MUNI ECON

WORKING PAPER

n. 2022-04
ISSN 2571-130X

Homophily in Voting Behavior: Evidence from Preferential Voting

Lucie Coufalová / Masaryk University, Faculty of Economics and Administration, Department of Economics
Štěpán Mikula / Masaryk University, Faculty of Economics and Administration, Department of Economics
Michal Ševčík / Masaryk University, Faculty of Economics and Administration, Department of Economics, Mendel University in Brno, Brno, Czech Republic

Homophily in Voting Behavior: Evidence from Preferential Voting

Abstract

Homophily in Voting Behavior: Evidence from Preferential Voting Abstract: Homophily is a strong determinant of many types of human relationships. It affects, for example, whom we marry and potentially also whom we vote for. We use data on preferential voting from Czech parliamentary elections in 2006, 2010, 2013, and 2017 matched with 2011 Census data to identify the effect of homophily on voting behavior. We find that a one percent increase in the share of the municipality's population that has the same occupation or education level as the candidate increases the number of preferential votes that candidate receives by 0.7% or 0.5%, respectively. We also find that candidates who live in the voters' municipality receive a substantially higher number of preferential votes.

Masaryk University
Faculty of Economics and Administration

Authors:
Lucie Coufalová / Masaryk University, Faculty of Economics and Administration, Department of Economics
Štěpán Mikula / Masaryk University, Faculty of Economics and Administration, Department of Economics
Michal Ševčík / Masaryk University, Faculty of Economics and Administration, Department of Economics, Mendel University in Brno, Brno, Czech Republic

Contact: stepan.mikula@econ.muni.cz

Creation date: 2022-03
Revision date:

Keywords: voting behavior; homophily; preferential voting; Czech parliamentary elections
JEL classification: D72, D91, P16

Citation:
Coufalová, L., Mikula, Š., Ševčík, M. 2022. *Homophily in Voting Behavior: Evidence from Preferential Voting*. MUNI ECON Working Paper n. 2022-04. Brno: Masaryk University.

Homophily in Voting Behavior: Evidence from Preferential Voting*

Lucie Coufalová[†a], Štěpán Mikula[‡a], and Michal Ševčík[§a,b]

[a]Masaryk University, Brno, Czech Republic
[b]Mendel University in Brno, Brno, Czech Republic

March 2, 2022

Homophily is a strong determinant of many types of human relationships. It affects, for example, whom we marry and potentially also whom we vote for. We use data on preferential voting from Czech parliamentary elections in 2006, 2010, 2013, and 2017 matched with 2011 Census data to identify the effect of homophily on voting behavior. We find that a one percent increase in the share of the municipality's population that has the same occupation or education level as the candidate increases the number of preferential votes that candidate receives by 0.7% or 0.5%, respectively. We also find that candidates who live in the voters' municipality receive a substantially higher number of preferential votes.

Keywords: voting behavior; homophily; preferential voting; Czech parliamentary elections
JEL Codes: D72, D91, P16

*Coufalová and Ševčík acknowledge financial support from the Czech Science Foundation (grant No. 20-23131S). Mikula acknowledges financial support as part of a joint grant from the Czech Science Foundation and Austrian Science Fund (GF20-18033L). We thank Lukáš Laffers for his comments and suggestions. All remaining errors are our own. The replication package for all regression tables is available at https://github.com/s-mikula/homophily.

[†]Masaryk University, Lipová 507/41a 602 00 Brno, lucie.coufalova@econ.muni.cz
[‡]Corresponding author, Masaryk University, Lipová 507/41a 602 00 Brno, stepan.mikula@econ.muni.cz
[§]Masaryk University, Lipová 507/41a 602 00 Brno, michal.sevcik@econ.muni.cz

1 Introduction

The empirical literature shows that voters tend to prefer candidates who demonstrate suitability for the office in question. Voters assess candidates' suitability on the basis of their characteristics and stereotypes associated with them (Carnes and Lupu 2016; Portmann and Stojanović 2019). An alternative explanation of voters' support for certain candidates assumes the existence of *homophily*, i.e. preference for candidates who share socio-demographic or attitudinal characteristics with the voter (McPherson et al. 2001).

Homophily strongly affects many types of human relationships. These include, among others, our choice of friends (Adida et al. 2015) and life partners (Skopek et al. 2011). We generally tend to keep in touch with people whose age and gender (Stehlé et al. 2013; Smith et al. 2014), race (Smith et al. 2014), religion (Adida et al. 2015), education (Smith et al. 2014), social class (McPherson et al. 2001), behavior (Boman IV and Mowen 2018) or political ideology (Boutyline and Willer 2017; Huber and Malhotra 2017) is similar to ours.

Homophilic preferences may also be expected even when it comes to voting. Voters may assume that politicians from other social classes and groups do not understand their problems and give priority to other topics. Consequently, they prefer candidates who are similar to them in terms of various socio-demographic characteristics and who they feel would better defend their interests (Kendall and Yum 1984; Sevi 2021; Wardt et al. 2021). The empirical literature has demonstrated the presence of homophilic preferences in voting behavior based on background characteristics, such as gender (Yalley 2021), age (Sevi 2021), ethnicity (Barreto et al. 2005), and religious affiliation (Adida et al. 2015).

The literature to date has made use of a range of approaches to identify the effect of homophily on voting behavior. Sevi (2021) uses post-election surveys from multiple countries to show that voter support for a particular candidate decreases as the age difference between voter and candidate increases. Based on survey data collected before the 2016 Ghanian parliamentary elections, Yalley (2021) shows that ethnicity is the most important driver of homophilic preferences.

However, survey data may suffer from social desirability bias (see e.g. Portmann and Stojanović 2019). Adida et al. (2015) address this issue by conducting a laboratory experiment that allows for random treatment assignment. Using a voting game, they show that religious similarity strongly determines voter choice. They find that none of the other tested characteristics (gender, age, education, ethnicity, and socio-economic similarity) have similar predictive power.

A third methodological approach analyses preferences revealed in actual voting behavior. Using data from the 2000 mayoral and 2001 presidential elections in the U.S., Barreto

et al. (2005) show that in districts with larger shares of Latino voters co-ethnicity increased voter turnout and Latino candidates received larger shares of the vote than in other districts.

In this paper we add to the literature on the role status (background) homophily plays in voting behavior. We identify the effect of homophily on voting behavior using data on preferential voting from Czech parliamentary elections in 2006, 2010, 2013, and 2017. The system of preferential voting used in the Czech Republic (see Section 2 for a description) is well suited for this identification because: (1) Voters primarily vote for a single party but may, if they wish, give preferential votes to up to four candidates on the ballot. This enables us to filter out the impact of party preferences and identify the effect of similarity between individual candidates and the municipal population on the number of preferential votes received by those particular candidates in the given municipality. (2) Secondary information about the candidates (gender, age, political party affiliation, education, occupation, place of residency) is listed on the ballot and is, therefore, available to all voters. (3) There are 14 constituencies in the Czech parliamentary elections with many municipalities (up to 1,144) in each of them. By focusing on preferential votes from small municipalities that cannot substantially affect elections results, we address potential endogeneity that may arise due to optimization of electoral list composition.

We find that a one percent increase in the share of the municipality's population that has the same occupation or education level as the candidate significantly (at 1% level) increases the number of preferential votes that candidate receives by 0.7% or 0.5%, respectively. We also find that candidates who live in their voters' municipality receive a substantially higher number of preferential votes.

2 Parliamentary elections and preferential voting

The Parliament of the Czech Republic has two houses: the upper house is known as the Senate and the lower house is known as the Chamber of Deputies. We focus on elections to the latter, to which 200 deputies are elected under a party-list proportional representation system. For these elections to the lower house there are 14 constituencies, which match the country's administrative regions (*"kraj"*). Parties draw up unique electoral lists for each region. In total they nominate up to 343 candidates[1]; the parties themselves nominate candidates for each constituency and determine the order of their lists via their own internal procedures. The order in which candidates are listed on the ballot is important, since it is also the order in which seats are subsequently allocated.

1. The constituencies (regions) and their maximum numbers of candidates as set by law are: Prague (36), Moravian-Silesian (36), Central Bohemian (34), South Moravian (34), Ústí nad Labem (26), Olomouc (23), South Bohemian (22), Zlín (22), Plzeň (20), Hradec Králové (20), Vysočina (20), Pardubice (19), Liberec (17), and Karlovy Vary (14).

Voters can, however, also influence the order in which seats will be allocated to candidates using their optional preferential votes.[2] In the Czech system of preferential voting each voter may assign one preferential vote to up to four candidates on the ballot. If a candidate receives preferential votes that exceed 5% of the total number of votes for his/her party, he/she is re-ranked to the top of the electoral list.

The system of preferential voting was introduced in 1990 for the first free elections after the fall of the communist regime and has since gained substantial popularity. In 2006 there were 0.4 preferential votes cast per ballot. The subsequent elections in 2010 set a record in our estimation sample with 0.7 preferential votes per ballot cast and the share of preferential votes did not then drop below 0.5 within our observation period. The preferential votes substantially affected the election outcomes: between 2006 and 2017, 14% of deputies were elected as a result of the preferential votes they received.

3 Data and estimation sample

The primary data set used in our empirical analysis is a an electoral database[3] containing the secondary information on candidates that is also available to voters on the ballot (name, age, political party affiliation, occupation, municipality of residency)[4] together with the electoral outcomes, including the number of preferential votes received.[5] We aggregate the electoral data at the level of municipalities[6] which results in a data set with one observation for each candidate, municipality and election.

The electoral database contains exact data on candidates' ages, which we classify into seven categories (18–29, 30–39, 40–49, 50–59, 60–69, 70–79, 80–89). Candidates list one or more occupations on the ballot. We manually encode these occupations into 11 categories according to the ISCO classification. In addition to ISCO level 1 we include a

2. Preferential voting systems are very common in Europe, although they take many forms. Casting preference votes can be optional, as is the case in the Czech Republic or Switzerland (Portmann and Stojanović 2019), or mandatory (Poland, Finland). In Finland, Poland and in state elections in Bavaria, voters cast just one preferential vote (Marcinkiewicz and Stegmaier 2015; Söderlund et al. 2021), whereas in the Czech Republic, Luxembourg, Lithuania, Greece, Slovakia and Switzerland voters may select multiple candidates, with or without a threshold (Marcinkiewicz and Stegmaier 2015). Some systems, such as the Swiss one, also allow voters to cast negative votes or even to choose candidates from more than one ballot list (*panachage*, Portmann and Stojanović 2019)

3. The electoral database is publicly available at https://volby.cz/opendata/opendata.htm.

4. For an example of a ballot see annotated Figure A.1 in the Appendix.

5. The electoral database does not contain any unique IDs that would enable us to match individual candidates across elections.

6. The raw data from the electoral database contains information at the level of polling stations. Unfortunately, the polling areas do not match census blocks. Therefore, we aggregate both the electoral and census data to the municipality level.

special category for "occupations" that are not listed in ISCO, such as *"student"*, *"mother"*, or *"retiree"*. When a candidate listed more than one occupation we use the first one.

Information on education and gender must be inferred from the name and candidate description. Tertiary education is signaled by the presence of academic titles, which are commonly used in formal communication (written as well as spoken) in the Czech Republic.

We infer the candidates' genders from their first names. We match the name listed in the electoral database with a database published by the Ministry of the Interior of the Czech Republic which includes the frequency of each name for each gender.[7] We assign each candidate the gender that is more frequent for his/her first name. This is a reliable measure since the Czech Republic is highly linguistically and ethnically homogeneous[8] and first names commonly given to both genders are rare.

The estimation sample contains voting data for elections to the Chamber of Deputies that took place in 2006, 2010, 2013, and 2017. The sample is restricted to elections that are close to the 2011 Census, which is our source of population data, and which were conducted according to the same set of rules.

From the 2011 Census we use individual level data and construct variables for each municipality and candidate characteristic (i.e. education, occupation, age category, gender) that capture each candidate's similarity with the population of the given municipality. These variables are defined as the percentage of the population that shares the specific variant of the given characteristic with the candidate (e.g. belongs to the same age group). It results in a continuous measure for all four characteristics with a theoretical range from 0 to 100%. We define an additional indicator variable that takes the value one for the municipality in which the candidate lives.

We adjust the estimation sample by excluding 17 municipalities whose boundaries changed during the period of interest and/or that are not represented in all elections – i.e. they did not exist or had zero turnout in at least one election. We also exclude the country capital Prague because Prague is the only single-municipality constituency and would therefore clearly violate our identification assumption (see Section 4). The electoral database and our estimation sample also do not contain municipality–party pairs in cases where a particular party received no votes. The resulting estimation sample contains data on 20,811 candidates from 6,240 municipalities.

7. This database is not available any more due to GDPR. We use the version that was released in 2015.

8. Over 91% of respondents who filled in their nationality in the 2011 Census self-assigned to Czech nationality or to nationalities related to the Czech lands (these nationalities do not differ in culture and language from the majority).

Table 1: Demographic characteristics and candidates' and voters' occupations

	Candidates	All municipalities	Small municipalities
	(1)	(2)	(3)
Age (years)	45.919	47.822	48.393
	(12.402)	(2.128)	(2.515)
Male (%)	72.125	49.783	50.249
		(2.363)	(2.945)
Tertiary educated (%)	45.510	7.713	6.706
		(4.177)	(3.779)
ISCO 0 (%)	0.053	0.172	0.166
		(0.344)	(0.407)
ISCO 1 (%)	27.370	2.447	2.151
		(1.478)	(1.582)
ISCO 2 (%)	23.190	5.175	4.464
		(2.580)	(2.475)
ISCO 3 (%)	19.110	7.719	7.170
		(2.803)	(3.013)
ISCO 4 (%)	3.633	2.316	2.095
		(1.249)	(1.441)
ISCO 5 (%)	4.320	6.146	5.760
		(2.210)	(2.636)
ISCO 6 (%)	0.947	1.858	2.432
		(1.765)	(2.106)
ISCO 7 (%)	3.662	9.321	9.391
		(2.788)	(3.290)
ISCO 8 (%)	3.373	8.177	8.488
		(2.825)	(3.273)
ISCO 9 (%)	0.207	2.066	2.090
		(1.705)	(2.284)
Unclassified occupations (%)	14.137	4.185	4.153
		(1.986)	(2.441)
Observations	20,811	6,240	3,119

Note: "Small municipalities" are defined as municipalities with populations below the median in 2011, i.e. 426 inhabitants. The table contains mean ages for candidates and for both categories of municipalities. For the other variables the table contains shares of candidates and mean shares for municipalities. Standard deviations are reported in parentheses. Values are calculated using data from the 2006, 2010, 2013, and 2017 parliamentary elections (Column 1) and the 2011 Census (Columns 2 and 3).

Table 2: Similarity between candidates and municipalities' populations

	All municipalities	Small municipalities
	(1)	(2)
Share of voters with the same education level as the candidate (%)	87.5856 (18.8336)	88.5563 (19.8213)
Share of voters with the same occupation as the candidate (%)	4.3725 (3.4373)	3.8503 (3.3101)
Share of voters of the same age as the candidate (%)	17.3224 (3.7809)	17.1332 (4.4263)
Share of voters of the same gender as the candidate (%)	49.8278 (2.1660)	50.0890 (2.8130)
Candidate lives in the municipality (=1)	0.0023 (0.0484)	0.0002 (0.0157)

Note: "Small municipalities" are defined as municipalities with populations below median in 2011. Standard deviations are reported in parentheses. Values are calculated using data from the 2006, 2010, 2013, and 2017 parliamentary elections and the 2011 Census.

Table 1 compares candidates with municipalities' populations. While there is only a little difference in age, there is a substantial overrepresentation of men and tertiary educated people among the electoral candidates. The population of candidates is also more skewed towards more qualified occupations. These differences tend to be more pronounced in small municipalities with populations below the median (i.e. fewer than 426 inhabitants in 2011). Table 2 focuses on variables that measure similarity between candidates and municipalities' populations and provides similar information: there is remarkable similarity in age and education level. However, the high similarity for education is driven by the low number of categories for this characteristic (tertiary educated, other). The sample of all municipalities also tends to show greater similarity with candidates than the subset of small municipalities.

4 Identification strategy and empirical specification

We identify the effect of homophily using data on preferential voting for individual candidates. As each party can field up to 343 candidates in a single election, with up to 36 candidates on each ballot paper, the voters are not likely to know the individual candidates personally (see Jurajda and Münich 2015). However, they may base their decision to support a candidate with a preferential vote on secondary information that is provided to each voter on the ballot paper. We test the hypothesis that voters are more likely to support

candidates that are similar to them in background characteristics. The political parties may be aware of the importance of such homophily and may optimize their candidate lists to match the population characteristics in the given constituency. This potential endogeneity of electoral list composition presents a challenge to identification.

In elections to the Chamber of Deputies the parties establish candidate lists at the regional (constituency) level, where each region contains up to 1,144 municipalities (mean = 480 when excluding Prague). We exploit this feature of the Czech electoral system in our identification of the effect of homophily on voting behavior. Our identification rests on the assumption that political parties do not optimize their candidate lists with respect to the population characteristics of individual municipalities. We estimate the effect of homophily using the following empirical specification:

$$\log(E(pv_{cpme})) = \boldsymbol{\gamma}\mathbf{H}_{cm} + \log(v_{pme}) + \theta_m + \theta_b + \theta_{ce} \tag{1}$$

where pv is the number of preferential votes received by candidate c running for party p in municipality m and election e. The vector $\mathbf{H}$ contains variables of interest: an indicator variable for the candidate being a resident in the given municipality and four continuous variables defined as the percentages of the municipality's population that share the candidate's education level (tertiary, lower than tertiary), occupation (ISCO1 or unclassified), age group (18–29, 30–39, 40–49, 50–59, 60–69, 70–79, 80–89), and gender.

Variable v is an offset variable that captures the number of votes cast for the respective party. It effectively expresses the maximum number of preferential votes a given candidate can receive. The empirical specification also includes a set of fixed effects. The municipality fixed effect (θ_m) controls for time-invariant municipality characteristics that may affect residents' propensity to use preferential votes—such as education level or other unobserved characteristics. We also include the fixed effect θ_b for the number of candidates on the ballot list, as a lower number of candidates increases *per se* the probability of each candidate receiving a preferential vote. Most importantly, our specification (1) contains a full set of candidate (θ_{ce}) fixed effects (including constants) that controls for the candidate's individual characteristics, the popularity of the party he/she is running for, and his/her position on the ballot list. These candidate fixed effects ensure that the homophily effect is estimated within candidate.

Due to the large share of zeros (80.3%) in the outcome variable we estimate regression (1) with Poisson Pseudo-Maximum Likelihood (Silva and Tenreyro 2011). For all estimates we cluster robust standard errors by ballot (i.e. using one cluster for each party, election, and constituency/region) and municipality. We exclude observations that have a fixed effect with perfect fit.

5 Results and robustness checks

In Table 3 we report the estimated coefficients for the variables of interest (γ) from Equation (1). The positive and statistically significant coefficients suggest that homophily is indeed present in voting behavior: candidates receive more preferential votes in municipalities where a larger share of the population shares their personal characteristics. However, whether the estimates from Table 3 can be interpreted causally is uncertain, as our identification assumption may not hold for the full sample of municipalities. Political parties may optimize their electoral lists with respect to the populations of large municipalities which have the potential to provide a substantial number of votes. Such optimization would make our similarity measures endogenous.

Table 3: Homophily and voting behavior

	Dependent variable: Number of preferential votes received					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of voters with the same education level as the candidate (%)	0.0098*** (0.0012)					−0.0024 (0.0021)
Share of voters with the same occupation as the candidate (%)		0.0236*** (0.0033)				0.0007 (0.0027)
Share of voters of the same age as the candidate (%)			0.0062*** (0.0012)			0.0031*** (0.0010)
Share of voters of the same gender as the candidate (%)				0.0024 (0.0019)		0.0019 (0.0021)
Candidate lives in the municipality (=1)					1.7694*** (0.0691	1.7710*** (0.0686)
Candidate fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Municipality fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Length of electoral list fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,856,299	7,856,299	7,856,299	7,856,299	7,856,299	7,856,299

Notes: Table contains estimates of coefficients γ from Equation (1). Standard errors clustered by ballot and municipality are reported in parentheses: *, ** and *** denote statistical significance at 10%, 5% and 1%. 40,561 observations were excluded due to having a fixed effect with perfect fit.

We address this issue in Table 4 where we restrict our estimation sample to small municipalities with populations below the median (426 people). Only 6.5% of voters lived in these municipalities in all four covered elections, which made them an unattractive target for electoral list optimization and thus make the composition of the electoral list credibly exogenous.

The estimates reported in Columns (1) to (4) are substantially smaller than the estimates obtained with the full sample of municipalities. However, they maintain their statistical significance, showing the positive effect of homophily. A one percent increase in the share of the municipality's population whose occupation or education level are the same as the

Table 4: Effect of homophily on voting behavior in municipalities with population below the median

	Dependent variable: Number of preferential votes received					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of voters with the same education level as the candidate (%)	0.0051*** (0.0012)					0.0049*** (0.0012)
Share of voters with the same occupation as the candidate (%)		0.0068*** (0.0018)				0.0064*** (0.0018)
Share of voters of the same age as the candidate (%)			0.0012* (0.0006)			0.0011* (0.0006)
Share of voters of the same gender as the candidate (%)				0.0018* (0.0010)		0.0018* (0.0010)
Candidate lives in the municipality (=1)					3.0133*** (0.0653)	3.0131*** (0.0651)
Candidate fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Municipality fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Length of electoral list fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,991,126	2,991,126	2,991,126	2,991,126	2,991,126	2,991,126

Notes: Table contains estimates of coefficients γ from Equation (1). Standard errors clustered by ballot and municipality are reported in parentheses: *, ** and *** denote statistical significance at 10%, 5% and 1%. 198,058 observations were excluded due to having a fixed effect with perfect fit.

candidate's increases the number of preferential votes that candidate receives by 0.7% and 0.5% respectively.

Apart from substantial educational and occupational homophily the results also indicate the presence of gender and age homophily. However, these effects are lower in magnitude: a one percent increase in the share of municipality's population belonging to the same age group as the candidate increases the number of preferential votes that candidate receives by 0.1%. Similarly, a one percent increase in the population of the same gender increases the number of preferential votes by 0.2%. In contrast to the estimates we obtained using the full sample of municipalities, the coefficients of interest maintain their statistical significance in the specification that includes all similarity measures (see Column 6).

Column (5) shows that being a resident of the observed municipality is a very strong predictor of the number of preferential votes a given candidate receives. The estimated coefficient implies that candidates who run for seats in their place of residence are more likely to obtain preferential votes. This may be because people believe that deputies from their own municipality would be more likely to share their interests or would increase the municipality's prestige. However, it is also true that—in small municipalities (i.e. with fewer than 426 inhabitants) especially—voters are more likely to know these candidates in person. The other secondary information available on the ballot might, therefore, be irrelevant in these cases as voters would not need any cues to cast their preferential votes.

As a robustness check we, therefore, exclude from the sample all observations where candidates run in their municipality of residence. The coefficients reported in Table A.1 in the Appendix are comparable in magnitude to the results with the full sample of candidates (Table 4). However, only educational and occupational homophily maintain statistical significance at the 1% level.

To account for variation in municipalities' populations and in support for political parties we also re-estimate the equation (1) on the sample of small municipalities with weights set to the number of voters (v_{pme}). The results reported in Table A.2 in the Appendix do not substantially deviate for occupational and educational homophily, nor for the impact of the candidate being a resident. However, similarities in age and gender no longer have any significant effect.

6 Conclusion

Using data on preferential voting in Czech parliamentary elections we identify the effect of homophily on voting behavior. The Czech parliamentary elections and preferential voting system in particular are well suited to the analysis of homophilic preferences, as they enable us to filter out preferences for political parties and focus solely on candidates' individual background characteristics, which are listed on the ballot and are thus readily available to voters. Focusing on small municipalities, which are abundant in the Czech Republic and which are an unattractive target for electoral list optimization, also enables us to address the problem of potential endogeneity.

We find that a one percent increase in the share of a municipality's population whose occupation or education level are the same as the candidate's increases the number of preferential votes a given candidate receives by 0.7% or 0.5% respectively. Our result also indicates the presence of homophily based on gender and age although these latter effects are smaller and less robust. We also find that living in the given municipality substantially increases the number of preferential votes a candidate receives.

References

Adida, Claire L, David D Laitin, and Marie-Anne Valfort. 2015. "Religious homophily in a secular country: Evidence from a voting game in France." *Economic Inquiry* 53 (2): 1187–1206.

Barreto, Matt A, Mario Villarreal, and Nathan D Woods. 2005. "Metropolitan Latino political behavior: Voter turnout and candidate preference in Los Angeles." *Journal of Urban Affairs* 27 (1): 71–91.

Boman IV, John H, and Thomas J Mowen. 2018. "Same feathers, different flocks: Breaking down the meaning of 'behavioral Homophily'in the etiology of crime." *Journal of Criminal Justice* 54:30–40.

Boutyline, Andrei, and Robb Willer. 2017. "The social structure of political echo chambers: Variation in ideological homophily in online networks." *Political Psychology* 38 (3): 551–569.

Carnes, Nicholas, and Noam Lupu. 2016. "Do voters dislike working-class candidates? Voter biases and the descriptive underrepresentation of the working class." *American Political Science Review* 110 (4): 832–844.

Huber, Gregory A, and Neil Malhotra. 2017. "Political homophily in social relationships: Evidence from online dating behavior." *The Journal of Politics* 79 (1): 269–283.

Jurajda, Štěpán, and Daniel Münich. 2015. "Candidate ballot information and election outcomes: The Czech case." *Post-Soviet Affairs* 31 (5): 448–469.

Kendall, Kathleen E, and June Ock Yum. 1984. "Persuading the blue-collar voter: Issues, images, and homophily." *Annals of the International Communication Association* 8 (1): 707–723.

Marcinkiewicz, Kamil, and Mary Stegmaier. 2015. "Ballot position effects under compulsory and optional preferential-list PR electoral systems." *Political Behavior* 37 (2): 465–486.

McPherson, Miller, Lynn Smith-Lovin, and James M Cook. 2001. "Birds of a feather: Homophily in social networks." *Annual Review of Sociology* 27 (1): 415–444.

Portmann, Lea, and Nenad Stojanović. 2019. "Electoral discrimination against immigrant-origin candidates." *Political Behavior* 41 (1): 105–134.

Sevi, Semra. 2021. "Do young voters vote for young leaders?" *Electoral Studies* 69:102200.

Silva, JMC Santos, and Silvana Tenreyro. 2011. "Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator." *Economics Letters* 112 (2): 220–222.

Skopek, Jan, Florian Schulz, and Hans-Peter Blossfeld. 2011. "Who contacts whom? Educational homophily in online mate selection." *European Sociological Review* 27 (2): 180–195.

Smith, Jeffrey A, Miller McPherson, and Lynn Smith-Lovin. 2014. "Social distance in the United States: Sex, race, religion, age, and education homophily among confidants, 1985 to 2004." *American Sociological Review* 79 (3): 432–456.

Söderlund, Peter, Åsa von Schoultz, and Achillefs Papageorgiou. 2021. "Coping with complexity: Ballot position effects in the Finnish open-list proportional representation system." *Electoral Studies* 71:102330.

Stehlé, Juliette, Francois Charbonnier, Tristan Picard, Ciro Cattuto, and Alain Barrat. 2013. "Gender homophily from spatial behavior in a primary school: A sociometric study." *Social Networks* 35 (4): 604–613.

Wardt, Marc van de, Arjen Van Witteloostuijn, Anthony Chambers, and Bram Wauters. 2021. "Birds of a feather flock together? The survival of underrepresented groups within parliamentary parties, 1991–2015." *European Journal of Political Research* 60 (2): 474–496.

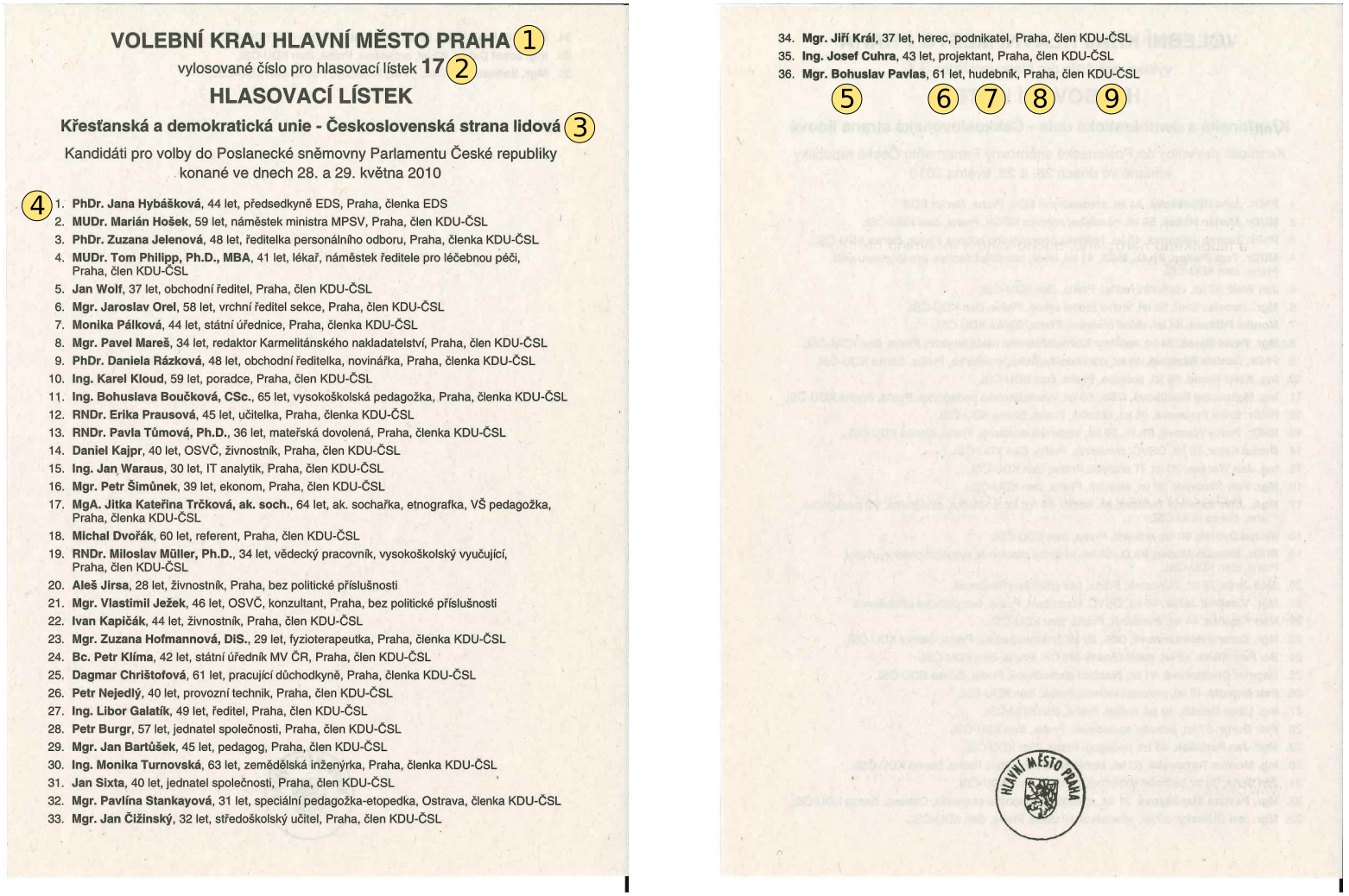
Yalley, Andrews Agya. 2021. "Voter and Political Candidate's Status Homophile: A Ghanaian Context." *Journal of Political Marketing* 20 (2): 175–197.

A Appendix

Table A.1: Effect of homophily on voting behavior in municipalities with population below the median, excluding candidates who reside in the municipality

	Dependent variable: Number of preferential votes received				
	(1)	(2)	(3)	(4)	(5)
Share of voters with the same education level as the candidate (%)	0.0050*** (0.0012)				0.0048*** (0.0012)
Share of voters with the same occupation as the candidate (%)		0.0068*** (0.0018)			0.0064*** (0.0018)
Share of voters of the same age as the candidate (%)			0.0010* (0.0006)		0.0010* (0.0006)
Share of voters of the same gender as the candidate (%)				0.0015 (0.0010)	0.0016 (0.0010)
Candidate fixed effect	Yes	Yes	Yes	Yes	Yes
Municipality fixed effect	Yes	Yes	Yes	Yes	Yes
Length of electoral list fixed effect	Yes	Yes	Yes	Yes	Yes
Observations	2,986,721	2,986,721	2,986,721	2,986,721	2,986,721

Notes: Table contains estimates of coefficients γ from Equation (1). Standard errors clustered by ballot and municipality are reported in parentheses: *, ** and *** denote statistical significance at 10%, 5% and 1%. 201,674 observations were excluded due to having a fixed effect with perfect fit.



VOLEBNÍ KRAJ HLAVNÍ MĚSTO PRAHA (1)

vylosované číslo pro hlasovací lístek **17** (2)

HLASOVACÍ LÍSTEK

Křesťanská a demokratická unie - Československá strana lidová (3)

Kandidáti pro volby do Poslanecké sněmovny Parlamentu České republiky
konané ve dnech 28. a 29. května 2010

(4)
1. **PhDr. Jana Hybášková**, 44 let, předsedkyně EDS, Praha, členka EDS
2. **MUDr. Marián Hošek**, 59 let, náměstek ministra MPSV, Praha, člen KDU-ČSL
3. **PhDr. Zuzana Jelenová**, 48 let, ředitelka personálního odboru, Praha, členka KDU-ČSL
4. **MUDr. Tom Philipp, Ph.D., MBA**, 41 let, lékař, náměstek ředitele pro léčebnou péči, Praha, člen KDU-ČSL
5. **Jan Wolf**, 37 let, obchodní ředitel, Praha, člen KDU-ČSL
6. **Mgr. Jaroslav Orel**, 58 let, vrchní ředitel sekce, Praha, člen KDU-ČSL
7. **Monika Pálková**, 44 let, státní úřednice, Praha, členka KDU-ČSL
8. **Mgr. Pavel Mareš**, 34 let, redaktor Karmelitánského nakladatelství, Praha, člen KDU-ČSL
9. **PhDr. Daniela Rázková**, 48 let, obchodní ředitelka, novinářka, Praha, členka KDU-ČSL
10. **Ing. Karel Kloud**, 59 let, poradce, Praha, člen KDU-ČSL
11. **Ing. Bohuslava Boučková, CSc.**, 65 let, vysokoškolská pedagožka, Praha, členka KDU-ČSL
12. **RNDr. Erika Prausová**, 45 let, učitelka, Praha, členka KDU-ČSL
13. **RNDr. Pavla Tůmová, Ph.D.**, 36 let, mateřská dovolená, Praha, členka KDU-ČSL
14. **Daniel Kajpr**, 40 let, OSVČ, živnostník, Praha, člen KDU-ČSL
15. **Ing. Jan Waraus**, 30 let, IT analytik, Praha, člen KDU-ČSL
16. **Mgr. Petr Šimůnek**, 39 let, ekonom, Praha, člen KDU-ČSL
17. **MgA. Jitka Kateřina Trčková, ak. soch.**, 64 let, ak. sochařka, etnografka, VŠ pedagožka, Praha, členka KDU-ČSL
18. **Michal Dvořák**, 60 let, referent, Praha, člen KDU-ČSL
19. **RNDr. Miloslav Müller, Ph.D.**, 34 let, vědecký pracovník, vysokoškolský vyučující, Praha, člen KDU-ČSL
20. **Aleš Jirsa**, 28 let, živnostník, Praha, bez politické příslušnosti
21. **Mgr. Vlastimil Ježek**, 46 let, OSVČ, konzultant, Praha, bez politické příslušnosti
22. **Ivan Kapičák**, 44 let, živnostník, Praha, člen KDU-ČSL
23. **Mgr. Zuzana Hofmannová, DiS.**, 29 let, fyzioterapeutka, Praha, členka KDU-ČSL
24. **Bc. Petr Klíma**, 42 let, státní úředník MV ČR, Praha, člen KDU-ČSL
25. **Dagmar Chrištofová**, 61 let, pracující důchodkyně, Praha, členka KDU-ČSL
26. **Petr Nejedlý**, 40 let, provozní technik, Praha, člen KDU-ČSL
27. **Ing. Libor Galatík**, 49 let, ředitel, Praha, člen KDU-ČSL
28. **Petr Burgr**, 57 let, jednatel společnosti, Praha, člen KDU-ČSL
29. **Mgr. Jan Bartůšek**, 45 let, pedagog, Praha, člen KDU-ČSL
30. **Ing. Monika Turnovská**, 63 let, zemědělská inženýrka, Praha, členka KDU-ČSL
31. **Jan Sixta**, 40 let, jednatel společnosti, Praha, člen KDU-ČSL
32. **Mgr. Pavlína Stankayová**, 31 let, speciální pedagožka-etopedka, Ostrava, členka KDU-ČSL
33. **Mgr. Jan Čižinský**, 32 let, středoškolský učitel, Praha, člen KDU-ČSL

34. **Mgr. Jiří Král**, 37 let, herec, podnikatel, Praha, člen KDU-ČSL
35. **Ing. Josef Cuhra**, 43 let, projektant, Praha, člen KDU-ČSL
36. **Mgr. Bohuslav Pavlas**, 61 let, hudebník, Praha, člen KDU-ČSL

(5) (6) (7) (8) (9)

HLAVNÍ MĚSTO PRAHA

(a) Front side (b) Back side

Figure A.1: Example of a ballot from 2010 parliamentary elections

Note: 1 – constituency; 2 – party number; 3 – party name; 4 – candidate's position on the ballot (ranking); 5 – candidate's name; 6 – age; 7 – occupation; 8 – municipality of residence; 9 – affiliation with a political party

Table A.2: Effect of homophily on voting behavior in municipalities with population below the median: results from weighted regression

	Dependent variable: Number of preferential votes received					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of voters with the same education level as the candidate (%)	0.0052*** (0.0017)					0.0050*** (0.0017)
Share of voters with the same occupation as the candidate (%)		0.0085*** (0.0026)				0.0082*** (0.0026)
Share of voters of the same age as the candidate (%)			0.0011 (0.0010)			0.0012 (0.0010)
Share of voters of the same gender as the candidate (%)				0.0019 (0.0015)		0.0017 (0.0015)
Candidate living in municipality (=1)					3.0212*** (0.0859)	3.0209*** (0.0853)
Candidate fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Municipality fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Length of electoral list fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,991,126	2,991,126	2,991,126	2,991,126	2,991,126	2,991,126

Notes: Table contains estimates of coefficients γ from Equation (1) weighted with the number of votes cast for the party (v_{pme}). Standard errors clustered by ballot and municipality are reported in parentheses: *, ** and *** denote statistical significance at 10%, 5% and 1%. 198,058 observations were excluded due to having a fixed effect with perfect fit.

MUNI Econ Working Paper Series (since 2018)

2022-04 Coufalová, L., Mikula, Š., Ševčík, M. 2022. *Homophily in Voting Behavior: Evidence from Preferential Voting*. MUNI ECON Working Paper n. 2022-04. Brno: Masaryk University.

2022-03 Kecskésová, M., Mikula, Š. 2022. *Malaria and Economic Development in the Short-term: Plasmodium falciparum vs Plasmodium vivax*. MUNI ECON Working Paper n. 2022-03. Brno: Masaryk University.

2022-02 Mladenović, D., Rrustemi, V., Martin, S., Kalia, P., Chawdhary, R. 2022. *Effects of Sociodemographic Variables on Electronic Word of Mouth: Evidence from Emerging Economies*. MUNI ECON Working Paper n. 2022-02. Brno: Masaryk University.

2022-01 Mikula, Š., Montag, J. 2022. *Roma and Bureaucrats: A Field Experiment in the Czech Republic*. MUNI ECON Working Paper n. 2022-01. Brno: Masaryk University.

2021-14 Abraham, E. D., Corazzini, L., Fišar, M., Reggiani, T. 2021. *Delegation and Overhead Aversion with Multiple Threshold Public Goods*. MUNI ECON Working Paper n. 2021-14. Brno: Masaryk University.

2021-13 Corazzini, L., Cotton, C., Longo, E., Reggiani, T. 2021. *The Gates Effect in Public Goods Experiments: How Donations Flow to the Recipients Favored by the Wealthy*. MUNI ECON Working Paper n. 2021-13. Brno: Masaryk University.

2021-12 Staněk, R., Krčál, O., Mikula, Š. 2021. *Social Capital and Mobility: An Experimental Study.*. MUNI ECON Working Paper n. 2021-12. Brno: Masaryk University.

2021-11 Staněk, R., Krčál, O., Čellárová, K. 2021. *Pull yourself up by your bootstraps: Identifying procedural preferences against helping others in the presence of moral hazard.*. MUNI ECON Working Paper n. 2021-11. Brno: Masaryk University.

2021-10 Levi, E., Sin, I., Stillman, S. 2021. *Understanding the Origins of Populist Political Parties and the Role of External Shocks.*. MUNI ECON Working Paper n. 2021-10. Brno: Masaryk University.

2021-09 Adamus, M., Grežo, M. 2021. *Individual Differences in Behavioural Responses to the Financial Threat Posed by the COVID-19 Pandemic.*. MUNI ECON Working Paper n. 2021-09. Brno: Masaryk University.

2021-08 Hargreaves Heap, S. P., Karadimitropoulou, A., Levi, E. 2021. *Narrative based information: is it the facts or their packaging that matters?*. MUNI ECON Working Paper n. 2021-08. Brno: Masaryk University.

2021-07 Hargreaves Heap, S. P., Levi, E., Ramalingam, A. 2021. *Group identification and giving: in-group love, out-group hate and their crowding out*. MUNI ECON Working Paper n. 2021-07. Brno: Masaryk University.

ISSN electronic edition 2571-130X

MUNI ECON Working Paper Series is indexed in RePEc:

https://ideas.repec.org/s/mub/wpaper.html