Political Parties and the Timeline of Elections

Will Jennings
University of Southampton
w.j.jennings@soton.ac.uk

Christopher Wlezien
University of Texas at Austin
Wlezien@austin.utexas.edu

Abstract:

Scholars are beginning to understand the evolution of electoral sentiment across countries. Recent research shows that early vote intention polls – from years before Election Day – contain substantial information about the final result but that they become increasingly informative over the election cycle. The degree to which this is true varies across countries in important and understandable ways given differences in political institutions, but the pattern is strikingly general. What we do not know is whether and how the characteristics of political parties matter. Do preferences evolve differently for niche and catch-all parties? For government and opposition parties? For new and old parties? This paper addresses these issues. We consider differences in political parties and how they might impact voter preferences over the course of the election cycle. We then outline an empirical analysis relating support for parties in pre-election polls to their final vote in legislative elections. The analysis relies on 23,000 vote intention polls in 31 countries since 1942, covering 212 discrete electoral cycles and encompassing 236 political parties. Our results indicate that party characteristics are important to the structure and evolution of preferences, and that the size and age of parties matter most of all. Preferences for smaller and older parties crystallize early and remain strikingly stable over the course of the election cycle by comparison with larger and newer parties. Though the patterns are as we expected, the details are somewhat surprising, as we will see.

Acknowledgements: Prepared for presentation at the Annual Meeting of the Midwest Political Science Association, Chicago, April 16-19, 2015. For poll data, we owe special thanks to Chris Anderson, Joe Bafumi, Graziella Castro, Jan Drijvers (TNS-Global), Robert Erikson, Johanna Laurin Gulled (Ipsos Public Affairs), Chris Hanretty, Michael Marsh, Michael Thrasher, Jack Vowles, and Francisco José Veiga,. Thanks also to the Roper Center for Public Opinion Research, Australian Social Science Data Archive, Norwegian Social Science Data Services, Data Archiving and Networked Services in the Netherlands, Canadian Opinion Research Archive, and GESIS/Leibniz Institute for the Social Sciences. We are indebted to Yulia Marchenko from StataCorp for assistance with programming of multiple imputation, and David Baker from High Performance Computing at the University of Southampton.

How do election outcomes come into focus over the electoral cycle? Are voters' preferences in place early? Or do they come into focus very late? Do preferences evolve in a patterned way? Answers to these question tell us a lot about the underlying structure of preferences and also the effects of election campaigns. If preferences are highly structured and in place early on, voters are less subject to influence during the official election campaign. By contrast, if preferences are not highly structured, voters may be influenced by everything that happens between elections. While we know a lot about what voters do on Election Day, we know comparatively little about the evolution of electoral sentiment over time (for reviews of the literature, see Jacoby 2010; Heath 2010; Dalton and Klingemann 2007).

What we do know mostly comes from the United States. Erikson and Wlezien (2012) found that, at the beginning of the election year, there is virtually no relationship between the results of presidential polls and the actual vote. By Election Day, poll results virtually match the final result. In between, polls become more and more accurate. Bafumi et al (2010) detect a similar pattern in US Congressional relying on the "generic" ballot, in which respondents are asked which party's candidate for Congress they would vote for in in their district. These preferences are more informative than presidential polls are (of the presidential vote) early in the election year but less informative at the end of the election cycle. Polls for parliamentary elections in the UK are informative much earlier still (Wlezien et al 2013), and begin to come into focus years before Election Day.

Very recent comparative analysis supports and extends these findings. Jennings and Wlezien (N.d.) examine the polls-vote relationship in over 300 election cycles in 45 countries and reveal a

general pattern: polls become increasingly informative about the vote over the election cycle but very early polls contain substantial information about the final result. There also demonstrate significant variation: the evolution of preferences differs across countries, reflecting differences in political institutions. First, preferences come into focus much later in presidential elections than parliamentary (and other legislative) elections. Second, preferences are more structured throughout the election cycle in party-centric electoral systems.

The results of previous research are intuitively satisfying. They indicate that voters' support for political parties do come into focus over time – and in seemingly sensible ways – and that the pattern differs across government and electoral institutions in understandable ways. The research does not consider differences in political parties themselves. Do electoral preferences evolve similarly for all parties? Or does the pattern differ depending on their characteristics? There are numerous possibilities here, most notable of which may be whether parties are in government or opposition, as is suggested by the literature on economic voting (e.g. Fiorina 1981; Anderson 1995; Duch and Stevenson 2008). Another is whether parties are catch-all or niche. The size and age of parties also could matter. The impact of party characteristics also could depend on the type of government and electoral institutions. Clearly, there are many possibilities.

In this paper, we consider whether and how the characteristics of political parties structure the evolution of electoral preferences in systematic ways. To begin with, we review the previous research and consider how differences in political institutions might impact voter preferences over the election cycle. We then describe our empirical analysis relating support for political parties in pre-election polls to their final Election Day vote. The results reveal that party

characteristics are important to the structure and evolution of preferences, and that the size and age of parties matter most of all. Preferences for smaller and older parties crystallize early and remain strikingly stable over the course of the election cycle by comparison with larger and newer parties. This may come as little surprise. What may surprise is that the difference between catch-all and niche parties matters little when taking into account party size.

Polls and the Vote over the Election Timeline

Consider the timeline of elections (following Erikson and Wlezien 2012; also see Wlezien and Erikson 2002). We start the timeline immediately after the previous election. We end it on Election Day. Many events occur over the timeline, some very prominent and others routine. We want to know whether these events have effects. We also want to know whether the effects last.

Ideally we would use time series analysis. That is, we would examine the relationship between polls at different points in time within the various election years taken separately or pooled together. For instance, we could estimate the following equation of the vote division (V_t) in the polls during a particular election cycle to be of the following form:

$$V_t = \alpha + \beta V_{t-1} + u_t, \quad (1)$$

where V_t is the vote percentage for a particular party and u_t is a series of random campaign shocks.[1] That is, preferences on one day are modeled as a function of preferences on the preceding day and the new effect of campaign events, broadly defined.

[1] These are assumed to be independent and drawn from a normal distribution.

In theory, this equation tells us exactly what we want—whether there are shocks to preferences (and how much) and whether changes to preferences decay or last. In practice, however, it is not so straightforward because of data limitations. There are two main problems. First, data frequently are commonly missing for daily and weekly periods and sometimes even for months. This has fairly obvious implications for what we can do with standard time series techniques. Second, the ratio of sampling variance to the true variance often is quite large. This has substantial, if less obvious, complications: the presence of sampling (and other survey) error makes it difficult to uncover the underlying process.

What can we do instead? Erikson and Wlezien (2012) proposed treating the data not as a series of time series but as a series of cross-sections—across elections—for each day of the election cycle. With the data organized as a series of cross sections, we can assess how polls and the vote across elections match up at different points in time. Specifically, we can estimate the following equation relating the Election Day vote across elections *j* to the polls across those same election years on each day *T*, which indicates the number of days before the election:

$$VOTE_j = a_T + b_T V_{jT} + e_{jT} \quad . \tag{2}$$

We are interested in seeing how the regression coefficients (b_T) and the root mean squared errors (RMSEs) evolve over time. Sampling error is not a problem for such an exercise; whereas error may swamp the variance from true change when observing within-election polls, it is dwarfed by

election-to-election (and party-to-party) differences in the cross-section.[2] However, the problem that pre-election polls are sometimes sparse and conducted at irregular intervals remains.

When readings of electoral preferences are missing, we can interpolate daily voter preferences from available polls. For any date without a poll, an estimate is created as the weighted average from the most recent date of polling and the next date of polling. Weights are in proportion to the closeness of the surrounding earlier or later poll.[3] Where we interpolate, we also introduce a random component based on the poll variance -- controlling for country, party and election -- to reflect the uncertainty associated with the imputed values. We thus are able include in our analysis any election cycle from the moment the first poll is conducted in that cycle. This would not be acceptable in conventional time series analysis, as interpolating would compromise the independence of observations. Given that the methodology is explicitly cross-sectional, there is

[2] Consider that, when measured across presidential elections in the US between 1942 and 2008, the variance in the vote exceeds the error variance by a factor of 50 or more. For instance, when the vote is measured as 30-day cross-sections, the minimum of the estimated reliabilities is 0.98, i.e., virtually all of the difference *across elections* is real.

[3] Specifically, given poll readings on days $t - x$ and $t + y$, the estimate for a particular day t is generated using the following formula:

$$\hat{V}_t = \{ [y * V_{t-x} + x * V_{t+y}] / (x + y) \} + \varepsilon,$$

where ε, is drawn from the following distribution: $\mu=0, \sigma=3.394$. Recall that for days in the timeline after the final poll before an election, we carry forward the numbers from the final poll. This has some consequence for the accuracy of poll predictions very close to Election Day, as we use polls from well before the end of the cycle in some cases.

no such problem—interpolating actually permits a more fine-grained analysis. The main drawback of the approach is that we cannot assess whether dynamics differ across particular elections. Importantly, the method allows us to assess patterns of correspondence across elections in different subsets of elections, e.g., across types of systems and political parties.

Research on Polls and the Vote

As discussed, there is evidence of lasting influences from pre-election polls in various countries. In these vote intention polls, survey organizations typically ask respondents which candidate they would vote for "if the election were held today." The results of these polls tell us only a little about the persistence of specific events, as it is difficult to even identify their effects.[4] The polls can tell us quite a lot about general patterns, however, as we can assess how poll results at different points of the election cycle match the final results. If polls are increasingly informative across the timeline, then we know that electoral preferences change and the some of the change lasts to impact the outcome. If polls are equally well informative across the timeline, then either (1) preferences do not change or else (2) preferences do change but these innovations do not

[4] It is difficult to characterize the effects of events for at least three reasons. First, the effects of most events are small, with exceptions such as party nominating conventions (e.g., Holbrook 1996; Shaw 1999; Erikson and Wlezien 2012) and possibly candidate debates (e.g., Johnston et al 1992; Holbrook 1996; Shaw 1999; Blais et al 2003). Second, survey error makes the effects of events hard to detect (see Wlezien and Erikson 2001; Zaller 2002). Third, the net effects of different campaign activities can cancel out. For additional details, see Erikson and Wlezien (2012).

persist, i.e., the fundamentals remain the same.[5]

Scholars have found that, at the beginning of the election year in the US, some 300 days before the balloting, there is virtually no relationship between the results of presidential polls and the actual vote. At the end of the cycle, by contrast, poll results virtually match the final result. In between, polls become more and more accurate. This is revealing about voter preferences. It tells us that they change over the election year and in meaningful ways: although much of the change that we observe is short-lived, and dissipates before Election Day, a *substantial portion carries forward to impact the final outcome*.[6]

Polls and the vote in US Congressional elections exhibit a similar pattern (Bafumi et al 2010). In the "generic" ballot, survey organizations ask respondents which party's candidate for Congress they would vote for in in their district. These measured preferences are more informative early in the election year. They also are less informative towards the end of the election cycle. Polls for parliamentary elections in the UK are informative earlier still (Wlezien et al 2013), starting to come into focus years before Election Day.

Very recent comparative analysis (Jennings and Wlezien, N.d.) supports and extends these findings based on analysis of polls and vote in over 300 election cycles in 45 countries. To begin with, they demonstrate that polls become more revealing about the outcome the closer the

[5] Where the latter is true, we may see a late uptick in the correspondence between polls and the vote owing to short-term effects that arrive late and do not fully decay before Election Day.

[6] Voters are, at least to some extent, "online processors," updating their preferences based on new information about the parties and candidates (see Lodge et al 1995).

election in all countries, but that early polls do contain substantial information about the final result. They also show that the pattern differs and that political institutions matter. First, electoral preferences come into focus much later in presidential elections than in legislative elections in parliamentary and presidential systems. Second, preferences are consistently more structured -- from early on in the election cycle until Election Day -- in party-centric systems, which mostly are in countries with proportional election rules. While the research teaches us quite a lot about the evolution of electoral preferences across countries, it leaves a lot unexplained. In particular, it reveals little about whether and how patterns differ across political parties.

Political Parties and the Polls and the Vote

Political parties are central to voters' electoral preferences, and this is true whether people actually voter for parties or candidates (see, e.g., Duverger 1954; Campbell et al 1960; Lipset and Rokkan 1967; for a review see Boix 2007). Parties differ in many ways, of course, and there is reason to expect that some of the differences matter for the structure and evolution of electoral preferences over time. There are a number of leading suspects in the literature on party systems and political behavior: catch-all vs. niche parties, small vs. large parties, government vs. opposition, and old vs. new parties. Let us onsider how these characteristics might influence the formation and stability of preferences across the timeline of elections.

Catch-all vs. Niche Parties

Much research, at least since Kircheimer (1966), recognizes the difference in the scope of parties. Some "catch-all" parties tend to be more mainstream and focus on a range of issues that

have been and are on the political agenda, and are less tied to the representation of specific sectors of society. Others "niche" parties are less mainstream and concentrate on particular issues that commonly are different from those on which mainstream parties compete (Meguid 2005; 2008). These differences are well-known, and there also is research considering their consequences for party behavior.

There are reasons to think that support for niche parties will be structured earlier in the election cycle. First, their constituency is narrower and concerned with particular niche issues, and so will be less likely switch to opposition (mainstream) parties, i.e., preferences should be fairly stable. Second, the behavior of niche parties also matters, as they tend to maintain a distinctive platform that encourages stable party support. Third, there may be more fluidity, or "switching," in voter choices among catch-all parties, than among niche parties or between niche and catch-all parties. Support for the latter thus should evolve significantly over time.

Research on party behavior tends to support this expectation. Most notable is Adams et al. (2006), which considers whether the type of party makes a difference to dynamic representation of public opinion in party platforms. They found that niche parties do not respond to shifts in public opinion while catch-all parties do. They also found that niche parties seem to have an incentive to not respond, as they are penalized at the ballot box when they do moderate their positions. This is not to say that there are no incentives for niche parties to change their positions (especially see Meyer and Wagner 2013), just that this is less likely than for catch-all parties. The research suggests that electoral preferences for niche parties, further from the

median, will be more durable and will be structured far earlier in the election cycle. Of course, we need to and can directly test the possibility.

Party Size

Political scientists also recognize that the size of parties differs and has consequences for voters (van der Brug et al 2007). This partly relates to the catch-all vs -niche scope of parties discussed just above, as niche parties tend to be small. It also relates to the attribution of responsibility to governing parties, which we consider below. We also can assess the independent effects of size, for instance, by seeing whether and how party size matters *among* catch-all parties, keeping in mind that niche parties are by definition fairly small.

Government vs Opposition Parties

Theories of electoral behavior suggest that voters' preferences are based on the record of parties or candidates. Whether parties are in the government or opposition thus may be important. The (conditional) retrospective voting model of Fiorina (1981) is an important point of theoretical reference here (also see Downs 1957; Key 1966). In the model, voters' preferences are based on the performance of the sitting government, a referendum judgment. The model finds considerable support in work on economic voting (e.g., Powell and Whitten 1993; Lewis-Beck and Stegmaier 2000; Duch and Stevenson 2008), where voters' preferences are a function of economic evaluations. Most of the literature shows (or assumes) that late economic conditions matter—the slope of the economy leading up to Election Day (see Wlezien N.d.). This is supported by broader research on valence politics (e.g., Clarke et al. 2004; 2009). There thus is reason to suppose that preferences for governing parties evolve over time, being less structured

early on and coming increasingly into focus leading up to and during the campaign as voters take stock of performance. Support for opposition parties, by contrast, should tend to more stable and structured as voters do not tend to adjust their preferences in response to new information about the government.

We know that not all governing parties are the same. As noted above, there is reason to suppose that evaluations of economic (and other aspects of) performance are more consequential for the main governing party (see van der Brug et al 2007). This implies that the structure and evolution of preferences will differ, evolving more for large parties in the coalition and less so for smaller ones. A similar difference may pertain to opposition parties, at least to the extent that shifts to and away from the largest governing party are especially felt by the largest opposition parties. Whether this is true can be assessed empirically.

Old vs New Parties

Converse (1969) argues that there is reason to think that the age of democracy is important to the formation and evolution of preferences, as partisan loyalties take time to take root. This led us previously to posit that the age of democracy may be important for the dynamics of electoral preferences, being more fluid in new democracies than in older ones (Jennings and Wlezien N.d.). As Converse's logic centers on partisan loyalties, there really are two expectations, one relating to the age of the party system and the other relating to the age of parties themselves. New parties can emerge in old systems after all, and we can assess whether and how much this matters for preference formation. More specifically, Converse suggests electoral preferences for

older parties will more structured than for younger parties, for which partisan loyalties are less developed.

Poll Data

Pollsters have sought to measure citizen's preferences for candidates or parties for almost three quarters of a century. While varying due to differences in context, most pre-election polls ask how citizens would vote "if the election were held today." We have compiled what we believe is the most extensive comparative dataset ever assembled of national polls of the vote intentions for presidential and legislative elections.[7] Note that exit polls are not included in our analysis. The dataset consists of 26,917 polls spanning the period from 1942 to 2013. (Supplementary Appendix S1 details the sources.) The data cover a total of 312 elections (including 22 run-off elections) in 45 countries, 13 of which are pure presidential systems, 28 of which are parliamentary systems, and 4 of which are mixed, including a president and a parliament. All told, we have poll data for presidential elections in 23 countries and legislative elections in 31 countries, summarized in Table 1. For this paper, we focus only on legislative elections (where we have 22,948 polls). For these 235 elections, we have 740 polls per country on average for approximately 8 elections per country, or about 98 polls per election cycle. Given the average

[7] In every poll in our dataset respondents were asked for which candidate or party they would vote; we ignore cross-national and within-country differences in question wording. Lau (1994) shows that in the US such differences matter little for poll results, McDermott and Frankovic (2003) demonstrate that some are consequential. To the extent wording does matter, it serves to introduce error into our measure of electoral preferences.

interval—1,143 days—between elections, we are missing polls on most dates and in many weeks and even months.

Table 1. Poll Data in 45 Countries, 1942-2013

Country	System	Election	Rule	First poll	Last election
Australia	Parliamentary	Legislative (1st Pref)	SMDP	1943	2013
		Legislative (2nd Pref)*	SMDP	1993	2013
Belgium	Parliamentary	Legislative	PR	2004	2010
Bulgaria	Parliamentary	Legislative	PR	2009	2013
Canada	Parliamentary	Legislative	SMDP	1942	2011
Croatia	Parliamentary	Legislative	PR	2008	2011
		Presidential	Majority	2009	2010
Czech Republic	Parliamentary	Presidential	Majority	2012	2013
Denmark	Parliamentary	Legislative	PR	1960	2011
Finland	Parliamentary	Legislative	PR	2010	2011
Finland	Parliamentary	Presidential	Majority	2006	2012
Germany	Parliamentary	Legislative	PR	1961	2013
Greece	Parliamentary	Legislative	PR	2007	2012
Hungary	Parliamentary	Legislative	PR	2009	2010
Iceland	Parliamentary	Legislative	PR	2009	2012
		Presidential	Plurality	2012	2012
Ireland	Parliamentary	Legislative	PR	1974	2011
Italy	Parliamentary	Legislative	PR	2012	2013
Japan	Parliamentary	Legislative	PR	1998	2012
Malta	Parliamentary	Legislative	SMDP	2012	2013
Netherlands	Parliamentary	Legislative	PR	1964	2012
New Zealand	Parliamentary	Legislative	SMDP/PR	1975	2013
Norway	Parliamentary	Legislative	PR	1964	2013
Poland	Parliamentary	Legislative	PR	2010	2011
		Presidential	Majority	2011	2011
Serbia	Parliamentary	Legislative	PR	2008	2012

Slovakia	Parliamentary	Legislative	PR	2010	2012
Slovenia	Parliamentary	Presidential	Majority	2012	2012
Spain	Parliamentary	Legislative	PR	1980	2011
Sweden	Parliamentary	Legislative	PR	2000	2010
Switzerland	Parliamentary	Legislative	PR	2010	2011
Turkey	Parliamentary	Legislative	PR	2010	2011
U.K.	Parliamentary	Legislative	SMDP	1943	2010
Argentina	Presidential	Presidential	Majority	2006	2011
Brazil	Presidential	Presidential	Majority	2002	2010
Chile	Presidential	Presidential	Majority	2008	2010
Colombia	Presidential	Presidential	Majority	2010	2010
Cyprus	Presidential	Presidential	Majority	2007	2013
Ecuador	Presidential	Presidential	Majority	2010	2013
Mexico	Presidential	Presidential	Plurality	2005	2012
Paraguay	Presidential	Presidential	Plurality	2013	2013
Peru	Presidential	Presidential	Majority	2006	2011
Philippines	Presidential	Presidential	Plurality	2010	2010
South Korea	Presidential	Legislative	PR	2011	2012
		Presidential	Plurality	2012	2012
U.S.	Presidential	Legislative	SMDP	1942	2012
		Presidential	Electoral College	1952	2012
Venezuela	Presidential	Presidential	Plurality	2006	2013
Austria	Semi-Presidential	Legislative	PR	2006	2013
		Presidential	Majority	2010	2010
France	Semi-Presidential	Presidential	Majority	1965	2012
Portugal	Semi-Presidential	Legislative	PR	1985	2011
		Presidential	Majority	2010	2011
Romania	Semi-Presidential	Legislative	PR	2008	2012
		Presidential	Majority	2009	2009

* Polls of two-party preferences under Australia's transferable vote electoral system are excluded from analysis to avoid double-counting.

There are several important points about these data. Firstly, we are compelled to work with vote intention figures that do not reflect consistent sampling or weighting strategies by different polling organizations or even by the same organization over time. Older polls are more likely to use face-to-face and quota samples, for example, whereas recent polls may include internet panels. While we ideally would like to work with data generated using a consistent methodology, assembling a time series that takes into account differences in weighting and sampling practices is impossible, as the required data are not available for most of the polls. We therefore use the headline figure vote intentions as the most consistent attainable time series of poll data—the numbers reflect the survey houses' best estimates of voter preferences at each point in time. Where a survey house changes their sampling or weighting strategies our poll data will reflect this change. Unfortunately, there is little alternative to using the headline figures, as these often are the only available data. It also is the norm in previous research.[8]

Second, survey organizations typically conduct polls over multiple days, which requires a number of coding decisions. To begin with, for organizations reporting moving averages from a tracking poll, we use non-overlapping results. Since most polls are conducted over multiple days, where possible we "date" each poll by the middle day of the period that the survey is in the field. For days when more than one poll result is recorded, we pool the results together into a

[8] These decision rules might seem innocuous but the poll universe and, especially, weighting have been shown to affect the reported headline figures, particularly in recent elections (Moore and Saad 1997; Wlezien and Erikson 2001). This does not appear to influence the evolving accuracy of reported polls, at least based on analysis of presidential and congressional polls in the US (Erikson and Wlezien 2004; 2012; Bafumi et al. 2010).

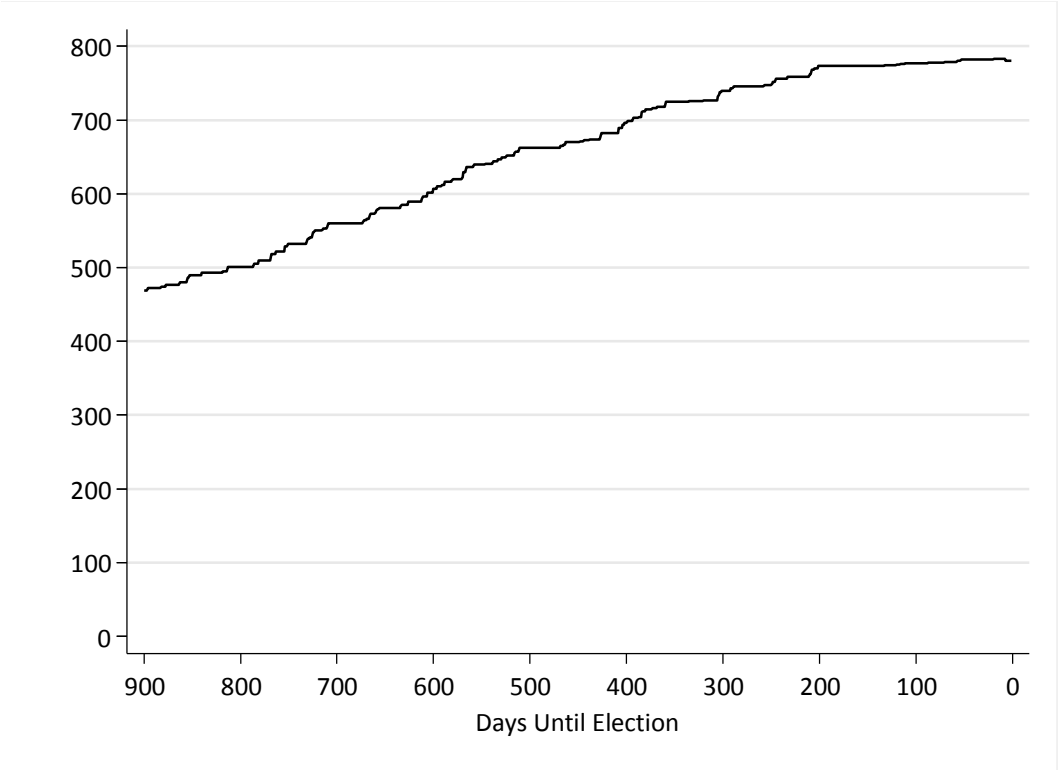
single poll of polls. During the later stages of the election cycle, we often have near day-to-day monitoring of vote intentions.[9]

Third, the length of election cycles vary considerably. Some presidential elections involve a five- or six-year time interval, as do some legislative elections, while run-off elections can span just a couple of weeks, resulting in a very short election cycle. Because pollsters ask hypothetically about vote intentions for run-off elections we are able to extend our analysis beyond this period in some cases. Even where the election cycles are long, polling may begin at different points in time in different cycles. These all mean that the number of election cycles and parties for which we have poll data increases the closer we are to Election Day. This can be seen in Figure 1.

In some countries there are legal restrictions on publication of poll results on or prior to Election Day (for a review see Spangenberg 2003). This means that in some cases we have missing data over the final days of the campaign. In such circumstances, we carry forward the results from the final poll until the very end of the cycle. Thus, our analysis understates the strength of the relationship between polls and the election outcome at that point in time.

[9] It is important to note that polls on successive days are not truly independent. Although they do not share respondents, they do share overlapping *polling periods*. Thus polls on neighbouring days will capture a lot of the same things, which complicates a conventional time-serial analysis.

Figure 1. Number of Parties for which there are Poll Data



Polls and the Vote

For our analysis, we also need data on the actual vote shares parties and candidates receive in the elections for which we have poll data, and for this we rely on a wide range of official sources and election data resources—details are reported in supplementary Appendix S2.

Methodology

Recall that our approach is to estimate a series of daily equations predicting the vote share from vote intentions for different parties (*j*) in different elections (*k*) across countries (*m*) from polls shares on each day of the timeline:

$$VOTE_{jkm} = a_{jmT} + b_T Poll_{jkmT} + e_{jkmT} \quad , \qquad (3)$$

where T designates the number of days before Election Day and a_{jmT} represents a separate intercept for each party j in country m. This is important to include because the level of electoral support in the polls and the vote varies systematically across parties. If our timeline covers the year before Election Day, we would estimate an equation using polls from 365 days before each election, and then do the same using polls from 364 days in advance, and so on up to Election Day itself. Recall that we are missing poll data on most of days, and so we are imputing a lot of data. Because of this, we employ multiple imputation (Rubin 1987), which averages the coefficients across the imputed data series and adjusts the standard errors to reflect noise both due to imputation and residual variance.[10] Using the resulting estimates, we can see whether and how preferences come into focus over time.

We are primarily interested in the explained variances and the regression coefficients (b_T) from these regressions. The former tells us how well the polls *predict* the variation in the vote. For our purposes, the root mean squared error (RMSE) is preferable to the *R*-squared because it allows us to compare different groups of elections or parties where the vote share variances differ. It indicates how much of the actual vote variance is unexplained, for instance, 3.5 percentage points on average for one set of parties by comparison with 1.5 points for another,

[10] Rubin (1987) shows that where γ is the rate of missing data, estimates based on m imputations have an efficiency that approximates to a value of $(1 + \frac{\gamma}{m})^{-1}$. Since polls are missing on around 92% of days we use 50 imputed data series, which implies a relative efficiency of 0.98 compared to an infinite number of imputations.

and so is a particularly useful measure when forecasting, which in effect is what we are doing here.[11] Thus, if the RMSE declines over time, we know that polls increasingly account for the vote the closer the election.[12] A decline in the RMSE would not necessarily mean that the polls themselves are increasingly accurate estimates of the vote. For this, we need the regression coefficient (b_t) relating the polls and the vote, which tells us how much of the poll division lasts to impact the outcome. As the coefficient approaches 1.0, the observed poll becomes the best estimate of the Election Day vote share. If the RMSE also tends toward 0, the polls increasingly *match* the vote.

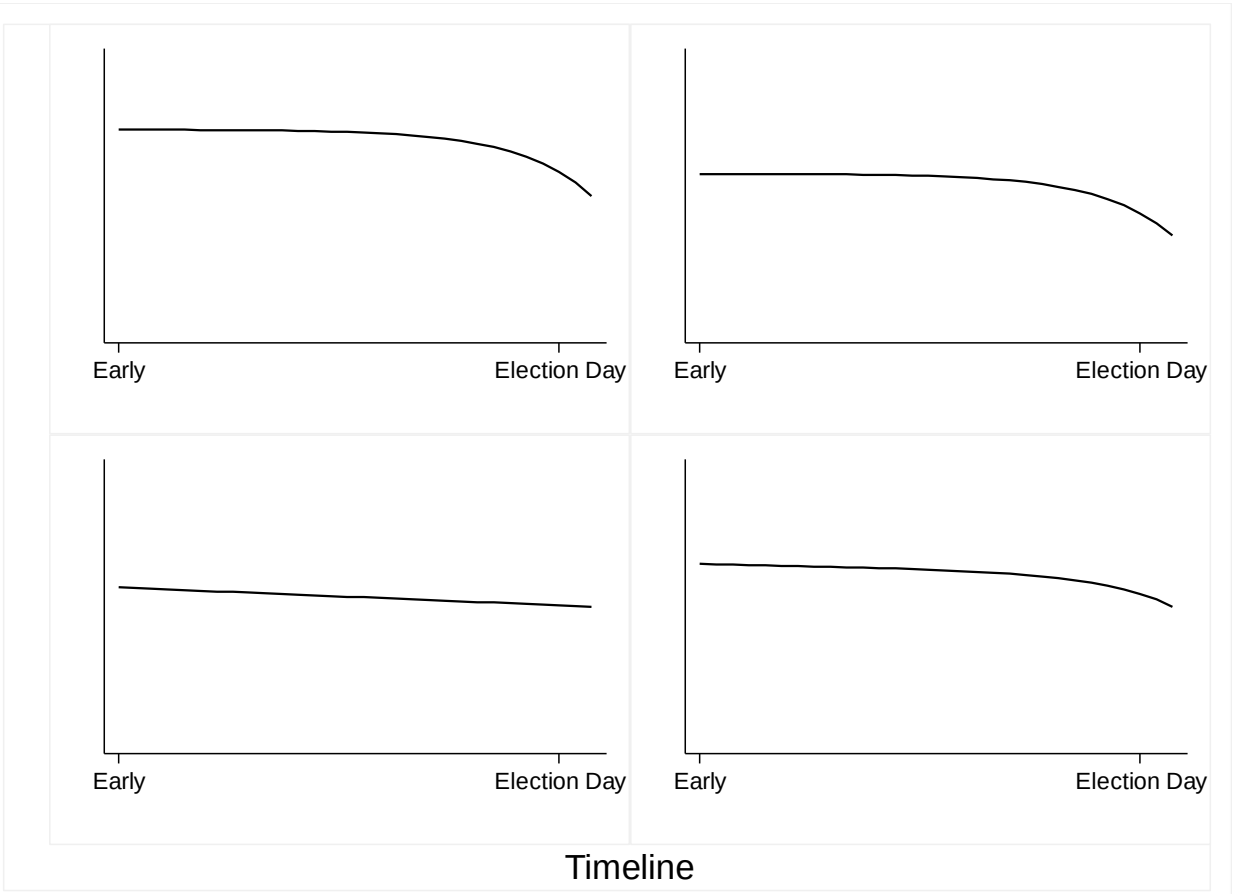
Now, we are interested in how the relationship between the polls and the vote evolves over time. Let use consider what we might observe, focusing on the RMSE. Clearly, if preferences evolve at all, the RMSE will go down over time. That is, the polls would increasingly predict the vote. As Erikson and Wlezien (2012) and Jennings and Wlezien (N.d.) posit, the exact functional form would depend on how much of the variance is due to long-term and short-term components, however. If most of the change in preferences is short-lived, then the vote would come into focus late, as in the upper-left hand frame of Figure 2, where the RMSE remains fairly flat and then drops sharply at the end of the campaign, as late-arriving effects (increasingly) do not fully

[11] As the *R*-squared also is informative (Krueger and Lewis-Beck 2007), it is worth noting that those estimates and the RMSEs *always* are negatively correlated at 0.99 or higher for all of the analyses that follow. This indicates that when the RMSE is lower, the *R*-squared is, almost without exception, higher.

[12] Note that the improvement in predictability will reflect the variance of the shocks and the proportion that persists, bearing in mind that some changes will not last.

decay before Election Day. By contrast, if most of the change in preferences lasts, then the vote would increasingly come into focus over the cycle, as in the lower left-hand frame of Figure 2. If both processes are at work and some of the change decays and some lasts, then the pattern would resemble what is depicted in the lower right-hand frame of the figure. This is the dominant pattern in previous research.

Figure 2. Different Functional Forms of the Evolution of Electoral Preferences



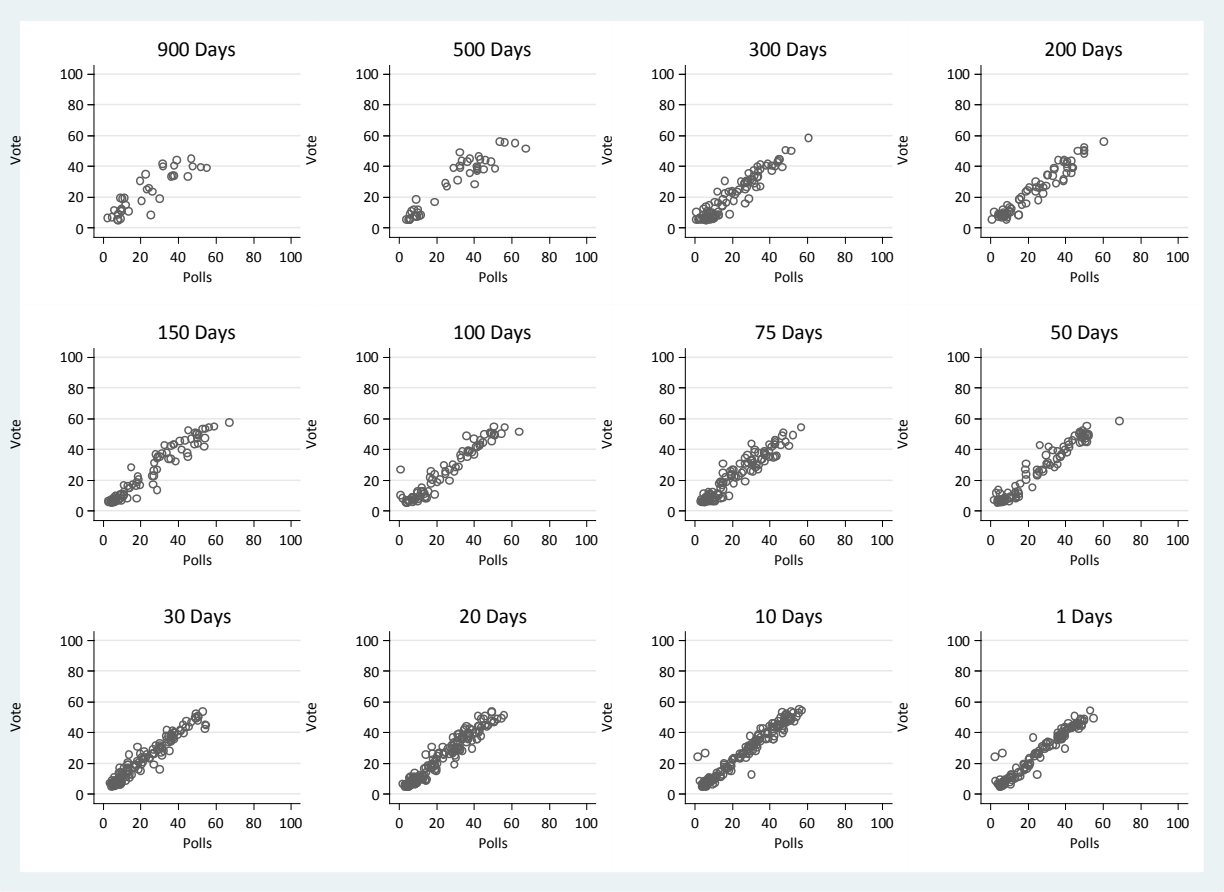
It may be that electoral preferences evolve over the election cycle in a similar way across political systems. It also may be that the pattern differs. Consider our discussion of political parties. There we posited differences between niche and catch-all parties, where electoral preferences are expected to come into focus more quickly for the latter. In terms of Figure 2, we

hypothesize that the RMSEs would be consistently lower for niche parties until the very end of the timeline, when support for catch-all parties come into focus. We also posited differences between opposition and governing parties, which we expect to show a similar, if less pronounced pattern, where preferences for the latter come into focus later. Recall that we think that this may be especially true for large parties. Finally, we posited possible, contrasting differences between new and old parties, where preferences for the former would be less structured early in the election cycle due to less partisan loyalties. These are our hypotheses. Now let us see what the data reveal.

Results

To begin the analysis, let us consider the scatterplot between polls and the vote at various points of the election cycle. This is shown in Figure 3. The figure displays the poll share for all parties or candidates in all elections and countries for which we have actual polls, i.e., excluding imputed polls numbers. In the upper left-hand panel of the figure, using polls that are available 900 days before the election, fully two and a half years before an election, we see that there already is a discernible pattern. That is, the poll share and the vote share are positively related, though there also is a good amount of variation. At that point in time, we have polls in the field in approximately 40% of our cases, and this increases fairly steadily, reaching 75% one year before Election Day. As we turn to polls later in the election cycle, moving horizontally and then vertically through the figure, a clearer pattern emerges; the poll share and final vote share line up. Simply, as we get closer to the election, the polls tell us more about the outcome. It is as one would expect if preferences change and a nontrivial portion lasts. But how much do preferences evolve? How does this depend on party characteristics?

Figure 3. Scatterplot of Party Vote Share by Party Poll Share for Selected Days of the Election Cycle—Pooling all Legislative Elections



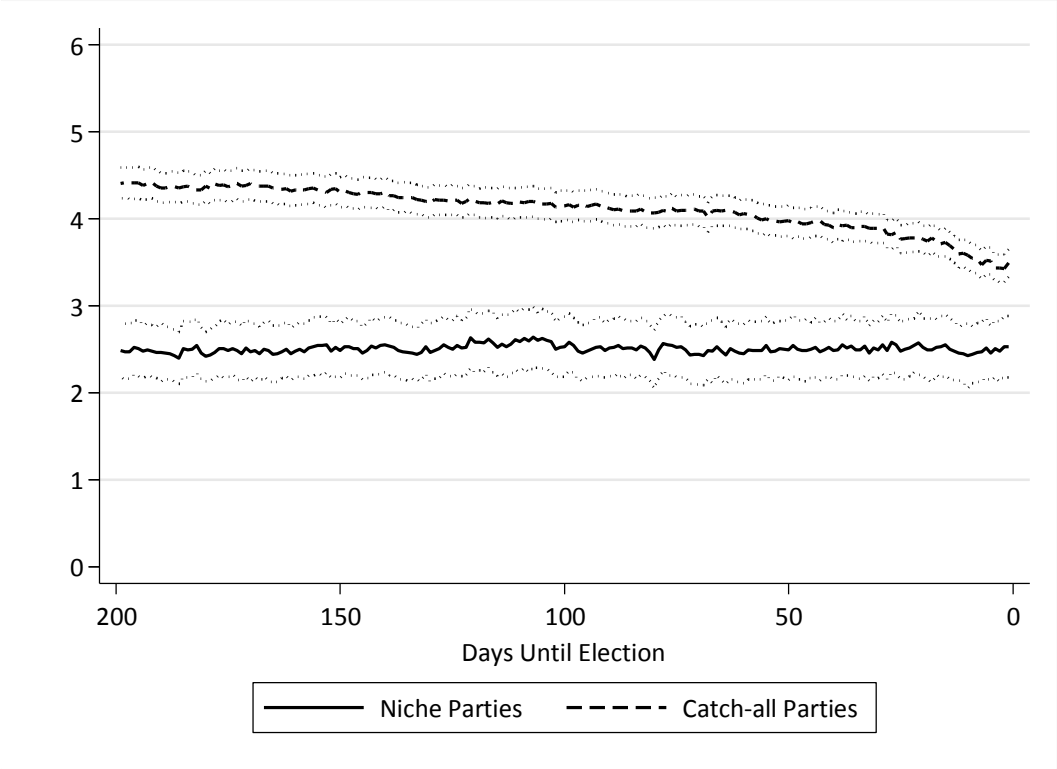
To provide a definitive assessment of party differences, we begin by estimating cross-sectional regressions of the vote division on the poll division for each date and set of parties and compare the relationship over time. Specifically, we focus on the root mean squared error (RMSE), which tells us whether the polls become more informative about the vote as the election cycle unfolds, i.e., the more the polls predict the vote, the lower the RMSE. To calculate the RMSE, one estimates the prediction errors from the regression, squares them, calculates the mean of those squared errors, and then takes the square root of the mean. The regression includes controls for different parties in different countries— which effectively accounts for differences across both

countries and parties, recalling that party variables are country-specific. The regression equation is bootstrapped to estimate the standard errors of the RMSEs, enabling us to determine whether the relationship between polls and the vote differs significantly across institutional settings.[13] For the comparison of RMSE's to be meaningful, we need to restrict our comparison to the set of cases, as the number of parties for which we have polls increases the closer we are to Election Day (see Figures 1 and 3).[14] For this analysis, we focus on all election cycles and parties when polls are in the field 200 days before Election Day, which leaves us with 212 elections and a total of 236 parties.

[13] In bootstrapping the regression, we assume that our sample distribution (a total of 775,703 party*poll *days*) is representative of the general population of polls of vote intentions. This is not an unreasonable assumption, as our data set likely contains the majority of available polls. To bootstrap the estimates, the regression is estimated for randomly drawn resamples (with replacement) of the data repeated 1,000 times for each day of the election cycle. The model is estimated as a linear regression with one categorical factor that allows the effects of party controls to be absorbed.

[14] The regression coefficient (*b*) from the equation relating polls and the vote offers additional, supporting information. The general pattern is that the coefficient grows over the timeline as the RMSE shrinks (Jennings and Wlezien N.d.). This tells us that an increasing portion of the polls lasts to impact the Election Day vote, i.e., the polls converge on the final result.

Figure 4. Root Mean Squared Errors for Niche and Catch-all Parties

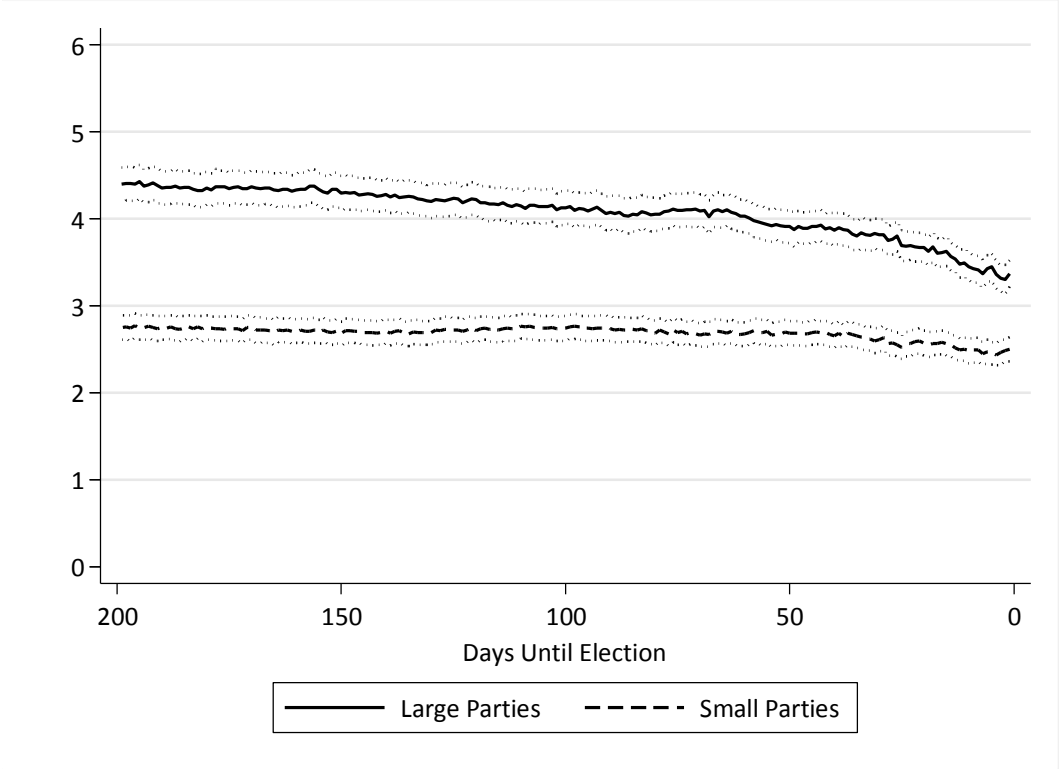


Let us begin with a comparison of catch-all and niche parties. Recall that we hypothesize that electoral preferences of the latter come into focus earlier in the election cycle. To test the hypothesis, we estimate separate equations relating poll and vote shares for the two types of parties. Our coding is based on the Manifesto Research on Political Representation (MARPOR) project database of party election programs which also includes classification of party families (Volkens et al 2014). Following Wagner (2012, p. 845), we classify niche parties as those "parties that compete primarily on a small number of non-economic issues." Our coding of niche parties therefore includes far-right, ethnic-regional, ecological and other special issue parties, while catch-all parties refer to mainstream left, right and center parties. Figure 4 plots the resulting RMSEs over the final 200 days of the election cycle, based on models including party controls. The patterns in the figure are consistent with our expectations. At the beginning of the

timeline, 200 days out, polls are much more informative for niche parties, with an RMSE of about 2.5 percentage points by comparison with 4.5 points for catch-all parties. As can be seen in the figure, the difference is statistically significant; that is, the confidence intervals do not overlap. The gap narrows over time, with preferences for catch-all parties coming increasingly into focus, though niche parties vote shares are more predictable even at the very end. Indeed, the predictability of the niche party vote doesn't change at all over the last 200 days of the cycle. This preferences evidently come into focus very early and stay that way. This is as hypothesized.[15]

[15] We have tested for differences between systems where the incumbent government is able to control the timing of legislative elections and those where it cannot (see Kayser 2005), and results suggest that preferences come into focus earlier in the former and remain so right up until the final days of the campaign. These differences hold – and do not vary significantly – across countries with different government and electoral institutions. For details, see Jennings and Wlezien (N.d.). This makes no difference to our analyses of party characteristics.

Figure 5. Root Mean Squared Errors for Large and Small Parties

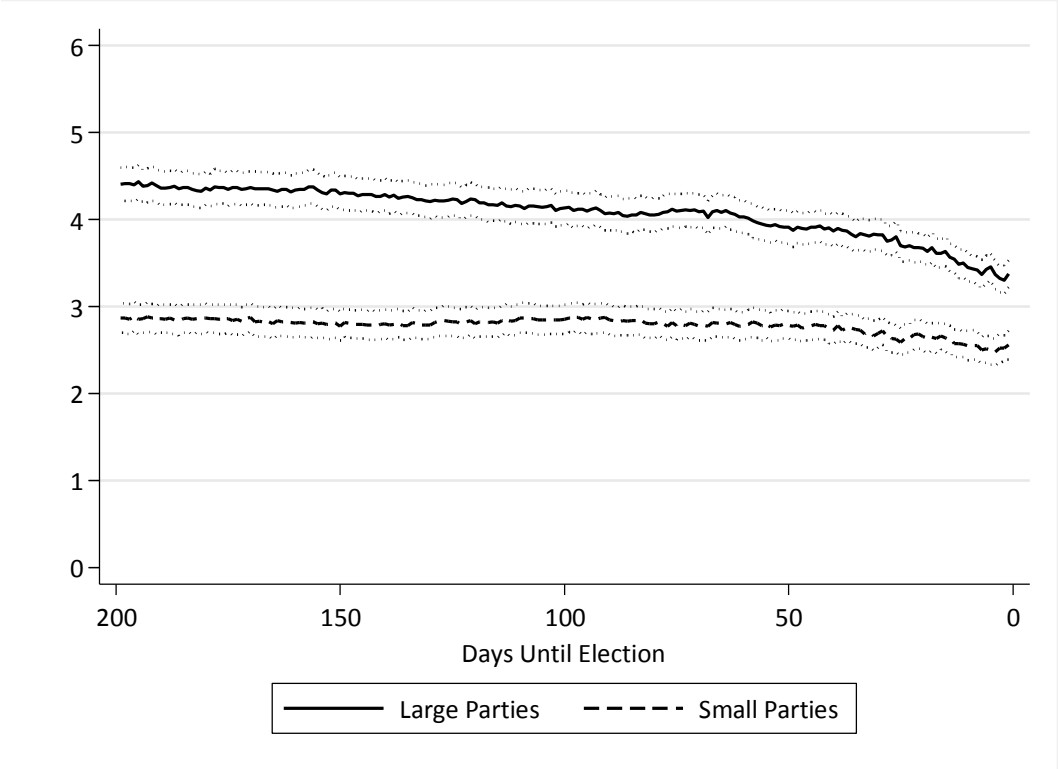


We next turn to differences in party size. As discussed, there is reason to think that electoral support for large parties crystallizes later in the election cycle. Figure 5 shows the RMSEs for regressions relating the polls and the legislative vote for large and small parties, using a 20% vote threshold.[16] The results in the figure indicate that party size makes a big difference in the structure and evolution of preferences. Indeed, the pattern is virtually identical to what we observed for catch-all and niche parties, though there is some evidence that preferences for small parties do evolve over time, by contrast with niche parties. The striking similarity in Figures 4 and 5 highlights the possibility that party size is the determining factor.

[16] Varying the vote threshold, specifically using 15% or 10% instead, makes no real difference.

To assess this possibility, we separate catch-all parties by party size and estimate a separate series of regressions for each. This allows us to see whether and how size matters, controlling for the scope of parties. The resulting RMSEs in Figure 6 look very much like those in Figures 4 and 5, which implies that size is what matters and not party type. Comparing the results for niche parties in Figure 4 and small catch-all parties in Figure 6 suggests slight differences, whereby preferences for the former remain largely unchanged over time and those for latter change over time. These differences are slight, however. The differences between large and small parties are not.

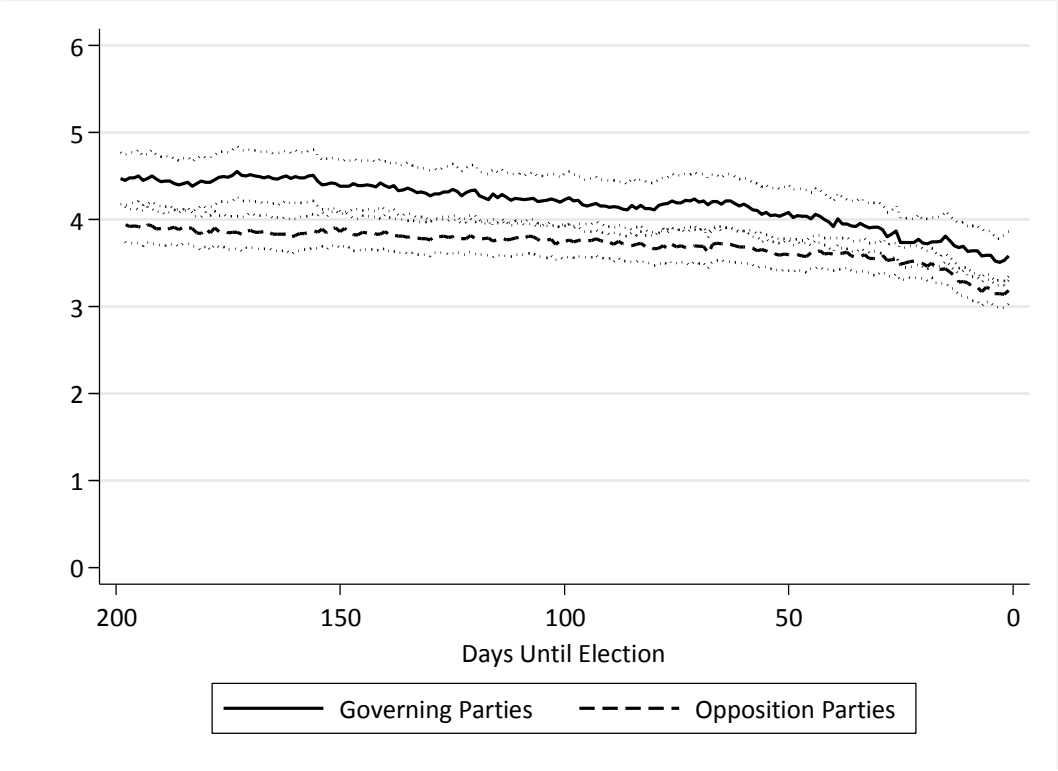
Figure 6. Root Mean Squared Errors for Large and Small Catch-all Parties



Earlier we posited that parties' participation in government also may make a difference for the evolution of electoral preferences. Our expectation is that support for governing parties comes into focus later in the election cycle than that for opposition parties. We also expect that party

size has an effect. Figure 7 shows RMSEs for regressions relating polls and the legislative vote for all government and opposition parties, where the former includes any parties that are part of a governing coalition. There we can see support for our main expectation, as preferences for the latter develop earlier and that preferences for government parties also become more predictive of the vote over time. A gap between the two types of parties remains even at the very end of the campaign, though the difference then is small and not always reliable.[17] This suggests that there are real, if small differences in the evolution of electoral support for governing and opposition parties.

Figure 7. Root Mean Squared Errors for Governing (Coalition) and Opposition Parties



[17] The estimates to some extent exaggerate the differences because we carry forward earlier polls on days when polls are missing in the closing days of campaigns.

As we already have discussed, governing parties are different and opposition parties are too and we expect party size to matter, particularly for the former. Figure 8 plots the RMSEs for large and small governing parties and Figure 9 does the same for opposition parties, again using 20% of vote share as the threshold. In Figure 8 we see that support for large governing parties is less structured throughout the election cycle, right up to Election Day. By contrast, preferences for small governing parties evolve much as for opposition parties in Figure 7, though they are consistently more stable. The results in Figure 9 indicate a similar, if less pronounced, pattern for large and small opposition parties.

Figure 8. Root Mean Squared Errors for Large and Small Governing Parties

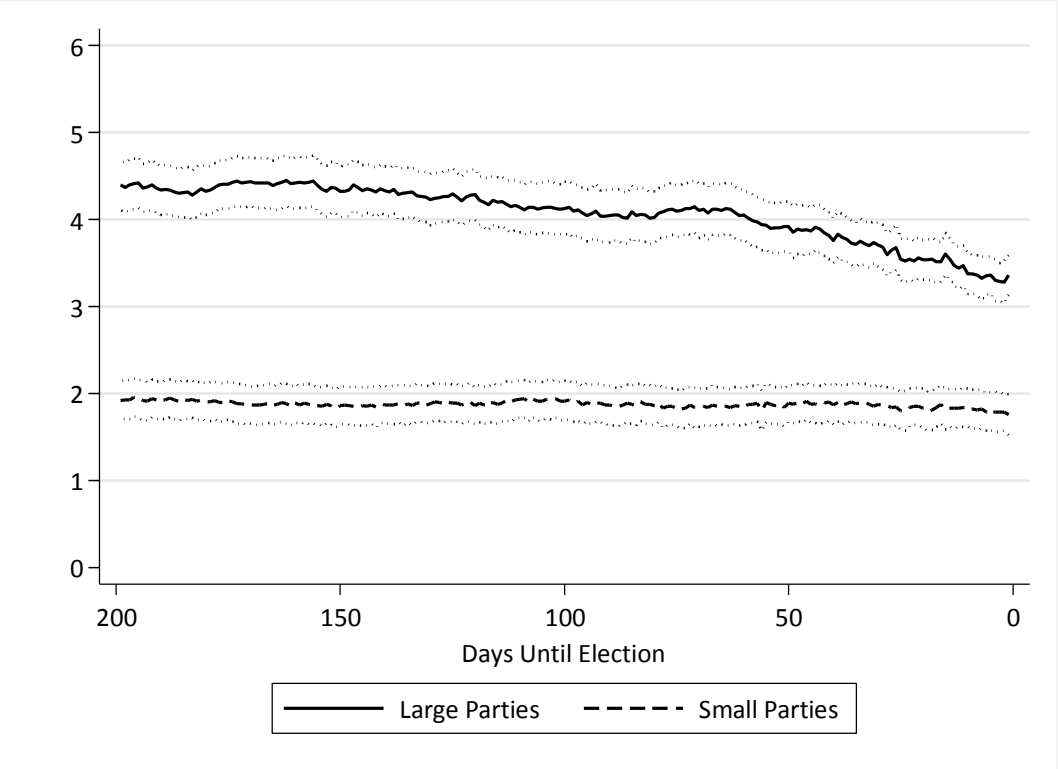
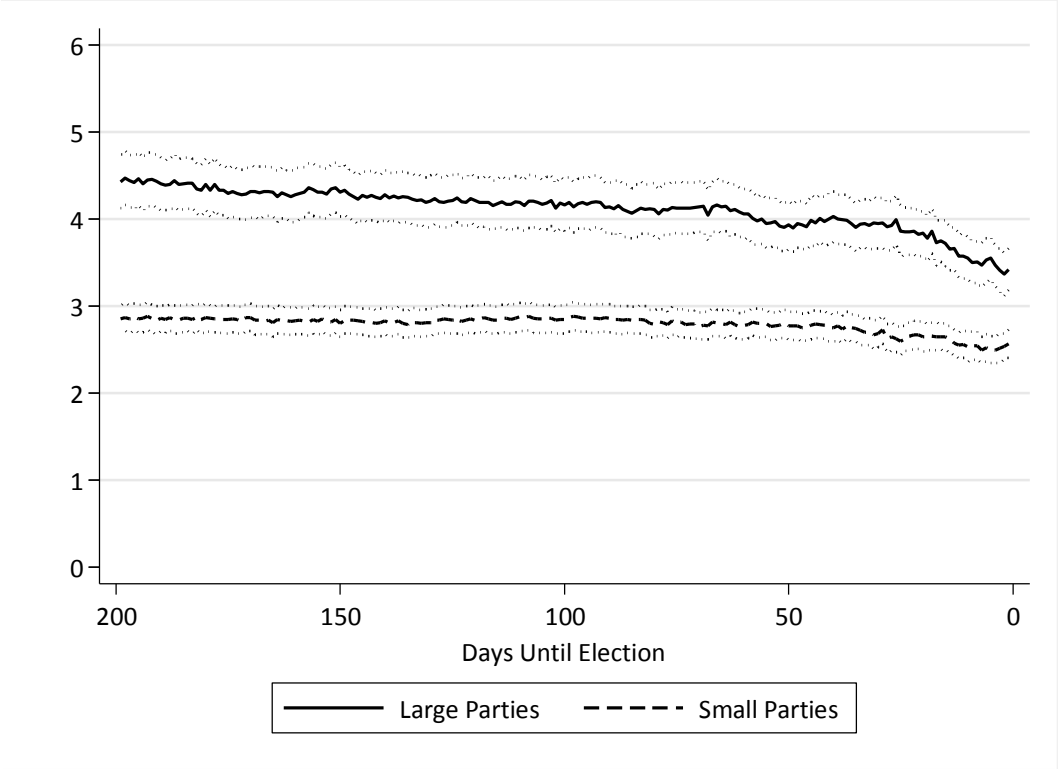


Figure 9. Root Mean Squared Errors for Large and Small Opposition Parties



Thus far, we have seen that party size is an important determinant of the structure and development of electoral preferences, and that it accounts for most of the effects of party scope and government-opposition differences. We have yet to consider the age of parties, however. As discussed, we expect that preferences for older parties remain more structured throughout the election cycle, due to stronger party loyalties. For our analysis, we begin by separating parties into those that formed before 1987 and those that formed in that year or thereafter. The RMSEs for these two groups are shown in Figure 10. It is clear from the figure that preferences for older parties are in place earlier than for newer parties and this remains true throughout the election cycle. This is as expected.

What the analysis does not allow us to distinguish is whether the pattern primarily reflects differences in the age of parties and not the age of representative democracies per se. To assess this possibility, we focus on older democracies -- specifically, democratic regimes established before 1976 -- and examine the differences in the polls-vote relationship for old and new parties in those systems. Figure 11 shows the resulting RMSEs for the two groups. Here we see clear evidence that support for older parties is much more structured early and throughout the election cycle even in older democracies. This confirms and underscores the finding in Figure 11.

Figure 10. Root Mean Squared Errors for Old and New Parties (Parties founded before 1987 or from 1987 onwards)

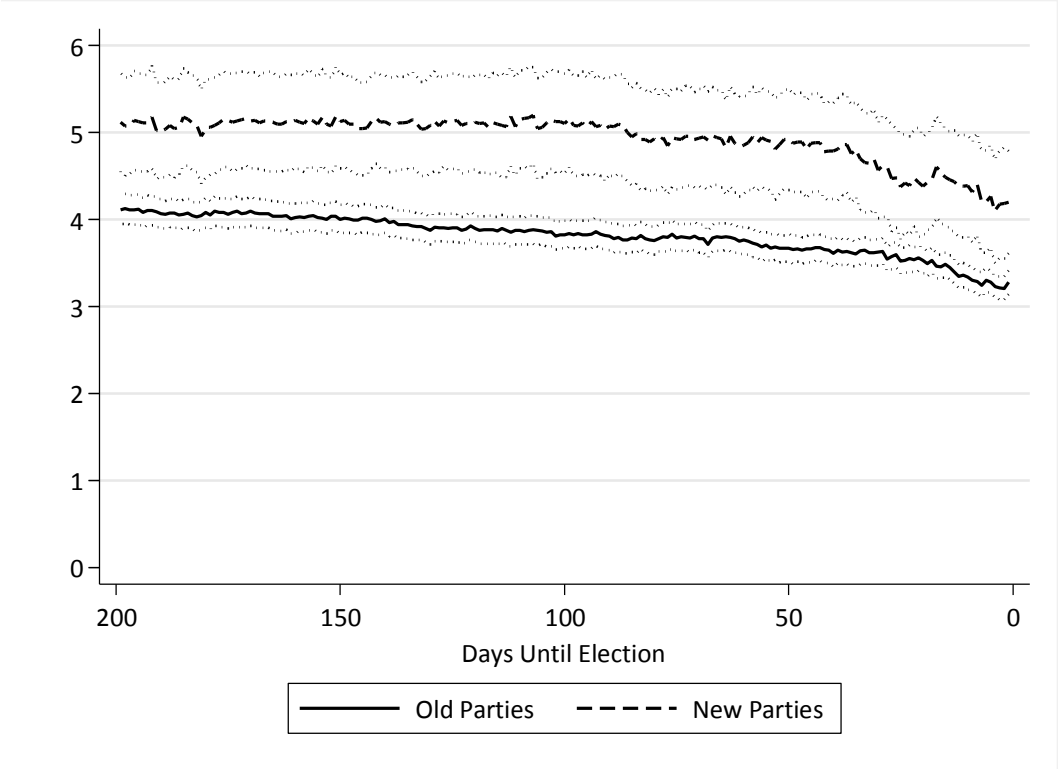
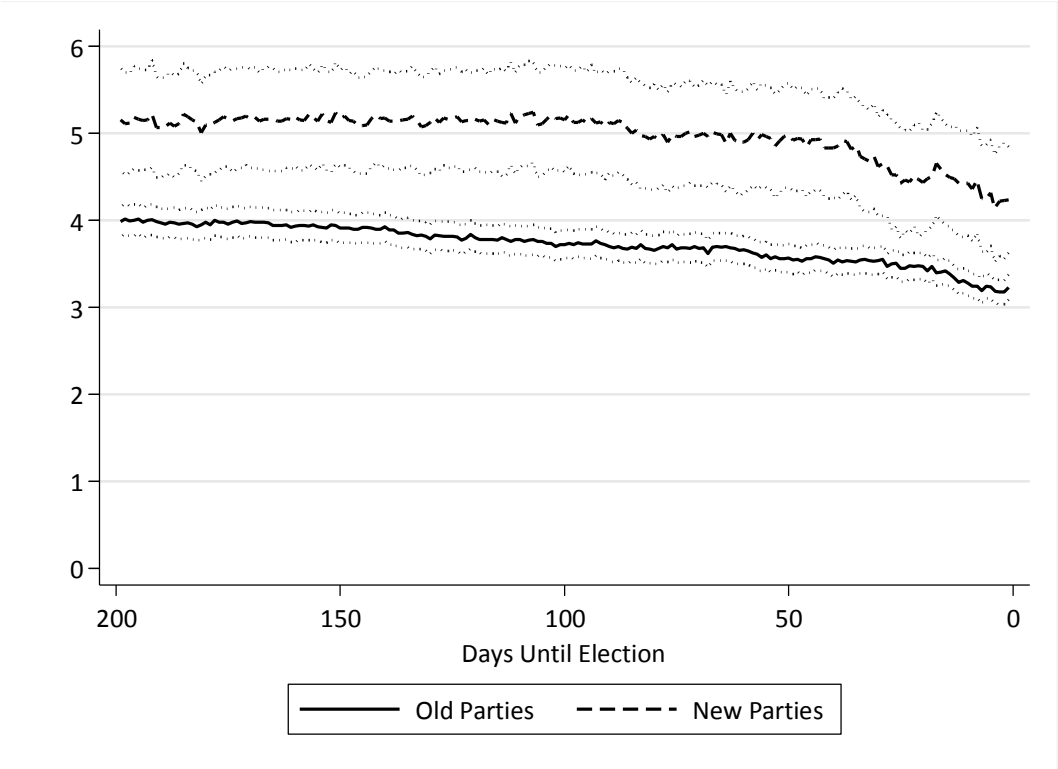
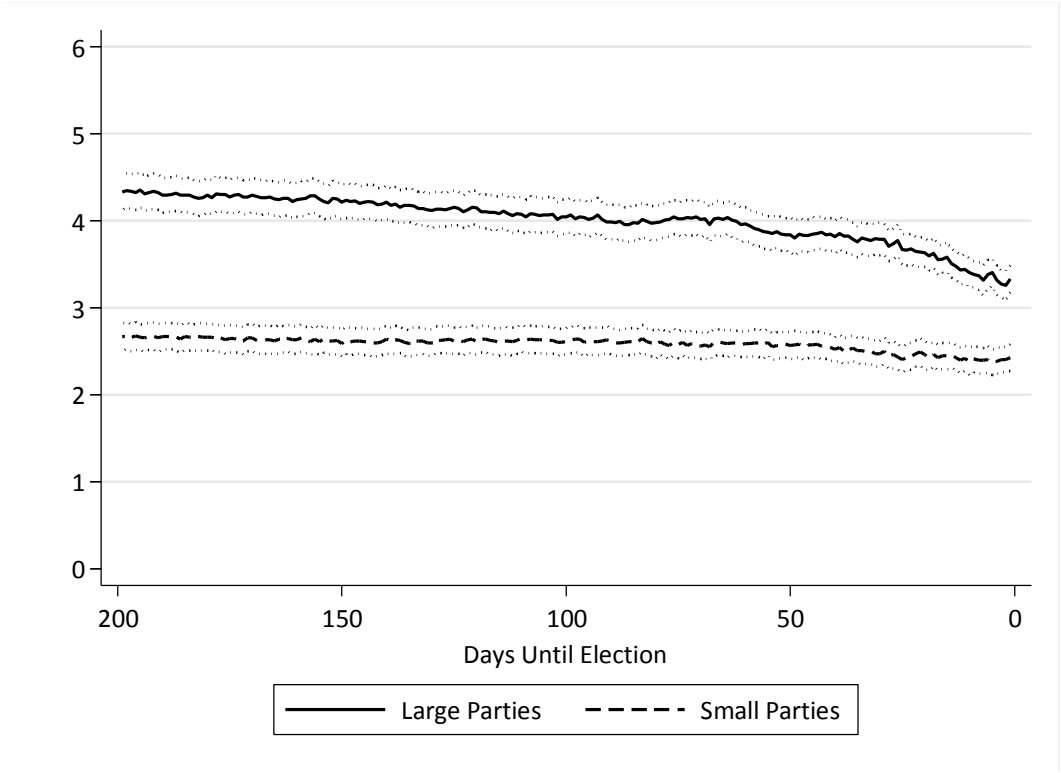


Figure 11. Root Mean Squared Errors for Old and New Parties (Parties founded before 1987 or from 1987 onwards) in pre-1976 democracies



It is worth considering whether and how party size moderates these effects. Figure 12 depicts RMSE's for regressions using large and small parties founded before 1987, as per Figure 10. It is clear in Figure 12 that party size has the effects we've seen throughout our analysis. (Focusing specifically on older democracies, as in Figure 11, produces virtually identical results.) By contrast with what we saw for the differences between niche and catch-all parties and, to a lesser extent, governing and opposition parties, party size does not account for the differences we observe between old and new parties. That is, electoral preferences for big older parties show much more structure than newer parties.

Figure 12. Root Mean Squared Errors for Large and Small Parties founded before 1987



Discussion and Conclusion

Voter preferences evolve in a systematic way over the election timeline in a wide range of representative democracies. There is structure to preferences well in advance of elections, indeed, years before citizens actually vote. That is, very early polls predict the vote, at least to some extent. This largely reflects differences in the equilibrium support of parties. Polls do become increasingly informative over time, however, pointing to real evolution of preferences. That this pattern holds across countries is important and points towards a general tendency in the formation of electoral preferences in legislative elections. But the pattern is not precisely the same for all parties.

Political parties structure the evolution of voters' preferences. The size and age of parties matter most of all. Preferences for smaller and older parties crystallize early and remain strikingly stable over the course of the election cycle by comparison with larger and newer parties. The results for smaller parties may not surprise given that many of those rely on constituencies that care about specific niche issues on which the party leaders concentrate. What may surprise is that the effect of party size holds even for catch-all parties; indeed, the evolution of preferences for niche and small catch-all parties is indistinguishable, which implies that the scope of the parties does not matter much at all. (That said, small parties are more likely to be niche parties and vice versa.) Size moderates the effect of other variables that do influence preference formation, including whether parties are in government or the opposition and the age of parties as well. That size matters may come as little surprise – prediction errors are larger for larger parties, after all – but it also is important to explain.

One possibility is error variance. Based on well-known research regarding sampling-induced variances of proportions, there is reason to expect that the error variance in observed poll shares for smaller parties will be lower than that for larger parties when the true levels of support remain unchanged.[18] The problem with this explanation is that sampling error is quite small by comparison with the prediction error variance, and taking it into account only slightly reduces the gap in the RMSEs, and actually leaves the ratios between RMSEs for large and small parties essentially the same. Another possibility is that small parties tend to be more ideologically extreme, support for which is in place early and enduring, much as we predicted for niche

[18] This expectation is not perfectly clear as it is based on theory regarding dichotomous proportions, which is rarely the case in our data.

parties. The problem with this explanation is that far right parties are niche parties in our analyses and the evolution of preferences for left and right parties are not very different, so the resulting patterns are indistinguishable from what we saw for niche and catch-all parties. In the end, it may simply be that large parties are competing for government control and votes and so their preferences are more elastic. This remains to be seen.

That preferences for older parties are structured early on is more straightforward, and precisely as we expected. After all, there is a now-massive literature demonstrating the importance of partisan loyalties and/or dispositions for electoral preferences and that these tendencies take time to develop. Electoral support for older parties should therefore be in place early and be less subject to change, which is what we observe. Our results show the age of parties matters almost identically when taking into account the age of the representative democracy itself, i.e., the effect of the former is not an artifact of the latter. All of this highlights the importance of parties as a central organizing institution in modern democratic polities.

We have only scratched the surface of the variation in context. To begin with, the effects of party characteristics may be conditioned by electoral and government institutions. Do the patterns hold equally in both plurality and proportional systems? In presidential and parliamentary systems? Clearly, much research remains to be done, and the methodology can guide the way. That said, we have learned something about the general pattern relating preferences and the vote over the election timeline and the structuring influences of political parties. We have seen that preferences are often in place far in advance of Election Day, particularly for smaller and older parties, and that they evolve slowly over time, in some cases, e.g., niche parties, not at all. Indeed, the final outcome is fairly clear in the polls before election

campaigns begin. This is not to say that campaigns do not matter, as they do, particularly for certain types of parties. Even there, it appears that the "long campaign" between elections matters most of all.

REFERENCES

Abramson, Paul R., John H. Aldrich, and David W. Rohde. 2007. *Change and Continuity in the 2004 and 2006 Elections*. Washington, D.C.: CQ Press.

Adams, James, Michael Clark, Lawrence Ezrow, and Garrett Glasgow. 2006. "Are Niche Parties Fundamentally Different from Mainstream Parties? The Causes and the Electoral Consequences of Western European Parties' Policy Shifts, 1976-1998." *American Journal of Political Science* 50(3):513-529.

Adams, James, Samuel Merrill, III, and Bernard Grofman. 2005. *A Unified Theory of Party Competition: A Cross-National Analysis Integrating Spatial and Behavioral Factors*. Cambridge: Cambridge University Press.

Anderson, Christopher. 1995. *Blaming the Government*. New York: M.E. Sharpe

Arceneaux, Kevin. 2005. "Do Campaigns Help Voters Learn: A Cross-National Analysis." *British Journal of Political Science* 36(1): 159-173.

Bafumi, Joseph, Robert S. Erikson, and Christopher Wlezien. 2010. "Balancing, Generic Polls and Midterm Congressional Elections." *Journal of Politics* 72(3):705-719.

Berelson, Bernard R., Paul F. Lazarsfeld, and William N. McPhee. 1954. *Voting: a study of opinion formation in a presidential campaign*. Chicago: University of Chicago Press.

Blais, Andre, Elisabeth Gidengil, Richard Nadeau, and Neil Nevitte. 2003. "Campaign Dynamics in the 2000 Canadian Election: How the Leader Debates Salvaged the Conservative Party." *PS: Political Science and Politics* 36(1): 45-50.

Boix, Carles. 2007. "Emergence of Parties and Party Systems." In Carles Boix and Susan Stokes (eds.) *The Oxford Handbook of Comparative Politics*. Oxford: Oxford University Press.

van der Brug, Wouter, Cees van der Eijk and Mark Franklin. 2007. *The Economy and the Vote: Economic Condition and Elections in 15 Countries*. Cambridge: Cambridge University Press.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. New York: Wiley.

Campbell, James E. 2000. *The American Campaign: U.S. Presidential Campaigns and the National Vote.* College Station, Texas: Texas A&M University Press.

Clarke, Harold, David Sanders, Marianne Stewart and Paul Whiteley. 2004. *Political Choice in Britain*. Oxford: Oxford University Press.

Converse, Philip E. 1969. "Of Time and Partisan Stability." *Comparative Political Studies* 2(2): 139-171.

Dalton, Russell J. and Hans-Dieter Klingemann (eds.). 2007. *Oxford Handbook of Political Behavior*. Oxford: Oxford University Press.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper.

Duch, Raymond and Randolph Stevenson. 2008. *The Economic Vote*. Cambridge: Cambridge University Press.

Duverger, Maurice. 1954. *Political Parties: Their Organization and Activity in the Modern State*. New York: Wiley.

van der Eijk, Cees and Mark Franklin. 1996. *Choosing Europe? The European Electorate and National Politics in the Face of Union*. Ann Arbor: University of Michigan Press.

Erikson, Robert S. and Christopher Wlezien. 2012. *The Timeline of Presidential Elections: How Campaigns do (and do not) Matter*. Chicago: University of Chicago Press.

Finkel, Steven E. 1993. "Reexamining the "Minimal Effects" Model in Recent Presidential Campaigns." *Journal of Politics* 55(1): 1-21.

Fiorina, Morris P. 1981. *Retrospective Voting in American Elections*. New Haven: Yale University Press.

Fournier, Patrick, Richard Nadeau, Andre Blais, Elisabeth Gidengil, and Neil Nevitte. 2004. "Time of Voting Decision and Susceptibility to Campaign Effects." *Electoral Studies* 23(4): 661-681.

Gelman, Andrew and Gary King. 1993. "Why are American Presidential Election Polls so Variable When Votes are so Predictable?" *British Journal of Political Science* 23(4): 409-451.

Heath, Anthony. 2010. "Perspectives on Electoral Behavior." In Dalton, Russell J. and Hans-Dieter Klingemann (eds.), *Oxford Handbook of Political Behavior*. Oxford: Oxford University Press.

Hillygus, D. Sunshine and Todd G. Shields. 2008. *The Persuadable Voter: Wedge Issues in Presidential Election Campaigns*. Princeton: Princeton University Press.

Holbrook, Thomas. 1996. *Do Campaigns Matter?* Thousand Oaks, California: Sage.

Jacoby, William. 2010. "The American Voter." In Jan E. Leighley (ed.), *The Oxford Handbook of American Elections and Political Behavior*. Oxford: Oxford University Press.

Jennings, Will and Christopher Wlezien. N.d. "The Timeline of Elections: A Comparative Perspective." *American Journal of Political Science*, forthcoming.

Johnston, Richard, Andre Blais, Henrey E. Brady, and Jean Crete. 1992. *Letting the People Decide: Dynamics of a Canadian Election*. Kingston, Canada: McGill-Queen's Press.

Kayser, Mark. 2005. "Who Surfs, Who Manipulates?" *American Political Science Review* 99(1): 17-27.

Key, V. O., Jr. 1966. *The Responsible Electorate*. New York: Vintage.

Kircheimer, Otto. 1966. "The Transformation of the Western European Party Systems." In Joseph LaPalombara and Myron Weiner (eds.), *Political Partiers and Political Development*. Princeton: Princeton University Press.

Krueger, James S., and Michael Lewis-Beck. 2007. 'Goodness-of-Fit: R-Squared, SEE and 'Best Practice'.' *The Political Methodologist* 15(1): pp.2-4.

Lau, Richard. 1994. "An Analysis of the Accuracy of 'Trial-Heat' Polls during the 1992 Presidential Election." *Public Opinion Quarterly* 58(1): 2-20.

Lewis-Beck, Michael. 1988. *Economics and Elections: The Major Western Democracies*. Ann Arbor: University of Michigan Press.

Lewis-Beck, Michael, and Mary Stegmaier. 2000. "Economic Determinants of Electoral Outcomes." *Annual Review of Political Science* 3(1):183-219.

Lewis-Beck, Michael, and Tom W. Rice. 1992. *Forecasting Elections*. Washington, DC: CQ Press.

Lipset, Seymour Martin, and Stein Rokkan. 1967. *Party Systems and Voter Alignments: Cross-National Perspectives*. New York: Free Press.

Lodge, Milton, Marco Steenbergen, and Shawn Brau. 1995. "The Responsive Voter: Campaign Information and the Dynamics of Candidate Evaluation." *American Political Science Review* 89(2): 309-326.

McDermott, Monika and Kathleen Frankovic. 2003. "Horserace Polling and Survey Method Effects: Analysis of the 2000 Presidential Campaign." *Public Opinion Quarterly* 67(2): 244-264.

Meguid, Bonnie. 2005. "Competition between Unequals: The Role of Mainstream Party Strategy and Niche Party Success." *American Political Science Review* 99(3): 347-360.

Meyer, Thomas M and Markus Wagner. 2013. "Mainstream or Niche? Vote-Seeking Incentives and the Programmatic Strategies of Political Parties." *Comparative Political Studies* 46:1246-1272.

Moore, David and Lydia Saad. 1997. "The Generic Ballot in Midterm Congressional Elections: Its Accuracy and Relationship to House Seats." *Public Opinion Quarterly* 61(4): 603-614.

Powell, G. Bingham, and Guy D. Whitten. 1993. "A Cross-National Analysis of Economic Voting: Taking Account of the Political Context." *American Journal of Political Science* 37(2): 391-414.

Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: J. Wiley & Sons.

Shaw, Daron R. 1999. "A Study of Presidential Campaign Event Effects from 1952 to 1992." *Journal of Politics* 61(2): 387-422.

Spangenberg, Frits. 2003. *The Freedom to Publish Opinion Poll Results: Report on a Worldwide Update*. Amsterdam: Foundation for Information.

Stevenson, Randolph and Lynn Vavreck. 2000. "Does Campaign Length Matter? Testing for Cross-National Effects." *British Journal of Political Science* 30(2): 217-235.

Vavreck, Lynn. 2009. *The Message Matters*. Princeton: Princeton University Press.

Volkens, Andrea, Pola Lehmann, Nicolas Merz, Sven Regel, Annika Werner, Lacewell Onawa Promise and Henrike Schultze. 2014. *The Manifesto Data Collection. Manifesto Project (MRG / CMP / MARPOR). Version 2014b*. Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB).

Wagner, Markus. 2012. "Defining and measuring niche parties." *Party Politics* 18(6): 845-864.

Wlezien, Christopher. N.d. "The Myopic Voter?" *Electoral Studies*, forthcoming.

Wlezien, Christopher and Robert S. Erikson. 2001. "Campaign Effects in Theory and Practice." *American Politics Research* 29(5): 419-437.

Wlezien, Christopher and Robert S. Erikson. 2002. "The Timeline of Presidential Election Campaigns." *Journal of Politics* 64(4): 969-993.

Wlezien, Christopher, Will Jennings, Stephen Fisher, Robert Ford, and Mark Pickup. 2013. "Polls and the Vote in Britain." *Political Studies* 61(S1): 66-91.

Zaller, John. 2002. "The Statistical Power of Election Studies to Detect Media Exposure Effects in Political Campaigns." *Electoral Studies* 21(2): 297-329.

Zaller, John. 1998. "Politicians as Prize Fighters: Electoral Selection and Incumbency Advantage." In John Geer (ed.), *Party Politics and Politicians*. Baltimore: Johns Hopkins University Press.

SUPPLEMENTARY MATERIALS

APPENDIX S1. POLL DATA

In this supplementary appendix we provide further details of the poll data collected for this study.

Non-Responses and Don't Knows

The norm in the polling industry is to adjust vote intention polls to exclude don't knows and non-responses. However in a small number of cases non-responses are included in the headline figures and we recalculate the poll numbers to ensure that the data are standardized. These sorts of adjustment are the exception but have been implemented consistently.

Dating the Polls

Since most polls are conducted over multiple days, where possible we "date" each poll by the middle day of the period that the survey is in the field. For surveys in the field for an even number of days, the fractional midpoint is rounded down to the earlier day. Information on fieldwork dates is not available for all polls and in those cases we follow careful procedures to calibrate the date assigned to each poll. The rules for poll dating are as follows, using the first possible option before moving onto the next when that possibility had been exhausted: (1) if both fieldwork dates available, the mid-point of the start and end dates is calculated, (2) if only one of the fieldwork dates is available, that date is used, (3) if only the date of publication of the poll in the media is available, that date is used, (4) if only information on the month or week of the poll is available, the mid-point of the corresponding month or week is used, (5) if only information on the month of the poll is available and is observed during the month of the election and is known to be prior to the election, the first of the month is used as the start date and the final day before

the election day is used as the end date (and if the poll asks about voting on "… next Monday [or other day]", then the start date is instead taken as seven days before the election).

Sources

Polls were compiled from a large number of sources, with additional cross-checks and triangulation conducted in the case of inconsistencies or missing data. Wherever possible, polls obtained from secondary poll aggregators were cross-checked and triangulated against other available sources, including the original cross-tabs or media reports. Some of our largest country datasets were collected from archival survey repositories. These included the Roper Center for Public Opinion Research's *iPoll* databank, the Norwegian Social Science Data Archive, the Australian Social Science Data Archive, the Netherlands' Data Archiving and Networked Services, Canadian Opinion Research Archive, and the GESIS/Leibniz Institute for the Social Sciences). A number of datasets were kindly shared with us by other scholars or pollsters. The sources of poll data for our largest poll collections are listed below.

- United States: presidential trial-heat polls are from Erikson and Wlezien (2012). Congressional poll data consist of 1,997 polls from Bafumi et al (2010), further supplemented with data from the Roper Center for Public Opinion Research's *iPoll* databank.

- United Kingdom: dataset of national surveys where respondents were asked about which party they would vote "if the election were held tomorrow" from Wlezien et al (2013), including data from Michael Thrasher, Mark Pack, Ipsos-MORI, YouGov, ICM Research Ltd, Gallup Political and Economic Index.

- Portugal: poll data kindly provided by Francisco José Veiga (see Veiga and Veiga 2004).

- Australia: historical data from the Australian Social Science Data Archive; additional data from Newspoll (www.newspoll.com.au) and Roy Morgan Research (http://www.roymorgan.com/).

- Ireland: poll data via Michael Marsh's Irish Opinion Poll Archive. (http://www.tcd.ie/Political_Science/IOPA/)

- Germany: Forschungsgruppe Wahlen "Politbarometer" data from GESIS/Leibniz Institute for the Social Sciences; additional poll data from the *Wahlrecht.de* website (http://www.wahlrecht.de/). Historical poll data from the Institut für Demoskopie, Allensbach were obtained from replication data for Christopher Anderson's (1995) *Blaming the Government*, via the Harvard Dataverse.

- Netherlands: the dataset "NIPO weeksurveys 1962-2000: NIWI/Steinmetz Archive study number P1654" from Data Archiving and Networked Services (DANS).

- Sweden: all companies' poll data from Johanna Laurin Gulled, Ipsos Public Affairs.

- Italy: all companies' poll data from Chris Hanretty and Graziella Castro.

- Norway: the following datasets from Norwegian Social Science Data Services -- "Respons Analyse AS, 2005-2012" (MMA0067), "ACNielsen, 1987-1994" (MMA0455), "Opinion, 1988-2003" (MMA0585), "Synovate (MMI), 1987-1998" (MMA0802), "TNS Gallup AS, 1964-2010" (MMA0952), "Opinion, 2007-2010" (MMA1119). More recent

poll data was obtained from the TV2 Partibarometeret poll aggregator (http://politisk.tv2.no/spesial/partibarometeret/).

- Canada: poll data from monthly Gallup reports (1942-2000); data via the Canadian Opinion Research Archive.

- France: historical poll data from the publication Gallup Organization (1976) *The Gallup International Public Opinion Polls: France, 1939, 1944-1975*; contemporary poll data from *TNS-Sofres* (http://www.tns-sofres.com) and from other sources.

- Spain: data from El Centro de Investigaciones Sociológicas (CIS) (http://www.cis.es/) and other sources.

REFERENCES

Anderson, Christopher. (1995). *Blaming the Government*. New York: M.E. Sharpe

Bafumi, Joseph, Robert S. Erikson, and Christopher Wlezien. 2010. "Balancing, Generic Polls and Midterm Congressional Elections." Journal of Politics 72:705-719.

Erikson, Robert S. and Christopher Wlezien. 2012. *The Timeline of Presidential Elections: How Campaigns do (and do not) Matter*. Chicago: University of Chicago Press.

Gallup Organization. (1976). *The Gallup International Public Opinion Polls: France, 1939, 1944-1975*. New York: Gallup Organization.

Moore, David and Lydia Saad. 1997. "The Generic Ballot in Midterm Congressional Elections: Its Accuracy and Relationship to House Seats." *Public Opinion Quarterly* 61(4): 603-614.

Spangenberg, Frits. 2003. *The Freedom to Publish Opinion Poll Results: Report on a Worldwide Update*. Amsterdam: Foundation for Information.

Veiga, Francisco José and Linda Gonçalves Veiga. 2004. "The Determinants of Vote Intentions in Portugal." *Public Choice* 118(3/4): 341-364.

Wlezien, Christopher and Robert S. Erikson. 2001. "Campaign Effects in Theory and Practice." *American Politics Research* 29(5): 419-437.

Wlezien, Christopher, Will Jennings, Stephen Fisher, Robert Ford, and Mark Pickup. 2013. "Polls and the Vote in Britain." *Political Studies* 61(S1): 66-91.

APPENDIX S2. ELECTION DATA

We rely on a wide range of official sources and election data resources. Official sources were preferred where possible. Where official sources were not readily available, resources such as the Election Guide database of the International Foundation for Electoral Systems (www.electionguide.org) were used as an alternative or were used to cross-check the reliability of data obtained from unofficial sources (such as the websites of opinion pollsters and academic or amateur poll spotters). Some of the older data is from Nohlen and Stöver (2010).

General Resources

The European Election Database of the Norwegian Social Science Data Services (NSD)

http://www.nsd.uib.no/european_election_database/

ElectionGuide, International Foundation for Electoral Systems

http://www.electionguide.org/

Political Database of the Americas: Electoral Systems and Data

http://pdba.georgetown.edu/elecdata/arg/arg.html

Election Resources

http://electionresources.org/

Nohlen, Dieter, and Philip Stöver. 2010. *Elections in Europe: A data handbook*. Baden-Baden: Nomos Verlagsgesellschaft.

Country-Specific Resources

Australian Politics and Elections Database at the University of Western Australia,

http://elections.uwa.edu.au/

Bundesministerium für Inneres, Austria,

http://www.bmi.gv.at/cms/BMI_wahlen/nationalrat/start.aspx

Federal Elections in Brazil, Brazil

http://electionresources.org/br/index_en.html

Bularian Parliament, Bulgaria

http://www.parliament.bg/bg/electionassembly

Elections Canada, Canada

http://www.elections.ca/home.aspx

Parliament of Canada, Canada

http://www.parl.gc.ca/parlinfo/compilations/electionsandridings/ResultsParty.aspx

Ministerio del Interior, Republica de Chile

http://historico.servel.cl/

Ministry of Interior, Cyprus

http://www.ekloges.gov.cy/

Consejo Nacional Electoral (National Electoral Council), Republic of Ecuador

http://resultados.cne.gob.ec/

Ministry of Justice, Finland

http://www.vaalit.fi/

Ministry of Interior, France

http://www.interieur.gouv.fr/Elections/Les-resultats

Der Bundeswahlleiter (the Federal Returning Officer), Germany

http://www.bundeswahlleiter.de/en/index.html

Ministry of the Interior, Greece

http://ekloges.ypes.gr/

Statistics Iceland

http://www.statice.is/Statistics/Elections/

Ministry of the Interior, Italy

http://elezioni.interno.it/

Ministry of Internal Affairs and Communications, Japan

http://www.soumu.go.jp/senkyo/senkyo_s/data/shugiin44/index.html

Government of Malta

http://www.gov.mt/en/Government/Government%20of%20Malta/Election%20Results/Pages/Elections-DOI-site.aspx

Instituto Federal Electoral, Mexico

http://www.ife.org.mx/portal/site/ifev2

Statistics Norway, Norway

http://www.ssb.no/a/english/kortnavn/stortingsvalg_en/tab-2009-10-15-02-en.html

Electoral Commission, New Zealand

http://www.electionresults.govt.nz/

National Office of Electoral Processes, Peru

http://www.onpe.gob.pe/inicio.php

Commission on Elections, Republic of the Philippines

http://www.comelec.gov.ph/

Comissão Nacional de Eleições, Portugal

http://eleicoes.cne.pt/sel_eleicoes.cfm?m=raster

Ministry of the Interior, Spain

http://www.infoelectoral.mir.es/min/

Election Authority, Sweden

http://www.val.se/in_english/previous_elections/index.html

Federal Office of Statistics, Switzerland

http://www.bfs.admin.ch/

National Electoral Council, Venezuela

http://www.cne.gob.ve/web/estadisticas/index_resultados_elecciones.php

Rallings, C. and Thrasher, M. 2007. *British Electoral Facts, 1832–2006, seventh edition*. Aldershot: Ashgate.

Rallings, C. and Thrasher, M. 2010. *Election 2010: The Official Results*. London: Biteback publishing.