

Dieter Stiers
Ruth Dassonneville
Michael S. Lewis Beck

**The Abiding Voter:
The Lengthy Horizon of Retrospective Evaluations**

Abstract

Although the theory of retrospective voting receives wide support in the literature on voting behavior, less agreement exists on voters' time horizon when assessing the government's performance – i.e., whether voters are *myopic*. Previous studies on voter myopia tend to focus on aggregate-level measures of the economy, or use an experimental approach. Using panel data, this paper offers the first investigation into voter myopia that uses individual-level evaluations of government performance in a representative survey at several points during the electoral cycle. Our study focuses on The Netherlands, but we also provide tests of the generalizability and robustness of our findings, and a replication in the U.S. context. The results indicate that voter satisfaction early in the government's term adds to explaining incumbent voting. Thus, rather than the *myopic* voter, we find evidence of the *abiding* voter – steady at her or his post, evaluating government performance over a long length of time.

Keywords: Retrospective voting; voter myopia; reward-punishment mechanism; performance voting; economic voting; accountability

This is the accepted version of an article published in the *European Journal of Political Research*. The publisher's version can be found here: <https://doi.org/10.1111/1475-6765.12360>

Introduction

The principle of electoral accountability presupposes that citizens hold governing parties to account with their vote, revealing one of the fundamental ways in which elections contribute to the process of democratic representation (Strøm 2000). This mechanism bases itself on voters evaluating the government's performance and voting according to this evaluation (Przeworski, Stokes, and Manin 1999). More specifically, the theory of retrospective voting hypothesizes that voters reward or punish incumbent parties based on their assessment of the government's performance (Key 1966). If voters are satisfied, they will vote for the governing parties, while dissatisfied voters are expected to support the opposition (Ferejohn 1986; Kramer 1971).

Previous studies on retrospective voting have debated the time horizon that voters use. In general, voters are assumed to be *myopic* – implying they do not look over the full duration of a government term, but only take conditions close to Election Day into consideration (Achen and Bartels 2016; Fiorina 1981; Healy and Lenz 2014; Huber, Hill, and Lenz 2012). If voters merely use a short time horizon, this raises incentives for incumbents to shirk, just caring about voter opinion near the end of their term. Myopic voters might not select the best leader but instead the best manipulator, one following a “political business cycle” with a peak performance close on the heels of the election, as “there is little or no electoral incentive for presidents to promote myopic voters’ well-being during much of their time in office” (Achen and Bartels 2016, 170). In contrast, if voters take into account government performance over the long term, they might force incumbents to be more responsive to their demands.

We build on earlier work about voter myopia, but take it a step further by investigating the voters’ time horizon in an extended real-world survey setting. More specifically, we deviate from previous research in two important ways. First, empirically, research on this question has mostly limited itself to studying patterns of retrospective voting on the aggregate level (Achen

and Bartels 2016; Paldam and Nannestad 2000; Wlezien 2015) or using an experimental approach (Healy and Lenz 2014; Huber, Hill, and Lenz 2012). While aggregate-level studies yielded useful – if inconclusive – results, they do not provide insights on voter myopia at the individual level. In addition, aggregate-level studies risk committing the ecological fallacy of finding relationships that do not hold at the individual level (Kramer 1983). On the other hand, experimental designs suffer from problems of external validity; while they show possible mechanisms in the link between voters’ (economic) perceptions and their vote, it is unclear whether previous findings from experimental studies apply in a real-world setting. Second, theoretically, using their respective approaches, previous studies implicitly assumed voters apply memory-based information processing, which might actually be too cognitively demanding (Zaller and Feldman 1992). We address these limitations by using repeated measures of citizens’ individual-level evaluations of government performance and party choice in a representative panel survey. Doing so, we argue in favor of an online information processing model, in which voters update their performance evaluations as a running tally (Lodge et al. 1989; 1995)

We employ panel data from The Netherlands, where party choices are plentiful. In the LISS panel, respondents are asked how satisfied they are with “what the government has been doing lately” at different points in the electoral cycle. Thus, we can assess whether only performance evaluations close to the Election Day matter for the vote, as compared to satisfaction throughout the cycle. Furthermore, we replicate our analyses using data from the United States.

We confirm that voters’ evaluation of government performance shortly before Election Day affects their party choice. But we also find that earlier assessments of the incumbent are at work. Additional analyses provide evidence for the generalizability and robustness of the findings. Thus, our study offers grounds for democratic optimism. Incumbents cannot afford

to simply offer performance peaks near the end of their term, since voters evaluate their performance during the whole term, and these evaluations all count toward their party choice. Rather than a myopic voter, we observe an *abiding* voter, one attentive to the government's performance in the long run as well as the short. We begin our examination below, with a review of relevant literature.

Retrospective voting: theory and literature

The most often investigated manifestation of retrospective voting is economic voting, where “economic conditions shape election outcomes in the world’s democracies. Good times keep parties in office, bad times cast them out” (Lewis-Beck and Stegmaier 2000, 183). A large body of literature supports this hypothesis (Duch and Stevenson 2008). However, the economy is not the only policy domain where the theory of performance voting operates. For instance, Ecker, Glinitzer, and Meyer (2016) find that perceptions of government corruption decrease the likelihood of voting for an incumbent party – a result supported by other research on corruption voting (Crisp, Olivella, Potter, and Mishler 2014; Shabad and Slomczynski 2011). Additional studies find performance voting in the realm of war and security (Holmes and Gutiérrez de Piñeres 2013; Hibbs 2000), the quality of government’s services (Boyne, James, John, and Petrovsky 2009), or satisfaction with the government’s performance “in general” (de Vries and Giger 2014). As Singer (2011, 285) states: “The economy represents only one of many possible bases for evaluating the incumbent’s competence.”

Are these retrospective evaluations of government performance myopic? If voters are myopic, the vote becomes a weaker tool for holding the incumbent accountable. Much research presumes that voters only use recent experiences to evaluate the performance of the government (Healy and Lenz 2014; Huber et al. 2012; Paldam and Nannestad 2000).¹ Further,

¹ For an alternative reading of the myopia thesis, see Healy and Malhotra (2009).

in their overview article on performance voting, Nannestad and Paldam (1994) state that voters generally do not look back for more than one quarter or even two months before the election to evaluate the government's performance. Likewise, Achen and Bartels (2016, 16) find that it is mostly the economic performance during the last six months of the incumbent's term that matters to voters, while the economic growth earlier in the term adds little or nothing to explaining the incumbent's electoral success: "voters are *myopic*, focusing almost entirely on income growth in the months just before each election."

Different reasons have been proposed for this myopia, e.g., voters might only become politically engaged and informed when the election comes close, and disregard previous information (Healy and Lenz 2014; Wlezien 2015). Or, more subtlety, voters can use the economic performance during the election year as a proxy, focusing "on the election-year economy because that attribute is an easily available substitute for the overall growth for which they are searching" (Healy and Lenz 2014, 32). The dominance of the myopia hypothesis in the field is also evident in the standard survey research approach of using a one-year lag when measuring voters' evaluations of the government's performance. For example, an individual-level economic voting survey generally includes a measure of the economic situation – either personal, or of the whole country – "over the past 12 months."

As Lewis-Beck and Stegmaier (2013, 378) observe, while the myopia assumption holds sway in the literature, "it wants hard testing in order [to] discover if it is indeed preferred." In fact, certain studies fail to support the standard myopic assumption. Wlezien (2015) argues that voters take into account the last two years before the election. Furthermore, Healy et al. (2017) find no evidence of voter myopia in pocketbook voting; their results even indicate that voters put most weight on income changes in the first year of a government's term. In the case that Healy et al. (2017) analyze – Sweden between 2002 and 2010 – voters' focus on the first year suggests rational voting behavior, as this is when the incumbent government implemented a

large-scale tax cut. Relatedly, Owen and Tucker (2010) find that voters can discriminate between long- and short-term economic evaluations.

Hence, even though the myopia thesis is a dominant perspective in work on retrospective voting, results as these suggest that voters' behavior does not always fit this expectation. There are, in fact, good theoretical reasons to assume that voters will – at least some of the time – take a longer time perspective when evaluating the incumbent.

Work on public opinion-formation distinguishes between different models of information-processing: a memory-based model and an online model (Hastie and Park 1986; Kim and Garrett 2012; Lodge et al. 1989; 1995; Zaller and Feldman 1992). The former model assumes that voters do not continuously form opinions, but store relevant information. Only when they have to form an opinion, such as deciding whom to vote for, do they retrieve the information and rely on what they recall from memory to form their opinion (Zaller and Feldman 1992). In contrast, according to the online model of information processing, citizens continuously form opinions and update their opinion, as a 'running tally', when exposed to new information – without necessarily recalling the considerations that formed their opinions (Lodge et al., 1989; 1995).

If it is assumed that citizens follow a memory-based approach when exposed to economic information, it seems sensible to follow the myopic voter hypothesis. Voting retrospectively is potentially less effortful than evaluating the policy proposals of all candidates and parties on the ballot (Fiorina 1981). However, it is cognitively quite demanding for voters to store and recall relevant economic statistics for a full electoral term. If this is the model of information processing, it is not surprising that scholars find that voters seek to rely on shortcuts, such as an end-year heuristic, when forming an opinion based on economic information (Healy et al. 2017; Healy and Lenz 2014). If they do so, their behavior will fit with the myopic voter hypothesis.

The work of Hellwig and Marinova (2014), however, casts doubt on the importance of the memory-model for explaining retrospective voting. Asking respondents to assess trends in economic conditions, they find that citizens are mostly wrong. Furthermore, citizens are wrong regardless of whether they are asked to evaluate short- or long-term changes in economic conditions. Based on these findings, Hellwig and Marinova (2014) conclude, voters are *not* myopic, but simply misinformed about the economy.

The online model of information processing, in contrast, is less cognitively demanding. Even voters who cannot accurately recall objective economic conditions in the long- or the short-term might be influenced by these conditions when forming an opinion about the government's performance. If this opinion subsequently serves as a running-tally – that is updated whenever new information becomes available – then economic information or other indicators of government performance from the beginning of the electoral term might feed into their vote choice.

To be sure, an online model of information processing does not exclude the possibility of voter myopia. The effects of events and conditions early in the electoral cycle on citizens' opinions likely decay, while more recent events and conditions gain weight (Chong and Druckman 2010). However, an online model of information processing acknowledges – more so than a memory-based model – that what happens throughout the government's term can influence voters' opinions on the incumbent's performance. What happens at the very beginning of the term might even be of particular importance, as this is when citizens form their initial opinion on the incumbent (Chong and Druckman 2010). Depending on whether voters process information by retrieving facts from memory or whether they follow an online model, we would thus expect to find more or less evidence of myopia in retrospective voting.

Finally, it needs to be noted that the online model of information processing can be cognitively challenging as well, and it could even be argued that constantly updating a running

tally actually is *more* cognitively demanding. However, in the context of retrospective voting specifically, the difference between the memory-based and the online model implies different requirements from voters. In the memory-based model, voters are expected to correctly recall objective (economic) indicators from early in the term, so they can create a well-informed evaluation. In the online model, in contrast, voters do not necessarily need to hold objective knowledge, but can rely on their (sociotropic) *perceptions* of the state of the country. In line with the study of Hellwig and Marinova (2014) referred to above, previous research has shown that voters do not necessarily hold factual knowledge about e.g. the economy, but that their perceptions of the evolution of the economy can still be rather accurate (Conover 1986). Furthermore, it is these perceptions that feed into the vote choice (Sanders 2000). Hence, even though we do not argue that the memory-based model is more cognitively demanding than the online model *per se*, in the case of retrospective voting, we believe that assuming a memory based model poses high requirements on the voter. These assumptions render voter myopia a logical consequence, while assuming an online processing model might reveal voters to be less myopic than commonly assumed.

Data

Investigating myopia in individual-level perceptions of the government's performance poses high data requirements. More specifically, it demands repeated measurements of satisfaction with the government's performance throughout the electoral cycle. Therefore, we make use of data from the LISS (Longitudinal Internet Studies for the Social Sciences) panel administered by CentERdata (Tilburg University, The Netherlands).² This data set meets these strict requirements. The initial sample was based on a probability sample of Dutch households drawn from the national population register. All participants were contacted offline, and subsequently

² More information about the LISS panel can be found at: www.lissdata.nl

asked to fill in questionnaires online.³ To make sure that the online data collection does not influence the representativeness of the data, households are provided with a computer that is easy to use, and internet connection if needed.⁴ We use the data of the Politics and Value Core Study, fielded yearly in December since more than ten years now, with a second round of surveys in January of the subsequent year.

The Netherlands offers a useful case study for our purposes, standing as an established Western democracy, where voter participation traditionally has been high (Howe 2006). Moreover, the country has a very proportional system, leading to a large number of political parties in Parliament (van Elsas, Lubbe, van der Meer, and van der Brug 2013). On the one hand, this might render the attribution of responsibility a difficult task. On the other hand, it also ensures a viable – i.e., ideologically close – alternative for virtually every party choice (van der Meer, van Elsas, Lubbe, and van der Brug 2015; van Elsas et al. 2013). If a voter is dissatisfied with the performance of one of the parties in government, there is most likely a nearby alternative.

Since the start of the data collection in 2007, three elections took place. We have adequate panel waves from two of the three cycles: four waves for the June 2010 election, and five waves for the March 2017 election. (For the September 2012 election we cannot carry out our panel evaluation, as we have insufficient waves.) Figure 1 clarifies which measures are derived from which survey waves, and the time-span between every wave. It should be noted that the notation ‘ $t-x$ ’ is used ascendingly, irrespective of the time-span between the waves. For instance, there was no survey wave at the end of 2014; thus, whereas the difference between $t-1$ and $t-2$ in the first electoral cycle is one year, in the second cycle the difference is two years.

³ More information about the representativeness of the panel can be found here: <https://www.lissdata.nl/about-panel/composition-and-response>

⁴ The average response rate of the panel amounts to 82%. Educational level is not a significant predictor of drop-out.

Figure 1: Survey waves and measures used⁵

First electoral cycle						
Survey	Wave 1	Wave 2	Wave 3		Wave 4	
Date of survey	12/07-03/08	12/08-01/09	12/09-01/10	06/10	12/10-01/11	
Measure used	Satisfaction $t-2$	Satisfaction $t-1$	Satisfaction t Covariates	ELECTION	Vote	

Second electoral cycle						
Survey	Wave 6	Wave 7	Wave 8	Wave 9		Wave 10
Date of survey	12/12-01/13	12/13-01/14	12/15-01/16	12/16-01/17	03/17	12/17-03/18
Measure used	Satisfaction $t-3$	Satisfaction $t-2$	Satisfaction $t-1$	Satisfaction t Covariates	ELECTION	Vote

Measures

Our dependent variable is the party choice of the respondent in elections for the Lower Chamber. In the survey waves immediately following the elections – i.e., survey wave 4 (fielded in December 2010) for the first cycle, and survey wave 10 (fielded in December 2017) for the second cycle – respondents were asked which party they voted for (see Figure 1 for information on the timing of different survey waves). The answers are coded “1” in case of a vote for a governing party and “0” in the case of an opposition party.⁶ Given the binary nature of the dependent variable, logistic regression models are estimated.⁷

Our main independent variable is the respondent’s evaluation of the government’s performance. Following previous research on retrospective voting (de Vries and Giger, 2014; Stiers, forthcoming), we use general performance measures. In every survey wave, respondents

⁵ Note that the dates in the Figure follow the notation MM/YY. Hence, 12/07, for instance, denotes December 2007.

⁶ Respondents who said they did not turn out or who cast a blank vote are excluded from the analyses. However, as abstention and blank voting can also be considered a way of punishing the incumbent, we also conducted the analyses coding abstention and blank voting as a non-incumbent vote (code 0). The results, reported in Appendix A, support the conclusions presented in the text.

⁷ As previous research has indicated that, in coalition governments, it is mainly the party of the Prime Minister that is credited or blamed for the government’s performance, we also conducted the analyses using as a dependent variable the vote for the PM party only. The results, reported in Appendix B, are largely in line with the results presented here, but also show some interesting differences. In the main model of the second electoral cycle under investigation, the performance measure at $t-3$ does not significantly predict voting for the PM party, while it is the only measure that does so in the economic voting model.

were asked: “How satisfied or dissatisfied are you, generally speaking, with what the government has done lately?” Respondents could answer using a scale ranging from 1 (very dissatisfied) to 5 (very satisfied). The wording of the question – asking to evaluate the government’s performance ‘lately’ – is particularly useful since it does not require a curtailed, 12-month time horizon, as is common with economic voting survey questions. Since the panel started in 2007, in the first cycle, we can trace the respondent’s evaluations up to 2.5 years before the 2010 election.⁸ In the second cycle, we trace the respondent’s evaluation up to 4 years before the 2017 election. Before proceeding, it is important to point out that there is substantial variation in the performance evaluations in the different survey waves, with the correlations between the different measures ranging from 0.23 to 0.62.

A possible drawback of using these measures is that we do not know what feeds into voters’ general performance evaluations. However, while previous research mostly focused on evaluations of the economy, other studies have shown that less than half of the voters indicate that the economy is the most important issue in their country – followed by issues such as social or foreign policy (de Vries and Giger, 2014; Singer, 2011). Hence, while we are agnostic about the determinants of general performance evaluations, we leave it up to the voters themselves to think of the issues that they deem most important. Nevertheless, to be sure that our results are not determined by the use of this measure only, and to connect our study to the large body of literature on economic voting, we also report robustness tests using evaluations of the economy.

Next to the variables of interest, we control for relevant individual characteristics for explaining voting for a party of the governing coalition (consisting of two Christian-democratic parties and a Social-democratic party in the first electoral cycle and a Conservative-liberal and

⁸ The general election preceding the election under investigation was held in November 2006. However, the coalition was only sworn in on 22 February 2007.

Social-democratic party in the second cycle). As with previous studies that have investigated the effects of individual-level characteristics on voting for these parties (Duncan 2017; Norrander and Wilcox 2008; Swank and Eisinga 1999), we control for age and sex. Age reflects the age of the respondent in the election year. Sex is included as a dummy variable with male voters as the reference category. We also control for the ideological position of the respondent, holding this is to be a “super issue” influencing the party choice (van der Eijk, Schmitt, and Binder 2005). We include left-right self-placement as a categorical variable, with respondents reporting a center position as the reference category.

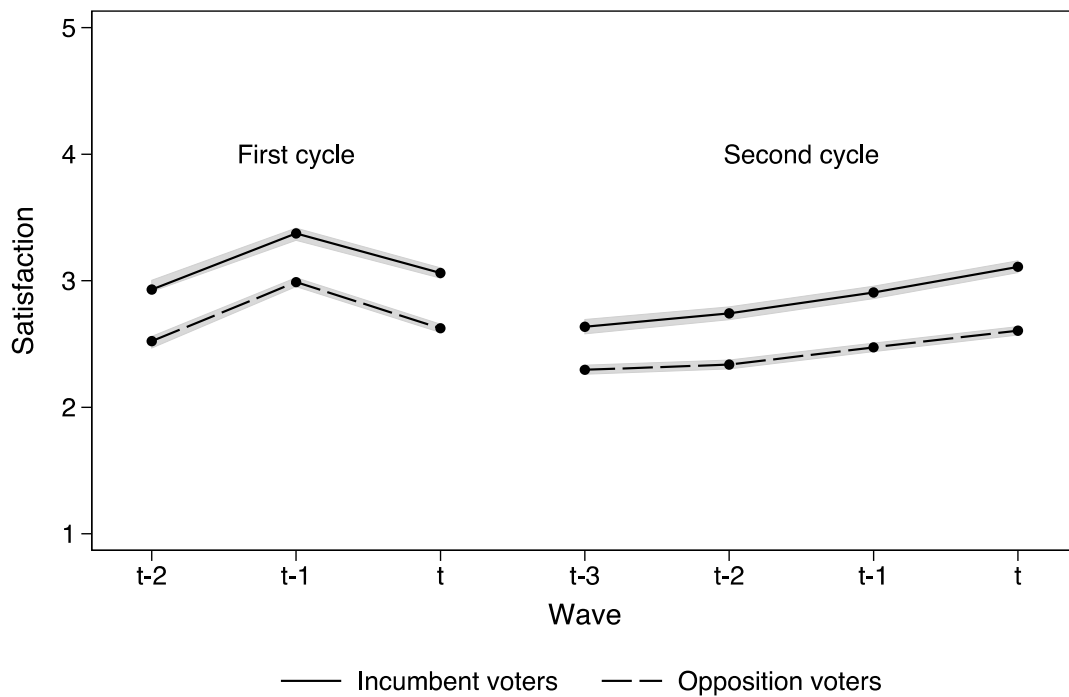
Finally, we control for the economic position of the respondents, as these economic conditions have been shown to influence the probability of voting for an incumbent party (Kinder and Kiewiet 1979). To this end, we include the employment status and income category of the respondents. Employment status is included as a variable with three categories: employed (reference category), unemployed, and other. To measure income, respondents were asked to select an income value from a list of categories with intervals of 500 euros, so forming a scale. (More information on the exact questions and coding of the variables, as well as descriptive statistics, can be found in Appendix C.)

Results

Main results

To get a first look at how incumbent and opposition voters differ in their evaluation of the government’s actions, we test in a bivariate way the difference in satisfaction between these two groups. Figure 2 displays the average satisfaction with the government’s performance for incumbent voters and opposition voters. The exact numbers, as well as the correlation between performance evaluations and the vote measure, are included in Appendix D.

Figure 2: Average satisfaction of incumbent and opposition voters



Note: The dots display average satisfaction for voters who – in the next election – voted for an opposition party (dashed line) or an incumbent party (solid line) respectively. The grey bands are 95% confidence intervals. T-tests indicated that the difference in satisfaction is significant at each time point, with $p < 0.001$.

As Figure 2 shows, those who vote for an incumbent party consistently report a higher satisfaction with the government’s performance than voters for an opposition party, and t-tests show that this difference is significant. Furthermore, the difference is already visible at the time of the first wave of surveys – i.e., 2.5 years before the election in the first electoral cycle, and 4 years in the second electoral cycle. Furthermore, while in the second electoral cycle the correlations decrease slightly with performance evaluations further away in time (see Table D.1 in Appendix D), this pattern cannot be observed in the first electoral cycle.

Thus, these initial bivariate tests suggest that voters’ performance evaluations early in the electoral cycle might be influencing the party choice. However, before conclusions are drawn, we need to control for other relevant factors that have been shown to correlate with performance evaluations. The results of the logistic regression models predicting the party choice are summarized in Table 1. In the first model, only the voter’s satisfaction with the government’s performance shortly before the election is included. As the theory of

retrospective voting predicts, the coefficient is positive and significant: the more satisfied the voter is with “what the government has been doing lately” close to Election Day, the more likely this voter will choose an incumbent party.

In the second model, the voter’s satisfaction with the government’s performance in the previous survey wave – in the field 1.5 years before the election – is included as well. The coefficient of this variable is positive and significant, suggesting that the evaluation of the voter more than a year before the election adds to the explanation of an incumbent vote. Finally, to put the myopic voter thesis to a fuller test, satisfaction 2.5 years before the election is included.⁹ As the results in the third model show, this distant satisfaction measure still helps to predict incumbent voting.

An important note needs to be made about the interpretation of the results. The myopic voter thesis prescribes that voters almost entirely rely on their opinions close to Election Day when they vote, and the expectation would therefore be that earlier performance evaluations do not have a substantial impact on the probability of voting for an incumbent party. Hence, finding evaluations from earlier in the term to feed into the vote choice as well shows that voters are not fully myopic. Importantly, this also holds when earlier evaluations exert a smaller effect on the vote than more recent ones; while such a result implies more recent evaluations to be more important than earlier ones, they still show early assessments also to be at play. Furthermore, it is important to note that the coefficient of satisfaction closest to the election is in fact not statistically different from the other satisfaction coefficients (i.e., using Wald tests, the difference between the coefficients is not significant in Model 2, $p=0.11$ or in Model 3, $p=0.15$). Thus, recent performance evaluations do not seem to add significantly more to

⁹ To avoid problems of multicollinearity, we do not include $satisfaction_t$, $satisfaction_{t-1}$ and $satisfaction_{t-2}$ in one model. However, as an additional test, we estimated models including all performance evaluations. In line with the results presented here, all performance evaluations measure throughout the cycle have a significant impact on the vote. The only exception is $satisfaction_{t-3}$ in the second electoral cycle under investigation, which loses its significance. These results are reported in Appendix E.

explaining the party choice than evaluations of earlier in the cycle. Taken together, these results offer further evidence against the claim that voters are fully myopic.¹⁰

¹⁰ Next to these models in which we include the performance measures at different points in the cycle as separate independent variables, we also estimated fixed-effects panel models. These analyses are in line with the retrospective voting theory, as they show that a positive (negative) change in satisfaction with the performance of the government leads to an increase (decrease) in the probability of voting for the incumbent parties. We do not report these analyses here, as they do not allow us to investigate the time horizon voters use when deciding which party to vote for. The results are included in Appendix F.

Table 1: Logistic regression models explaining an incumbent vote measure

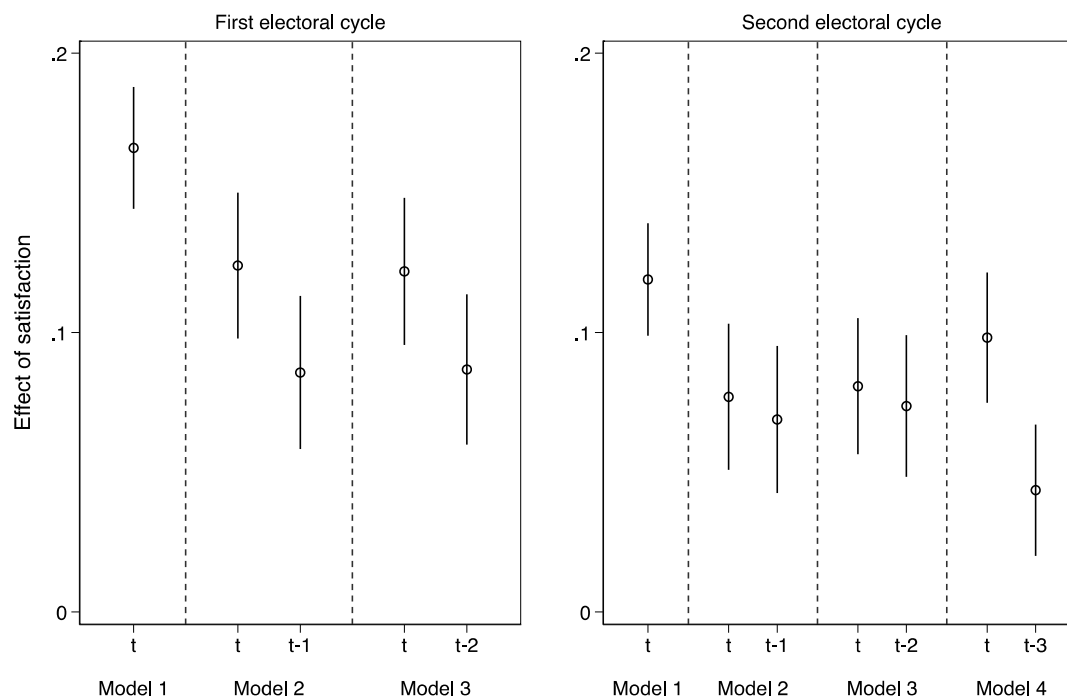
	First electoral cycle			Second electoral cycle			
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 4
Sex (ref.=male)	-0.000 (0.109)	0.018 (0.110)	0.003 (0.110)	0.232* (0.111)	0.224* (0.112)	0.194 (0.112)	0.220* (0.111)
Age	0.031*** (0.004)	0.030*** (0.004)	0.031*** (0.004)	0.005 (0.005)	0.006 (0.005)	0.005 (0.005)	0.005 (0.005)
Employment: employed (ref.)							
Employment: unemployed	0.091 (0.319)	0.115 (0.323)	0.162 (0.322)	-0.571 (0.329)	-0.494 (0.331)	-0.550 (0.331)	-0.574 (0.330)
Employment: other	0.285* (0.120)	0.302* (0.121)	0.251* (0.122)	-0.201 (0.139)	-0.164 (0.141)	-0.191 (0.140)	-0.212 (0.140)
Income	-0.019 (0.030)	-0.028 (0.030)	-0.025 (0.030)	0.119*** (0.029)	0.113*** (0.029)	0.118*** (0.029)	0.116*** (0.029)
Ideology: middle (ref.)							
Ideology: left	0.384** (0.118)	0.364** (0.119)	0.382** (0.119)	0.283* (0.132)	0.307* (0.132)	0.298* (0.133)	0.252 (0.132)
Ideology: right	-0.559*** (0.114)	-0.567*** (0.115)	-0.584*** (0.115)	0.827*** (0.112)	0.824*** (0.113)	0.827*** (0.113)	0.818*** (0.112)
Satisfaction t	0.835*** (0.065)	0.634*** (0.073)	0.624*** (0.073)	0.659*** (0.062)	0.432*** (0.076)	0.454*** (0.072)	0.547*** (0.069)
Satisfaction $t-1$		0.438*** (0.073)			0.386*** (0.077)		
Satisfaction $t-2$			0.445*** (0.072)			0.414*** (0.074)	
Satisfaction $t-3$							0.243*** (0.067)
Constant	-4.556*** (0.316)	-5.317*** (0.349)	-5.127*** (0.337)	-3.931*** (0.338)	-4.350*** (0.354)	-4.398*** (0.354)	-4.169*** (0.347)
N	2222	2222	2222	2295	2295	2295	2295
Pseudo R^2	0.116	0.128	0.129	0.093	0.102	0.104	0.098

Note: Entries are log-odds coefficients, standard errors reported in parentheses. Data: LISS Panel. Significance levels: *: p<0.05; **: p<0.01; ***: p<0.001.

We proceed in the same way for evaluating the role of performance evaluations on the vote in the second electoral cycle. As is clear from the results in Table 1 (right columns), the patterns are largely consistent with what we found for the 2010 legislative election: voters do not seem to be fully myopic, as also their evaluation of the government's performance at the earlier points in the electoral cycle helps in predicting their vote for an incumbent party. Furthermore, as the time-span between the measures is longer during the second electoral cycle than during the first (see Figure 1), these results provide strong evidence against voter myopia. Comparing the coefficients of different performance indicators, the differences in Model 2 and Model 3 are not significant ($p=0.74$ and $p=0.75$ respectively). Only in Model 4, with a performance evaluation measured more than four years before the election, the difference with the evaluation some months before the election is significant ($p=0.009$).

The results presented in Table 1 provide evidence against the argument that voters are fully myopic, as in each case also earlier performance evaluations help in explaining the vote. However, even though performance evaluations measured several years before the election seem to exert a positive impact on the vote, the coefficients do not allow us to interpret their substantive effect sizes. Therefore, we calculate the average marginal effects of these variables – as these have been argued to be well-suited to discuss effect sizes of logistic regression models (Mood, 2010). The average marginal effects are shown in Figure 3.

Figure 3. Average marginal effects of general performance evaluations



Note: The Figure depicts the average marginal effects and 95% confidence intervals of the economic evaluation variables, based on the models presented in Table 1.

The results displayed in Figure 3 provide further evidence against voter myopia. The average marginal effects show that a one-unit increase in satisfaction with the government's performance increases the probability of voting for an incumbent part with 9.5 percentage points on average. Even though there is some variation in the size of the coefficients, this means that performance evaluations over the whole duration of the electoral term exert a substantial effect on the vote.

Generalizability of the results

Our results suggest that voters are not fully myopic – contradicting much previous research. In considering whether our results are generalizable, we wish to point out that our analyses differ in three important aspects from most previous research. First, we investigate voters' time horizon in The Netherlands, a context unlike that of the United States, where most of the relevant investigations have been carried out. Second, while most previous research focused

on economic indicators, we use retrospective performance evaluations of the government ‘in general’. Third, whereas many previous studies used aggregate-level data, we investigate retrospective voter myopia at the individual level. While we argue that the individual-level focus explains the difference between our findings and previous research, other possible factors are yet to be ruled out. Therefore, we go on to test whether these alternative factors can explain our findings.

First, with respect to contextual differences, it is possible that Dutch voters are simply not myopic – irrespective of whether the investigation takes place at the aggregate or at the individual level. One way of testing this argument is to examine whether the conclusion of voters being myopic holds when investigating aggregate economic indicators in The Netherlands. Finding evidence of voter myopia in aggregate-level models in The Netherlands would show that, at least in this respect, it is not an exceptional case for our study.

Therefore, we carry out an aggregate-level study in The Netherlands, in order to see whether it supports the myopic voter thesis. To do this, we regress the economic situation at different points in the electoral term on incumbent vote shares. As indicator of the economy, we use yearly GDP growth, as this allows us to go back in time several decades. Using the data from The Conference Board (2017), this measure is available from 1951 onward – leaving us with data on 20 electoral cycles. Analogous to the individual-level models, we include the evolution of the economy with one-year lags in the model.¹¹ Given the nature of the data and the necessary precautions (Beck and Katz 1995), we tested for autocorrelation in the models, and the Durbin-Watson statistic of initial models indicated problems of autocorrelation. To

¹¹ It needs to be noted, however, that in the Netherlands, the exact date of an election within a calendar year is not fixed. Therefore, following Bélanger and Gélinau (2010, 98), we use the following formula to calculate weighted economic indicators: $\rho = \left[\rho_{(t-1)} * \left(12 - \frac{\sigma(t)}{12} \right) \right] + \left[\rho_{(t)} * \left(\frac{\sigma(t)}{12} \right) \right]$ – in which ρ is the annual GDP growth, σ is the election month, and t is the election year. For instance, for an election held in May 2002, we multiply the GDP growth of 2001 by 7/12 and add it to the GDP growth of 2002 multiplied by 5/12. Using this coding scheme results in GDP growth indicators with each time a lag of one year, regardless in which month the election was held (see also Dassonneville and Hooghe 2017).

correct for this, we present Prais-Winsten regressions using the Cochrane-Orcutt transformation. The results are summarized in Table 2.

Table 2: Aggregate-level model of voter myopia in the Netherlands

	Model 1	Model 2	Model 3	Model 4
	B	B	B	B
	(s.e.)	(s.e.)	(s.e.)	(s.e.)
GDP growth $t-1$	2.618*	4.798**	4.343**	6.011***
	(1.287)	(1.779)	(1.588)	1.752)
GDP growth $t-2$		-0.953		
		(2.004)		
GDP growth $t-3$			-0.803	
			(1.672)	
GDP growth $t-4$				-3.664
				(2.307)
Constant	47.642***	41.227***	42.472***	48.218***
	(5.285)	(5.962)	(7.568)	(8.212)
N	19	15	14	11
R^2	0.196	0.468	0.412	0.595

Note: Entries are unstandardized coefficients, standard errors reported in parentheses. Data: The Conference Board (2017). Significance levels: *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$. As we have the specific expectation of a positive effect of GDP growth on incumbent vote shares, we use one-tailed significance tests.

The results show that relying on an aggregate-level model to study myopia in The Netherlands leads to the conclusion that voters are myopic: In every model, the GDP-growth in the year before the election significantly influences incumbent voting, while the economic situation in the years before does not add to explaining incumbent vote shares. These results seem to rule out the possibility that our individual-level findings are determined by the focus on the Dutch case. i.e., like most other democracies, The Netherlands generates a “myopic” result, when the economic voter model is estimated at the aggregate level. Thus, the results of Table 2 lend support to the generalizability of our findings.

Second, our measure of performance evaluations differs from most previous research. Instead of using economic indicators, we use a more general measure of satisfaction with what the government has done. It may even be considered a measure of ‘popularity’ instead of

‘performance’. To the extent it represents more of a popularity measure, it might also mimic respondents’ vote choice. While we argue that our question measures performance rather than affect, there could be a troubling connection between them. We conducted an additional test to investigate whether our findings are determined by our reliance on a general performance measure. For doing so, we replicate the analyses using the most often used performance evaluation: the evaluation of the economic situation in the country. While this domain is widely investigated, evaluations of the economy also provide a more ‘factual’ measure than the general performance evaluations: while respondents could affectively answer a general measure that does not incentivize them to take into account any objective indicators, probing the economy forces them to take into account the actual performance of the government at least to some extent.

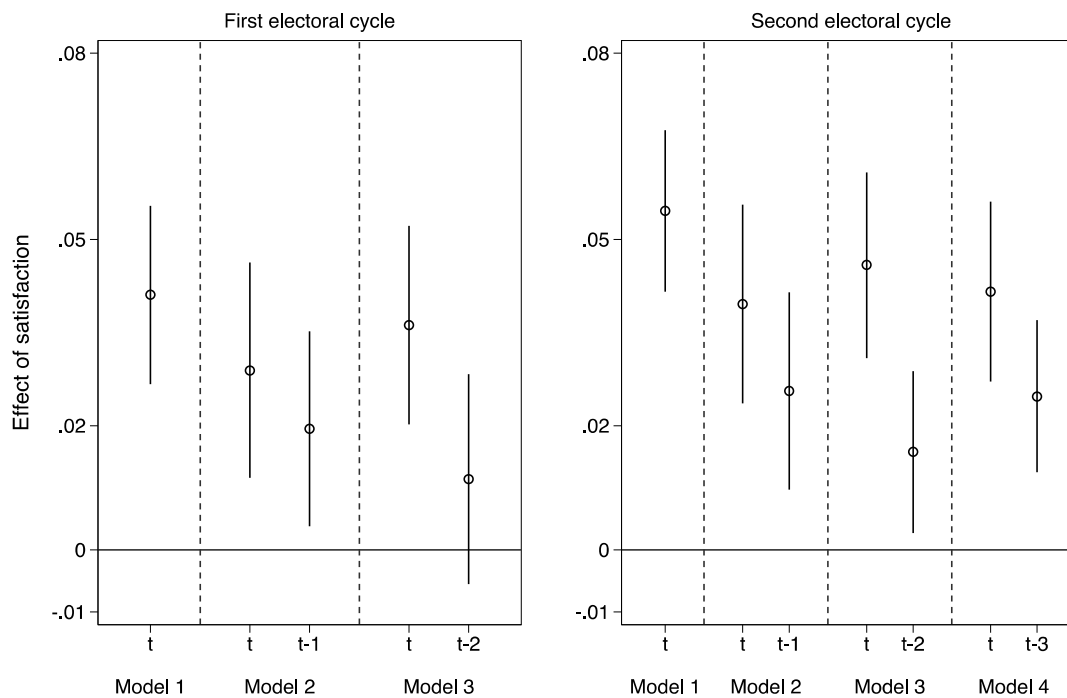
Every year, the respondents of the LISS panel were asked: “How satisfied are you with the way in which the following institutions operate in the Netherlands?” – and for the economy respondents could indicate their satisfaction on a scale ranging from 0 (very dissatisfied) to 10 (very satisfied). We include this evaluation – measured in the respective survey waves – in models that are built analogous to those presented above. The results of the analyses are summarized in Table 3. The average marginal effects are displayed in Figure 4.

Table 3: Replication of Table 1 using economic evaluations

	First electoral cycle			Second electoral cycle			
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 4
Sex (ref.=male)	0.045 (0.109)	0.041 (0.109)	0.050 (0.109)	0.317** (0.115)	0.316** (0.116)	0.316** (0.115)	0.317** (0.116)
Age	0.026*** (0.004)	0.027*** (0.004)	0.026*** (0.004)	0.004 (0.005)	0.005 (0.005)	0.004 (0.005)	0.004 (0.005)
Employment: employed (ref.)							
Employment: unemployed	0.144 (0.315)	0.169 (0.316)	0.149 (0.315)	-0.584 (0.338)	-0.603 (0.339)	-0.581 (0.339)	-0.600 (0.341)
Employment: other	0.369** (0.120)	0.376** (0.120)	0.368** (0.120)	-0.238 (0.144)	-0.236 (0.145)	-0.234 (0.145)	-0.238 (0.145)
Income	-0.003 (0.030)	-0.005 (0.030)	-0.005 (0.030)	0.136*** (0.030)	0.130*** (0.030)	0.132*** (0.030)	0.129*** (0.030)
Ideology: middle (ref.)							
Ideology: left	0.335** (0.118)	0.347** (0.119)	0.335** (0.118)	0.229 (0.136)	0.232 (0.136)	0.241 (0.136)	0.233 (0.136)
Ideology: right	-0.533*** (0.114)	-0.536*** (0.114)	-0.543*** (0.114)	0.742*** (0.115)	0.737*** (0.115)	0.731*** (0.115)	0.722*** (0.116)
Economy' t	0.191*** (0.035)	0.135** (0.042)	0.169*** (0.039)	0.295*** (0.037)	0.215*** (0.045)	0.249*** (0.042)	0.227*** (0.041)
Economy' $t-1$		0.091* (0.038)			0.139** (0.044)		
Economy' $t-2$			0.053 (0.040)			0.085* (0.036)	
Economy' $t-3$							0.135*** (0.035)
Constant	-3.129*** (0.317)	-3.326*** (0.329)	-3.327*** (0.352)	-3.911*** (0.375)	-4.259*** (0.397)	-4.068*** (0.383)	-4.169*** (0.385)
N	2059	2059	2059	2126	2126	2126	2126
Pseudo R^2	0.063	0.065	0.064	0.076	0.080	0.078	0.082

Note: Entries are log-odds coefficients, standard errors reported in parentheses. Data: LISS Panel. Significance levels: * : p<0.05; ** : p<0.01; *** : p<0.001.

Figure 4. Average marginal effects of economic evaluations



Note: The Figure depicts the average marginal effects and 95% confidence intervals of the economic evaluation variables, based on the models presented in Table 3.

The results of the analyses that include economic evaluations are very similar to those including general performance evaluations, as the measures earlier in the term also significantly add to explaining whether or not the voter votes for an incumbent party. The single exception is the first cycle, in which the evaluation measure 2.5 years before the election does not add significantly to explaining the vote. However, in the second cycle we can trace back voters' evaluations even further, and here the *all* evaluations significantly affect the vote. Furthermore, the dynamics of the results are the same, with the coefficients of the performance evaluations in the beginning and at the end of the cycle to be largest – although also in these models the differences in size are not significant.¹² Moreover, while with 3.2 percentage points the average marginal effect is lower than those of general evaluations, the effects are still substantial. Thus, overall, these analyses seem confirm our findings, and show that the conclusions drawn above

¹² The only exception is Model 3 of the second electoral cycle, where the difference is significant, with $p=0.015$.

are not determined by the use of general performance evaluations. Also using individual-level evaluations of the economy, voters do not seem to be myopic.

Self-evidently, our analyses are subject to some limitations. One important limitation of the analyses is the possible reciprocal relationship between performance evaluations and party support. While we argue that performance evaluations influence the vote, it is also possible that these perceptions are, in their turn, influenced by party preferences (Evans and Andersen 2006; Evans and Pickup 2010). To verify whether our results are subject to this reciprocity, we tested the robustness of our models against controlling for autocorrelation and endogeneity in perceptions of performance and the vote. We first did so by estimating models that include a lagged dependent variable (Nadeau, Lewis-Beck and Bélanger 2012). The results of this test are summarized in Appendix H and provide support for the conclusions here. Including the lagged dependent variable in the model as a very strong control still yields the conclusion that voters do not seem to be myopic, as the evaluations throughout the electoral cycle all add to explaining incumbent voting. Furthermore, also in these analyses, the differences in size of the coefficients are non-significant. The only exception is that the coefficient of the evaluation more than four years before the election in the second electoral cycle is not significant. As an additional test, we also exogenized the performance evaluations using the instrumental variables technique. The results of these tests are reported in Appendix I, and provide further support for the findings presented here, as the general performance evaluations at all points during the term add to explaining the vote. Furthermore, in these analyses, there are only small differences in coefficients between recent evaluations and earlier assessments.

Replication using U.S. data

The results of the additional test do not show any reason to assume that the results presented here would be any different in other countries. However, the best test of the generalizability of the results is to replicate the analyses using data from another country. While our models impose strict data requirements that are hard to meet, we are able to replicate the analyses in the U.S. context – the case under investigation in most previous studies on voter myopia.

In our replication, we use the data of The American Panel Survey (TAPS). TAPS is an online panel survey based on a national probability sample of the population of the United States using an address-based sampling frame. As was the case for the LISS survey, respondents that did not have online service, were provided with a computer and internet access by TAPS. The surveys were conducted for the Weidenbaum center at Washington University.¹³ Using these data, we can replicate our models for the electoral term 2012-2016.

For comparability, we build our models as similar as possible to those in the Netherlands. The dependent variable is a binary indicator of whether the respondent voted for the incumbent party (i.e. the Democrats, code 1) in the presidential election of 2016, or not (code 0). The independent variable of interest is the respondent's perception of the change in economic conditions in the country. In every survey wave – i.e., every month – respondents were asked: 'Do you think that the economic conditions in the country as a whole are getting better, getting worse, or not changing much.' Respondents could answer using a scale ranging from 'getting much better' and 'getting somewhat better' over 'not changing much', to 'getting somewhat worse' and 'getting much worse'. Building our models as similar as possible to those in the Netherlands, we include each time the evaluation of the economy right before the election (i.e., October 2016), and include the evaluation in the month of December each time one year earlier respectively.¹⁴ One drawback of the monthly surveys, however, is that some respondents

¹³ For more information, see: <https://taps.wustl.edu/>

¹⁴ Hence, in our models, t refers to October 2016, $t-1$ to December 2015, $t-2$ to December 2014, $t-4$ to December 2013, and $t-4$ to December 2012.

do not take part in some survey waves. This leads to a high number of missing values on our key variables. As a solution, if the economic evaluation from December of a certain year is lacking, we substitute it with the evaluation from November of the same year (September in 2016).¹⁵

In line with the models in the Netherlands, we add controls for age, sex, employment status (1=is working at a job for pay), income (household income), and ideology (5-point liberal-conservative scale). More details on these variables and how they are coded, is included in Appendix G. The results of the replication, are summarized in Table 4.

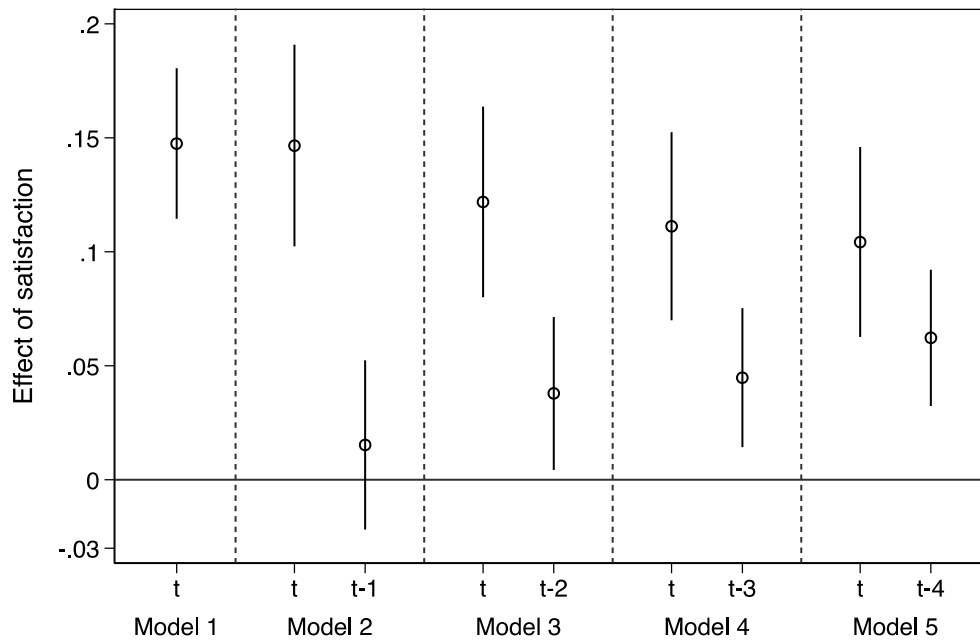
Table 4: Logistic regression models explaining an incumbent vote measure in the U. S.

	Model 1	Model 2	Model 3	Model 4	Model 5
Sex (ref.=male)	0.689** (0.211)	0.648** (0.234)	0.839*** (0.226)	0.781*** (0.218)	0.679** (0.232)
Age	0.010 (0.008)	0.008 (0.009)	0.016 (0.009)	0.009 (0.009)	0.015 (0.009)
Employment: paid job (ref.)					
Employment: no paid job	0.287 (0.249)	0.108 (0.274)	0.439 (0.267)	0.396 (0.259)	0.374 (0.275)
Income	0.034 (0.029)	0.032 (0.032)	0.034 (0.030)	0.034 (0.030)	0.032 (0.032)
Ideology	-1.686*** (0.143)	-1.702*** (0.160)	-1.710*** (0.152)	-1.605*** (0.145)	-1.507*** (0.158)
Economy t	1.222*** (0.162)	1.277*** (0.217)	1.080*** (0.203)	0.936*** (0.188)	0.891*** (0.192)
Economy $t-1$		0.133 (0.165)			
Economy $t-2$			0.335* (0.153)		
Economy $t-3$				0.377** (0.134)	
Economy $t-4$					0.532*** (0.136)
Constant	-1.641 (0.870)	-1.772 (0.972)	-2.569** (0.938)	-1.997* (0.909)	-2.759** (0.987)
N	794	678	753	760	682
pseudo R^2	0.448	0.473	0.480	0.455	0.462

Note: Entries are log-odds coefficients, standard errors reported in parentheses. Data: TAPS. Significance levels: * : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$.

¹⁵ In all respective years, approximately half of the evaluations are from December, and the other half from November.

Figure 5. Average marginal effects of performance evaluations in the U.S.A.



Note: The Figure depicts the average marginal effects and 95% confidence intervals of the economic evaluation variables, based on the models presented in Table 4.

The results in Table 4 provide support for the conclusions we draw from the Dutch data. Also investigating the U.S. case, we find that voters are also influenced by past perceptions. With the exception of the evaluation of the economy 11 months before the election (i.e., December 2015 – $t-1$ in the table), all economic evaluations significantly add to explaining a vote for the incumbent party. It needs to be noted, however, that recent evaluations exert a larger effect on the vote than earlier assessments in all cases, and other than in the analysis of the LISS panel, in this case, the differences in coefficient between the measures at time t and $t-x$ are significant (model 2: $p < 0.001$; model 3: $p = 0.017$; model 4: $p = 0.043$). Interestingly, the only exception is model five ($p = 0.189$). Hence, this result indicates that the evaluation of the economy right after the start of the term of the new president adds to predicting incumbent voting to the same extent as the evaluation just before the new election. Furthermore, as explained above, even though the effects of earlier assessments are smaller, their significant effects provide evidence against

the view of fully myopic voters. This finding is further corroborated by the average marginal effects that are presented in Figure 5. With an average effect of 8.8 percentage points, the results displayed in this figure provide further support for the generalizability of the findings in the Netherlands.

Conclusion

We set out to investigate the extent to which voters use performance evaluations throughout the electoral cycle. Earlier studies that investigated retrospective voting have raised doubts about the voter's ability to comprehensively assess the government's track record throughout the whole term, and to use this information in deciding which party to vote for. More specifically, voters are claimed to be myopic, relying almost exclusively on recent performance evaluations in making their party choice. However, previous research has been mostly based on aggregate-level data or experimental approaches. Instead, we relied on individual-level evaluations of the government's performance in order to investigate the extent to which individual voters keep the track record of the government in mind across the electoral cycle. Investigating individual-level perceptions, we argue that some voters might be using an online information processing model as opposed to a memory-based model. Following the online model, we argued that voters do not store objective information throughout the electoral cycle, but regularly update their overall evaluation of the incumbent government as a running tally. It is the state of the running tally at the end of the term that ultimately influences the vote. Hence, voters need not to remember accurately the objective state of the country throughout the electoral cycle to be able to take into account their evaluations of past government performance during the whole term in their vote.

Using panel data from The Netherlands, the results of our analyses indicate that performance evaluations from early in the government's term, as well as later, explain the party

choice: performance evaluations of 2.5 and 4 years before the election respectively add to explaining incumbent voting. We also underlined that these individual-level results from The Netherlands need not be suspect, since the country generates standard myopic economic voter results when tested at the aggregate level and equivalent results can be obtained using a standard economic evaluation measure. Our replication in the U.S. context that shows evidence against voter myopia in American presidential elections also corroborate these arguments. Hence, our results are consistent with an online-model of information processing. However, we acknowledge that, for now, it is largely speculative, and later research could focus specifically on the question whether voters use a memory-based or online processing model when they vote retrospectively, or whether they combine both in a hybrid model (Kim and Garrett 2012).

Our analyses indicate that voters are not myopic, so going against the common belief that voters almost exclusively take into account the performance of the government close to Election Day. Rather than a *myopic* voter, we find evidence of an *abiding* voter – a voter who is attentive, evaluating the government’s performance throughout the whole term, bringing all these evaluations to bear before voting. Challenging the common belief, these findings hold strong implications for the broader study of voting behavior. Voter myopia has been argued to be a threat to successful electoral control over the incumbent parties, as it would lead politicians to focus their efforts only on election years (Achen and Bartels 2016). Hence, politicians would only need to care about their constituents’ opinions when Election Day is close, while they were free to act in the first part of the term without being held accountable for their behavior. Our study – the first of its kind to use individual-level perceptions of the government’s performance at different points in the electoral cycle – shows that voters might be better able to act as a principal over their agents than was assumed so far. Indeed, an abiding voter is steady at her or his post, evaluating the incumbent’s actions and decisions during the whole term.

As a final note, we would like to point out that, even though we provide many tests of generalizability as well as a replication in the U.S. context, our two cases under investigation are both Western old democracies, and it can be expected that this helps voters to abide. In countries with a long democratic tradition, voters might have learned about punishing and rewarding incumbents on Election Day. Hence, investigating these cases makes it more likely to find voters abiding – as they are familiar with the way in which they can hold rulers to account. Other factors as well can be expected to facilitate or inhibit the abiding voter. It might be easier, for instance, to abide when there is a high level of clarity of responsibility, while low clarity might make voters more likely to resort to using recent information only, or heuristics.¹⁶ As we do not have sufficient cases to investigate the effects of contextual factors on voter myopia, this remains speculation, and future research could investigate which contextual factors facilitate or inhibit the abiding voter.

This unique study shows a voter who abides, attentive to the government's performance, rewarding and punishing accordingly in an additive calculation across the electoral cycle. A task for future research is to investigate why aggregate-level analyses fail to expose this mechanism. For now, we keep to the optimistic finding that performance assessments throughout the whole term in office matter on Election Day.

¹⁶ On this point, note, however, that the Netherlands has a long tradition of coalition governments, which decreases the clarity of responsibility.

References

- Achen, Christopher, and Bartels, Larry. 2016. *Democracy for Realists. Why Elections Do Not Produce Responsive Government*. Princeton: Princeton University Press.
- Beck, Nathaniel, and Katz, Jonathan K. 1995. "What to Do (and Not to Do) with Time-Series Cross-Section Data." *American Political Science Review* 89 (3): 634–647.
- Bélanger, Éric, and Gélinau, François. 2010. "Does Perceived Competence Matter? Political Parties and Economic Voting in Canadian Federal Elections." *Journal of Elections, Public Opinion and Parties* 20 (1): 83-101.
- Boyne, George A., James, Oliver, John, Peter, and Petrovsky, Nicolai. 2009. "Democracy and Government Performance: Holding Incumbents Accountable in English Local Governments." *Journal of Politics* 71(4): 1273–1284.
- Chong, Dennis, and Druckman, James N. 2010. "Dynamic Public Opinion: Communication Effects over Time." *American Political Science Review* 140(4): 663-680.
- Conover, Pamela J., Feldman, Stanley, and Knight, Kathleen. 1986. "Judging Inflation and Unemployment. The Origins of Retrospective Evaluations." *Journal of Politics* 48 (3), 565-588.
- Crisp, Brian F., Olivella, Santiago, Potter, Joshua D., and Mishler, William. 2014. "Elections as Instruments for Punishing Bad Representatives and Selecting Good Ones." *Electoral Studies* 34: 1–15.
- Dassonneville, Ruth, and Hooghe, Marc. 2017. "Economic Indicators and Electoral Volatility: Economic Effects on Electoral Volatility in Western Europe, 1950-2013." *Comparative European Politics* 15 (6): 919-943.
- de Vries, Catherine, and Giger, Nathalie. 2014. "Holding Governments Accountable? Individual Heterogeneity in Performance Voting." *European Journal of Political Research* 53 (2): 345–362.
- Duch, Raymond, and Stevenson, Randolph T. 2008. *The Economic Vote. How Political and Economic Institutions Condition Election Results*. New York: Cambridge University Press.
- Duncan, Fraser. 2017. "Gender, Voting and Christian Democratic Parties." *European Politics and Society* 18 (2): 218–244.
- Ecker, Alejandro, Glinitzer, Konstantin, and Meyer, Thomas M. 2016. "Corruption Performance Voting and the Electoral Context." *European Political Science Review* 8 (3): 333–354.
- Evans, Geoffrey, and Pickup, Mark. 2010. "Reversing the Causal Arrow: The Political Conditioning of Economic Perception in the 2000-2004 Presidential Election Cycle." *Journal of Politics* 72(4): 1236-1251.

- Evans, Geoffrey, and Andersen, Robert. 2006. "The Political Conditioning of Economic Perceptions." *Journal of Politics* 68 (1): 194-207.
- Ferejohn, John. 1986. "Incumbent Performance and Electoral Control." *Public Choice* 50(1): 5-25.
- Fiorina, Morris P. 1981. *Retrospective Voting in American National Elections*. New Haven: Yale University Press.
- Hastie, R., and Park, B. 1986. The Relationship Between Memory and Judgment Depends on Whether the Judgment Task is Memory-Based or On-Line. *Psychological Review*, 93 (3) 258-268.
- Healy, Andrew, and Lenz, Gabriel S. 2014. "Substituting the End for the Whole: Why Voters Respond Primarily to the Election-Year Economy." *American Journal of Political Science* 58 (1): 31-47.
- Healy, Andrew, and Malhotra, Neil. 2009. "Myopic Voters and Natural Disaster Policy." *American Political Science Review* 103 (3): 387-406.
- Healy, Andrew, and Malhotra, Neil. 2013. "Retrospective Voting Reconsidered" *Annual Review of Political Science* 16: 285-306.
- Healy, Andrew, Persson, Mikael, and Snowberg, Erik. 2017. "Digging into the Pocketbook: Evidence on Economic Voting from Income Registry Data Matched to a Voter Survey." *American Political Science Review* 111(4): 771-785.
- Hellwig, Timothy, and Marinova, Dani M. 2015. "More Misinformed than Myopic: Economic Retrospections and the Voter's Time Horizon." *Political Behavior* 37 (4): 865-887.
- Hibbs, Douglas A. 2000. "Bread and Peace Voting in U.S. Presidential Elections." *Public Choice* 104 (1): 149-180.
- Holmes, Jennifer S., and Gutiérrez de Piñeres, Sheila Amin. 2012. "Security and Economic Voting: Support for Incumbent Parties in Colombian Presidential Elections." *Democratization* 20 (6): 1117-1143.
- Howe, Paul. 2006. "Political Knowledge and Electoral Participation in the Netherlands: Comparisons with the Canadian Case." *International Political Science Review* 27 (2): 137-166.
- Huber, Gregory A., Hill, Seth J., and Lenz, Gabriel. 2012. "Sources of Bias in Retrospective Decision Making: Experimental Evidence on Voters' Limitations in Controlling Incumbents." *American Political Science Review* 106 (4): 720-741.
- Key, Valdimer Orlando. 1966. *The Responsible Electorate. Rationality in Presidential Voting 1936-1960*. Cambridge: Belknap Press of Harvard University Press.

- Kim, Young. M., and Garrett, Kelly (2012). "On-Line and Memory-Based: Revisiting the Relationship Between Candidate Evaluation Processing Models." *Political Behavior*, 34 (2), 345-368.
- Kinder, Donald R., and Kiewiet, Roderick D. 1979. "Economic Discontent and Political Behavior: The Role of Personal Grievances and Collective Economic Judgments in Congressional Voting." *American Journal of Political Science* 23 (3): 495–527.
- Kramer, Gerald H. 1971. "Short-Term Fluctuations in U.S. Voting Behavior, 1896-1964." *American Political Science Review* 65 (1): 131–143.
- Kramer, Gerald H. 1983. "The Ecological Fallacy Revisited: Aggregate-Versus Individual-Level Findings of Economics and Elections, and Sociotropic Voting." *American Political Science Review* 77 (1): 92–111.
- Lewis-Beck, Michael, and Stegmaier, Mary. 2000. "Economic Determinants of Electoral Outcomes." *Annual Review of Political Science* 3: 183–219.
- Lewis-Beck, Michael, and Stegmaier, Mary. 2013. "The VP-Function Revisited: A Survey of the Literature on Vote and Popularity Functions After Over 40 Years." *Public Choice* 157 (3): 367–358.
- Lodge, Milton, McGraw, Kathloon M., and Stroh, Patrick. 1989. An Impression-Driven Model of Candidate Evaluation. *American Political Science Review*, 83 (2), 399-419.
- Lodge, Milton, Steenbergen, Marco R., and Brau, Shawn. 1995. "The Responsive Voter: Campaign Information and the Dynamics of Candidate Evaluation." *American Political Science Review* 89(2): 309-326.
- Manin, Bernard, Przeworski, Adam, and Stokes, Susan, C. 1999. Elections and Representation. In Adam Przeworski, Susan C. Stokes, and Bernard Manin (Eds.), *Democracy, Accountability, and Representation* (pp. 29–54). Cambridge: Cambridge University Press.
- 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, Richard, Lewis-Beck, Michael S., and Bélanger, Eric. 2012. "Economics and Elections Revisited." *Comparative Political Studies*, 46 (5), 551-573.
- Nannestad, Peter, and Paldam, Martin. 1994. "The VP-Function: A Survey of the Literature on Vote and Popularity Functions After 25 Years." *Public Choice* 79 (3): 213–245.
- Norrander, Barbara, and Wilcox, Clyde. 2008. "The Gender Gap in Ideology." *Political Behavior* 30 (4): 503–523.
- Owen, Andrew, and Tucker, Joshua A. 2010. "Past is Still Present: Micro-Level Comparisons of Conventional vs Transitional Economic Voting in Three Polish Elections." *Electoral Studies* 29 (1): 25–39.

- Paldam, Martin, and Nannestad, Peter. 2000. "What do Voters Know About the Economy? A Study of Danish Data, 1990-1993." *Electoral Studies* 19 (2): 363–391.
- Przeworski, Adam, Stokes, Susan C., and Manin, Bernard. (Eds.). 1999. *Democracy, Accountability, and Representation*. Cambridge: Cambridge University Press.
- Sanders, David. 2000. "The Real Economy and the Perceived Economy in Popularity Functions. How Much do Voters Need to Know? A Study of British Data, 1974-97." *Electoral Studies* 19 (2-3): 275-294.
- Shabad, Goldie, and Slomczynski, Kazimierz M. 2011. "Voters' Perceptions of Government Performance and Attributions of Responsibility: Electoral Control in Poland." *Electoral Studies* 30 (2): 309–320.
- Singer, Matthew M. 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.
- Stiers, D. forthcoming. "Beyond the Distinction Incumbent-Opposition. Retrospective Voting on the Level of Political Parties." *Party Politics*, accepted.
- Strøm, Kaare. 2000. "Delegation and Accountability in Parliamentary Democracies." *European Journal of Political Research* 37 (3): 261–290.
- Swank, Otto H., and Eisinga, Rob. 1999. "Economic Outcomes and Voting Behaviour in a Multi-Party System: An Application to the Netherlands." *Public Choice* 101 (3): 195–213.
- The Conference Board. 2017. Total Economy Database. <https://www.conferenceboard.org/data/economydatabase/> (Accessed 20 September 2017).
- Tilley, James, Garry, John, and Bold, Tessa. 2008. "Perceptions and Reality: Economic Voting at the 2004 European Parliament Elections." *European Journal of Political Research* 47 (5): 665–686.
- van der Eijk, Cees, Schmitt, Hermann, and Binder, Tanja. 2005. Left-Right Orientations and Party Choice. In Jacques Thomassen (Ed.), *The European Voter. A Comparative Study of Modern Democracies* (pp. 167-191). Oxford: Oxford University Press.
- van der Meer, Tom W. G., van Elsas, Erika, Lubbe, Rozemarijn, and van der Brug, Wouter. 2015. "Are Volatile Voters Erratic, Whimsical or Seriously Picky? A Panel Study of 58 Waves Into the Nature of Electoral Volatility (The Netherlands 2006-2010)." *Party Politics* 21 (1): 100-114.
- van Elsas, Erika, Lubbe, Rozemarijn, van der Meer, Tom W. G., and van der Brug, W. 2013. "Vote Recall: A Panel Study on the Mechanisms That Explain Vote Recall Inconsistency." *International Journal of Public Opinion Research* 26 (1): 18–40.
- Wlezien, Christopher. 2015. "The Myopic Voter? The Economy and US Presidential Elections." *Electoral Studies* 39: 195–204.

Zaller, John R. and Feldman, S. (1992). A Simple Theory of the Survey Response: Answering Questions versus Revealing Preferences. *American Journal of Political Science* 36 (3), 579-616.