Accountability and alternation: how wholesale and partial alternation condition retrospective voting

Otjes, S.P.; Stiers, D.

Accountability and alternation. How wholesale and partial alternation condition retrospective voting

Simon Otjes

Leiden University, the Netherlands; Groningen University, the Netherlands

Dieter Stiers

Centre for Political Science Research, KU Leuven, Belgium

Abstract

Holding the government accountable is a crucial function of elections. The extent to which voters can actually do so depends on the political system. One element that may influence the likelihood that voters hold the government accountable is the difference between wholesale and partial alternation. Prominent political scientists like Mair, Bergman and Strøm and Pellegata and Quaranta propose that in countries with wholesale alternation voters are better able to hold governments accountable because in essence voters have the choice to keep their current government or 'throw the rascals out'. However, this relationship has not been tested. We examine the relationship between partial and wholesale alternation and retrospective voting in a large-N cross-country study. We show that the association between government satisfaction and vote choice is stronger in countries with wholesale alternation than in systems with partial alternation.

Keywords

electoral accountability, retrospective voting, wholesale/partial alternation

Introduction

Holding the government accountable is a crucial function of elections. The extent to which voters can actually do so depends on the political system (Anderson, 2000; Hobolt et al., 2013). One element that may influence the likelihood that voters hold the government accountable is the difference between wholesale and partial alternation (Bergman and Strøm, 2011; Mair, 2008; Pellegata and Quaranta, 2018). In countries such as the United Kingdom, where party competition is characterised by wholesale alternation, it becomes obvious quickly after the votes are counted whether the current government party holds its majority and can continue to govern or whether the opposition takes over. The same is true for Denmark, except that here two blocs of parties instead of two individual parties vie for power. In countries with partial alternation like Belgium and Germany, election results say little about the upcoming government. In these systems, backroom negotiations rather than the election results form the crucial forum that

determines who gets into government (Mair, 2011). In partial alternation systems, after elections, some parties leave the government, some parties enter government, and usually at least one party stays in government. The outcome of this process is not yet clear on election night.

Bergman and Strøm (2011) propose that democratic accountability is greater in systems with wholesale alternation. We understand accountability as the extent to which voters are able to sanction the government if its actions are contrary to their interests (Manin et al., 1999). In systems

with wholesale alternation, voters have a clear choice: they can support the government or they can discharge it if they do not like what it has done. The idea that electoral accountability is greater in systems with wholesale alternation has not been tested empirically. Therefore, we will set out examine *to what extent voters in systems characterised by wholesale alternation are more likely to hold their government accountable compared to voters in systems with partial alternation.*

In the political science literature, electoral accountability is often understood as retrospective voting (Fiorina, 1978; Key, 1966; Lewis-Beck and Stegmaier, 2000). The underlying idea is that voters who are satisfied with the government's accomplishments will support the governing party or parties in the elections. There is increasing attention on how political institutions condition retrospective voting (Silva and Whitten, 2017). The crucial idea is that the alternative options voters can choose from, determine the likelihood that they will punish the incumbent (Anderson, 2000). With our study, we contribute to the literature studying clarity of political responsibility – i.e. the literature finding that the clarity of responsibility of a political system affects the ability to attribute responsibility for the country's state of affairs to one party or the other (Powell and Whitten, 1993). Initially, authors like Powell and Whitten (1993), Schwindt-Bayer and Tavits (2016), and Whitten and Palmer (1999), tended to focus strongly on formal institutional rules when examining clarity of responsibility. More recently, Hobolt et al. (2013) moved the perspective away from the static institutional rules to the composition of government. They argue that the composition and cohesiveness of the ruling government, rather than stable institutional characteristics, are important to understand why retrospective voting is more common in some countries than in others. They label this *government clarity*. Since then, a number of researchers have employed more dynamic and more specific indicators of government clarity (e.g., Stiers and Dassonneville, 2020). With our present study, we aim to contribute to this development in the literature by arguing that the nature of party competition in terms of wholesale and partial alternation may shape the level of retrospective voting. We suggest that it is important for voters to be able see the effects of their vote (Otjes and Willumsen, 2019).

On the one hand, in a wholesale alternation system, voters are more aware of their options to either support the incumbent government or to vote against it and replace it entirely. In other words, voters' motivation is higher since their vote may have a considerable effect on the composition of the next government. On the other hand, in partial alternation systems, voters are less certain that their vote will influence the composition of the government. One of our main contributions is hence that we do not just look at cohesiveness of the incumbent government; we add to this by looking at the extent to which voters can expect on the

basis of previous government alternation that, if enough voters cast a vote against the government as they do themselves, that government will not return. This clarity of the options that are presented to the voters should make them more likely to make the effort to monitor incumbent performance and vote accordingly.

Mair (1997, 2001, 2011) provided a theoretical framework on the concept of wholesale and partial alternation, which has been used by a number authors to understand policy outcomes (Campos and Horvárth, 2012a, 2012b; Green-Pedersen, 2002; Horowitz et al., 2009; Milanovic et al., 2010) and patterns of political cooperation within the political elite (Anthonsen and Lindvall, 2009; Green-Pedersen, 2004; Louwerse et al., 2017; Meyer-Sahling and Veen, 2012; Otjes and Rasmussen, 2017). Yet, this distinction has not played a major role in the study of party-vote relations (but see Otjes and Willumsen, 2019 and Pellegata and Quaranta, 2018). The purpose of this study is to assess the effect of wholesale and partial alternation on retrospective voting. We do not claim to offer a completely new theory of retrospective voting. Rather, we believe that within the existing literature wholesale and partial alternation can be a small but significant addition that can help to understand how contextual variables impact the relationship between satisfaction with the government and voting behaviour.

This article has the following structure: firstly, we discuss what we know about wholesale and partial alternation and retrospective voting in greater detail. Consequently, we bring those two literatures together. Next, we look at our data and our modelling strategy. We use the Comparative Study of Electoral Systems which covers 55 elections in 30 countries. From there we move on to discuss our results. The final section will draw a number of conclusions about the effect of the patterns of party competition on retrospective voting.

Wholesale and partial alternation

The idea that there is a difference between party systems where voters have a choice between 'alternative teams of governors' and force a 'clear and abrupt' change in government composition or between systems where the relationship between the election results and government composition is weaker date back to the 1970s (Finer, 1975: 31; Rokkan, 1970: 93). Despite prominent political scientists acknowledging this difference (Ieraci, 2012; Lundell, 2011; Mair, 1997), it has not been a prominent element in political science theory about the functioning of democracy. Yet, what sets a party system apart from a collection of parties is the way in which these parties interact when competing for government (Sartori, 1976: 39). This in turn shapes the ability of voters to influence the composition of the government and hold it accountable.

Wholesale alternation (sometimes called perfect government turnover) has often been observed in two-party systems. It is common in Commonwealth systems such as the United Kingdom. Sartori (1976: 165) even considers the alternation of government as the definitive feature of 'twopartism'. Yet, wholesale alternation can also occur in multiparty systems. For example, in Denmark a left-wing and a right-wing bloc compete for power. Another example is Spain, where the major party of either the left or the right often governs as a minority government supported by smaller parties. In other countries, government formation can be characterised as partial alternation or limited government turnover. This occurs in Germany and countries that have a similar party system in Western Europe, such as Luxembourg, Austria, the Netherlands and Belgium, but also in countries like Finland and Estonia. During the French Fourth Republic, government alternation was also partial; when it moved to the Fifth Republic, one of the most remarkable changes was away from partial alternation to wholesale alternation. A similar change can be observed in Italy when comparing the 'First' and 'Second' Republic (De Giorgi and Marangoni, 2015; Verzichelli and Cotta, 1999). Wholesale alternation occurs in two-party systems but also in multiparty systems, in particular in those that are strongly polarised. Partial alternation tends to occur in systems with a large number of parties but a low level of polarisation (Otjes, 2018).

Government alternation is closely related to the quality of democracy and democratic stability (Casal Bértoa and Enyedi, 2016; Cheibub et al., 2009; Kaiser et al., 2002; Lundell, 2011; Otjes and Willumsen, 2019). A key aspect of democracy as 'government of the people' is the ability of voters to hold the governing parties accountable (Manin et al., 1999). In a system with wholesale alternation voters have the ability to 'throw the rascals out' (Mair, 2008). In a system where partial alternation is the norm, voters may punish the governing parties at the polls, but these may still return to government through the byzantine process of government formation as a result of partial alternation (Mair, 2011). For instance, in 2002 the Austrian Freedom Party lost almost two-thirds of its seats but still stayed in government, and in 2010 the Dutch Christian-Democrats lost nearly half of their seats but continued to govern. The weak link between election outcomes and government formation can be an important source of voters' frustration with the political system (Irwin and Van Holsteyn, 2011).

We know surprisingly little about how this difference between wholesale and partial alternation shapes the relationship between voters and parties. Pellegata and Quaranta (2018) find indeed that during periods of economic downturn, wholesale alternation is more likely. Otjes and Willumsen (2019) show that the difference between wholesale and partial alternation affects turnout, boosting turnout in systems with proportional electoral systems but decreasing turnout in systems with disproportional electoral systems.

However, these two studies do not rely on individual-level data but look at aggregate election results instead.

Retrospective voting

In essence, the theory of retrospective voting proposes that whether or not voters vote for the government parties depends on their assessment of the government's performance (Fiorina, 1978; Key, 1966; Lewis-Beck and Stegmaier, 2000; Strøm, 2000). The more satisfied voters are with the government, the more likely they are to vote for one of the incumbent parties. Authors like Fiorina (1981), Healy and Malhotra (2013) and Lewis-Beck and Stegmaier (2007) have investigated retrospective voting and convincingly demonstrated that the assessment of government performance plays a role in voting behaviour. Most of these studies focused on the state of the economy (Lewis-Beck and Stegmaier, 2013; Nannestad and Paldam, 1994; Stubager et al., 2014). Yet, other policy domains are important to voters as well (Singer, 2011). This is supported by other studies that have found evidence of retrospective voting in a variety of domains (Crisp et al., 2014; Ecker et al., 2016; Fowler and Hall, 2018; Gasper and Reeves, 2011; Karol and Miguel, 2007). De Vries and Giger (2014) and Stiers (2019) move away from government specific performance evaluations and instead look at voters' assessment of government performance 'in general'.

While the relationship between government performance evaluations and the vote is quite strong, there are both individual and country-level factors that influence voters' ability to hold their government accountable (Anderson, 2007a). At the country level, institutional complexity may blur the lines of responsibility and may therefore weaken the role of performance evaluation in voting (Silva and Whitten, 2017). If voters are convinced that a particular party has total control over policy-making, they are more likely to assign responsibility for policy outcomes to that party (Powell and Whitten, 1993: 398). Hobolt and colleagues (2013) distinguish between two different dimensions of the political system: the more static 'institutional clarity', and the more dynamic 'government clarity'. Institutional clarity, on the one hand, is the extent to which government power is concentrated according to the formal institutional rules. Research has since examined the effect of power concentration on the relationship between the government performance and the attribution of responsibility. Studies have examined the 'vertical' distribution of power between different levels of government and the 'horizontal' distribution of power between the executive and the legislative, (Bengtsson, 2004; Cutler, 2008; Nadeau et al., 2002; Whitten and Palmer, 1999). Government clarity, on the other hand, is the extent to which the government is cohesive. Hobolt et al. (2013) propose that voters will be inclined to punish or reward a political party for the state of the country. If one party or a cohesive coalition

holds power, it is easier to assign blame than if the government consists out of many, ideologically distinct parties. Overall, they indicate that retrospective voting is more common when government clarity is evident.

What all these studies have in common is that they look at the extent to which voters are able to attribute responsibility to the government. What they fail to take into account however, is the extent to which citizens have a meaningful choice at the elections. Anderson (2000) proposes that the clarity of available alternatives shapes the extent to which voters are able to effectively hold the government accountable. He argues that the larger the number of parties that run in an election, the less clear the coalition which can be formed after the election. If there are only two alternatives, the choice voter have, is between continuing the current government or choosing a new team (Anderson, 2000: 156). Anderson (2000) and Bengtsson (2004) provide empirical support for the link between retrospective voting and the fractionalisation of parliament. Recently, Stiers and Dassonneville (2020) found that increasing polarisation between the government and the opposition helps voters to distinguish parties, which in turn positively moderates retrospective voting. Hence, if it is clear to voters that they can alter government policies in the future by voting for an ideologically distinct party. Thus, the accountability function of elections is strengthened (Hellwig, 2010). All these studies emphasise the important role of dynamic contextual factors in retrospective voting.

Bringing retrospective voting and government alternation together

The literature on wholesale and partial alternation proposes that different countries have different norms and expectations about how government composition. On the one hand there are countries where the expectation is that if the government loses its majority, the opposition will come to power. On the other hand, there are countries where at least one of the governing parties stays in office after a months-long formation, even if the government loses its majority.

Citizens are likely to use the latest government formation as a heuristic, a judgmental shortcut to predict future government formations (Sniderman et al., 1991). Crucial here is that citizens' experiences with the political system, more than theoretical possibilities, shape their expectations about the system they live in (Anderson, 2007b; Mishler and Rose, 2002). Specifically, citizens' expectations about the functioning of their democratic system, are shaped by their own experience of the system (Heyne, 2018: 58). That is: people learn the rules of game by seeing how it is played.

From the perspective of retrospective voting, the main distinction in the political landscape is between the government and the opposition (Key, 1966; Stiers, 2019). The extent to which voters' perceptions match this schema is

crucial. If the latest government was formed by a government displacing the sitting government, voters are more likely to expect that the current election will offer a similar choice, and hence that they can vote the incumbent government out of office if they would want to do so. If, in contrast, the latest government was formed by some parties from the previous government and some from the opposition, voters are less likely to expect that they can punish the incumbent very strongly, because government composition may not change completely. These expectations may feed into voter behaviour (Otjes and Willumsen, 2019): the motivation to hold the government accountable is possibly stronger when citizens expect – on the basis of previous experience – that wholesale alternation occurs. Citizens who have not experienced the possibility to 'send the rascals out', the notion that by voting they have an effect on government formation is more foreign.

We expect that a previous experience with wholesale alternation boosts the probability that citizens take government evaluations into account. In summary:

Wholesale-Partial Retrospective Voting Hypothesis: If the most recent government alternation after elections was wholesale alternation, the effect of retrospective performance evaluations on voting for government parties is stronger than after partial alternation.

Data and methods

To test our hypothesis empirically, we use the data of the Comparative Study of Electoral Systems (CSES). The CSES data consist of a collection of national election surveys including a common set of questions. Hence, using these data, we employ electoral surveys from a variety of different countries and contexts. This makes the CSES data well suited for our purposes here, as we need sufficient variation in government alternation patterns.

Our hypothesis brings forward two variables of interest. First, as we investigate the extent to which voters hold incumbents accountable in elections, we need a measure of the voter's retrospective evaluation of the government's performance. Performance indicators of different domains have been used to investigate retrospective voting, and the most commonly investigated indicator is a voter's evaluation of the evolution of the economy (Lewis-Beck and Stegmaier, 2013; Nannestand and Paldam, 1994). However, as Singer (2011) indicates, voters not only care about the economy and other domains can be important to voters as well. Therefore, we use a general performance measure that was included in the second and third module of the CSES: '*Now thinking about the performance of the [government in [capital]/president] in general, how good or bad a job do you think the government/president in [capital]] has done over the past [number of years since the last government took office] years? Has it done a very good job? A good job? A bad job? A very bad job?*' This question

hence leaves it up to the voter to think about a domain relevant to them in evaluating the incumbent's performance, and previous research has indicated that these general evaluations are important to explain voting behaviour (De Vries and Giger, 2014; Stiers, 2019; Stiers and Dassonneville, 2020). This question is available in the second and third module of the CSES (CSES 2015a, 2015b). However, given the prevalence of economic voting specifically, we will report a robustness test using retrospective evaluations of the economy as well, which we derive from the CSES module 1 and 4.

The second variable of interest is an indicator of government alternation. Different indicators can be used to measure alternation. While we use these indicators to investigate the same characteristic of political systems, they imply a different understanding of politics by its citizens (Otjes and Willumsen, 2019). Therefore, we will test the conditioning effect of two different measures of alternation. Our first measure (*Latest Government Change*) is a dichotomous indicator of whether the most recent change of the government composition after elections was wholesale (code 1) or partial (code 0).[1] We expect that the underlying mechanism is when voters experience one kind of government alternation, they will expect the same kind of government alternation in the future (Otjes and Willumsen, 2019). Therefore, we will focus on the effects of recent alternation on voters' willingness to hold incumbents accountable for their performance. It seems reasonable that voter refer to the latest change of government as a heuristic: voters have relative short memories (Achen and Bartels, 2016).[2]

We additionally test the robustness of the results using two alternative operationalisations of this measure. The first alternative operationalisation looks at the share of government changes in the previous 10 years, whether it was partial or wholesale (*Share Government Changes*), instead of just the latest change in government. This measure differs from the main measure in how it assumes voters develop expectations about the future: in the main measure we expect that the latest election overrides all previous experiences, in the second case we expect that voters take a number of different events into account. The second alternative specification looks at the latest change but only when it occurred within a period of 10 years (*Latest Government Change Limited*). The main measure only looks at a change in government, even if this occurred a long time ago, e.g. before the 1997 UK election, the previous government change was in 1979. It is rather unlikely to expect voters to take events from 18 years earlier into account when voting. These three measures all conceptualise partial and wholesale alternation as a dichotomy.

Our second main measure is the Government Turnover Index (Ieraci, 2012), which is the share of parties that are new in government. In the case of wholesale alternation, this is always 100 per cent. The Government Turnover

Index differentiates between small and large parties entering government. For example, the CDU and CSU joining the SPD in German government, as senior partner after the 2005 elections was a larger change in government composition then when the FDP joined the CDU and CSU in government after the 2009 election. The Government Turnover Index takes this into account by looking at the share of seats taken by parties that were previously not in government. This was 50 per cent in 2005 (the CDU/CSU had 226 seats compared to the 222 of the SPD that continued in government) and 28 per cent in 2009 (the FDP had 93 seats compared to the 239 of the CDU/CSU). The Government Turnover Index assumes a different assumption of the political system by voters. It assumes that the feeling that they can send the rascals out scales with the share of the government that was sent home the previous time a government changed. Consequently, a two third change of government affects voters twice as much compared to a one third change, and in case the entire government changes, the effect is three times higher.

Similar to the measures of wholesale-partial dichotomy, we first look at the latest change in government. We also report robustness tests for the average of government changes in the previous 10 years (*Average Government Turnover Index*), and the most recent government change if it was within a period of 10 years (*Government Turnover Index Limited*). The combination of these results will provide a robustness test of the effects of government alternation on retrospective voting, and which aspects of alternation are most important.[3]

Besides these variables of interest, we include some additional variables to control for their impact (Stiers and Dassonneville, 2020). More specifically, we include the standard socio-demographic variables. Gender binary self-identification is included with male respondents as reference category. Age is included as the age of the respondent in the election year. Educational level is a categorical variable distinguishing four categories: (1) voters that had no, primary, or incomplete secondary education (reference category); (2) voters with secondary education; (3) voters with post-secondary training but no completed university training; (4) voters with university training. We also include the ideological position of the respondent as self-placement on the ideological continuum ranging from 'left' (value 0) to 'right' (value 10). Descriptive statistics of these variables are included in Appendix A.

Besides these covariates, we conduct an additional robustness test in Appendix E. As our argument strongly builds on the literature on clarity of responsibility, it is important to test whether our results are determined by government alternation or a mere reflection of clarity of responsibility. To test this, we estimate additional models, including an indicator of the clarity of responsibility of the election under investigation. To do so, we replicate the measure of 'government clarity' as operationalised by

Table 1. Multilevel models explaining incumbent voting in CSES module 2 and 3.

Model	(1)	(2)	(3)	(4)
Gender (ref. = male)	0.027	0.039*	0.027	0.039*
	(0.018)	(0.019)	(0.018)	(0.019)
Age	0.008***	0.009***	0.008***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)
Education (ref. = no secondary)				
Education: secondary	0.014	−0.002	0.014	−0.002
	(0.027)	(0.028)	(0.027)	(0.028)
Education: post-secondary	−0.003	−0.011	−0.003	−0.011
	(0.030)	(0.031)	(0.030)	(0.031)
Education: university	0.116***	0.106***	0.116***	0.106***
	(0.027)	(0.028)	(0.027)	(0.028)
Ideological position	0.036***	0.042***	0.036***	0.042***
	(0.004)	(0.004)	(0.004)	(0.004)
Performance evaluation	1.303***	1.157***	1.303***	0.801*
	(0.016)	(0.191)	(0.016)	(0.348)
Latest Government Change	−0.541*	−0.613**		
	(0.215)	(0.212)		
Latest Government Change X Performance evaluation		0.615*		
		(0.240)		
Government Turnover Index			−0.719^{+}	−0.823*
			(0.372)	(0.370)
Turnover Index X Performance evaluation				0.935*
				(0.411)
Constant	−0.328^{+}	−0.357*	−0.099	−0.090
	(0.173)	(0.170)	(0.316)	(0.314)
N (individuals)	63559	63559	63559	63559
N (elections)	55	55	55	55
Var(constant)	0.583***	0.567***	0.609***	0.599***
	(0.114)	(0.112)	(0.119)	(0.118)
Var(Evaluation)		0.709***		0.726***
		(0.139)		(0.143)
AIC	70882.263	67493.165	70884.642	67497.442
BIC	70972.860	67601.882	70975.239	67606.159

Note: Entries are log-odds coefficients, standard errors reported in parentheses. Data: CSES module 2 and 3. Significance levels: $^{+}$: $p < 0.10$; *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

Hobolt et al. (2013) – as this was the measure they found to matter most for retrospective voting. Their measure of clarity of responsibility includes indicators of (1) coalition government or not; (2) cohabitation in semi-presidential regimes or not; (3) ideological cohesion of the government; (4) dominance of the main governing party. As clarity of responsibility is expected to moderate retrospective voting, in our robustness test, we include this indicator in interaction with performance evaluations.

As the dependent variable is binary – distinguishing a vote for an incumbent party from an opposition vote – we estimate logistic regression models. However, using data from different election studies, the observations within the respective election studies cannot be assumed to be independent of each other. To account for this clustering, we estimate multilevel models with random intercepts (Gelman and Hill, 2007). We are interested in the conditioning effect of government alternation for retrospective voting, and consequently include interactions in our models (Brambor et al., 2006). As these are cross-level interactions, we include random slopes for the individual-level performance evaluations (Heisig and Schaeffer, 2019). Finally, all continuous individual-level variables are group-mean centred so that we are left with a pure estimate of their effects (Enders and Tofighi, 2007).

Results

Table 1 summarises the results of the models explaining incumbent voting. Model 1 and 2 focus on the wholesale or partial alternation of the latest government; Model 3 and 4 focus on the Government Turnover Index for the latest change in government.

First, the results presented in Table 1 support the traditional retrospective voting theory. As the coefficients of performance evaluations in Model 1 and Model 3 show, the more positive a voter's evaluation of the government's performance, the higher their probability to support an incumbent party. On itself there is a negative relationship between wholesale alternation and the likelihood of voting for the government (Model 1). We, however, will investigate the extent to which the effect of performance evaluations is conditioned by past experiences with government alternation. To test our hypothesis, we include interactions between performance evaluations and our respective measures of government alternation. First, in Model 2, we interact voter perceptions of performance with the dummy-indicator of whether the previous government alternation was wholesale or partial. The coefficient of the interaction is positive and significant, indicating that the coefficient of retrospective evaluations is larger in countries with wholesale government alternation than in countries with partial alternation. However, as interactions in logistic regression coefficients are hard to interpret, we calculate the predicted probability of voting for an incumbent party at different levels of satisfaction for partial and wholesale alternation systems respectively. These are displayed in Figure 1.

The results in Figure 1 confirm the finding of Table 1: wholesale government alternation positively moderates retrospective voting. At lower levels of satisfaction, voters living in systems with wholesale alternation are less likely to support an incumbent than dissatisfied voters in systems with partial alternation. For very satisfied voters, this is reversed, as satisfied voters in wholesale alternation systems are more likely to support the incumbent than voters in partial alternation systems. Hence, overall the effect of satisfaction on incumbent voting is larger in systems with wholesale alternation. This also becomes clear looking at the average marginal effects (Mood, 2010), which we display in Figure D.1 in Appendix D: when the previous government alternation was partial, a one-unit increase in performance evaluation increases the probability of voting for an incumbent party with 24.87 percentage points. For voters having a wholesale alternation in mind, the average marginal effect is 31.51 percentage points. Hence, previous wholesale alternation increases the effect of performance evaluations on the vote with more than 25% compared to partial alternation.

Our second indicator of government alternation is the Government Turnover Index. The interaction coefficient – presented in Model 4 of Table 1 – is also positive and significant, showing that this indicator also positively moderates retrospective voting. To get a better view on the effect, we plot the predicted probability of voting for an incumbent party at different levels of satisfaction at the first and third quartile of the Government Turnover Index respectively in Figure 2.
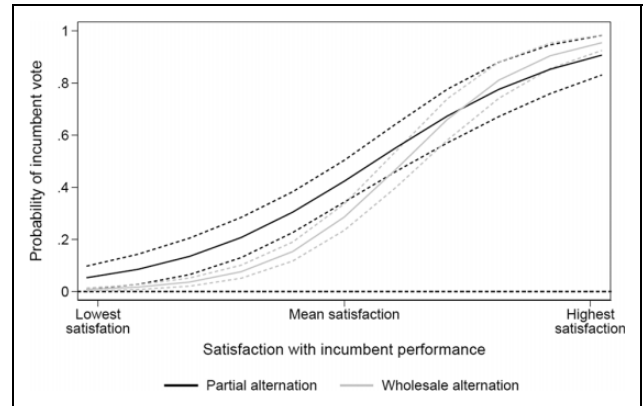

Figure 1. Predicted probabilities of incumbent voting. *Note*: Figure 1 shows predicted probabilities and 95% confidence intervals of voting for an incumbent party at different levels of government satisfaction for systems with partial alternations (black line) and systems with wholesale alternation (grey line) respectively, based on the coefficients in Model 2 of Table 1.
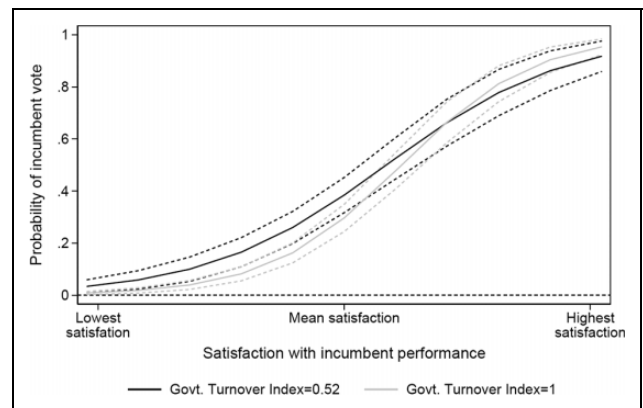

Figure 2. Predicted probabilities of incumbent voting. *Note*: Figure 2 shows predicted probabilities and 95% confidence intervals of voting for an incumbent party at different levels of government satisfaction for systems with a Turnover Index at the first quartile (0.52; black line) and systems with Turnover Index at the third quartile (1; grey line) respectively.

The results in Figure 2 are fully in line with those in Figure 1. Again, dissatisfied voters are less likely to vote for an incumbent party in wholesale alternation systems, while satisfied voters are more likely to support the incumbent in these systems. The average marginal effects (Figure D.2 in Appendix D) also show that retrospective performance evaluations have a stronger effect on voting behaviour when a larger share of the government has changed in the latest election. More specifically, the effect is almost 70% higher in full alternation systems compared to a minimal alternation. This moderation effect is comparable to that found in other studies (see, for instance, Stiers and Dassonneville (2020)).

Robustness test

To test the robustness of the findings presented here, we estimated a series of alternative models. Firstly, it is possible that at least a part of this result could be explained by the difference in clarity of responsibility between these different systems. To test for this, we estimate the models again, this time including an interaction between performance evaluations and the indicator of government clarity designed by Hobolt et al. (2013). The results are summarised in Table E.1 in Appendix E. When we control for the interaction between performance evaluations and clarity of responsibility, the conclusion is that government alternation significantly conditions retrospective voting. Furthermore, the interaction between performance evaluations and clarity of responsibility is no longer significant. However, it needs to be noted that there is a substantial correlation between government alternation and government clarity, so these results should be interpreted with caution.

Secondly, while the focus is on the government alternation of the election previous to the election under investigation, we also tested alternative indicators spanning the previous 10 years before the election. Hence, these indicators measure the share of full government alternations in the 10 years before the election and the average turnover index in the previous 10 years respectively. The results of these tests are included in Table B.1 in Appendix B, and show support for the hypothesis and the findings presented under this specification. That is, the interactions between performance evaluations and the indicators of government alternations are significant at the 0.1-level. These results support our conclusions, but indicate that recent experiences with government alternation play a greater role in the minds of voters than earlier elections. Besides these measures indicating the averages for our indicators over the previous 10 years, we also estimated the models using the same measures as those presented in Table 1, but with a 10-year time limit. These measures take into account that it might be unrealistic to expect a government formation of more than 10 years ago to influence current elections. The results of these models are summarised in Table B.2 in Appendix B. We still find a significant effect of the interaction between performance evaluations and the binary indicator, that is the previous election being wholesale alternation or not. The interaction between performance evaluations and the Government Turnover Index is not supported.

Thirdly, our measure of retrospective performance evaluations differs from the indicator used by a number of earlier studies – i.e. evaluations of the state of the economy. While we think it is preferable not to use evaluations from one domain in particular, it is useful to test how our results relate to the economic voting literature. We use the data of the first and fourth CSES module to supplement our analyses with economic evaluations, as the second and third module of the CSES do not include evaluations of the economy. We build our models as similar as possible to the models presented here. The results, summarised in Appendix C, show further support for our findings. While there is strong evidence for the conclusion of a positive moderation effect for the first binary indicator of alternation, the moderation by share of government change reaches the 0.1-level of significance.

Conclusion

Central in this paper is the question whether democratic accountability is greater in systems characterised by wholesale alternation compared to systems characterised by partial alternation. We find that this is indeed the case; voters are more likely to consider government performance while voting if the last time the government changed, all government parties were new. Depending on the indicator employed, the effect of the government evaluations on voting increases by 25% to even 70%. We find that this effect persists when looking at different operationalisations of wholesale alternation, a different time horizon, and an alternative measure of government performance.

Our results refer to two literatures: the literature on retrospective voting and the literature on wholesale and partial alternation. On the one hand, this study contributes to increasing interest in the effect of institutional factors in the study of retrospective voting (Anderson, 2000; Hobolt et al., 2013; Silva and Whitten, 2017). We show that party system characteristics, and in our case the way party competition for government office is structured, affects extent to which evaluations of government influence how citizens vote. If the most recent elections resulted in a wholesale government change, citizens are more likely to take government performance into account while voting. If the most recent government was the result of backroom talks rather than the election results, these considerations are considerably weaker.

Our results support a long-held belief by those who are interested in differences between wholesale and partial alternation systems (Bergman and Strøm, 2011; Mair, 1997): in systems with wholesale alternation democratic accountability is significantly greater compared to systems with partial alternation. Citizens are more likely to 'vote the rascals out'. Systems with wholesale alternation reflect more closely the model of democracy that Schumpeter (1943) envisioned: citizens can choose between different teams of governors. If they are happy with the current government, they can back one of the government parties, otherwise they can support the opposition. The electoral feedback mechanism that is crucial in traditional models of democratic functioning, such as Easton's (1965) appears to be somewhat stronger in systems with wholesale alternation.

Democratic accountability seems to be greater in systems with wholesale alternation. However, this may come at the expense of other desired aspects of a well-functioning democratic system. In particular, greater democratic accountability may weaken the quality of policy representation (Quinn, 2016). Voters that take policy performance into account may be less likely to vote for the party which is closest to their policy preference. Future research may want to delve deeper into the role of policy considerations in voting comparing systems characterised by wholesale and partial alternation.

Supplemental material

Supplemental material for this article is available online.

Notes

1. In our view when voters consider what their vote can affect, they are likely to think about what previous elections have affected. To us, it seems unlikely that they take into account changes that occurred at other times (that is when a government changed without elections). Those cases were unaffected what voters did.
2. There also is a mathematical reason why a latest-election measure works better than an average of the previous decade of elections. Consider the situation where there is a deviation from the norm in 1 year, Say, there is wholesale alternation in a system where there have been decades of partial alternation. And that citizens' expectations are not affected by this aberration and voters still expect partial alternation, this deviation will only provide an 'incorrect' measure for 1 year using the latest-election measure. But it will distort the values for a decade of elections for the other measure.
3. There is one alternative measure that we did not employ and that is a measure of ideological alternation (e.g. Horowitz et al., 2009; Pellegata, 2016; Tsebelis, 2002; Tsebelis and Chang, 2004; Zucchini, 2011). This is a measure of the ideological distance between successive governments. This brings an element of left-right voting into the analysis. We do not think that this is relevant to our argument because we look at retrospective voting in terms of a respondents' assessment of the job the government did (i.e., a valence

consideration), not the ideological proximity of the policy output (i.e., a spatial consideration).

References

Achen CH and Bartels LM (2016) *Democracy for Realists: Why Elections do not Produce Responsive Government*. Princeton, NJ: Princeton University Press.

Anderson CJ (2000) Economic voting and political context: a comparative perspective. *Electoral Studies* 19(2): 151–170.

Anderson CJ (2007a) The end of economic voting? Contingency dilemmas and the limits of democratic accountability. *Annual Review of Political Science* 10: 271–296.

Anderson CJ (2007b) The interaction of structures and voter behavior. In: Dalton RJ and Klingemann HD (eds) *Oxford Handbook of Political Behavior*. Oxford: Oxford University Press, pp. 589–606.

Anthonsen M and Lindvall J (2009) Party competition and the resilience of corporatism. *Government and Opposition* 44(2): 167–187.

Bengtsson A (2004) Economic voting: the effect of political context, volatility and turnout on voters' assignment of responsibility. *European Journal of Political Research* 43(5): 749–767.

Bergman T and Strøm K (2011) Nordic Europe in comparative perspective. In: Bergman T and Strøm K (eds) *The Madisonian Turn. Political Parties and Parliamentary Democracy in Nordic Europe*. Ann Arbor, MI: University of Michigan Press, pp. 35–65.

Brambor T, Clark WR and Golder M (2006) Understanding interaction models: improving empirical analyses. *Political Analysis* 14(1): 63–82.

Campos NF and Horváth R (2012a) Reform redux: measurement, determinants and growth implications. *European Journal of Political Economy* 28(2): 227–237.

Campos NF and Horváth R (2012b) On the reversibility of structural reforms. *Economics Letters* 117(1): 217–219.

Casal Bértoa F and Enyedi Z (2016) Party system closure and openness: conceptualization, operationalization and validation. *Party Politics* 22(3): 265–277.

Cheibub JA, Gandhi J and Vreeland JR (2009) Democracy and dictatorship revisited. *Public Choice* 143(12): 67–101.

Comparative Study of Electoral Systems (CSES) (2015a) The Comparative Study of Electoral Systems (www.cses.org). CSES module 2 full release [dataset]. 15 December 2015 version. DOI: 10.7804/cses.module2.2015-12-15.

Comparative Study of Electoral Systems (CSES) (2015b) The Comparative Study of Electoral Systems (www.cses.org). CSES module 3 full release [dataset]. 15 December 2015 version. DOI: 10.7804/cses.module3.2015-12-15.

Crisp BF, Olivella S, Potter JD, et al. (2014) Elections as instruments for punishing bad representatives and selecting good ones. *Electoral Studies* 34: 1–15.

Cutler F (2008) One voter, two first-order elections? *Electoral Studies* 27: 492–504.

De Giorgi E and Marangoni F (2015) Government laws and opposition parties' behaviour in parliament. *Acta Politica* 50: 64–81.

De Vries CE and Giger N (2014) Holding governments accountable? Individual heterogeneity in performance voting. *European Journal of Political Research* 53(2): 345–362.

Easton D (1965) *A framework for political analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Ecker A, Glinitzer K and Meyer TM (2016) Corruption performance voting and the electoral context. *European Political Science Review* 8(3): 333–354.

Enders C and Tofighi D (2007) Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological Methods* 12(2): 121–138.

Finer SE (1975) Adversary politics and electoral reform. In: Finer SE (ed) *Adversary Politics and Electoral Reform*. London: Anthony Wigram, pp. 3–35.

Fiorina MP (1978) Economic retrospective voting in American national elections: a micro-analysis. *American Journal of Political Science* 22(2): 426–443.

Fiorina MP (1981) *Retrospective Voting in American National Elections*. New Haven, CT: Yale University Press.

Fowler A and Hall AB (2018) Do shark attacks influence presidential elections? Reassessing a prominent finding on voter competence. *The Journal of Politics* 80(4): 1423–1437.

Gasper JT and Reeves A (2011) Make it rain? Retrospection and the attentive electorate in the context of natural disasters. *American Journal of Political Science* 55(2): 340–355.

Gelman A and Hill J (2007) *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge, MA: Cambridge University Press.

Green-Pedersen C (2002) *The Politics of Justification. Party Competition and Welfare-State Retrenchment in Denmark and the Netherlands from 1982 to 1988*. Amsterdam: Amsterdam University Press.

Green-Pedersen C (2004) Center parties, party competition, and the implosion of party systems: a study of centripetal tendencies in multiparty systems. *Political Studies* 52(2): 324–341.

Healy A and Malhotra N (2013) Retrospective voting reconsidered. *Annual Review of Political Science* 16: 285–306.

Heisig JP and Schaeffer M (2019) Why you should *Always* include a random slope for the lower-level variable involved in a cross-level interaction. *European Sociological Review* 35(2): 258–279.

Hellwig TT (2010) Context, information, and performance voting. In: Dalton RJ and Anderson CJ (eds) *Citizens, Context, and Choice: How Context Shapes Citizens' Electoral Choices*. New York, NY: Oxford University Press, pp. 149–175.

Heyne L (2018) The making of democratic citizens. How regime-specific socialization shapes Europeans' expectations of democracy. *Swiss Political Science Review* 25(1): 40–63.

Hobolt S, Tilley J and Banducci S (2013) Clarity of responsibility: how government cohesion conditions performance voting. *European Journal of Political Research* 52(2): 164–187.

Horowitz S, Hoff K and Milanovic B (2009) Government turnover: concepts, measures and applications. *European Journal of Political Research* 48(1): 107–129.

Ieraci G (2012) Government alternation and patterns of competition in Europe: comparative data in search of explanations. *West European Politics* 35(3): 530–550.

Irwin G and Van Holsteyn J (2011) Sterft Gij oude vormen en gedachten? Over kiezers, verkiezingen en kiesstelsel in Nederland. In: Andeweg RB and Thomassen JAA (eds) *Democratie Doorgelicht. Het Functioneren van de Nederlandse Democratie*. Amsterdam: Amsterdam University Press, pp. 335–348.

Kaiser A, Lehnert M, Miller B, et al. (2002) The democratic quality of institutional regimes: a conceptual framework. *Political Studies* 50(2): 313–331.

Karol D and Miguel E (2007) The electoral cost of war: Iraq casualties and the 2005 U.S. Presidential Election. *The Journal of Politics* 69(3): 633–648.

Key V (1966) *The Responsible Electorate. Rationality in Presidential Voting 1936-1960*. Cambridge, MA: Belknap Press of Harvard University Press.

Lewis-Beck MS and Stegmaier M (2000) Economic determinants of electoral outcomes. *Annual Review of Political Science* 3: 183–219.

Lewis-Beck MS and Stegmaier M (2007) Economic models of voting. In: Dalton RJ and Klingemann H-D (eds) *The Oxford Handbook of Political Behavior*. Oxford: Oxford University Press, pp. 518–537.

Lewis-Beck MS and Stegmaier M (2013) The VP-function revisited: a survey of the literature on vote and popularity functions after over 40 years. *Public Choice* 157(3): 367–3858.

Louwerse T, Otjes S, Willumsen DM, et al. (2017) Reaching across the Aisle: explaining government-opposition voting in Parliament. *Party Politics* 23(6): 746–759.

Lundell K (2011) Accountability and patterns of alternation in pluralitarian, majoritarian and consensus democracies. *Government and Opposition* 46(2): 145–167.

Mair P (1997) *Party System Change. Approaches and Interpretations*. Oxford: Clarendon Press.

Mair P (2001) The green challenge and political competition: How typical is the German experience? *German Politics* 10(2): 99–116.

Mair P (2008) Electoral volatility and the Dutch party system: a comparative perspective. *Acta Politica* 43(2–3): 235–253.

Mair P (2011) Is governing becoming more contentious? In: Rosema M, Denters B and Aarts K (eds) *How Democracy Works: Political Representation and Policy Congruence in Modern Societies: Essays in Honour of Jacques Thomassen*. Amsterdam: Pallas Publications, pp. 77–88.

Manin B, Przeworski A and Stokes SC (1999) Elections and representation. In: Przeworski A, Stokes SC and Manin B (eds) *Democracy, Accountability, and Representation*. Cambridge, MA: Cambridge University Press, pp. 29–52.

Meyer-Sahling J-H and Veen T (2012) Governing the post-communist state: government alternation and senior civil

service politicisation in central and eastern Europe. *East European Politics* 28(1): 4–22.

Milanovic B, Hoff K and Horowitz S (2010) Turnover in power as a restraint on investing in influence: evidence from the postcommunist transition. *Economics and Politics* 22(3): 329–361.

Mishler W and Rose R (2002) Learning and re-learning regime support: the dynamics of post-communist regimes. *European Journal of Political Research* 41(1): 5–36.

Mood C (2010) Logistic regression: why we cannot do what we think we can do, and what we can do about it. *European Sociological Review* 26(1): 67–82.

Nadeau R, Niemi RG and Yoshinaka A (2002) A cross-national analysis of economic voting: taking account of the political context across time and nations. *Electoral Studies* 21(3): 403–423.

Nannestad P and Paldam M (1994) The VP-function: a survey of the literature on vote and popularity functions after 25 years. *Public Choice* 79(3): 213–245.

Otjes S (2018) How fractionalisation and polarisation explain the level of government turn-over. *Acta Politica* 55(1): 41–66.

Otjes S and Rasmussen A (2017) Collaboration between interest groups and political parties in multi-party democracies: party system dynamics and the effect of power and ideology. *Party Politics* 23(2): 96–109.

Otjes S and Willumsen DM (2019) How does turnover affect turnout? Government alternation and voter participation in parliamentary democracies. *Electoral Studies* 59: 49–63.

Pellegata A (2016) Assessing the complex relationship between government alternation and ideological congruence. *International Political Science Review* 37(1): 51–65.

Pellegata A and Quaranta M (2018) Accountability through government alternation: economic performance and the conditional role of political institutions in fifty countries, 1990–2015. *International Political Science Review* 40(3): 419–436.

Powell BG and Whitten GD (1993) A cross-national analysis of economic voting: taking account of the political context. *American Journal of Political Science* 37(2): 391–414.

Quinn T (2016) Throwing the rascals out? Problems of accountability in two-party systems. *European Journal of Political Research* 55(1): 120–137.

Rokkan S (1970) *Citizens, Elections, Parties; Approaches to the Comparative Study of the Processes of Development*. New York, NY: McKay.

Sartori G (1976) *Parties and Party Systems: Volume 1: A Framework for Analysis*. Cambridge, MA: Cambridge University Press.

Schumpeter JA (1943) *Capitalism, socialism and democracy*. Londen: George Allen and Unwin.

Schwindt-Bayer LA and Tavits M (2016) *Clarity of Responsibility, Accountability, and Corruption*. Cambridge, MA: Cambridge University Press.

Silva T and Whitten GD (2017) Clarity of responsibility and vote choice. In: Arzheimer K, Evans J and Lewis-Beck MS (eds) *The Sage Handbook of Political Behaviour*. London: Sage, pp. 80–91.

Singer MM (2011) Who says "It's the Economy"? cross-national and cross-individual variation in the salience of economic performance. *Comparative Political Studies* 44(3): 284–312.

Sniderman PM, Brody RA and Tetlock PE (1991) *Reasoning and Choice: Explorations in Political Psychology*. New York, NY: Cambridge University Press.

Stiers D (2019) Beyond the distinction incumbent-opposition: retrospective voting on the level of political parties. *Party Politics* 25(6): 805–816.

Stiers D and Dassonneville R (2020) Retrospective voting and the polarization of available alternatives. *Canadian Journal of Political Science* 53(1): 99–115.

Strøm K (2000) Delegation and accountability in parliamentary democracies. *European Journal of Political Research* 37(3): 261–290.

Stubager R, Botterill NW, Lewis-Beck MS, et al. (2014) Scope conditions of economic voting: the Danish exception? *Electoral Studies* 34: 16–26.

Tsebelis G (2002) *Veto Players: How Political Institutions Work*. Princeton, NJ: Princeton University Press.

Tsebelis G and Chang EC (2004) Veto players and the structure of budgets in advanced industrialized countries. *European Journal of Political Research* 43(3): 449–476.

Verzichelli L and Cotta M (1999) Italy: from 'Constrained' coalitions to alternating governments? In: Müller S and Strøm K (eds) *Coalition Governments in Western Europe*. Oxford: Oxford University Press.

Whitten GD and Palmer HD (1999) Cross-national analyses of economic voting. *Electoral Studies* 18(1): 49–67.

Zucchini F (2011) Government alternation and legislative agenda setting. *European Journal of Political Research* 50(6): 749–774.

Author biographies

Simon Otjes is an assistant professor of political science at Leiden University and researcher at Groningen University. His research focuses on parties and party systems.

Dieter Stiers is a PhD researcher of the Research Foundation Flanders (FWO) at the Centre for Citizenship and Democracy of KU Leuven. He is mainly interested in electoral behaviour, with a specific focus on the retrospective voting theory and the quality of democratic representation.