

ARTICLE

Sharing is caring: the role of voter-candidate similarities in intra-party electoral competition

Patrick F. A. van Erkel*

Department of Political Science, University of Antwerp, Antwerp, Belgium

*E-mail: patrick.vanerkel@uantwerpen.be

(Received 19 September 2017; revised 6 July 2018; accepted 17 September 2018; first published online 7 December 2018)

Abstract

Previous studies have found similarities with presidential candidates or party leaders to be an important factor in explaining voting behaviour. However, with the exception of gender, few studies have structurally studied voter-candidate similarities in intra-party electoral competition. This study investigates the Belgian case and argues that voter-candidate similarities play a role in the decision-making process of citizens when casting preferential votes. Moreover, it investigates whether underrepresented groups, and especially women, are more guided by these voter-candidate similarities than overrepresented groups. To achieve this aim voter and candidate characteristics are modelled simultaneously. This enables an investigation of the decision-making process of voters while taking into account structural inequalities at the supply side. The results demonstrate that citizens are indeed more likely to cast preferential votes for candidates similar to themselves and that these effects are stronger for underrepresented groups. Hence, preferential voting could ultimately pave the way for better descriptive representation.

Keywords: preferential voting; representation; electoral behaviour; elections

Introduction

To what extent do citizens vote for candidates that ‘are like them’? Over the years, a number of studies have provided evidence that the propensity to vote for a candidate increases when candidate and voter resemble each other on one or multiple characteristics (Sigelman and Sigelman, 1982; Piliavin, 1987; Cutler, 2002; McDermott, 2009). They demonstrate that women are more likely to vote for female candidates (Plutzer and Zipp, 1996; Dolan, 1998; Cutler, 2002) and African-American voters are more likely to select African-American candidates (e.g. Sigelman and Sigelman, 1982; Sigelman *et al.*, 1995). They also find that citizens identify more with candidates from the same region (Lewis-Beck and Rice, 1983; Gimpel *et al.*, 2008). However, almost all these studies have been conducted in majoritarian, single-member district systems, most notably in the United States. Research in open- or flexible-list proportional (PR) systems is more limited, although Banducci and Karp (2000) and Cutler (2002) find that in PR systems sharing traits with a party leader does increase the likelihood that citizens vote for that party.

The aim of this study is to investigate whether voter-candidate similarities also play a role in the intra-party electoral competition of many European countries. Quite a few European countries – such as Austria, Belgium, Denmark, Switzerland, and the Netherlands – have systems where voters not only need to select a political party, but subsequently can cast a (preferential) vote for one or more candidates within that party. Thus, not only do parties compete with each other, but also within each party there is an electoral competition between candidates for votes, the so-called ‘intra-party competition’ (Katz, 1985). These individual votes matter as they partly determine who will occupy a seat for the party. They also act as a resource for politicians and can

help to obtain legislative or executive mandates, or get a better ballot list position in subsequent elections (André *et al.*, 2017).

This study expects voters to be more prone to cast preferential votes for candidates who are more similar to them. However, not that much research has examined the effect of voter-candidate similarity on candidate voting in the intra-party context. Only with regard to gender do we find a handful of studies, but their findings are mixed with some finding evidence that gender similarities increase the propensity to vote for a candidate, although not necessarily for all groups equally (Holli and Wass, 2010; Giger *et al.*, 2014; Erzeel and Caluwaerts, 2015; Marien *et al.*, 2017; Erzeel *et al.*, 2018), while others find no evidence for this effect at all (McElroy and Marsh, 2010). These contrasting results could be explained by the fact that some studies use voter data, whereas others study the level of political candidates. Moreover, gender is unlikely to be the only sociodemographic characteristic on which voters base their decision. It can be expected that also other characteristics, such as age and the municipality of a candidate, will matter.

Investigating the effect of voter-candidate similarities is not only useful for an insight into why voters cast preferential votes for certain candidates and not for others, but also provides insight into the role of sociodemographic factors on voting behaviour outside single-member district systems. Also it is important from the viewpoint of descriptive representation. If citizens, and especially underrepresented groups such as women and ethnic minorities, are indeed guided by similarities with candidates in their vote decision, this could ultimately be a mechanism to overcome underrepresentation. Through a better descriptive representation this may ultimately lead to better substantive representation (Mansbridge, 1999).

Specifically, this study makes three contributions. First, it studies the influence of voter-candidate similarities on electoral behaviour in intra-party competition, focusing not only on gender, but also examining factors such as age and geographical proximity. Second, it investigates whether especially underrepresented groups, most notably regarding gender, are inclined to vote for candidates similar to them. Finally, this study makes a methodological contribution by simultaneously including voter and candidate characteristics in one model. Most studies on preferential voting either focus on political candidates or on voters, but almost never bring the two together in one model. Studies on gender-based voting, for instance, often use survey data, asking respondents whether they voted for a male or female. Yet, with few exceptions (see Marien *et al.*, 2017) they do not sufficiently take into account factors at the candidate level, such as the fact that female politicians are less likely to be positioned first on the ballot list and often still occupy less executive mandates, biasing the results of these studies. By looking at dyadic voter-candidate pairs, I aim to build further on previous work conducted on this topic and investigate both the supply and demand side. This makes it possible to model the role of voter-candidate similarities in the decision-making process around preferential voting. I focus on the case of Belgium, specifically Flanders, an institutional context in which citizens are relatively unrestrained to vote on the basis of sociodemographic cues.

The influence of voter-candidate similarities on voting

In general, the literature identifies two reasons why citizens may be inclined to vote for candidates similar to them. Both explanations find their roots in social identity theory (Tajfel and Turner, 1979) and are based on the notion that citizens are guided by social group membership in their vote choice (Pomper, 1975; Plutzer and Zipp, 1996). First, because of their social identity, which can be based on characteristics such as gender, race, age, or locality, citizens may feel more solidarity and affection with candidates from their ‘in-group’ and be more inclined to vote for them. In this sense, casting a preferential vote can be seen as a symbolic action to support the own group.

Second, citizens may vote for candidates resembling them because they expect these candidates to share similar experiences and ideas and, therefore, to be the best option to represent their interests (Erzeel and Caluwaerts, 2015), following a more instrumental logic. This idea can be

traced back to scholars working on descriptive representation (Pitkin, 1967; Mansbridge, 1999). It is argued that especially in low information elections, where citizens lack information on policy stances of candidates, they may be guided by similarities with candidates in order to fill this information gap and use cues such as gender, age, and location to extrapolate the ideological position of candidates (McDermott, 1997; Cutler, 2002). Both explanations are in line with the identity logic of preferential voting introduced by André *et al.* (2013), who find that certain groups are more inclined to cast preferential votes, because they have a stronger feeling of identity and vote on the basis of this identity.

It is also argued that the two mechanisms of symbolic support and interest representation especially motivate underrepresented groups, such as female and ethnic minorities, to base their vote on similarities with the political candidate. Because these groups still face barriers in society, identity may be more salient for them, making them more likely to vote based on it (Erzeel and Caluwaerts, 2015). Moreover, they may be more concerned that their interests are underrepresented in parliament and try to change the status quo by voting for someone from the in-group (Sanbonmatsu, 2002; Dolan, 2008).

Empirical evidence in single-member district systems supports the idea that citizens vote for candidates on the basis of shared characteristics. In an experimental design Sigelman and Sigelman (1982) pitted fictional candidates against each other in a two-candidate race, each time changing their race, age, and gender. They asked undergraduate students who they would vote for and find that similarity leads to more support for a candidate. Five years later the experiment was replicated by Piliavin (1987) with a more diverse group of respondents, reaching similar results. Also, in a more recent study, McDermott (2009) shows that both in the US elections of 1958 and 2004 similarities mattered, although she does find a small decrease of the effects over the years and also shows that not all types of similarities have the same effect. Banducci and Karp (2000) and Cutler (2002) demonstrate that similarities are also important outside the single-member district systems and find that voters in PR systems are more likely to vote for a party when the party leader resembles them more.

Studies that examine whether voter-candidate similarities also impact preferential voting are scarce. Only regarding gender do we find a handful of studies investigating whether women are more inclined to cast a preferential vote for a female candidate, but their findings are mixed at best (McElroy and Marsh, 2010; Erzeel and Caluwaerts, 2015; Marien *et al.*, 2017; Erzeel *et al.*, 2018). Teney *et al.* (2010) also find evidence that candidates with an ethnic minority background are more likely to cast a vote for an ethnic minority candidate.

It can be expected that voter-candidate similarities play a role in explaining why a voter selects a candidate over others when casting a preferential vote. First, unlike single-member district systems, in this context party preference does not prevent citizens from basing their vote on sociodemographic characteristics. Whereas in single-member district systems citizens may decide not to vote for a female/young/local candidate because this candidate belongs to the other party, this is not the case in systems with intra-party competition where parties present long lists of candidates, usually with mixed demographic backgrounds. Thus, citizens can first decide for which party they will vote and subsequently vote for a candidate of that party based on sociodemographic characteristics. There is no longer a trade-off between basing the vote on partisanship or on sociodemographic cues. Second, in systems with intra-party competition ballots are more diverse, as political parties aim to construct balanced ballot lists consisting of candidates with various backgrounds (Gallagher and Marsh, 1988). This provides citizens actually with the opportunity to base their vote on candidate similarities, which is not always the case in single-member constituency systems where certain types of candidates may not run in some districts. A third reason why voter-candidate similarities may exert a larger influence in intra-party elections is that it is a low information context. Parties present long lists of candidates – in Belgium these lists range between 12 and 33 candidates – meaning that voters can choose between many candidates. Therefore, it is impossible for voters to gather information on

all these candidates and consequently they may instead search for other cues in order to fill this information gap, such as familiarity, ballot list position (van Erkel and Thijssen, 2016) and of course the sociodemographic characteristics of the candidates (Sanbonmatsu, 2002).

In sum, voter-candidate similarities can be expected to play a role in the decision-making process of voters when casting preferential votes. In the next section I focus on three such similarities: gender, age, and location. The reason to select these three characteristics is that they are key features in the representation literature and all three are used by Belgian parties when constructing the ballot list (De Winter, 1988). Ideally, I would also include ethnicity, but unfortunately I have no reliable information about the ethnicity of the voters in the sample.

Gender, age, and location

So far, a handful of studies have examined the role of same-gender based voting in intra-party electoral competition. Most of these studies provide evidence for same-gender-based voting. Holli and Wass (2010) and Giger *et al.* (2014) examine Finland and find that men are more likely to vote for men and women for women, although the effect for the former group seems to be stronger than for the latter group. They conclude that 'gender-based voting appears to be an important factor that affects electoral outcomes' (Holli and Wass, 2010: 624). Erzeel and Caluwaerts (2015), Erzeel *et al.* (2018), and Marien *et al.* (2017) find similar results for Belgium. However, not all studies reach the same conclusion. McElroy and Marsh (2010) investigate in Ireland whether women tend to vote more for female political candidates, but fail to find any gender effect. Also studies in the United States and the United Kingdom do not always support a same gender effect, or find a weak effect at best (McDermott, 2009; Campbell and Cowley, 2014). Despite these mixed findings, theoretically same gender can be expected to be a cue voters base their decision on. Gender may be the basis of a social identity, with women (or men) having the feeling that they should give support to candidates from the same gender and cast a symbolic vote. It may also lead to a vote because citizens may feel that someone from their own gender is better able to represent them. Research has shown that the gender composition of parliament indeed influences which issues are on the political agenda (Mansbridge, 1999; Schwindt-Bayer, 2006).

I am also interested in whether women are more inclined to base their vote decision on the gender cue than men. As stated before, underrepresented groups, such as women, who traditionally faced historical and systematic political barriers, may have more incentives to vote for candidates similar to them. They have a stronger sense of identity and more to win by getting represented. Erzeel and Caluwaerts (2015), Holli and Wass (2010), and Giger *et al.* (2014) have the same expectation, but surprisingly find the opposite, demonstrating that men are more likely to vote on the basis of gender than women. This finding could be explained by the design they use. All three studies use voter data, not really taking into account characteristics of the candidates apart from gender. Yet, taking other characteristics of candidates into account is important as there still exist strong gender differences in the position on the ballot list candidates occupy and in their political experience. For instance, in recent Belgian elections only 21 of the 70 Flemish lists had a female candidate positioned first. Similarly, in the Belgian parliament, 40% of the MPs are women. Also with regard to media coverage we still do find a gender bias, even when taking into account alternative explanations (Vos, 2013). Since voters are inclined to vote for candidates with more political experience and for candidates on the first position on the list (van Erkel and Thijssen, 2016), and since these candidates tend to be more often male, the finding that men vote more on male candidates might actually be a construct of the fact that they tend to vote for experienced candidates with a higher ballot list position who get more coverage in the media, rather than that they base their decision on the gender cue. This is in line with previous studies that focus on the candidate level and aim to explain why certain candidates receive more preferential votes. While in empty models with gender they find that men receive more preferential votes than women, this effect changes to women receiving more votes than men when controlling for structural inequalities (Thijssen and Jacobs, 2004).

A recent study by Marien *et al.* (2017) focuses on voters and does take into account the ballot list position of candidates. While their main findings are in line with Erzeel and Caluwaerts (2015) and Holli and Wass (2010), demonstrating that men are more likely to vote for male candidates than women for female candidates, this difference between men and women disappears when focusing only on the first candidate. Men are not more likely to cast a vote for the first candidate on the list when this is a male than women are when this is a female (Marien *et al.*, 2017: 16). This demonstrates the importance of including the supply side. I expect that once one controls for additional characteristics at the supply side that create structural inequalities, a stronger effect of the same gender cue among women will be found, being in line with the theoretical expectation. By bringing demand and supply together, thereby controlling for structural inequalities between male and female candidates, it is possible to better isolate same-gender-based voting and more rigorously test whether women base their vote more often on a same gender cue than men.

Hypothesis 1a: Voters are more likely to cast preferential votes for candidates with the same gender.

Hypothesis 1b: The effect of same gender on the likelihood of casting a preferential vote for a candidate is stronger among female voters than among male voters.

Whereas a number of studies investigate the role of shared gender in the decision-making process of voters, research on same-age voting is scarce. To the best of my knowledge, none of the studies on PR systems have taken age into account, and even in the United Kingdom and the United States studies on this topic are scarce (but see Sigelman and Sigelman, 1982; Campbell and Cowley, 2014). Nevertheless, it can be expected that, similar to gender, some citizens are more inclined to cast a preferential vote for candidates close to their own age. Especially given the longer lifespans of people and the influence of this demographic change on social spending, intergenerational issues are becoming more prominent in politics (Joshi, 2013). This manifests itself for instance by the existence of several pensioners parties in Western Europe. Hence, age may increasingly become a basis around which citizens form a social identity, and consequently may be an important cue for citizens in deciding whom to vote for. I thus expect the likelihood of a preferential vote for a candidate to increase when voter and candidate are from the same age group. I especially expect this effect for the younger and elder, as these groups are numerical underrepresented in parliament (Norris and Franklin, 1997) and may therefore have more reason to support candidates from their own age group.¹

Hypothesis 2a: Voters are more likely to cast preferential votes for candidates from the same age group.

Hypothesis 2b: The effect of same age on the likelihood of casting a preferential vote for a candidate is stronger among younger and older voters than among middle age voters.

Finally, voters may also look for candidates that share local ties. Already in 1949 Key suggested that local ties are of great importance in explaining voting behaviour. This was later confirmed by other studies showing that local voting plays an important role in presidential and parliamentary elections in majoritarian systems, with candidates having a clear home state advantage (Lewis-Beck and Rice, 1983; Gimpel *et al.*, 2008; Campbell and Cowley, 2014). Górecki and Marsh (2012, 2014) also find evidence for local voting in Ireland, demonstrating that the likelihood to rank a candidate on the ballot increases when the geographical distance

¹Note that underrepresentation in the case of age means purely in numbers that they are descriptively less represented in parliament. In the case of gender, underrepresentation also refers to the fact that women historically and systematically have encountered barriers in politics. I go into more detail on this difference in the discussion section.

between voter and candidate decreases. Like gender and age, voting for a candidate from the same municipality may follow an identity logic, as citizens may feel an affection with the place they live and therefore be more inclined to support local candidates. They may have the feeling that candidates from their region are best able to represent their (local) interests. However, unlike gender and age, voting for a candidate from the same location may also be based on a proximity logic (André *et al.*, 2012). Citizens may know a local candidate personally or from shared networks and for that reason cast a preferential vote for that candidate. This type of voting is therefore also sometimes labelled ‘friends and neighbor’ voting (Górecki and Marsh, 2014).

Outside the single transferable vote system of Ireland, almost no studies have investigated the role same-municipality voting plays in the decision-making process of preferential voting. Two exceptions are Put and Maddens (2015) in Belgium and Tavits (2010) in Estonia, but neither of these studies find significant effects. This may be due to the fact that they use an aggregate design, examining whether municipality size influences the relative share of preferential votes of a candidate. However, this effect may be neutralized at the aggregate level, as Lewis-Beck and Rice (1983) demonstrate that, at least in the United States, local ties play a more important role for voters from smaller regions. Both the identity logic and the proximity logic may be stronger in smaller municipalities where voters tend to be more strongly connected to their region, and hence may develop a stronger local identity or be more likely to know the candidate personally. Hence, I expect that when investigating the micro level, by matching the location of voters and candidates, there will be an effect of shared municipality, with a stronger effect for citizens of smaller municipalities.

Hypothesis 3a: Voters are more likely to cast preferential votes for candidates from the same municipality.

Hypothesis 3b: The effect of living in the same municipality on the likelihood to cast a preferential vote for a candidate will be stronger for voters from smaller municipalities.

Case selection

This study uses data from the general 2014 Belgian elections. Belgium is a flexible-list proportional system and has a political constellation with three more or less similar, but separate party systems in the regions of Wallonia, Brussels, and Flanders. This study focuses on Flanders – the Dutch-speaking region of Belgium, reflecting over 60% of the population. In Belgium parties determine the order of the ballot list beforehand, but voters can change this order using preferential votes. More specifically voters can cast two types of votes. Either they vote for the political party thereby agreeing with the list order, a list vote, or they vote for one or more candidates, a preferential vote. Citizens can cast as many preferential votes as they want, as long as candidates are from the same party. Preferential votes are used to determine how many seats each party receives, but also influence the composition of the parliament. Candidates receiving enough preferential votes to surpass a quota get immediately elected. Other candidates can complement their pool of votes using list votes. These votes first go to the number one on the list, until (s)he has enough votes to reach the quota, then to the number two, etc. Once half of the list votes are distributed, the last empty seats, if still any left, go to the remaining candidates with the most preferential votes.

The case of Belgium enables an investigation of the role of voter-candidate similarities in a context with few institutional restrictions to cast such a vote and also analytically forms a good case. The combination of long ballot lists – ranging from 12 to 33 candidates – and ticket-balancing incentives for parties ensures that each ballot list has a wide variation of candidates, with candidates from different geographical locations within the district and different age groups (De Winter,

1988). For gender there is even a perfect balance on each ballot list, as laws introduced in 2002 stipulate gender equality with 50% of the candidates on the list being women, as well as a gender parity rule for the first two candidates on the list. These diverse ballot lists are ideal as it means that voters actually have the opportunity to vote for candidates of the same gender/age/location, since these are always apparent on each ballot. In addition, the fact that voters can cast multiple preference votes makes Belgium analytically suitable. The reason is that there is less dependency between candidates. In systems with single preferential voting, a vote for candidate X automatically implies that the other candidates on that list no longer receive a vote from that citizen. Thus, for their votes candidates are strongly dependent on the strength and electoral success of the other candidates on the list. In systems with multiple preferential voting, such as in Belgium, this is less the case as citizen can cast a vote for as many candidates on the party list as they want. Thus, a candidate can receive a preferential vote from a citizen even when this citizen has also cast a vote for another candidate on that list. In practice, the problem of dependency is not completely averted in Belgium, as most voters still limit the number of preferential votes they cast, but because of multiple preferential voting, dependency is at least minimized. All in all, the context of Belgium, with its long lists of candidates from different sociodemographic backgrounds and the option of multiple preferential voting, allows an investigation of whether citizens base their preferential vote (s) on voter-candidate similarities in a context with few institutional restrictions to cast such a vote.

Research design

Most studies on preferential voting either focus on the level of political candidates (supply side) or on voters (demand side). While studies at the candidate level are useful and necessary when one wants to explain why certain candidates are electorally more successful than others, they do not allow the researcher to model the decision-making process of voters. By focusing on the candidate level, inferences on how voters reach a decision can only be made indirectly and furthermore this approach is not able to investigate how certain factors impact some voters more than others. Since this study is interested in the role of voter-candidate similarities in the decision-making process, focusing on the level of the voter is more appropriate. Yet, even solely focusing on the voter level does not provide the full picture as one has information on the 'demand' side, but not on the 'supply' side. As illustrated in the theoretical section, when testing whether shared gender, age, and location play a role in the decision-making process for preferential voting, one should also control for structural inequalities between candidates. Therefore, I propose to model voter and candidate characteristics together by looking at dyadic relationships. This way the vote decision-making process can be modelled while taking into account characteristics of the available choices.

In order to link voters and candidates, data were gathered on both. The focus is on the 2014 elections that were held for the Federal and Flemish parliaments; both first-order elections. Voter data were gathered using the PartiRep survey, a two-wave election survey in Belgium that followed from a cooperation between a number of Belgian universities. The sample of respondents was drawn on the basis of the National Registry, in order to enhance its representativeness. The most important part of the PartiRep survey that is used here is the simulation ballot. After a face-to-face interview in the pre-election round ($N = 1001$; response rate of 44% in Flanders) all respondents were given a simulation ballot, similar to the real ballot, which they had to fill in on election day. After the elections, they were asked by telephone, using the simulation ballot, how they voted. This makes it possible to measure which candidate(s) they gave a preferential vote to. Since I am interested in motivations behind a preferential vote, only respondents that gave a preferential vote to one or more candidates are included in the analyses.² A total of 816 Flemish

²Voters casting a list vote were excluded for substantial reasons as the subjects needed to have cast a vote for a candidate in order to explain why they voted for this particular candidate. In this sense the population I want to make a claim about are also those voters who cast one or more preferential votes.

respondents (76% of the respondents from the first wave responded in this second wave) took part in the survey. Of these respondents, 394 cast one or more preferential votes in at least one of the elections. This percentage of 48.3% of the respondents is more or less comparable to the national average of voters who cast a preferential vote (55%).³

Data on candidate characteristics were gathered using official documents. For information on the candidates' political party, district and ballot list position, official electoral lists were used. Moreover, data were collected on their gender, age, political experience, residence area, and media coverage. Gender was coded by the researcher using the names of candidates, although in almost all cases background checks were made via the websites of the candidates or their political party. In order to retrieve information on their political experience, operationalized as the political functions they previously occupied or still occupy, websites were used that keep track of all political mandates in Belgium (<http://www.cumuleo.be> and <http://directory.wecitizens.be>). Finally, media data were retrieved using GoPress, a database which archives all Belgian newspapers, to count the number of newspaper articles in which each candidate was mentioned. In the search eight paid Flemish newspapers were included together with the free daily *Metro*.⁴ I gathered data both on the media attention during the campaign (the month before Election day), as well as during the year before the elections.

As said before, I link voter and candidate data. Thus, respondents (voters) are linked to all candidates that ran in their district for the political party they voted for. This means that instead of using voters or candidates as units of analysis, a stacked dataset matrix is used in which the units of analysis are respondent-candidate dyadic pairs. Table 1 gives an (hypothetical) example of the data set for some of the included variables where respondents can vote for three candidates. Ultimately, the data consists of 14,454 dyadic pairs nested in 1310 candidates and 394 voters. It should be noted that one of the assumptions is that voters first select a party and subsequently vote for one or more candidates from this party. This assumption follows from previous literature (Van Holsteyn and Andeweg, 2010) and from the Belgian institutional rules, where preference votes can only be given to candidates belonging to the same party. Of course the assumption may not always hold. Especially party presidents may motivate voters to switch to a different party, thus the votes for these party presidents may involve an inter-party element. However, in general the assumption holds for the other candidates in most cases. In addition, even if the assumption does not always hold, it is unlikely to strongly bias the results as I focus on characteristics that are common on all lists. It is, for instance, unlikely that voters switch parties because they want to vote for a woman, as half of the candidates on the list from their own party are also female.

Since respondent-candidate dyadic pairs are nested simultaneously in respondents and candidates, and given the binary dependent variable, a cross-classified multi-level logistic model is used. This way, variables at the dyad level and controls for candidate characteristics can be modelled simultaneously. I also take into account that the dyadic respondent-candidate pairs are situated in political parties, electoral districts and at the federal or regional level.⁵ In order to remove this variance, fixed controls for parties, districts, and a dummy for the federal elections are added (depicted in Appendix B).

The main independent variables are same gender, same age group, and same municipality, all situated at the lowest level, since they are dyadic in nature. Same gender is a dummy coded 1 if the respondent and the candidate are both male or both female. Same age group is a dummy coded 1 if voter and candidate are from the same age group. I distinguish three age groups:

³The most likely explanation for it being slightly lower is that some respondents who indicated they cast a preferential vote, did not provide further details in the simulation ballot.

⁴These newspapers are *De Morgen*, *De Standaard*, *De Tijd*, *Het Laatste Nieuws*, *Het Nieuwsblad*, *Het Belang van Limburg*, *De krant van West-Vlaanderen*, and *De Gazet van Antwerpen*.

⁵I tested whether results differ between the regional and federal level, but this is not the case.

Table 1. Example of data matrix

Voter	Candidate	Preference vote (0=no, 1=yes)	Gender voter (0=f, 1=m)	Gender candidate (0=f, 1=m)	Gender congruence	Age voter	Age candidate	Same age group	Ballot list position
1	1	0	1	0	0	18	33	0	1
1	2	1	1	1	1	18	56	0	2
1	3	0	1	1	1	18	20	1	3
2	1	1	0	0	1	58	33	1	1
2	2	1	0	1	0	58	56	1	2
2	3	0	0	1	0	58	20	0	3
3	1	0	0	0	1	32	33	1	1
3	2	0	0	1	0	32	56	1	2
3	3	1	0	1	0	32	20	0	3

younger (<30), middle age (31–65), and elder (65+). Given that the choice of these borders for the age groups can impact the results, I also run robustness tests, using alternative ways to operationalize the age groups. In order to determine whether candidate and respondent are from the same municipality, their zip code is used. Belgian zip codes have four digits, with the first three digits indicating the municipality. Thus, same municipality is a dummy variable which is coded 1 if the first three digits of the candidate and the respondent match.⁶ In order to take into account the supply side, I control for several variables that the literature has identified as important explanations for a political candidate's success; ballot list position (van Erkel and Thijssen, 2016), political experience (Put and Maddens, 2015), gender, age (McElroy and Marsh, 2010), and media coverage (Maddens *et al.*, 2006). In addition, dummies for the first and last candidates on the lists are added, as research points out that these candidates receive an additional bonus above the normal ballot list position effect (van Erkel and Thijssen, 2016). Finally, I expect that once more candidates from a municipality are on the same ballot list, voters may vote for only one of these candidates. This could bias the results for candidates from larger municipalities, where there are likely to be more candidates from the same municipality. By adding a variable that measures how many other candidates from the same municipality are on a party ballot list, the model controls for this potential bias.

Characteristics at the voter level are not included in the main model, since they do not directly affect the choice for certain candidates – for a voter these characteristics are stable over the different choices. One exception is the number of preferential votes cast by the respondent, as the likelihood that a respondent voted for a candidate increases when this respondent casts more preferential votes.⁷ However, as speculated in Hypothesis 1b, 2b, and 3b, gender, age, and municipality impact which criteria citizens use to reach their vote decision and therefore moderate certain effects. In later models I therefore do include these variables, but in interaction with the dyadic variables. Appendix A provides a descriptive overview of the main variables.

Results

Table 2 presents the results of the models. In the first model the three main effects of same gender, same age, and same municipality are included. The model supports Hypothesis 1a and shows a positive significant effect of candidates and voters sharing the same gender. This gives an indication that, *ceteris paribus*, women are more likely to vote for female candidates, whereas men are more likely to vote for male candidates. Interesting, however, is that once one looks at the effect size, this claim should be somewhat nuanced. When estimating the predicted

⁶In cities such as Antwerp and Ghent, the first three correspond with the districts within these cities. Hence, for the large cities I investigate whether voters are more likely to vote for candidates from their own district. This is a conservative test of the hypothesis.

⁷The number of preferential votes cast is also linked to political interest. I therefore test what happens when political interest is included in the model, but this effect is not significant. The model can be found in Appendix D.

Table 2. Cross-classified multilevel logit models explaining when voters cast a preferential vote for a candidate

	Model 1 [b (SE)]	Model 2 [b (SE)]	Model 3 [b (SE)]	Model 4 [b (SE)]
Level 1 (dyadic)				
Same gender	0.402 (0.08)**	1.444 (0.14)**	0.402 (0.08)**	0.435 (0.09)**
Same age	0.060 (0.09)	0.044 (0.00)	0.097 (0.20)	0.053 (0.09)
Same municipality	3.346 (0.18)**	3.388 (0.18)**	3.356 (0.18)**	3.903 (0.21)**
Candidate characteristics				
Ballot list position	0.098 (0.01)**	0.096 (0.01)**	0.098 (0.01)**	0.100 (0.01)**
First candidate	2.602 (0.24)**	2.676 (0.23)**	2.610 (0.25)**	2.564 (0.25)**
Last candidate	1.189 (0.29)**	1.268 (0.28)**	1.197 (0.29)**	1.187 (0.30)**
Flemish parliament	0.535 (0.17)**	0.419 (0.16)*	0.543 (0.17)**	0.566 (0.17)**
Federal parliament	0.807 (0.17)**	0.666 (0.16)**	0.818 (0.17)**	0.789 (0.17)**
Mayor	0.016 (0.16)	0.272 (0.15)	0.021 (0.16)	0.049 (0.16)
Minister	0.001 (0.27)	-0.047 (0.26)	-0.007 (0.28)	0.135 (0.27)
Media coverage(log)	0.051 (0.03)*	0.075 (0.03)**	0.050 (0.03)*	0.049 (0.03)
Media coverage t-1(log)	0.032 (0.03)	0.076 (0.03)*	0.033 (0.03)	0.032 (0.03)
Other candidates in municipality	0.071 (0.08)	0.075 (0.08)	0.076 (0.08)	0.088 (0.08)
Number of preferential votes	0.488 (0.01)**	0.490 (0.02)**	0.486 (0.01)**	0.491 (0.02)**
Interactions				
Male voter		1.128 (0.15)**		
Same gender × male voter		-2.119 (0.24)**		
Age [ref = middle age (31–65)]				
Young voter (≤30)			-0.125 (0.23)	
Elder voter (65+)			0.179 (0.22)	
Same age group × young voter			0.727 (0.43)	
Same age group × elder voter			0.474 (0.73)	
Municipality size (per 1000 residents, centred)				-0.000 (0.00)
Same municipality × municipality size				-0.033 (0.01)**
Constant	-4.5382 (0.27)**	-5.301 (0.29)**	-4.687 (0.32)**	-4.647 (0.28)**
σ^2_{voters}	0.078	0.112	0.046	0.097
$\sigma^2_{candidates}$	0.888	0.783	0.896	0.897

N (dyadic pairs): 14,454. N (candidates): 1310. N (voters): 394.
 Controlling for party and district dummies.
 *P < 0.05, **P < 0.01.

probabilities, keeping all other variables fixed on their mean, the results show that the probability of voting for a candidate is 2.8% for candidates from the other gender and 4.2% for candidates from the same gender. Thus, the model points towards a same gender effect, but this effect is moderate and only one of many factors that play a role in the decision-making process.

However, as posited in Hypothesis 1b, differences may exist between male and female voters. Therefore, in model 2 an interaction between same gender voting and the gender of voters is added. This interaction is significant and shows that for women sharing the same gender with a candidate is indeed an important factor in the decision-making process. When calculating the margins it shows that the likelihood that women vote for a candidate is 2.2% if the candidate is male and 8.8% if the candidate is female, *ceteris paribus*. This supports Hypothesis 1b. Interesting is that for men gender cues do not seem to influence their decision-making process. The effect of same gender becomes insignificant and even negative for men. The analysis thus indicates that men do not vote more for male candidates because of gender-motivated reasons, but that rather this is a construct of citizens being more likely to vote for incumbent politicians on a high ballot list position, candidates who often happen to be male due to structural inequalities. The fact that the coefficient even becomes negative (but not significant) rather suggests that some men prefer to vote for a women, perhaps as a symbolic action because they feel women should be better represented. Hence, the mechanisms of identity voting on the one hand, and symbolic/ ‘sympathy’ voting on the other hand, may balance each other out, resulting in an insignificant effect of the same-gender cue for men. These findings contrast previous studies that found the opposite (e.g. Holli and Wass, 2010) and demonstrate the importance of taking the supply side into account when modelling the preferential vote decision-making process.

When focusing on the effect of same age group in model 1, no significant effect is found. In other words, in contrast to gender, there is no evidence that citizens are more inclined to cast preferential votes for candidates from the same age group. Hypothesis 2a should therefore be rejected. Model 3 finds no support for Hypothesis 2b either, which stated that only younger or elder citizens vote on the basis of same age. None of the interaction terms between same age and respectively younger and elder voters are significant at the $P < 0.05$ level, although there may be a small effect for younger voters as the effect is significant at the $P < 0.1$ level and statistical power at the voter level is low. Given that there may be debate on the classification of the age groups, I performed extra robustness tests of the age effect in Appendix C using alternative classifications. I also test what happens when rather than focusing on same age groups, a continuous measure of age similarity is used. This is operationalized as the absolute distance between voter and candidate, where 0 means that voter and candidate have exactly the same age and 20 that they differ by 20 years. The Appendix shows that the results for same age do not change when using different age borders or when splitting the middle-age groups in two groups. Notable is that when the continuous measure of age similarity is used, there is again a significant, but small effect of this measure in interaction with young voters. This finding, together with the significant effect at the $P < 0.1$ level in the main model, could indicate that perhaps for younger voters age does play a minor role, and that they may be less likely to vote for candidates much older than themselves.

Finally, I examine the effect of living in the same municipality. Do candidates have a home region advantage? Model 1 indicates that there is indeed a positive significant effect of candidates and voters living in the same municipality. Thus, the idea of a preferential vote as a local vote, as theorized by Hypothesis 3a, is supported. In other words, an important determinant in the decision-making process of voters when deciding whether they cast a preferential vote for a candidate, is whether this candidate is from the same area. Whereas the probability that a respondent casts a vote for a candidate that is not from the same municipality is 3.2%, keeping all other variables at the mean or at a score of 0, this probability increases to 48.4% when the candidate is from the same region. Shared municipality thus seems to play a much larger role in the voters' mind than shared gender or age. Its effect is even stronger than the effect of voting for the first candidate on the list, where the predicted probabilities increase from 3.1% to 29.9%. Moreover, I find strong evidence for Hypothesis 3b. Model 4 includes an interaction between shared municipality and the size of the municipality (or district of the municipality) per 1000 respondents, centred around the grand mean. The interaction is negative, indicating that local voting is strongest in smaller municipalities. This is plotted in Figure 1. The plot shows that for voters living in a municipality with 20,000 residents, the likelihood to vote for a candidate increases from 3.1% to 61.1% when that candidate is from the same municipality. For voters from municipalities with 5000 residents it increases from 3.1% to 71.9%, whereas for voters in municipalities with 100,000 residents it only increases from 2.9% to 5.6%. These findings strongly support Hypothesis 3b.

Conclusion and discussion

This study investigated to what extent citizens are inclined to vote for candidates that are like them. It did so in an institutional context in which citizens are relatively unconstrained to vote on the basis of sociodemographic cues and argued that voter-candidate similarities play a role in explaining why voters select some candidates over others. Moreover, the study investigated whether underrepresented groups, especially women, are more guided by voter-candidate similarities in their preferential vote decision-making process than overrepresented groups. To do so, a new approach was introduced which simultaneously models voter and candidate characteristics.

The findings of this study support the claim that citizens are more likely to cast votes for candidates when these candidates share certain traits with them. Voter-candidate similarities play

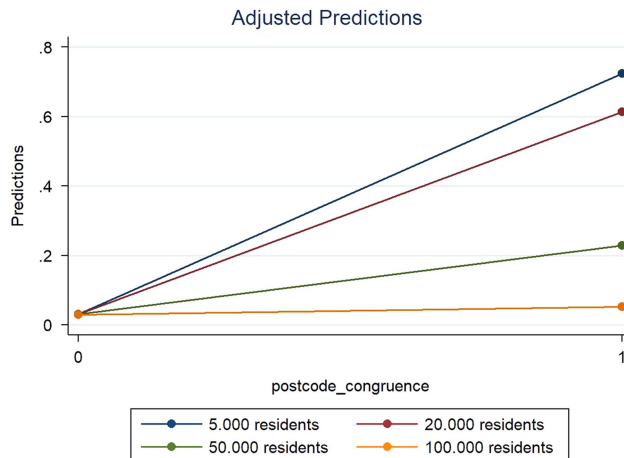


Figure 1. Predicted probabilities of living in the same municipality by residence size.

an important role in the decision-making process around preferential voting, especially for underrepresented groups, such as women. The results show that female voters are more likely to be guided by the gender of candidates than men. This can be either an instrumental vote, as a way to increase their substantive representation in parliament, or can be more symbolic, as a way of showing support.

However, there are differences between traits. Whereas there is a strong effect of shared municipality and a moderate effect of same gender, there is no effect, or possibly only a small effect for younger voters, of same age. I can only speculate about these differences, but one explanation may be variation in knowledge about the traits. Gender can in most cases be inferred from the names on the ballot list. For citizens it is also not difficult to gather information on which candidates are from their municipality, as candidates are more likely to campaign in their own municipality, get coverage in the local news and citizens may recognize them from local elections or local political functions. The age of candidates is more difficult to know since this cannot be inferred from the ballot list and is a characteristic that does not play a central role in most campaigns. An alternative explanation is that the mechanisms behind each of these traits work slightly differently. The fact that there is an effect for gender, but not for age, may be because, unlike gender, age has never been a basis on which groups have encountered historically and systematic barriers in politics. This could explain the finding that women vote more for female candidates, as the need for women to cast a symbolic or instrumental vote may simply be stronger. Additionally, whereas age and gender are linked to identity, location is also based on proximity as citizens may know local candidates personally or from shared networks (André *et al.*, 2012). This suggests that the proximity logic may be stronger than the identity logic. Further research is required to better tease out the mechanisms behind each of these traits.

This study also illustrates the need to control for the supply side when investigating the influence of candidate-voter similarities on voting behaviour. Given that many factors at the supply side are still unequally distributed, such as ballot list position, political experience, and media coverage, and given that these characteristics matter in the decision-making process of voters, results will be biased when not controlling for them. This makes it difficult to isolate voter-candidate similarities from other factors. Substantially, it shows that due to structural inequalities at the supply side, identity votes will not necessarily lead to a better representation of certain groups. Since identity voting is only one factor in the decision-making process around preferential voting, and since citizens are still guided by other factors, a more equal distribution between groups on the ballot list, in political functions and in media coverage, is necessary in order for identity voting to really be a mechanism to overcome inequality.

The findings leave some questions open that ought to be investigated in future research. First, this study focused on gender, age, and geographical proximity. Nevertheless, other similarities may also play a role in the decision-making process of voters, such as education, social class, and especially ethnicity. Unfortunately, due to data limitations I could not include these characteristics.

Second, this study focused on Belgium, a context where, because of its diverse lists and multiple preferential voting, there are few institutional restrictions to base a vote on similarities. While this case demonstrates that in such an unconstrained context similarities between voters and candidates play a role when voters decide who to cast a preferential vote for, it cannot give an answer to whether this also is the case in more restricted systems. Although quite a few countries with an open- or flexible-list system have multiple preferential voting similar to Belgium, such as Luxembourg, Norway, and Switzerland, other countries, such as Austria, Finland, and the Netherlands, have systems where voters are only allowed to cast a single preferential vote. In this situation voters are more restricted and face a tradeoff between different decision criteria, as they can only cast a vote for one candidate. Future studies should investigate whether also in the more restricted context of a single preferential vote, there is an effect of voter-candidate similarities on the decision-making process of voters. This would provide a deeper insight into the relative strength of factors such as gender, age and location in comparison to other factors such as ballot list position, political experience, or media attention. In addition, it would provide more insights on how electoral rules may enhance or temper the influence of voter-candidate similarities.

All in all, the findings of this study show that preferential votes are used by underrepresented groups, most notably women, as an instrument to support their own members and thereby as a way to reach more descriptive representation. However, due to structural differences between groups at the supply side, this potential for descriptive representation is often not realized. Moreover, the fact that parties determine the order of the ballot list and hold strong control over who gets elected, limits the influence of identity votes. The results of this study do suggest that once systems start to adopt more open lists, and once a more equal distribution between groups on the ballot list and in the coverage of the media is reached, voting on the basis of shared similarities could lead to a better descriptive composition of parliament. This could deliver a more equal representation of underrepresented groups, increasing their substantive representation, and in the long run can help to increase the trust of these groups in politics and ultimately democracy as a whole.

Acknowledgements. The author thanks Yves Dejaeghere, Kirsten Van Camp, Nadja Wehl, and the anonymous reviewers for their useful comments on earlier versions of the manuscript.

Funding. This work was supported by the Research Foundation Flanders under grant number G026513N.

Supplementary material. To view supplementary material for this article, please visit <https://doi.org/10.1017/S175577391800022X>

References

- André A, Wauters B and Pilet J-B (2012) It's not only about lists: explaining preference voting in Belgium. *Journal of Elections, Public Opinion & Parties* 22, 293–313.
- André A, Pilet J-B, Depauw S and Van Aelst P (2013) De partij, de leider of een “gewone” kandidaat? Het gebruik van de voorkeurstem bij de gemeenteraadsverkiezingen van 2012. in R Dassonneville, M Hooghe, S Marien and J-B Pilet (eds), *De lokale kiezer: Het kiesgedrag bij de Belgische gemeenteraadsverkiezingen van oktober 2012*, Brussel: ASP, pp. 119–142.
- André A, Depauw S, Shugart MS and Chytilek R (2017) Party nomination strategies in flexible-list systems: do preference votes matter? *Party Politics* 23, 589–600.
- Banducci SA and Karp JA (2000) Gender, leadership and choice in multiparty systems. *Political Research Quarterly* 53, 815–848.
- Campbell R and Cowley P (2014) What voters want: reactions to candidate characteristics in a survey experiment. *Political Studies* 62, 745–765.

- Cutler F** (2002) The simplest shortcut of all: sociodemographic characteristics and electoral choice. *The Journal of Politics* **64**, 466–490.
- De Winter L** (1988) Belgium: democracy or oligarchy? in M. Gallagher and M. Marsh (eds) *Candidate Selection in Comparative Perspective*, London: Sage, pp. 20–46.
- Dolan K** (1998) Voting for women in the “year of the woman”. *American Journal of Political Science* **42**, 272–293.
- Dolan K** (2008) Is there a “gender affinity effect” in American politics? Information, affect, and candidate sex in U.S. house elections. *Political Research Quarterly* **61**, 79–89.
- Erzeel S and Caluwaerts D** (2015) Is it gender, ideology or resources? Individual-level determinants of preferential voting for male or female candidates. *Journal of Elections, Public Opinion and Parties* **25**, 265–283.
- Erzeel S, de Leeuw S, Marien S and Rihoux B** (2018) Gender-based voting, in K. Deschouwer (ed) *Mind the Gap: Political Participation and Representation in Belgium*, London: Rowman & Littlefield International/ECPR Press, pp. 227–250.
- Gallagher M and Marsh M** (eds) (1988) *Candidate Selection in Comparative Perspective: The Secret Garden of Politics*, London: Sage.
- Giger N, Holli AM, Lefkofridi Z and Wass H** (2014) The gender gap in same-gender voting: the role of context. *Electoral Studies* **35**, 303–314.
- Gimpel JG, Karnes KA, McTague J and Pearson-Merkowitz S** (2008) Distance-decay in the political geography of friends-and-neighbors voting. *Political Geography* **27**, 231–252.
- Górecki MA and Marsh M** (2012) Not just “friends and neighbours”: canvassing, geographic proximity and voter choice. *European Journal of Political Research* **51**, 563–582.
- Górecki MA and Marsh M** (2014) A decline of “friends and neighbours voting” in Ireland? Local candidate effects in the 2011 Irish “earthquake election. *Political Geography* **41**, 11–20.
- Holli AM and Wass H** (2010) Gender-based voting in the parliamentary elections of 2007 in Finland. *European Journal of Political Research* **49**, 598–630.
- Joshi DK** (2013) The representation of younger age cohorts in Asian parliaments: do electoral systems make a difference? *Representation* **49**, 1–16.
- Katz RS** (1985) Intraparty preference voting, in B. Grofman and A. Lijphart (eds) *Electoral Laws and their Political Consequences*, New York: Agathon Press, pp. 85–103.
- Key VO** (1949) *Southern Politics in State and Nation*, New York: Knopf.
- Lewis-Beck MS and Rice TW** (1983) Localism in presidential elections: the home state advantage. *American Journal of Political Science* **27**, 548–556.
- Maddens B, Wauters B, Noppe J and Fiers S** (2006) Effects of campaign spending in an open list PR system: the 2003 legislative elections in Flanders/Belgium. *West European Politics* **29**, 161–168.
- Mansbridge J** (1999) Should blacks represent blacks and women represent women? A contingent “yes”. *The Journal of Politics* **61**, 628–657.
- Marien S, Wauters B and Schouteden A** (2017) Voting for women in Belgium’s flexible list system. *Politics & Gender* **13**, 305–335.
- McDermott ML** (1997) Voting cues in low-information elections: candidate gender as a social information variable in contemporary United States elections. *American Journal of Political Science* **41**, 270–283.
- McDermott ML** (2009) Voting for myself: candidate and voter group associations over time. *Electoral Studies* **28**, 606–614.
- McElroy G and Marsh M** (2010) Candidate gender and voter choice: analysis from a multimember preferential voting system. *Political Research Quarterly* **63**, 822–833.
- Norris P and Franklin M** (1997) Social representation. *European Journal of Political Research* **32**, 185–210.
- Piliavin JA** (1987) Age, race, and sex similarity to candidates and voting preference. *Journal of Applied Social Psychology* **17**, 351–368.
- Pitkin HF** (1967) *The Concept of Representation*, Berkeley, CA: University of California Press.
- Plutzer E and Zipp JF** (1996) Identity politics, partisanship, and voting for women candidates. *Public Opinion Quarterly* **60**, 30–57.
- Pomper G** (1975) *Voter’s Choice: Varieties of American Electoral Behavior*, New York: Dodd, Mead.
- Put G-J and Maddens B** (2015) The effect of municipality size and local office on the electoral success of Belgian/Flemish election candidates: a multilevel analysis. *Government and Opposition* **50**, 607–628.
- Sanbonmatsu K** (2002) Gender stereotypes and vote choice. *American Journal of Political Science* **46**, 20–34.
- Schwindt-Bayer LA** (2006) Still supermadres? Gender and the policy priorities of Latin American legislators. *American Journal of Political Science* **50**, 570–585.
- Sigelman CK, Sigelman L, Walkosz BJ and Nitz M** (1995) Black candidates, white voters: understanding racial bias in political perceptions. *American Journal of Political Science* **39**, 243–265.
- Sigelman L and Sigelman CK** (1982) Sexism, racism, and ageism in voting behavior: an experimental analysis. *Social Psychology Quarterly* **45**, 263–269.

- Tajfel H and Turner J** (1979) An integrative theory of intergroup conflict, in Austin WG and Worchel S (eds) *The Social Psychology of Intergroup Relations*, Chicago, IL: Nelson-Hall, pp. 33–47.
- Tavits M** (2010) Effect of local ties on electoral success and parliamentary behaviour the case of Estonia. *Party Politics* **16**, 215–235.
- Teney C, Jacobs D, Rea A and Delwit P** (2010) Ethnic voting in Brussels: voting patterns among ethnic minorities in Brussels (Belgium) during the 2006 local elections. *Acta Politica* **45**, 273–297.
- Thijssen P and Jacobs K** (2004) Determinanten van voorkeurstemproporties bij (sub-)lokale verkiezingen: de Antwerpse districtraadsverkiezingen van 8 oktober 2000. *Res Publica* **4**, 460–485.
- Van Erkel PFA and Thijssen P** (2016) The first one wins: distilling the primacy effect. *Electoral Studies* **44**, 245–254.
- Van Holsteyn JJM and Andeweg RB** (2010) Demoted leaders and exiled candidates: disentangling party and person in the voter's mind. *Electoral Studies* **29**, 628–635.
- Vos D** (2013) The vertical glass ceiling: explaining female politicians' underrepresentation in television news. *Communications: The European Journal of Communication Research* **38**, 389–410.