of elementary schools has declined and their mean size has increased substantially in Finland during the time period of our analysis as can be seen from the left panel of Figure 1. The closures are related to declining enrollment, mostly due to migration from rural to urban areas. The migration has been driven especially by younger households so that the number of school-aged children has declined in many municipalities and remote neighborhoods. At the same time, school closures have taken place all over Finland as is evident from the right panel of Figure 1. In more urban areas, school closures have taken place in search for economies of scale.

[8] See Hyytinen et al. (2018b) for more details.

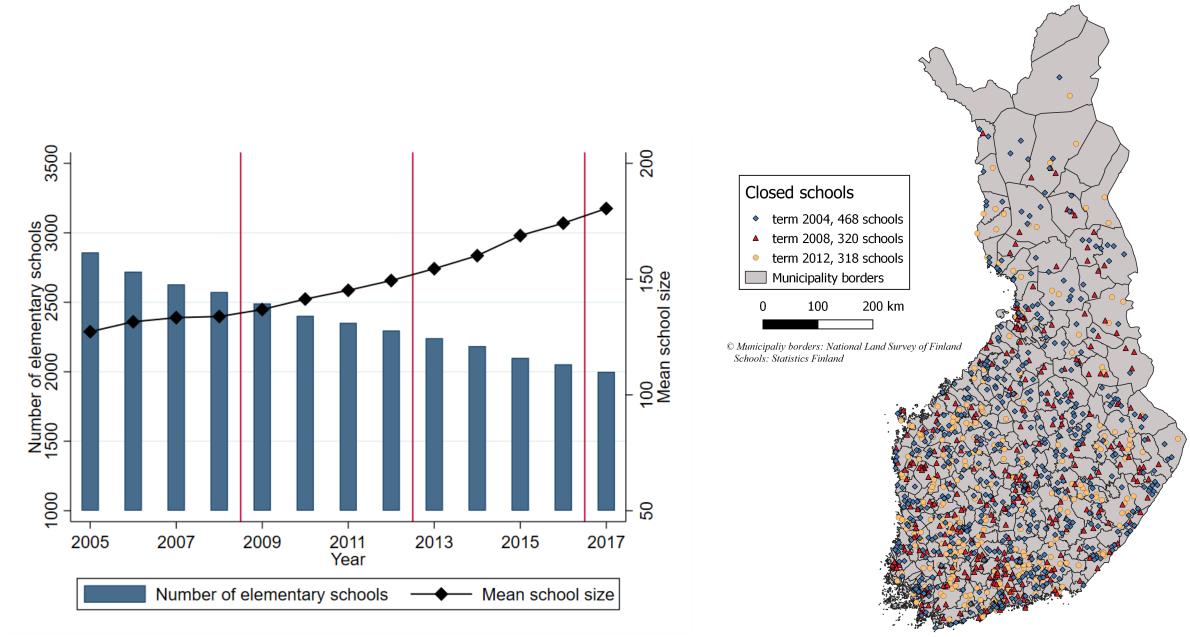

Figure 1: Number of elementary schools and geography of school closures.

Notes: On the left panel the blue bars depict the number of elementary schools during each year (left axis), the black line the mean size of schools (right axis) and the red vertical lines the election timing. The right panel depicts a map of school closures by election term.

2.3 Municipal task and finances

Finnish municipalities are multi-task jurisdictions. In addition to elementary schools, they are responsible for providing health and social care and other usual local public goods, such as public transport and waste management.

The most important revenue source is the flat municipal income tax which the municipalities can set freely. There is also a property tax, but importantly property tax revenue is not ear-marked for financing schools or in other ways to benefit the neighborhoods from which the taxes are collected, as is the case with US school districts. A central government grant system, consisting of 20% of total revenue, is used to equalize local cost and revenue disparities.

3 Data

3.1 Data sources

Our data come from several sources. Each data set comes with coordinate information so that we are able to combine them using GIS tools.

Candidates. The first source is the election data provided by the Ministry of Justice containing candidate-level information on the candidates' age, gender, party affiliation, the number of votes they received, their election outcomes, and incumbency status. We have linked these data to Statistics Finland's data on the candidates' education, occupation, and socioeconomic status. Finally, we obtained addresses and coordinates for the candidates' residential buildings from the Digital and Population Data Services Agency.[9] Thus, we have candidate-level election data with a rich set of background characteristics and residential location for elections held from 2004 to 2012.

Most municipalities are divided into multiple polling districts and our election data also contain information on the number of votes received by the candidate from each of the polling districts in the municipality. Each polling district has a specific polling station where its residents go to vote. Importantly, polling districts have no other role in the elections as they are held at-large. Together with information in which polling district candidates live in, we can examine whether candidates receive more votes from their home polling district compared to other districts.[10]

Candidates' neighborhoods. Our data on the characteristics of the electorate in candidates' neighborhoods come from Statistics Finland's Grid and Zipcode Databases. These geo-referenced data contain information on age, education and income of the residents in 250 m x 250 m sized grids and zipcodes covering the whole of Finland. In the analysis, we will aggregate the grids to square kilometer level. The upshot of the grid level data is its spatial precision allowing us to focus on small neighborhoods, but due to confidentiality reasons grids with less than ten persons over the age of 18 do not contain socio-economic information. This means that sparsely populated areas are left out when using this data, whereas with zipcode level data the geographic coverage is

[9] We could not match coordinate data for 668 out of the 115,540 candidates in these three elections .

[10] We can match a candidate's residence to a polling district because each building has a polling district code, which matches the code in the election results data. Unfortunately, this information is available only for candidates in the 2012 elections.

comprehensive.

Unfortunately, we are unable to perfectly match the polling districts to these neighborhood data as we do not have geo-referenced data on polling districts. However, since we know the zipcode and the home polling district for all candidates, we can approximate socio-economic make-up of the polling districts by averaging zipcode information within each polling district. This of course results in some measurement error.[11]

Elementary schools. Our school data come from the Register of Educational Institutions maintained by Statistics Finland. The data cover all elementary schools and include annual information on the number pupils in different grades, coordinates for the school buildings and the year the school was closed. We use this data from 2004 to 2017 in our analysis.

Since we have coordinate information for all schools and candidates, we can assign the closest school for each candidate at the time of the election. With this distance based approach, we assume that the candidates represent the school that is closest to them.[12] Figure A2 in the Online Appendix illustrates our strategy for one election in one municipality with three schools.

4 Results

4.1 Geographic representation and local support base

We start by examining geographic representation within municipalities. We first ask, whether the extent of representation is associated with the socio-economic makeup of the neighborhood. In this analysis, a neighborhood refers to a square kilometer sized grid or a zipcode. The grids are much smaller units than the zipcodes with mean populations of 66 and 1830, respectively. After that we ask, whether candidates receive more votes from their own neighborhoods compared to other neighborhoods. This analysis sheds light on whether candidates have a strong local support base, and thus, electoral incentives to cater to the voters in their neighborhood. In this analysis, a neighborhood refers to a

[11] Polling districts are typically somewhat larger geographically than zipcodes. Figure A1 in the Online Appendix illustrates this for one municipality we were able to obtain polling districts in GIS format.

[12] An alternative approach, would be to use the schools' catchment areas. However, we do not have data on the catchment areas for all municipalities for our time period.

polling district, which is the smallest geographic unit we have in the election data. In most municipalities, polling districts are geographically similar in size to zipcodes. In this analysis we use data from the 2012 elections.

To answer the first question, we run simple OLS regressions where we regress measures of neighborhood level representation on the income and education levels of the neighborhood's residents:

$$Representation_{im} = \alpha * Popshare_{im} + \beta * Sosecon_{im} + \gamma_m + u_{im}, \quad (1)$$

where $Rerepsentation_{im}$ refers to either the prevalence of candidates or councilors in neighborhood i (grid or zipcode) in municipality m. We use four measures: the candidate and councilor shares in the neighborhood, whether the neighborhood has any councilors and the election rate from the neighborhood, i.e. the share of the candidates running from the neighborhood that get elected. We control for the neighborhood's share of municipal population ($Popshare_{im}$) and municipality fixed effects (γ_m). $Sosecon_{im}$ is either the income or education level of the neighborhood's residents.

The results are presented in Table 1 where we report 16 separate regression results in total.[13] Panel A reports the results for the grids and Panel B for the zipcodes. We highlight three findings. First, there seems to be no robust pattern with respect to neighborhood candidate share (models (1),(5), (9) and (13)). Second, the neighborhoods with higher average incomes and better educated electorates have more representatives both in the extensive and intensive margins. This is true at both neighborhood scales. Finally, candidates' electoral success is better in these higher income and better educated neighborhoods. The associations are also quantitatively substantial. For example, at the grid level a one standard deviation increase in average income is associated with roughly 5.2%-point increase (from 40%) in the propensity to have at least one councilor.

The key takeaway from Table 1 is that neighborhoods with more affluent and better educated electorates are more strongly represented in municipal councils. It also seems that the geographic differences in candidacy are not the main explanation for these differences in representation, but rather how candidates from these types of neighborhoods perform in the elections.

[13]Mean income and education levels of a neighborhood are highly correlated, which is why we do not include them in the same regressions. The interpretation of the models with both measures would also be somewhat problematic.

Table 1: Political representation and socio-economic structure of neighborhoods.

Outcome:	Candidate share	Councilors (0/1)	Councilor share	Councilors/ candidates
	Panel A: Square km grid			
	(1)	(2)	(3)	(4)
Mean income	0.0001	0.052***	0.001***	0.018***
	(0.0001)	(0.016)	(0.0003)	(0.007)
	(5)	(6)	(7)	(8)
Share highly-educated	0.010**	1.054***	0.036***	0.417***
	(0.003)	(0.103)	(0.005)	(0.098)
Outcome mean	0.021	0.401	0.019	0.238
N	5733	5733	5733	4172
	Panel B: Zipcode			
	(9)	(10)	(11)	(12)
Mean income	-0.002**	0.048***	0.003**	0.012*
	(0.001)	(0.015)	(0.001)	(0.006)
	(13)	(14)	(15)	(16)
Share highly-educated	0.008	1.670***	0.062**	0.068
	(0.013)	(0.342)	(0.019)	(0.162)
Outcome mean	0.108	0.791	0.108	0.307
N	2653	2653	2653	2653
Municipality FE	yes	yes	yes	yes

Notes: The table presents results from regressions where the unit of observation is either a square kilometer grid or a zipcode. Grids with less than ten persons over 18 are omitted due to data confidentiality. All the models include the neighborhood's share of municipal population as a control. Standard errors are presented in parentheses and are clustered at the municipality level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

These results can be driven by a number of different mechanisms related to voter or candidate behavior and the mechanisms are not mutually exclusive. Pinpointing the exact mechanism is beyond the scope of this paper and our data, but turnout is one likely candidate. As it is well documented that turnout is higher among high-income and highly-educated voters (e.g. Lindgren et al. 2019, Lahtinen et al. 2019, Akee et al. 2020 and Cantoni and Pons 2020), neighborhoods with more affluent and educated residents may have higher turnout. If voters vote for local candidates at roughly similar rates, the results we observe could be driven by neighborhood differences in turnout. In Table A1 in the Online Appendix, we present results from polling district level regressions where we regress the district's turnout on mean income and education level of the district while controlling for municipality fixed effects. Turnout is clearly higher in the polling districts with higher income and more educated voters. These correlations are rather strong given that we are measuring the socio-economic attributes of the polling district with some error as explained earlier. These result support the hypothesis of turnout differences being important, but of course we cannot rule out other important mechanisms.

Next we turn to the question of candidates' local support base and their electoral incentives to cater to local voters, as opposed to all voters in the municipality. The election results data at the candidate-polling district level allow us to compute for all candidates in the 2012 elections the share of votes that the candidate received out of all the votes given in a particular polling district. Using this vote share, we estimate the following regression model:

$$\begin{aligned} Votesha_{ip} = \beta_0 Own_p + \beta_1 Own_p * Residence_i + \beta_2 Own_p * Incumbent_i \\ + \beta_3 PartyShare_i + \beta_4 Similarity_i + \gamma_i + \delta_p + u_{ip}. \quad (2) \end{aligned}$$

That is, we regress candidate i's vote share out of all the votes given in polling district p ($Votesha_{ip}$) on a dummy variable indicating the polling district where the candidate herself lives (Own). In addition, we interact this dummy with the residence spell (in years) in the polling district ($Residence$) and with incumbency status ($Incumbent$). The first serves as a proxy for the strength of local ties, whereas the latter proxies overall candidate quality and recognizability. We rely on within candidate variation across the polling districts by including candidate fixed effects (γ_i) and we also control for polling district fixed effects (δ_p).

Furthermore, we add two variables to control for potential confounders related to voter segregation across neighborhoods. $PartyShare$ measures the vote share of the candidate's own party in the polling district, which controls for residential segregation with respect to party affiliation.[14] $Similarity$ is a dummy variable that equals one if the candidate is highly-educated and if the share of highly-educated in the candidate's own district

[14] As an extreme example, consider a municipality with only two districts and two parties. If the voters of the parties are perfectly segregated into different districts, we would not be able to separate the own

is above the median of all districts. In other words, it equals one if a highly-educated candidate lives in a relatively highly-educated district.

The results, presented in Table 2, indicate that candidates' vote share is roughly 1.5%-points higher in their own district compared to other districts in the municipality. Given that the mean vote share from all districts is on average 0.3%, the own district vote premium is indeed substantial. As the regressions include candidate fixed effects, we are perfectly controlling candidate quality as well as their overall campaigning ability and effort. However, these fixed effects do not capture possible campaigning efforts that are targeted towards their own neighborhoods.

In column (2), we add the party and voter similarity controls. Both of the measures have a positive sign and are statistically highly significant, but the own district vote premium is virtually unaffected by these additional controls of voter segregation. This supports the interpretation that the own district vote premium is indeed related to voter preferences for local representation and is not a result of similar candidates and voters sorting into same neighborhoods.

In column (3), we add the two interaction terms. The results indicate that the own district vote premium is higher both for candidates who have resided in the district longer and for incumbents. One explanation for these results could be that candidates with longer residence spells and incumbents are more easily recognizable as locals. Moreover, a longer residence spell may signal to voters that a candidate is more likely to stay in the neighborhood and continue to promote the interests of the residents. The findings are also in line with the notion that long-standing neighborhood ties are important for cultivating personal votes (e.g., Shugart et al. 2005 and Jankowski 2016).[15] They are less in line with targeted campaigning being the reason behind the vote premium.

Taken as a whole, the results in this section show that the poorer and less-educated neighborhoods are left behind in geographic representation. Furthermore, the results

district vote premium from partisanship as candidates would only get votes from their home district. E.g. Brown and Enos (2021) report substantial segregation of voters with respect to party affiliation in the US.

[15]In Table A2 in the Online Appendix, we show that these results are not an artifact of the number of polling districts in the municipality. We divide the data into two samples based on the median number of polling district in all municipalities. The results in both samples are qualitatively similar, although there are some quantitative differences.

indicate that candidates enjoy a large local vote premium, and thus, should have electoral incentives to cater to their local support base.

Table 2: Candidates' vote shares from own polling district.

	(1)	(2)	(3)	(4)
Own district	1.554***	1.522***	0.574***	0.552***
	(0.098)	(0.096)	(0.052)	(0.051)
Own district ×			0.048***	0.047***
residence spell			(0.003)	(0.003)
Own district ×			2.258***	2.239***
incumbent			(0.141)	(0.141)
Own party vote share		0.015***		0.015***
		(0.002)		(0.001)
Voter similarity		0.032***		0.032***
		(0.008)		(0.007)
Outcome mean	0.299	0.299	0.299	0.299
N	679,690	679,690	679,690	679,690
R-squared	0.41	0.41	0.45	0.46
Candidate FE	yes	yes	yes	yes
Polling district FE	yes	yes	yes	yes

Notes: The table presents results from candidate-polling district level regressions where the dependent variable is the candidate's vote share (%) of the polling district. Only municipalities with more than one polling district are included. Columns (1) and (2) include all candidates from these municipalities. In columns (3) and (4), the data is divided into two samples based on the median number of polling district in all municipalities (4). Standard errors are presented in parentheses and are clustered at the municipality level. $^{*}\, p < 0.10$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

4.2 Effects of representation on school closure

We have shown that political representation is unequal in geographic terms and that candidates have a strong support base in their neighborhood. In this section, we examine whether this matters for policy outcomes, i.e. whether there is a causal channel running from geographic political representation to geographic policy outcomes. To achieve this, we ask whether electing a given candidate has an effect on the probability that the elementary school closest to that candidate is closed during the council term.

Our main model specification can be written as:

$$Closure_{it} = \alpha + \delta * Elected_{it} + \beta' X_{it} + u_{it}, \tag{3}$$

where $Closure_{it}$ is a dummy variable that equals one if the school closest to candidate i was closed during council term t. The dummy variable $Elected_{it}$ is equal to one if candidate i was elected into the municipal council for election term t. X_{it} include school and candidate level control variables and council term fixed effects, and u_{it} is the error term.

Our interest lies on the parameter δ, which measures the effect of increasing the representation of the school in question by one councilor on the probability that the school is closed down. A simple OLS estimation of equation (3) would likely suffer from omitted variable bias as many other factors besides political representation affect school closures. Many of these factors are unobservable to us and likely correlated with the election status as the same economic, demographic and social factors influence both the demand for schools in neighborhoods and candidate selection and voting decisions.

To estimate δ consistently we resort to randomly assigned election outcomes by including in our analysis only the candidates whose election status was decided by a lottery. This makes sure that *Elected* was randomly assigned. Concentrating on within party vote ties also ensures that the party composition of the council remains the same regardless of which candidate is elected. Moreover, we omit lotteries where all the involved candidates have the same closest school as these lotteries do not provide useful identifying variation (results are robust to including them).[16]

In order to increase precision and to examine the robustness the results, we create an additional sample by including the candidates whose election status was decided by a margin of a single vote. This sample is constructed in the following way. For each party list, we define the pivotal number of votes as the average of the maximum number of votes among the non-elected candidates and the minimum number of votes among the elected candidates. The distance to getting elected for each candidate is the number of votes of the candidate minus the pivotal number of votes of her party list. We include

[16] We include all municipalities with at least two schools in the estimations. In the elections between 2004 and 2012, 419 candidates had an eligible tie (different schools) within their party list for the last seat(s).

in this alternative sample those candidates whose distance was less or equal to one. As long as these candidates were not able to precisely manipulate the number of votes they got and the within-party election threshold, the election status for these candidates is as-good-as-random. This increases our sample size to 1540.[17] Also for this larger sample, we use the same eligibility criteria that the candidates involved in the close elections and ties are assigned to different schools.

Balance tests reported in Table A3 in the Online Appendix verify that both the candidate and school level covariates are balanced across the control and treatment groups in both the lottery sample and the sample that combines the lottery and the one vote margin candidates. Importantly, also the distance to the closest school is balanced across the groups suggesting that the groups are comparable in geographic terms. The last two variables in Table A3 show that the treatment group has more representation both at the intensive and extensive margins than the control group. These also reveal that the average school seat share is roughly 20% and that almost all schools have at least some representation. This means that even though the candidates in these samples get the very last seats in their party, they typically increase the already sizable seat share of the schools, which may be quite important in terms of within council coalition formation and bargaining related to school closures (see Baron and Ferejohn 1989).

In Figure A3 and Table A4 in the Online Appendix, we compare the municipalities with close elections (lotteries and one vote margin) to all municipalities. Figure A3 shows that close elections have taken place all over Finland in almost all municipalities. Table A4 further illustrates that municipalities with and without close elections are indeed very similar on average.[18]

Table 3 reports our main causal results from OLS regressions using the lottery sample (Panel A) and the one vote margin sample (Panel B). According to the results using the lottery sample, there is a clear and rather large effect of representation on school closure. Column (1) presents the most parsimonious model for the lottery sample, from which we see that, on average, 20% of the schools of the lottery losers are closed down, whereas for the winners school closures happen only half as frequently. Moving to the right of Panel

[17] We do not make use of a regression discontinuity design, because we would need to omit the lottery sample as it constitutes a mass point exactly at the treatment cutoff.

[18] The reason why the occurrence of close elections is not correlated with municipal population is due to larger number of council seats and parties, and larger party lists in the larger municipalities.

A, we see that adding school and candidate level controls and council term fixed effects has virtually no effect on the point estimates, which is consistent with the balance tests reported in Table A3.

In Panel B, we present the results including the candidates where the last seat(s) was decided by at most a one vote margin. In this case, the effect of representation is slightly smaller. However, the differences in the point estimates in Panel A and B are not statistically significantly different from each other at conventional significance levels. Moreover, the baseline closure rate for the close elections losers is lower in this sample (column (5)) making the relative effect more comparable to the lottery sample. Again the point estimates remain the same after adding controls and council term fixed effects.

Table 3: Effect of representation on school closure.

Panel A: Lottery	(1)	(2)	(3)	(4)
Constant	0.204***	0.282***	0.329***	0.305***
	(0.030)	(0.038)	(0.100)	(0.099)
Elected	-0.108***	-0.101***	-0.099***	-0.097***
	(0.035)	(0.035)	(0.035)	(0.035)
N	419	419	419	419
R-squared	0.023	0.064	0.096	0.108
Panel B: One vote margin	(5)	(6)	(7)	(8)
Constant	0.186***	0.260***	0.207***	0.213***
	(0.016)	(0.022)	(0.053)	(0.054)
Elected	-0.062***	-0.061***	-0.065***	-0.064***
	(0.020)	(0.020)	(0.020)	(0.020)
N	1540	1540	1540	1540
R-squared	0.007	0.045	0.063	0.064
P-value for effect difference	0.122	0.172	0.247	0.258
School controls	No	Yes	Yes	Yes
Candidate controls	No	No	Yes	Yes
Election term FE	No	No	No	Yes

Notes: The table presents results from linear probability models where the outcome is school closure (0/1). The school controls include the number of pupils in school. The candidate controls include age, sex, children in the family or not, incumbency and occupation status. The p-values are for the test of the statistical significance of the difference of the coefficients for Elected in the two samples with the same model specification. Standard errors are presented in parentheses and are clustered at the municipality level. $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

We have explored the robustness of our results by concentrating on candidates who do not move during the council term and by concentrating on only on those municipalities that had school closures. First, roughly 20% of the candidates moved during the council term. We re-estimated the models using only those candidates who did not move during the council term (non-movers). These additional results are reported in Table A5 in the Online Appendix and they are very similar to those presented in Table 3. Second, in Table A6 in the Online Appendix, we report results of models using only those municipalities that had school closures. In this sample of municipalities, the overall closure rate is higher, but in relative terms the magnitude of the effect is similar compared to the main results.

The results so far provide clear revealed preference evidence that local politicians want to prevent school closures in their neighborhoods. To understand what motivates the candidates and how they can achieve their objectives, we study the heterogeneity of the effect in various dimensions. This allows us also to evaluate whether the effects are larger in circumstances we would expect them to be. These results are presented in Tables A7–A10 in the Online Appendix. In each table, we split the sample with respect to the variable of interest and report the effect and control group mean of the outcome for each sample. It is important to compare the effect to the outcome mean because the school closures rates vary substantially depending on the sample, and thus, simply comparing point estimates would be misleading.

We start by asking how a single candidate can influence policy. We estimated separate models for municipalities with a council size below and above the sample median. According to Table A7, the effect is larger relative to outcome mean in the sample of municipalities with small councils and the effect is statistically significant only in this sample. This result could be driven by the fact that in smaller councils an individual councilor (or her party) is more likely to be pivotal in the council decision-making. Moreover, in smaller councils, a single councilor can be more persuasive in informal within-council discussions. Unfortunately, without more information on what goes on in the council and other informal meetings, it is difficult to know conclusively what is driving these differences.

We hypothesize that the electoral incentives and other forms of accountability are related to two aspects: how important the school is to the neighborhood's residents and how many residents or voters are affected by the potential school closure. We use school size as a proxy measure for both aspects. The absolute size of the school (number of pupils) directly measures how many citizens are affected, and thus, may lobby the candidate to preserve the school. Moreover, it serves as a proxy for school quality.[19] We

[19] As some of the schools we are analyzing schools that are quite small, it could be, for example, that in these schools different age cohorts have to share a classroom and the teacher. Larger schools are also able to provide a larger variety of subjects. Of course, this only a proxy measure and some parents may prefer smaller schools.

use the relative size of the school (school's share of all pupils in the municipality) as a proxy for the share of voters who are affected by the closure of the school. We divide our sample based on the median of absolute and relative school size and estimating separate models for these samples. According to Table A8, the effect relative to the baseline closure rate is systematically larger when the school is large in absolute terms, whereas the the relative effect size does not vary substantially with respect to relative schools size. This result is consistent, although not conclusive, with the idea that representation matters more when the policy choice is more important to the residents.

In addition to school size, the importance of the local school may depend on how inconvenient or costly school closure would be in terms of school commute. We measure this cost simply as the distance between two schools: the school closest to the candidate and the closest school to this school in the same municipality. For example, in Figure A2 for candidates assigned to school C we measure the distance between schools C and A. We divide the sample based on the median of this distance measure (roughly 4km in both the lottery and one vote margin samples) and estimate separate models for schools where the distance is either below or above the sample median. From Table A9, we see that the effect relative to baseline is larger when the distance is larger, but only slightly so.

Finally, we analyze whether the effects are larger when the school closure is likely to directly affect the politicians' own families. This will inform us whether their actions are motivated by personal interests. To this end, we split our sample based on whether the candidates themselves have young children (12-year-old or younger) who are currently or in the near future likely to attend the local school. Roughly 35% of the candidates in the lottery and one vote margin samples have young children. According Table A10, the effect does not vary systematically depending on whether the candidates have young children or not. Therefore, we do not find evidence suggesting that direct personal gains from the local school motivate the councilors' decisions regarding school closures.

Taken together, these additional results suggest that local politicians have more influence in small councils and that they want to prevent school closures in their neighborhood to cater to local residents rather than to obtain personal non-electoral gains. Although these results together with the substantial vote premium that candidates receive from their own neighborhoods (see Table 2) point to an important role for electoral incentives, we cannot rule out other mechanisms, such as better knowledge about local needs compared to the needs of other neighborhoods or non-electoral accountability related to frequent day-to-day encounters with their neighbors.

4.3 School closures and residential re-sorting

In this section, we examine residential sorting near closed schools using a staggered DID (event study) design for school closures that took place in 2013, 2014 and 2015.[20] We concentrate on these closures because our neighborhood data is available from 2010 to 2018 and we want to have enough pre- and post-treatment observations for testing pre-treatment trends and to have enough time for residential re-sorting to take place.

In this analysis, we use zipcode level panel data using only those zipcodes that have or had just one school during the analysis period. We make this choice because our neighborhood measures are at zipcode level and it is unclear what would happen in zipcodes with more than one school. For example, would the pupils of the closed school transfer to another school in the same neighborhood or to a school in another neighborhood? When there is only one school, the treatment is more clear cut. Furthermore, we use nearest-neighbour matching based on school size and zipcode level population of the last pre-treatment year.[21] This both increases the comparability of the treatment and the control groups in terms of school size and the number of people affected by the closures and facilitates clearer interpretation in our heterogeneity analysis below.

Using the matched data set, we estimate the following event study specifications:

$$y_{kt} = \sum_{\tau=-4}^{3} \delta_\tau D_{\tau,kt} + \gamma_t + \alpha_k + u_{kt}, \tag{4}$$

where y_{st} is the outcome of interest for zipcode k at year t. The dummy variable, $D_{\tau,st}$, indicates the year relative to year of closure of school s. The negative values of τ indicate the pre-closure years and the positive values indicate the post-closure periods. At $\tau = 0$, the school no longer exists. The specification includes zipcode and year fixed effects and we cluster the standard errors at the zipcode level. We use the Sun and Abraham (2021)

[20] Note that we cannot use the randomly elected candidates as an IV for school closures because it does not satisfy the exclusion restriction. That is, increasing representation may have effects on other neighborhood level public goods besides schools.

[21] We match two control units for each treated unit. The results are robust to using one or three control units (not reported).

method, which is robust to treatment effect heterogeneity (eventstudyinteract command in STATA).

The outcome we are interested in is the prevalence of different income groups in the zipcodes. In particular, we want to understand whether the high-income families move away from the neighbourhood as local public services deteriorate. We make use of the Zipcode Database by Statistics Finland where the adult population is divided into three groups based on income deciles constructed annually at the national level: low-income (deciles 1 and 2), middle-income (deciles 3 through 8) and high-income (deciles 9 and 10).

We focus on effect heterogeneity with respect to school size. School size captures two important aspects. First, the larger the school, the more people in the neighborhood are affected by the closure (due to our matching procedure, both in absolute and relative terms), and second, as explained earlier school size may be correlated with school quality. We use a cutoff of 90 pupils in the school to divide the sample into small and large schools. With 90 pupils, we would expect the class size at each grade to be 15 pupils. This cutoff is somewhat *ad hoc* and we analyse robustness with respect to this choice.

The event study plots (point estimates and 95% confidence intervals for δ_τ's) based on equation (4) are presented in Figure 2 for the small schools (under 90 pupils) and in Figure 3 for the large schools (at least 90 pupils). Starting from Figure 2, there is a slight downward trend in the total population in the zipcodes with a school closure both pre- and post-treatment. This is not surprising given that diminishing enrollment is one of the main reasons for school closures. However, individual point estimates are not statistically significant after the closure. This slight downward trend is visible with the middle-income residents (deciles 3 through 8). There seems to be no systematic pattern in the development of low- (deciles 1 and 2) and high-income groups (deciles 9 and 10), although these estimates are quite noisy. It seems that the small drop in total population is driven by the middle-income group (deciles 3 through 8). In sum, there seems to be no clear pattern of residential re-sorting after school closures when the schools are small.

Results are different for the sample of larger schools. According to Figure 3, there again seems to be a downward trend in total population in pre- and post-treatment periods that is driven by middle-income residents. Furthermore, there is a decline in the number of high-income residents after closure and this decline is also statistically significant. This group also exhibits a clean common pre-trend three years prior to closure. The decline in high-income residents appears to be accompanied by a slight increase in the number of low-income residents, although the point estimates are not statistically significant for this group. Thus, when the closed school is relatively large and possibly of higher quality, high-income residents vote with their feet suggesting that school closures reinforce residential segregation.[22]

[22] In Figure A4, we report event study plots for different school size cutoffs. The results are very similar to those in Figure 3.

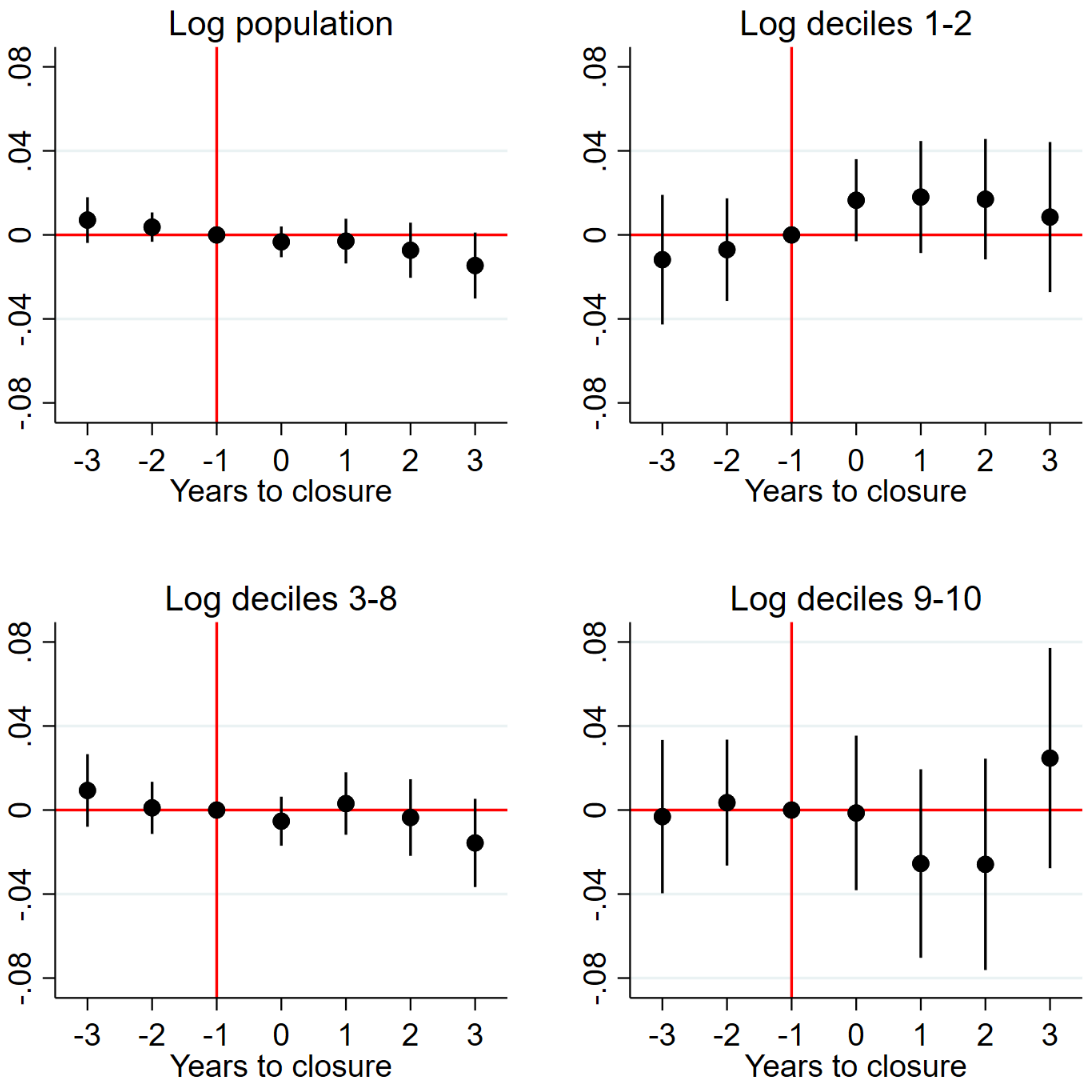

Figure 2: Effect of school closures on residential sorting, small schools.

Notes: The Figures plot the coefficient estimates and 95% confidence intervals from event study regressions using zipcode level panel data. $N = 2941$. The outcomes are the logs of total population and the number of neighborhood residents in the national level income deciles (1 and 2, 3 through 8 and 9 and 10). The schools in the sample had under 90 pupils prior to closure. Event time is the year relative to the year of school closure. School closures included in the analysis took place in 2013, 2014 and 2015. The omitted period is -1. The regressions include year and zipcode fixed effects. Standard errors used for confidence intervals are clustered at the zipcode level.

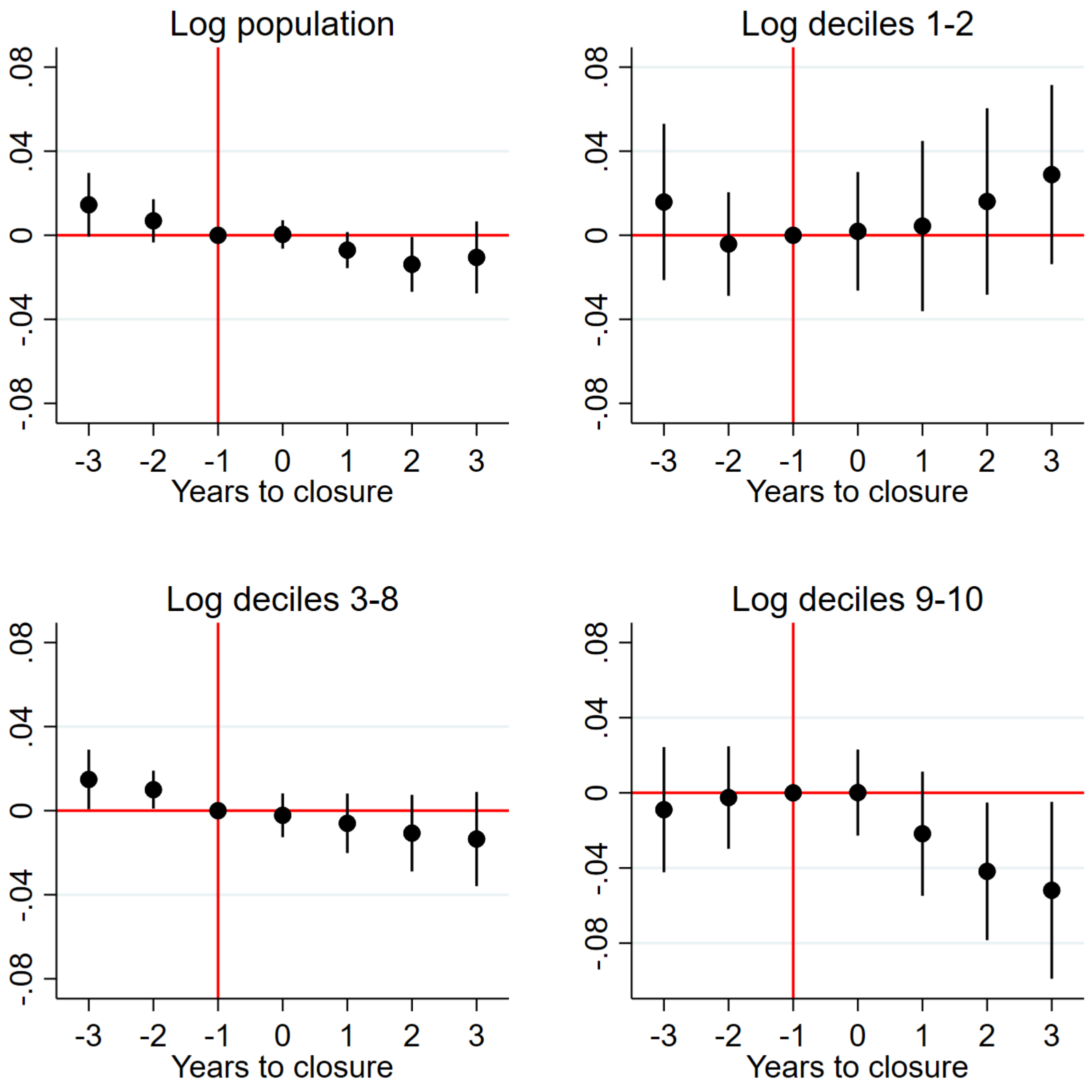

Figure 3: Effect of school closures on residential sorting, large schools.

Notes: The Figures plot the coefficient estimates and 95% confidence intervals from event study regressions using zipcode level panel data. $N = 532$. The outcomes are the logs of total population and the number of neighborhood residents in the national level income deciles (1 and 2, 3 through 8 and 9 and 10). The schools in the sample had at least 90 pupils prior to closure date. Event time is the year relative to the year of school closure. School closures included in the analysis took place in 2013, 2014 and 2015. The omitted period is -1. The regressions include year and zipcode fixed effects. Standard errors used for confidence intervals are clustered at the zipcode level.

5 Conclusions

We have studied the link between geographic political representation and geographic distribution of local public goods within local jurisdictions using detailed geo-coded micro data spanning three municipal election terms. We have produced four novel findings. First, neighborhoods with poorer and less-educated electorates are under-represented relative to their share of municipal population. Second, candidates have a strong local support base suggesting that they have electoral incentives to cater to voters in their neighborhood, even though the elections are held at-large. Third, based on randomly decided election outcomes, geographic representation has a causal effect on the geographic distribution of local public goods. Finally, deterioration of local public goods, in our case through school closures, seems to lead to residential re-sorting with respect to income of the residents. We conclude with some thoughts on the external validity and policy implications of our findings and highlight interesting avenues for future research.

It is important to note some limitations of our findings in terms of external validity. Although open-list PR systems are prevalent worldwide, the details in the systems and the tasks assigned to local jurisdictions may influence the link between geographic representation and local public good provision in nuanced ways. Finland is also a country with relatively low income inequality and residential segregation. Problems related to underpresentation may be exacerbated in countries and cities with higher levels of segregation. Moreover, our causal results come from a subset of candidates who were involved in close races for the last party seat(s). Concentrating on these marginal candidates facilitates causal inference, but provides us with a local effect that may not capture more general effects across a wider distribution of politicians. On the one hand, the candidates who occupy, in a sense, the last seat of the party may have less power to influence the council decisions. On the other hand, these marginal candidates are electorally vulnerable, and thus, may have stronger electoral incentives to cater to local voters compared to electorally safer candidates.

The important policy question that our results raise is how can we make political representation geographically more balanced. We consider three factors: voter mobilization, election system reforms and housing policies that aim to affect residential sorting.

First, if under-representation of poorer and less-educated neighborhoods is mainly due to lower turnout by the socio-economic groups in these neighborhoods (e.g., Lindgren et al. 2019, Akee et al. 2020, Hall and Yoder 2020 and Yoder 2021), policies designed to mobilize voters may be important. However, evidence on the effectiveness of interventions mobilising nonvoters (so-called get-out-the-vote) is not encouraging in this respect. For example, Enos et al. (2014) summarize the findings from 27 experimental mobilization interventions. They find that on average, these mobilization strategies actually widen disparities in participation and representation by mobilizing well-represented citizens more than the under-represented. On the other hand, increase in turnout and representation of disadvantaged groups due to major enfranchisement reforms typically result in large shifts in policies (e.g., Husted and Kenny 1997, Cascio and Washington 2014 and Fujiwara 2015). Whether mobilization interventions in countries that already have a broad franchise can be designed more effectively to alleviate the geographic differences in participation should remain high in the research agenda.

Second, when it comes to comparing election systems in terms of balanced geographic representation, perhaps the most interesting comparison is between at-large and ward systems. At a first glance, a ward system seems to offer clear benefits as it guarantees that all neighborhoods get at least some representation. Recent evidence also indicates that moving from at-large to a ward system indeed improves minority representation (e.g. Abott and Magazinnik 2020). However, the comparison between at-large and ward systems involves subtle trade-offs. For example, Mast (2020) and Hankinson and Magazinnik (2021) show that the ward system worsens the so-called not-in-my-neighborhood problems and suppresses local housing supply. Moreover, the effects of moving to a ward system on geographic representation crucially depends on districting as the large literature on optimal districting and gerrymandering indicates (e.g., Coate and Knight 2007, Trebbi et al. 2008 and Gul and Pesendorfer 2010). Moving to wards in a PR system may also have drastic effects on the number of parties due to Duverger's law (Duverger 1959). In addition, according Beath et al. (2016), at-large systems tend to select higher quality candidates compared to ward systems. Analyzing these important trade-offs should also prove a fruitful avenue for future research.

As the underlying cause of unequal geographic representation is residential segregation, it is important to also consider policies that directly address it. These can be either place-based social mixing policies where the main tool is the location of public housing across neighborhoods, or tenant-based programs where housing vouchers would somehow be conditioned on neighborhood characteristics (e.g., Collinson and Ganong 2018 and Davis et al. 2020). At the same time, the details of the program matter because the results of social mixing policies from different countries are quite mixed and the programs tend to be expensive (e.g., Eerola and Saarimaa 2018, Verdugo and Toma 2018, Diamond and McQuade 2019 and Bergman et al. 2019). Moreover, land use and housing supply regulation more generally have been shown to affect segregation at least along racial lines (Trounstine 2020). Future research should focus on gaining a better understanding of the relative merits of place- and tenant-based housing programs and land use regulation in creating socially mixed neighborhoods.

Finally, our results show that the effects of school closures are more far-reaching than policy-makers may have thought. Their indirect effects on neighbourhood segregation and geographic balance of political representation constitute an important consideration when deciding on school closures or the geography of local public good provision more generally.

References

Abott, C., Magazinnik, A., 2020. At-large elections and minority representation in local government. American Journal of Political Science 64, 717–733.

Akee, R., Copeland, W., Holbein, J.B., Simeonova, E., 2020. Human Capital and Voting Behavior across Generations: Evidence from an Income Intervention. American Political Science Review 114, 609–616.

Almagro, M., Domınguez-Iino, T., 2020. Location Sorting and Endogenous Amenities: Evidence from Amsterdam. Mimeo .

Ansolabehere, S., Gerber, A., Snyder, J., 2002. Equal votes, equal money: Court-ordered redistricting and public expenditures in the american states. American Political Science Review , 767–777.

Baron, D.P., Ferejohn, J.A., 1989. Bargaining in legislatures. American Political Science Review 4, 1181–1206.

Beach, B., Jones, D.B., Twinam, T., Walsh, R., 2019. Minority representation in local government. NBER Working Paper No. 25192 .

Beath, A., Christia, F., Egorov, G., Enikolopov, R., 2016. Electoral rules and political selection: Theory and evidence from a field experiment in afghanistan. The Review of Economic Studies 83, 932–968.

Bergman, P., Chetty, R., DeLuca, S., Hendren, N., Katz, L.F., Palmer, C., 2019. Creating moves to opportunity: Experimental evidence on barriers to neighborhood choice. Technical Report. National Bureau of Economic Research.

Beuchert, L., Humlum, M.K., Nielsen, H.S., Smith, N., 2018. The short-term effects of school consolidation on student achievement: Evidence of disruption? Economics of Education Review 65, 31–47.

Brollo, F., Nannicini, T., 2012. Tying your enemy's hands in close races: The politics of federal transfers in brazil. American Political Science Review 4, 742–761.

Broockman, D.E., Skovron, C., 2018. Bias in perceptions of public opinion among political elites. American Political Science Review 112, 542–563.

Brown, J.R., Enos, R.D., 2021. The measurement of partisan sorting for 180 million voters. Nature Human Behaviour .

Brummet, Q., 2014. The effect of school closings on student achievement. Journal of Public Economics 119, 108–124.

Campbell, R., Cowley, P., Vivyan, N., Wagner, M., 2019. Why Friends and Neighbors? Explaining the Electoral Appeal of Local Roots. The Journal of Politics 81, 937–951.

Cantoni, E., Pons, V., 2020. Does Context Outweigh Individual Characteristics in Driving Voting Behavior? Evidence from Relocations within the U.S. Working Paper 27998. National Bureau of Economic Research.

Carey, J.M., Shugart, M.S., 1995. Incentives to cultivate a personal vote: A rank ordering of electoral formulas. Electoral Studies 14, 417 – 439.

Carozzi, F., Repetto, L., 2019. Distributive Politics Inside the City? The Political Economy of Spain's Plan E. Regional Science and Urban Economics 75, 85–106.

Cascio, E.U., Washington, E., 2014. Valuing the vote: The redistribution of voting rights and state funds following the voting rights act of 1965. The Quarterly Journal of Economics 129, 379–433.

Chattopadhyay, R., Duflo, E., 2004. Women as policy makers: Evidence from a randomized policy experiment in india. Econometrica 5, 1409–1443.

Chetty, R., Friedman, J., Hendren, N., Jones, M.R., Porter, S., 2020. The opportunity atlas: Mapping the childhood roots of social mobility. Working paper .

Chetty, R., Hendren, N., 2018. The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. The Quarterly Journal of Economics 133, 1107–1162.

Chetty, R., Hendren, N., Katz, L.F., 2016. The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. American Economic Review 106, 855–902.

Chyn, E., 2018. Moved to opportunity: The long-run effects of public housing demolition on children. American Economic Review 108, 3028–56.

Chyn, E., Haggag, K., 2019. Moved to Vote: The Long-Run Effects of Neighborhoods on Political Participation. Working Paper 26515. National Bureau of Economic Research.

Coate, S., Knight, B., 2007. Socially optimal districting: a theoretical and empirical exploration. The Quarterly Journal of Economics 122, 1409–1471.

Collinson, R., Ganong, P., 2018. How do changes in housing voucher design affect rent and neighborhood quality? American Economic Journal: Economic Policy 10, 62–89.

Couture, V., Gaubert, C., Handbury, J., Hurst, E., 2020. Income growth and the distributional effects of urban spatial sorting. Mimeo .

Couture, V., Handbury, J., 2020. Urban revival in america. Journal of Urban Economics 119, 103267.

Cox, G.W., 2009. Swing voters, core voters, and distributive politics. In Political Representation, ed. Shapiro, I., S. Stokes, E. Wood and A. Kirschner. Cambridge University Press.

Damm, A.P., Dustmann, C., 2014. Does growing up in a high crime neighborhood affect youth criminal behavior? The American Economic Review 104, 1806–1832.

Davis, M., Gregory, J., Hartley, D., Tan, K., et al., 2020. Neighborhood effects and housing vouchers. Technical Report. Mimeo.

De Haan, M., Leuven, E., Oosterbeek, H., 2016. School consolidation and student achievement. The Journal of Law, Economics, and Organization 32, 816–839.

Deutscher, N., 2020. Place, peers, and the teenage years: Long-run neighborhood effects in australia. American Economic Journal: Applied Economics 12, 220–49.

Diamond, R., 2016. The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. American Economic Review 106, 479–524.

Diamond, R., McQuade, T., 2019. Who wants affordable housing in their backyard? an equilibrium analysis of low-income property development. Journal of Political Economy 127, 1063–1117.

Dixit, A., Londregan, J., 1996. The determinants of success of special interests in redistributive politics. Journal of politics , 1132–1155.

Dragu, T., Rodden, J., 2011. Representation and redistribution in federations. Proceedings of the National Academy of Sciences 21, 8601–8604.

Duverger, M., 1959. Political parties, their organization and activity in the modern state. Methuen.

Eerola, E., Saarimaa, T., 2018. Delivering affordable housing and neighborhood quality: A comparison of place-and tenant-based programs. Journal of Housing Economics 42, 44–54.

Enos, R.D., 2017. The space between us: Social geography and politics. Cambridge University Press.

Enos, R.D., Fowler, A., Vavreck, L., 2014. Increasing Inequality: The Effect of GOTV Mobilization on the Composition of the Electorate. The Journal of Politics 76, 273–288.

Fiva, J.H., Halse, A.H., 2016. Local favoritism in at-large proportional representation systems. Journal of Public Economics 143, 15–26.

Fiva, J.H., Halse, A.H., Smith, D.M., 2020. Local representation and voter mobilization in closed-list proportional representation systems. Quarterly Journal of Political Science forthcoming .

Fogli, A., Guerrieri, V., 2019. The End of the American Dream? Inequality and Segregation in US Cities. Working Paper 26143. National Bureau of Economic Research.

Folke, O., Martén, L., Rickne, J., Dahlberg, M., 2021. Politicians' neighborhoods: Where do they live and does it matter? Mimeo .

Fujiwara, T., 2015. Voting technology, political responsiveness, and infant health: Evidence from brazil. Econometrica 83, 423–464.

Gul, F., Pesendorfer, W., 2010. Strategic redistricting. American Economic Review 100, 1616–41.

Hall, A.B., Yoder, J., 2020. Does Homeownership Influence Political Behavior? Evidence from Administrative Data. Journal of Politics forthcoming .

Hankinson, M., Magazinnik, A., 2021. The supply-equity trade-off: The effect of spatial representation on the local housing supply. Mimeo .

Hodler, R., Raschky, P.A., 2014. Regional favoritism. The Quarterly Journal of Economics 129, 995–1033.

Husted, T.A., Kenny, L.W., 1997. The effect of the expansion of the voting franchise on the size of government. Journal of Political Economy 105, 54–82.

Hyytinen, A., Meriläinen, J., Saarimaa, T., Toivanen, O., Tukiainen, J., 2018a. Public employees as politicians: Evidence from close elections. American Political Science Review 1, 68–81.

Hyytinen, A., Meriläinen, J., Saarimaa, T., Toivanen, O., Tukiainen, J., 2018b. When Does Regression Discontinuity Design Work? Evidence from Random Election Outcomes. Quantitative Economics 9, 1019–1051.

Jankowski, M., 2016. Voting for locals: Voters' information processing strategies in open-list pr systems. Electoral Studies 43, 72–84.

Knight, B., 2008. Legislative representation, bargaining power and the distribution of federal funds: Evidence from the US congress. The Economic Journal 118, 1785–1803.

Lahtinen, H., Martikainen, P., Mattila, M., Wass, H., Rapeli, L., 2019. Do surveys overestimate or underestimate socioeconomic differences in voter turnout? Evidence from administrative registers. Public Opinion Quarterly 83, 363–385.

Laliberté, J.W., 2021. Long-term contextual effects in education: Schools and neighborhoods. American Economic Journal: Economic Policy 13, 336–377.

Lindgren, K.O., Oskarsson, S., Persson, M., 2019. Enhancing electoral equality: can education compensate for family background differences in voting participation? American Political Science Review 113, 108–122.

Mast, E., 2020. Warding off development: Local control, housing supply, and nimbys. WE Upjohn Institute for Employment Research .

Nakamura, E., Sigurdsson, J., Steinsson, J., 2021. The Gift of Moving: Intergenerational Consequences of a Mobility Shock. Review of Economic Studies forthcoming.

Pande, R., 2003. Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India. American Economic Review 4, 1132–1151.

Saarimaa, T., Tukiainen, J., 2016. Local representation and strategic voting: Evidence from electoral boundary reforms. European Journal of Political Economy 41, 31–45.

Scartascini, C., Cruz, C., Keefer, P., 2018. The database of political institutions 2017. https://publications.iadb.org/en/database-political-institutions-2017-dpi2017 .

Shugart, M.S., Valdini, M.E., Suominen, K., 2005. Looking for locals: Voter information demands and personal vote-earning attributes of legislators under proportional representation. American Journal of Political Science 49, 437–449.

Su, Y., 2021. The rising value of time and the origin of urban gentrification. American Economic Journal: Economic Policy forthcoming .

Sun, L., Abraham, S., 2021. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of Econometrics forthcoming .

Taghizadeh, J.L., 2020. Are students in receiving schools hurt by the closing of low-Performing schools? Effects of school closures on receiving schools in Sweden 2000–2016. Economics of Education Review 78.

Trebbi, F., Aghion, P., Alesina, A., 2008. Electoral rules and minority representation in US cities. The Quarterly Journal of Economics 123, 325–357.

Trounstine, J., 2010. Representation and accountability in cities. Annual Review of Political Science 13, 407–423.

Trounstine, J., 2020. The geography of inequality: How land use regulation produces segregation. American Political Science Review 114, 443–455.

Verdugo, G., Toma, S., 2018. Can public housing decrease segregation? Lessons and challenges from non-European immigration in France. Demography 55, 1803–1828.

Yoder, J., 2021. Does Property Ownership Lead to Participation in Local Politics? Evidence from Property Records and Meeting. American Political Science Review forthcoming .

Online Appendix

Love Thy (Elected) Neighbor? Residential Segregation, Political Representation and Local Public Goods

Oskari Harjunen, Tuukka Saarimaa and Janne Tukiainen

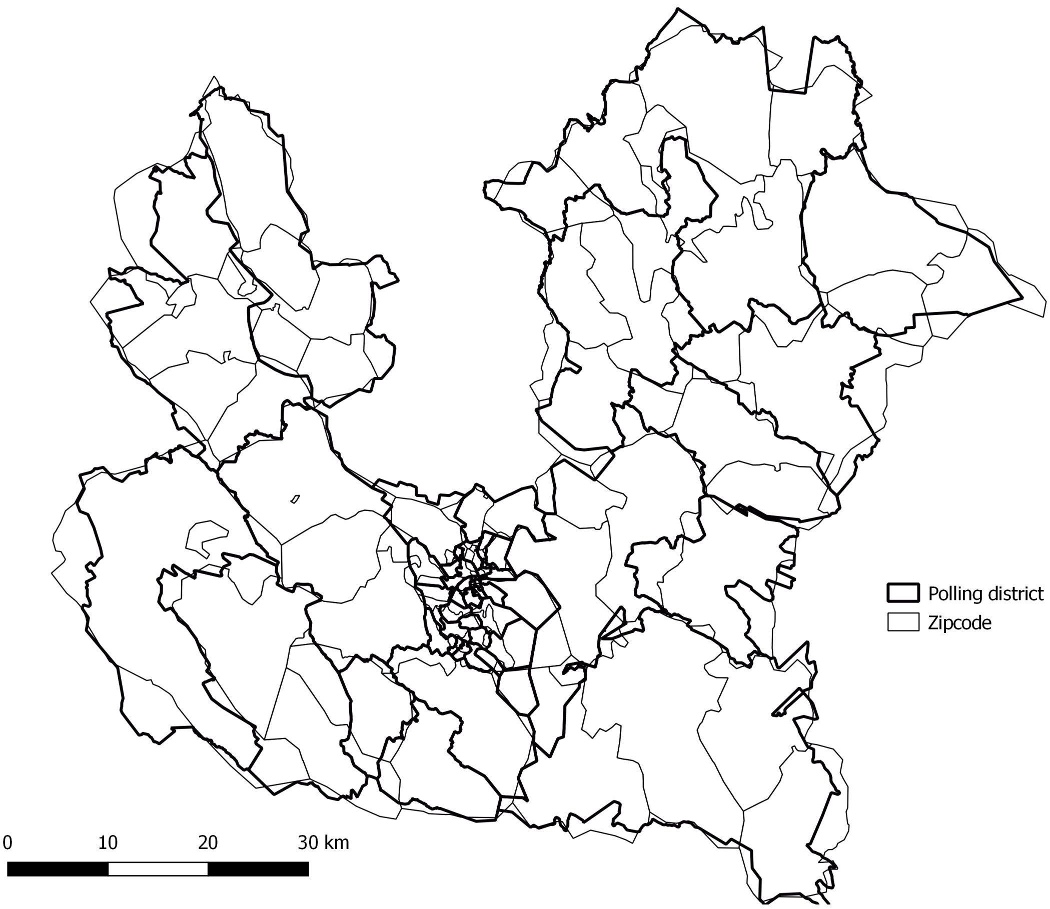

Figure A1: Polling districts and zipcodes in the municipality of Kuopio in 2019.

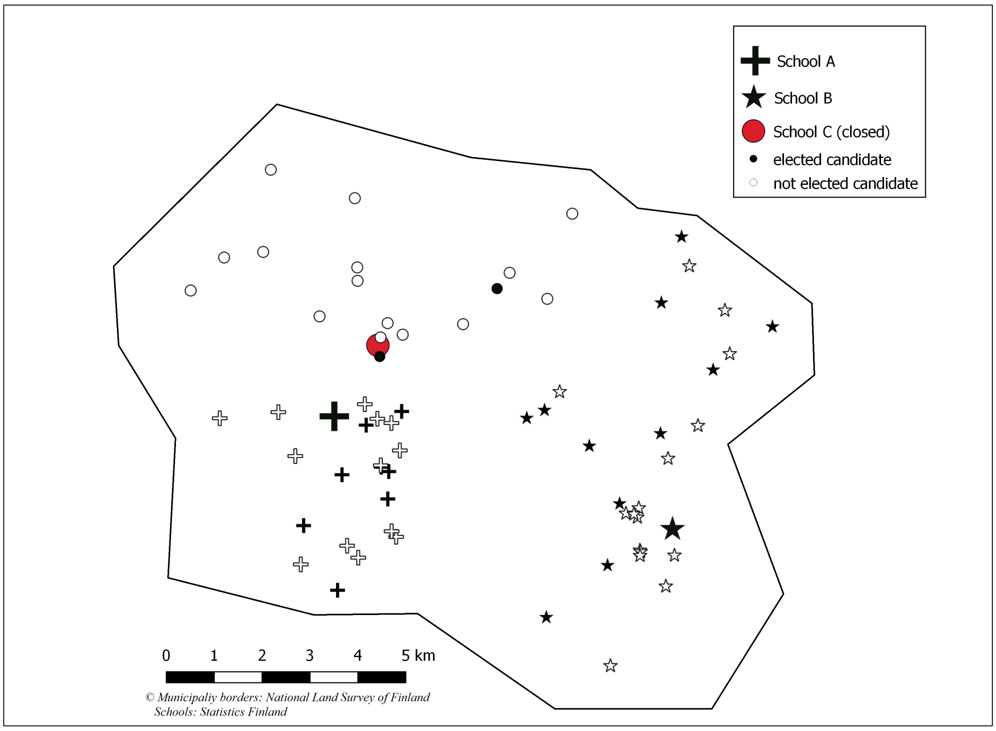

Figure A2: Assigning candidates to schools.

Notes: The figure illustrates how we assign candidates to their closest school in one municipality with three schools. The schools are marked with a circle, a cross and a star (large symbols). The red circle means that this school was closed during the election term. The smaller symbols refer to the candidates' places of residence in the municipality. In the figure, a particular school is closest to the candidate when the candidate's symbol matches the school's symbol. Hollow symbols mean that the candidate was not elected while the filled symbols indicate elected candidates. In this municipal election, there were in total 64 candidates of which 21 were elected. The seat shares of the schools are 10%, 33% and 57%, respectively, and the school with the lowest seat share is the one that was closed.

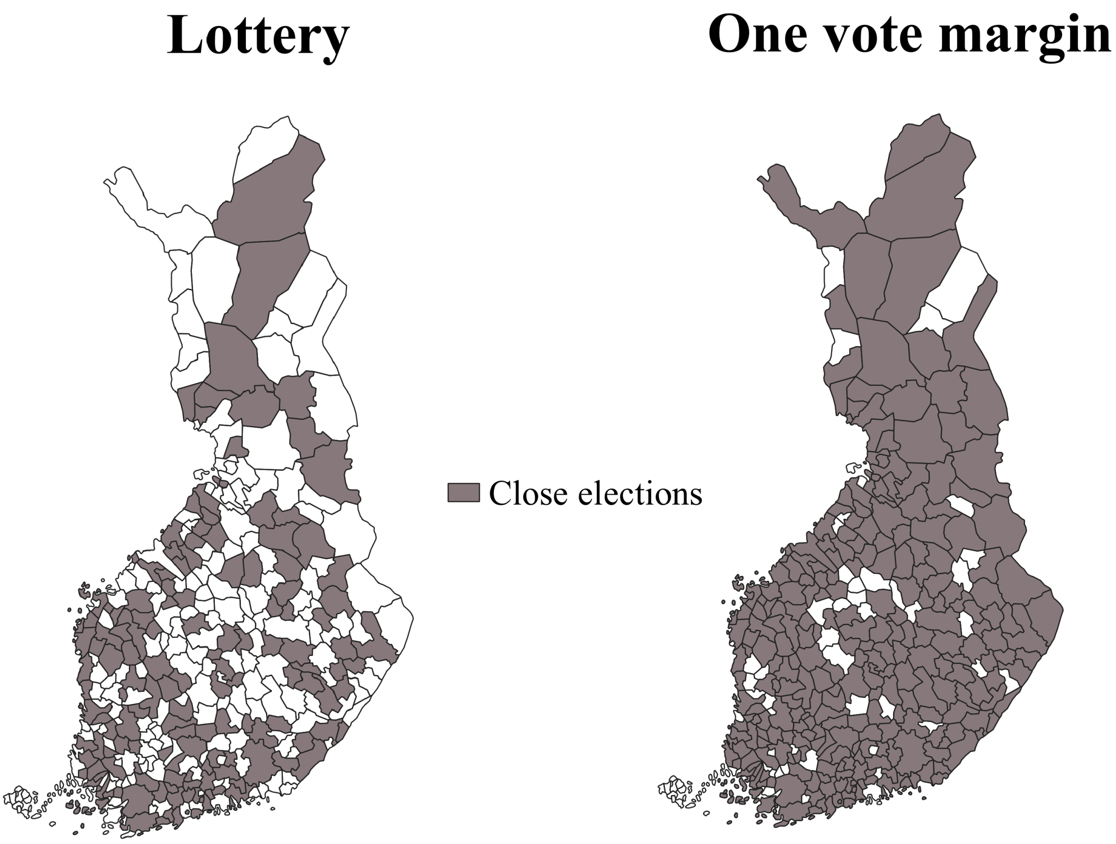

Figure A3: Municipalities with close elections.

Notes: The shaded municipalities had at least one close election (decided by a lottery or by one vote margin) in the 2004-2012 elections. Municipal borders are from 2012.

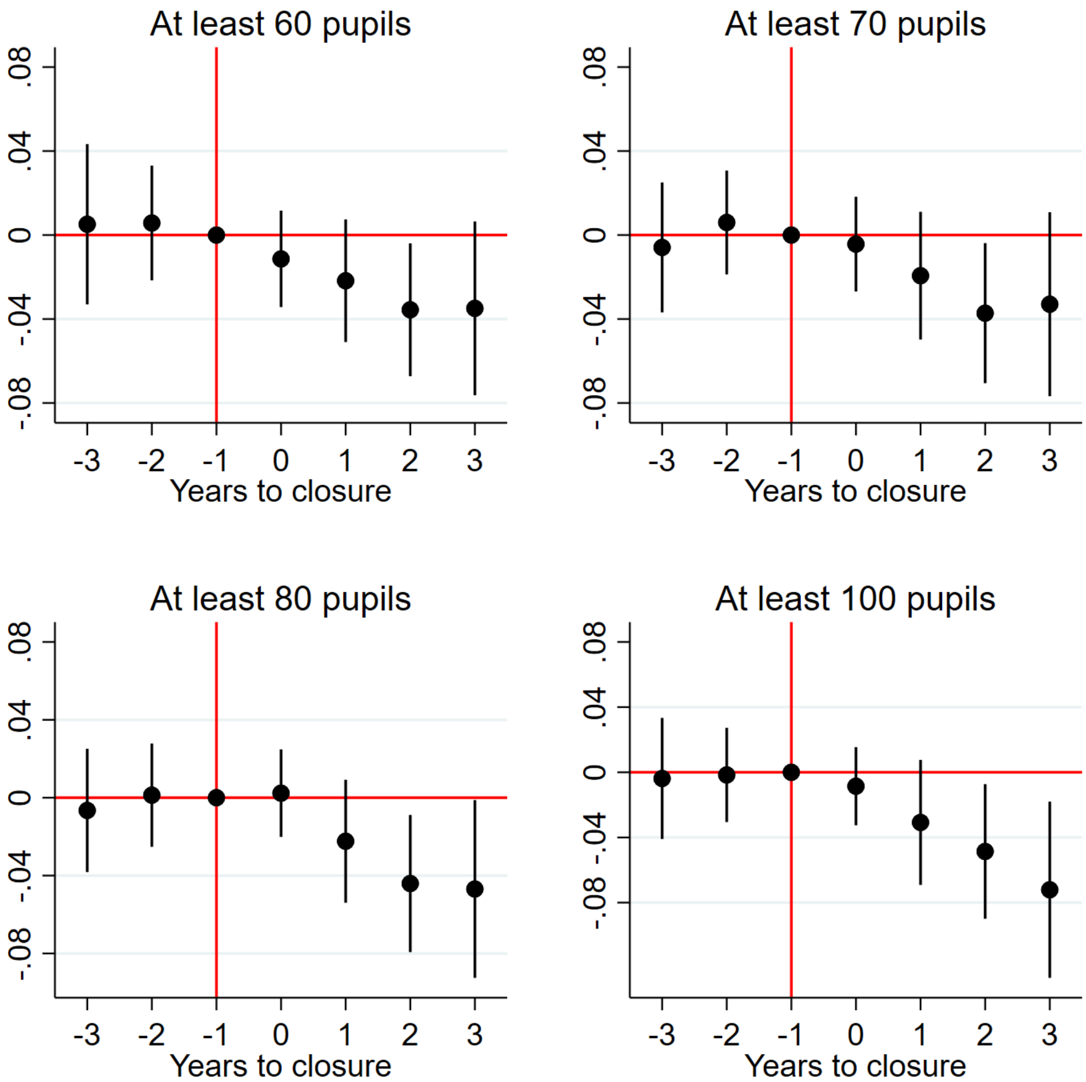

Figure A4: Sensitivity of effect of school closures on residential sorting with respect to school size cutoff.

Notes: The Figures plot the coefficient estimates and 95% confidence intervals from event study regressions using zipcode level panel data. Only zipcodes with one school are included in the analysis. The sample sizes are 700 for top-left, 650 for top-right, 593 for bottom-left and 407 for bottom-right, respectively. The outcomes are the logs of total population and the number of neighborhood residents in the national level income deciles (1 and 2, 3 through 8 and 9 and 10). The schools in the sample had at least 60, 70, 80 or 100 pupils prior to closure date. Event time is the year relative to the year of school closure. School closures included in the analysis took place in 2013, 2014 and 2015. The omitted period is -1. The regressions include year and zipcode fixed effects. Standard errors used for confidence intervals are clustered at the zipcode level.

Table A1: Turnout and socio-economic structure of polling districts.

	(1)	(2)
Mean income	2.584***	
	(0.194)	
Share highly-educated		57.43***
		(8.334)
N	2062	2062
Municipality FE	Yes	Yes

Notes: The table presents results from regressions where the unit of observation is a polling district. All the models include the polling district population as a control. Standard errors are presented in parentheses and are clustered at the municipality level. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table A2: Candidates' vote shares from own polling district.

	#-districts		#-districts	
	≤ 4	> 4	≤ 4	> 4
	(1)	(2)	(3)	(4)
Own district	1.706***	1.441***	0.911***	0.418***
	(0.145)	(0.098)	(0.110)	(0.042)
Own district x			0.024***	0.051***
residence spell			(0.004)	(0.004)
Own district x			1.603***	2.330***
incumbent			(0.155)	(0.167)
Own party vote share			0.038***	0.014***
			(0.003)	(0.001)
Voter similarity			0.691	0.017***
			(0.468)	(0.006)
Outcome mean	1.255	0.273	1.255	0.273
N	17,443	662,247	17,443	662,247
R-squared	0.54	0.37	0.58	0.43
Candidate FE	yes	yes	yes	yes
Polling district FE	yes	yes	yes	yes

Notes: The table presents results from candidate-polling district level regressions where the dependent variable is the candidate's vote share (%) of the polling district. Only municipalities with more than one polling district are included. Columns (1) and (2) include all candidates from these municipalities. In columns (3) and (4), the data is divided into two samples based on the median number of polling district in all municipalities (4). Standard errors are presented in parentheses and are clustered at the municipality level. $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table A3: Covariate balance at candidate level.

	Lottery (N = 419)		One vote margin (N = 1540)	
Variable	Control mean	Difference	Control mean	Difference
Age	46.79	-0.109	47.98	-0.240
		(1.329)		(0.674)
Male (0/1)	0.569	0.032	0.605	0.024
		(0.046)		(0.022)
Children (0/1)	0.313	0.067	0.332	0.026
		(0.048)		(0.025)
Number of children	0.725	0.111	0.725	0.041
		(0.132)		(0.064)
Incumbent (0/1)	0.332	0.014	0.339	0.042*
		(0.042)		(0.023)
Student (0/1)	0.038	-0.009	0.026	-0.002
		(0.016)		(0.008)
Unemployed (0/1)	0.043	-0.009	0.023	0.016*
		(0.018)		(0.009)
Entrepreneur (0/1)	0.156	0.065*	0.192	0.025
		(0.038)		(0.020)
High professional (0/1)	0.204	0.017	0.192	0.015
		(0.045)		(0.020)
Distance to school (km)	3.334	-0.573	2.829	-0.110
		(0.525)		(0.205)
Number of pupils in school	129.4	10.85	127.9	1.498
		(10.49)		(6.928)
School seat share	0.202	0.021	0.172	0.033***
		(0.016)		(0.010)
School has	0.948	0.052***	0.950	0.050***
representation (0/1)		(0.017)		(0.009)

Notes: Each row in the table refers to a separate bivariate regression where the dependent variable is reported on the first column and the explanatory variable is election status (*Elected*). The control mean refers to the constant in these models and the difference refers to the coefficient on the explanatory variable *Elected*. Standard errors are presented in the parentheses and are clustered at the municipality level. $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table A4: Descriptive statistics for municipalities with and without close elections.

Sample:	Lottery		One vote margin		Other municipalities	
	Mean	SD	Mean	SD	Mean	SD
Population	16,669	44,682	16,451	35,243	17,135	43,149
Share under 15 (%)	17.80	3.373	17.65	3.691	17.21	3.586
Income per capita	23,368	4009	23,218	3655	22,976	3935
Turnout (%)	63.99	5.708	62.66	6.12	62.73	6.057
Council size	32.58	10.60	32.788	11.31	32.02	11.75
Number of schools	8.576	9.639	8.395	8.549	8.168	9.730
N (municipalities)	198		605		913	

Notes: The descriptive statistics are calculated over three election years: 2004, 2008 and 2012.

Table A5: Effects of representation on school closure (non-movers).

Panel A: Lottery	(1)	(2)	(3)	(4)
Constant	0.200***	0.286***	0.348***	0.326***
	(0.031)	(0.041)	(0.108)	(0.109)
Elected	-0.091**	-0.085**	-0.080**	-0.080**
	(0.037)	(0.037)	(0.037)	(0.038)
N	379	379	379	379
R-squared	0.016	0.063	0.115	0.122
Panel B: One vote margin	(5)	(6)	(7)	(8)
Constant	0.188***	0.263***	0.199***	0.207***
	(0.018)	(0.023)	(0.060)	(0.061)
Elected	-0.060***	-0.060***	-0.062***	-0.062***
	(0.021)	(0.021)	(0.021)	(0.021)
N	1397	1397	1397	1397
R-squared	0.007	0.044	0.066	0.068
School controls	No	Yes	Yes	Yes
Candidate controls	No	No	Yes	Yes
Election term FE	No	No	No	Yes

Notes: The table presents results from linear probability models where the outcome is school closure (0/1). The school controls include the number of pupils in school. The candidate controls include age, sex, children in the family or not, incumbency and occupation status. Standard errors are presented in parentheses and are clustered at the municipality level. $^{*}\, p < 0.10$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

Table A6: Effect of representation on school closure (municipalities with closures).

Panel A: Lottery	(1)	(2)	(3)	(4)
Constant	0.339***	0.438***	0.554***	0.504***
	(0.045)	(0.052)	(0.180)	(0.175)
Elected	-0.171***	-0.162***	-0.154***	-0.157***
	(0.057)	(0.057)	(0.059)	(0.059)
N	246	246	246	246
R-squared	0.038	0.086	0.129	0.164
Panel B: One vote margin	(5)	(6)	(7)	(8)
Constant	0.301***	0.410***	0.313***	0.301***
	(0.024)	(0.029)	(0.081)	(0.083)
Elected	-0.091***	-0.091***	-0.095***	-0.095***
	(0.031)	(0.031)	(0.032)	(0.032)
N	929	929	929	929
R-squared	0.011	0.069	0.085	0.087
School controls	No	Yes	Yes	Yes
Candidate controls	No	No	Yes	Yes
Election term FE	No	No	No	Yes

Notes: The table presents results from linear probability models where the outcome is school closure (0/1). The school controls include the number of pupils in school. The candidate controls include age, sex, children in the family or not, incumbency and occupation status. Standard errors are presented in parentheses and are clustered at the municipality level. $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table A7: Effect heterogeneity with respect to council size.

	Lottery		One vote margin	
	Council size ≤ 27	Council size > 27	Council size ≤ 27	Council size > 27
	(1)	(2)	(3)	(4)
Elected	-0.163***	-0.046	-0.106***	-0.033
	(0.060)	(0.042)	(0.034)	(0.022)
Control group outcome mean	0.307	0.109	0.269	0.113
N	203	216	706	834
R-squared	0.168	0.167	0.082	0.041

Notes: The table presents results from linear probability models where the outcome is school closure (0/1). The models include election term fixed effects and the following controls: number of pupils in school, candidates' age, sex, if they have children or not, incumbency and occupation status. Standard errors are presented in parentheses and are clustered at the municipality level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A8: Effect heterogeneity with respect to school size.

	Lottery		One vote margin	
Panel A: Absolute school size	Small	Large	Small	Large
	(1)	(2)	(3)	(4)
Elected	-0.087	-0.097***	-0.098***	-0.048**
	(0.056)	(0.035)	(0.031)	(0.022)
Control group outcome mean	0.294	0.108	0.278	0.093
N	210	209	773	767
R-squared	0.245	0.207	0.148	0.038
Panel B: Relative school size	Small	Large	Small	Large
	(5)	(6)	(7)	(8)
Elected	-0.128**	-0.065	-0.089***	-0.049*
	(0.057)	(0.043)	(0.029)	(0.025)
Control group outcome mean	0.259	0.146	0.235	0.137
N	209	210	770	770
R-squared	0.156	0.127	0.083	0.057

Notes: The table presents results from linear probability models where the outcome is school closure (0/1). The median absolute school sizes are 92 in the lottery and 76 in the one vote margin samples whereas the median relative school sizes are 14% and 13%, respectively. The models include election term fixed effects and the following controls: number of pupils in school, candidates' age, sex, if they have children or not, incumbency and occupation status. Standard errors are presented in parentheses and are clustered at the municipality level. $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table A9: Effect heterogeneity with respect to school distance.

	Lottery		One vote margin	
	Below median	Above median	Below median	Above median
	(1)	(2)	(3)	(4)
Elected	-0.057	-0.129**	-0.044*	-0.083**
	(0.046)	(0.054)	(0.025)	(0.033)
Control group outcome mean	0.131	0.268	0.136	0.240
N	209	210	791	849
R-squared	0.133	0.128	0.058	0.071

Notes: The table presents results from linear probability models where the outcome is school closure (0/1). The median distance in both samples is roughly 4km. The models include election term fixed effects and the following controls: number of pupils in school, candidates' age, sex, if they have children or not, incumbency and occupation status. Standard errors are presented in parentheses and are clustered at the municipality level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A10: Effect heterogeneity with respect to candidates having children under the age of 12.

	Lottery		One vote margin	
	Has children	No children	Has children	No children
	(1)	(2)	(3)	(4)
Elected	-0.090**	-0.082	-0.045*	-0.100***
	(0.042)	(0.060)	(0.023)	(0.033)
Control group outcome mean	0.207	0.197	0.169	0.219
N	274	145	1009	531
R-squared	0.108	0.109	0.066	0.065

Notes: The table presents results from linear probability models where the outcome is school closure (0/1). The models include election term fixed effects and the following controls: number of pupils in school, candidates' age, sex, incumbency and occupation status. Standard errors are presented in parentheses and are clustered at the municipality level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The **Aboa Centre for Economics (ACE)** is a joint initiative of the economics departments of the Turku School of Economics at the University of Turku and the School of Business and Economics at Åbo Akademi University. ACE was founded in 1998. The aim of the Centre is to coordinate research and education related to economics.

Contact information: Aboa Centre for Economics, Department of Economics, Rehtorinpellonkatu 3, FI-20500 Turku, Finland.

www.ace-economics.fi

ISSN 1796-3133