2023

Online appendix for “Studying the interplay of party support and turnout; Chapter 5 in The Problem of Governing: essays for Richard Rose edited by Michael Keating, Ian McAllister, Edward Page and Guy Peters. London: Palgrave Macmillan

Mark N. Franklin

Rose festschrift chapter

CONTENTS of online appendices

ONLINE APPENDICES

A. Data

Respondent's left-right position, party's left-right position and which party respondents voted for were all coded at the individual respondent level in the Integrated Module Dataset (IMD) of the Comparative Study of Electoral Systems (Quinlen et al. 2018).

Questions giving rise to these measures were:

1. ["Which party did you vote for?"] (Question varies from country to country).

2. ["In politics people sometimes talk of left and right. Where would you place yourself on a scale from 0 to 10 where 0 means the left and 10 means the right?"] (Modules 1 and 2) "Where would you place yourself on this scale?" (Modules 3 and 4). Data recoded 0-1.

3. "Now, using the same scale, where would you place [Party A-F]?" (all modules).

To get a measure of left-right proximity I reshaped the individual-level data to the response level, where each response pertains to a separate party that respondents could vote for (or not) and for which they provide an estimate its left-right location. In political science we normally call this a stacked dataset, following Eijk et al. (2006). These stacked data were then collapsed (with values averaged across categories of variables defining the target level), either to the party level or to the party-birthyear level.[1]

Table A.1 displays univariate statistics, where they make sense, for variables employed in the main text at each level of aggregation used there. Note that generic party variables (variables having to do with parties in general ather than with specific named parties) and

[1] Note that the party level of analysis subsumes the country level since each country has parties with unique codes – codes that are not repeated for parties of any other country. For respodents whose response was missing for any particular party, party left-right location at that level was plugged with the mean location across the non-missing responses of other respondents.

Table A.1 Univariate statistics for variables employed in this paper's anayses

	R's left-right location	Party 1 's left-right location	R's vote for Party*	Generic party's left-right location	R's prox-imity to generic party	R's vote for generic party
Respondent level						
N of cases	120,015	116,409	114,207			
Minimum	0	0	360001 *			
Maximum	10	10	8400004 *			
Mean	5.41	5.79				
Std deviation	2.44	2.99				
Response level (stacked)						
N of cases	779,790	667,791		667,791	725,229	926,895
Minimum	0	0		0	0	0
Maximum	10	10		10	10	1
Mean	5.40	5.03		5.03	7.13	0.12
Std deviation	2.46	2.95		2.95	2.48	0.32
Birthyear-party level						
N of cases	34,228	31,662 +		31,950	33,237	33,237
Minimum	0	0		0	0	0
Maximum	10	10		10	1	0.53
Mean	5.45	5.10		7.07	0.	0.06
Std deviation	1.17	2.06		1.18	0.08	0.08
Party level						
N of cases	526	++		493	493	503
Minimum	4.09			0.60	0	0
Maximum	6.98			9.02	1	0.53
Mean	5.44			5.02	0.65	0.11
Std deviation	0.58			1.81	0.16	0.12

Notes:* Party ID code. ** Measured at country-year level and duplicated onto birthyear/party levels.
+ N = 12.596 with appropriate lags. ++ N = 130 with appropriate lags.

variables created from those generic party variables, only exist in stacked (response-level) data and levels of analysis (party or birthyear cohort) derived from stacked data.

Table A.2 lists all the elections conducted between 1996 and 2016 in each country that contributed at least 3 surveys to the CSES. Timepoints producing data for this paper are boldfaced. Australia, Israel and Japan each contributed four election studies to the IMD, but those were separated by additional elections rendering them non-contiguous, so no election studies from these countries are boldfaced. The final column counts the number of included studies.

Table A.2 Elections included in the span of time covered by the CSES IMD data, with boldfacing for adjacent elections yielding data included in this paper's analyses

Sequence in analysis	1			2		3		4	5	Total included
ustralia	1996	1998	2001	2004		2007	2010	2013		0
anada	1997		2000	2004	2006	**2008**		**2011**	**2015**	3
zech Republic	1996	1998		**2002**		**2006**		**2010**	**2013**	4
ermany	**1998**			**2002**		**2005**		**2009**	**2013**	5
eland	**1999**			**2003**		**2007**		**2009**	**2013**	5
elend	1997			**2002**		**2007**		**2011**		3
rael	1996		1999	2003		2006	2009	2013		0
pan	1996		2000	2004	2005	2007	2009	2013		0
epublic of Korea	**2000**			**2004**		**2008**		**2012**		4
exico	**1997**			**2000**		**2006**	2009	2012		3
ew Zealand	1996			2002	2005	**2008**		**2011**	**2014**	3
orway	**1997**			**2001**		**2005**		**2009**	**2013**	5
eru	**2000**			**2001**		**2006**		**2011**	**2016**	5
oland	**1997**			**2001**		**2005**		**2007**	**2011**	5
omania	1996		2000	**2004**		**2009**		**2014**		3
olvenia	1996		2000	**2004**		**2008**		**2011**		3
veden	**1998**			**2002**		**2006**	2010	2014		3
vitzerland	**1999**			**2003**		**2007**		**2011**		4
aiwan	1996	1998	2001	**2004**		**2008**		**2012**		3
nited States	1996		2000	**2004**		**2008**		**2012**		3

Bibliography

Quinlen, Steven, Christian Schimpf, Katharina Blinzler, Slaven Zivkovic, Bojan Todosijevic (2018) The Comparative Study of Electoral Systems (CSES) Integrated Module Dataset (IMD). GESIS Leibniz Institute for the Social Sciences, Colone Germany.

B. Robustness checks

B1 Level of analysis

A major concern for some scholars perusing this chapter might be the fact that analyses are conducted using data aggregated to a higher level of analysis (two different higher levels, in fact) than the level of aggregation at which the data were collected.[2] As explained in the main text, conceptually these are the levels of analysis relevant to the theorizing set out in the chapter; and few scholars will balk at seeing analyses relevant to understanding the behavior of political parties that are conducted at the party level of analysis, even if the data were originally collected at some other level of aggregation. Despite the fact that I can write the exact same phrase regarding birthyear cohorts, I know from bitter experience with journal reviewers that many of them do balk at passing on analyses conducted at the birthyear-cohort level. Generally they give no reason for doing so, treating the problem as self-evient; but some do mention the possibility that composition effects might threaten the findings. Indeed this is true, and also in regard to party-level analyses with aggregated data; but not because the aggregate-level findings are spurious. Rather it would be because the individual-level data is adding non-random noise that would need to be totally controlled for if the aggregate-level findings are to be replicated with individual-level data. And totally controlling for individual-level effects is hard to do. It requires that EVERY individual-level variable correlated with the outcome of interest be known and included in the analysis, a virtually impossible task. The truth is that, for aggregate-level effects truly governed by aggregate-level processes, using data that has been aggregated to the appropriate level is the safest approach, automatically removing whatever spurious effects would have threatened individual level findings. Said in another way, analyses at the theoretically-defined level removes the need to control for spurious effects at the original level at which the data were collected.

[2] That level is the response level of aggregation, not the respondent level, since the questions at the center of my analysis were asked about each party separately. Answers originally occupied multiple variables for each respondent. When reshaped (stacked) each party-regarding variable became a separate case in the response-level dataset (Google search for "De Sio stackMe").

This can be demonstrated by comparing findings from aggregate-level data with findings from individual-level data into which the aggregate-level indicators have been injected by merging the two datasets. Following such a merge, all members of a given birthyear cohort in the individual-level data will register the same value for each time-series indicator but will also be associated with the full spectrum of values for all avaliable individual-level variables that might be responsible for variance in those turnout values.

Table B1a uses merged response-level data to compare effects of party-level inputs (constant across individual members of each birthyear cohort) with the same effects when controlling for individual-level covariates; while matching both of these with the effects found with party-level taken alone (models from Table 1 in the main text). Table B1b does the same thing for birthyear-

Table B1a Individual-level (fixed effects) Error Correction Models of party support (IMD data; Greek letter Δ labels each differenced variable: $X_t - X_{t-1}$) where over time effects, derived from party-level data, have been merged into the response-level data

Origin for timevars:	Party-level data			Birthyear cohort-level data		
	Model A From main text	Model A1 Individual lvl w'out controls	Model A2 Individual lvl with controls	Model B From main text	Model B1 Individual lvl w'out controls	Model B2 Individual lvl with controls
Concept:	Representatn	Representatn	Representatn	Feedback	Feedback	Feedback
Outcome:	**Δ.**Support	**Δ.**Support	**Δ.**Support	**Δ.**Proximity	**Δ.**Proximity	**Δ.**Proximity
Inputs:	Coef. (s.e.)	Coef. (s.e.)	Coef. (s.e.)	Coef. (s.e.)	Coef..(s.e.)	Coef. (s.e.)
1) Lagged outcome (ECP)	-1.07 (0.06)	-1.06 (0.00)	-1.24 (0.00)	-1.55 (0.14)	-1.59 (0.00)	-1.52 (0.00)
2) **Δ**.left-right proximity $_t$	0.23 (0.08)	0.27 (0.00)	0.28 (0.00)			
3) Left-right proximity $_{t-1}$	0.33 (0.12)	0.37 (0.00)	0.37 (0.00)			
6) **Δ**.left-right proximity $_{t-1}$				-0.19 (0.14)ns	-0.16 (0.00)	-0.35 (0.00)
7) Left-right proximity $_{t-2}$				-0.49 (0.21)	-0.43 (0.00)	-0.70 (0.00)
8) Individual level covariates †	NO	NO	YES	NO	NO	YES
Intercept	-0.07 (0.07)	0.11 (0.00)	0.06 (0.00)	1.05 (009)	1.07 (0.00)	1.02 (0.00)
R-squared	0.64	0.62	0.68	0.65	0.64	0.69
Observations	354	665,360	196,523	188	372,046	107,355
Number of cntry-birthyrs	168	165	130	115	115	161

Notes: All coefficients significant at $p<0.001$, one-tailed unless marked n.s. (not significant).

† Individual-level covariates are age, age^2, gender, married, religion, knowledge, efficacy, partisan, and satisfaction with democracy.

level data. The most important thing to take away from these comparisons is that, to the extent that individual-level findings differ from corresponding aggregate-level findings, (while failing to control for individual-level covariates of the relevant dependent variable) the corresponding individual-level effects that *do* control for individual-level covariates generally differ in a direction that reduces the uncontrolled divergence.

This tendency is considerably more evident in Table B1b, which focuses on birthyear cohort processes. Indeed, this table includes an instance where uncontrolled individual-level effects (in Model D1) are actually reversed from those at the birthyear level (Model D) only to have their signs corrected when individual level covariates are introduced in Model D2.

Table B1b Individual-level (fixed effects) Error Correction Models of turnout (IMD data; Greek letter Δ labels each differenced variable: X_t - X_{t-1}), where over time effects, derived from birthyear cohort data, have been merged into the response-level data

Origin for timevars:	Birthyear cohort data			Birthyear cohort data		
	Model C	Model C1 w'out controls	Model C2 with controls	Model D	Model D1 w'out controls	Model D2 with controls
Concept:	Competition	Competition	Competition	Feedback	Feedback	Feedback
Outcome:	**Δ.**Turnout	**Δ.**Turnout	**Δ.**Turnout	**Δ.**Turnout.	**Δ.**Turnout	**Δ.**Turnout
Inputs:	Coef. (s.e.)	Coef. (s.e.)	Coef. (s.e.)	Coef. (s.e.)	Coef. (s.e.)	Coef. (s.e.)
1) Lagged outcome (ECP)	-1.34 (0.02)	-0.84 (0.00)	-0.98 (0.00)	-1.34 (0.02)	-0.83 (0.00)	-0.99 (0.00)
2) **Δ.**left-right proximity $_t$	-0.26 (0.04)	-0.41 (0.00)	-0.35 (0.00)			
3) Left-right proximity $_{t-1}$	-0.40 (0.07)	-0.42 (0.00)	-0.35 (0.00)			
6) **Δ.**left-right proximity $_{t-1}$				0.03 0.09)ns	-0.03 (0.00)	0.14 (0.01)
7) Left-right proximity $_{t-2}$				0.27 (0.11)	-0.07 (0.00)	0.07 (0.02)
8) Individual level covariates†	NO	NO	YES	NO	NO	YES
Intercept	1.29 (0.05)	0.94 (0.00)	0.75 (0.00)	0.87 (0.11)	0.69 (0.00)	0.44 (0.01)
R-squared	0.66	0.51	0.54	0.63	0.44	0.60
Observations	4,544	834.888	262,640	2,659	519,383	143,577
Number of cntry-birthyrs	165	268	217	101	216	161

Notes: All coefficients significant at p<0.01, one-tailed, unless marked ns (not significant).
† Individual-level covariates are age, age^2, gender, education, married, religion, income, urban, union, knowledge, efficacy, partisanship, and satisfaction with democracy.

These contrasts support the reasoning set out earlier. Individual-level data will only yield correct effects for causal processes that are aggregate in nature to the extent that individual-level covariates are fully controlled. We only get modest confiration for this reasoning from the analyses performed for this chapter because of the rather small number of individual-level correlates of turnout available in the IMD data. However, I should stress that *all* four sets of models end with an individual-level model that has significant effects with signs that match those hypothesized for the first, aggregate-level, model. See Franklin (2022) for additional findings relevant to this concern.

B2 Do turnout findings at birthyear-cohort level reflect actual turnout evolution?

Meanwhile we need to address a question glossed over in the main text for lack of space: whether birthyear cohort findings regarding the knock-on effects of feedback for party support actually correspond to meaningful effects on turnout when all birthyear cohorts present for a particular election are taken together as a single case. This question is easily addressed by predicting differenced turnout at the birthyear cohort level for the data employed for Table B1b and using the predicted values derived from Models C2 and D2; then aggregating the data (including the newly predicted values) to the country-election level and estimating actual differenced turnout at that level from the two sets of predicted values. Results are in Table B1c.

Table B1c. Country-election level turnout explained by predicted turnout from Models C2,D2

Outcome:	Differenced turnout
Inputs:	Coef. (s.e.)
1) Δ.turnout predicted by Model C2 $_{(t-1)}$	0.25 (0.95)ns
2) Δ.turnout predicted by Model D2 $_{(t-2)}$	1.23 (0.99)
Intercept	1.29 (0.05)
R-squared	0.93
Observations	31
Number of country-elections	20

Note: Coefficients significant at p<0.20, one-tailed, unless marked ns (not significant).

As can be seen, the two measures of differenced turnout (estimated at the birthyear cohort level and aggregated to the country level) explain over 90 percent of the variance in differenced turnout at that level of analysis for the rather few cases that remain after aggregation.[3] The effective coefficient in this analysis is only signifiacant at the 0.2 level, given the small number of available cases, but this is enough to support the supposition made in the main text, that synchronization would be achieved because competition is contageous. Evidently this topic needs further attention.

B3 Endogeneity problems when estimating persuasion and learning effects

As pointed out in the main text, attempting to estimate the relative contribution of each component of a proximity measure, as I do in Table 2 in the main text, evidently yields grave risks of findings contaminated by endogeneity. And, as mentioned there, the means chosen to create the proximity measures used in that table are virtually the only ones available that do not show endogeneity artifacts. Here we elaborate on that assertion.

Left-right proximity measures are constructed by taking the absolute value of the difference between measures of party location respondent location. Self-assessed respondent locations present some problems discussed in the main text. Here we address problems found in measures of party location. The measure originally employed by Franklin and Lutz (2020), the protopype for the research reported here, is straightforward. Principle Investigators for each survey were asked to code their country's parties appropriately and, from this measure along with respondent self-assessed location, measures of proximity were constructed. However, while widely used in research of this kind, such a measure is not the one we really want when persuasion effects are expected (which was not the case with the Franklin-Lutz research). This is because resulting proximity measures cannot include persuasion effects (respondents play no part in producing the party component of resulting

[3] Recall that Model D2 results in a prediction at only a single election for the large number of estimates calling for three timepoints and at most two elections otherwise.

proximities). But persuasion effects play a critical part in our theorizing, as explained in the main text.

The most common alternative to expert-judged party locations is to employ respondent-judged party locations. These are also the more appropriate measures if possible persuasion effects are to be discovered. But when respondent judgements are used in place of expert judgements, the question arises what to do about missing judgements? The most widely employed solution is to "plug" the missing party location judgements with the average value assigned by respondents who answered the party location question. But if some respondents are positioning a party on the basis of projection then the plugging value will reflect the most widespread projection effect – probably a bias towards the largest party. So an alternative strategy when finding an average plugging value is to ignore judgements that are the same as the respondent's own self-evaluated left-right position. These will be referred to in what follows as "diff-plugged" party locations, to differentiate them from "all-plugged" locations where the means are based on judgements derived from all non-missing responses. This reduces the number of respondents responsible for the party placements but increases the reliability of the responses obtained.

Table B2 compares results of analyses such as those presented in Table 2 of the main text when expert-assessed party locations are compared with expert-assessed and mean-plugged respondent-assessed models. Model B1 demonstrates what happens when party supporters have no opportunity to demonstrate projection effects because they play no part in positioning the parties (rows 2&3). There we see that party supporters are entirely responsible for the extent of left-right congruence.

The complete lack of any role for parties comes as a surprise, until we remember that the experts assessing party locations are election study Principal Investigators, who are largely the same individuals from election to election and may not very quickly revise their opinions about where parties stand.

Table B2 Change in left-right proximity due to change in respondent vs party left-right location (birthyear cohort analyses with expert vs respondent party placements)

Outcome: Inputs:	Model B Expert-assessed **Δ.**Proximity Coef. (s.e.)	Model B2 Mean-plugged Resp-assessed **Δ.**Proximity Coef. (s.e.)	Model B3 Diff-plugged Resp-assessed **Δ.**Proximity Coef. (s.e.)
1) Lagged outcome (ECP)	-1.27 (0.02)	-1.58 (0.02)	-1.32 (0.02)
2) **Δ.P**arty l-r location $_t$	-0.01 (0.02)ns	-0.26 (0.04)	0.23 (0.03)
3) Party left-right location $_{t-1}$	0.01 (0.02)ns	-0.25 (0.06)	0.33 (0.04)
4) **Δ.**Supporter l-r location $_t$	0.06 (0.02)	0.14 (0.02)	-0.07 (0.01)
5) Supporter l-r location $_{t-1}$	0.13 (0.03)	0.33 (0.03)	-0.05 (0.02)*
6) Constant	0.68 (0.02)	0.30 (0.04)	0.77 (0.03)
R-squared	0.66	0.72	0.65
Observations	4,357	4,469	4,544
Country-birthyr cohorts	1,961	1,962	1,971

Notes: All coefficients significant at the p<0.01 level, one-tailed, except as marked "*" for significance at the 0.05 level. Note that expert assessments pertain only to parties.

Endogeneity may also play some role. Model B2 shows such effects more obviously. It explains most variance of the three models and shows a larger long-term contribution from respondent shifts than from party shifts in left-right location, contrary to expectations.[4] That coefficient (of 0.33) includes endogeneity by construction, since effects take account of party positions that are assigned by respondents who place the parties where they place themselves.

What about effects shown in Model B3? What role do persuasion/learning and the other possible contaminants (projection, assimilation) play in those findings? Model B3 replicates the analysis presented in Model B of Table 2 in the main text, which rules out projection by eliminating positions assigned by respondents who place the parties where they place themselves. This "lobotomization" of voter assessments certainly removes any possible projection effects, but it will also have removed

[4] That model also switches the signs of effects to accord with the dominant influence found (see footnote 11 in the main text). Note that none of these models were evaluated by Franklin and Lutz (2020) who focused uniquely on party-level measures where the deficiencies of expert-ratings were not apparent.

assessments from voters who placed the party where they place themselves for reasons other than projection. So it conducts a very stringent test which is still passed, if only at the p=0.05 level of statistical significance. The true effects of party supporters on measured proximity must lie somewhere between the effects shown in Models B1 and B3.

Turning to persuasion/learning effects, we use as our laboratory the analyses reported in Table 1 of the main text. We start (model C1) by repeating that model, using the (difference-plugged respondent-assessed) measures employed in the main text. We then compare those, in Model C2, with all-plugged versions of the same measures and, in Model C3, with the same (expert-assessed) measures that were used in Table B2. Model C3 makes it clear why expert assessed measures were not used in Table 1: its (expected) long-term negative effect in that model is not statistically significant. Clearly the findings shown in the main text (and repeated in Model C2 below) are due in some measure to effects that were not present in expert assessments. Likely reasons why expert-assessed party positions would fail to track (or to be tracked by) respondent positions do not include any of the contaminants mentioned so far (projection, assimilation or learning), all of which would increase observed effects beyond those that would be seen with no contamination.

Table B3 Replications of Table 1, Model C, effects on turnout using various measures of respondent-party left-right proximity

Outcome: Inputs:	Model C1 Diff-plugged resp-assessed Coef. (s.e.)	Model C2 Mean-plugged resp-assessed Coef. (s.e.)	Model C3 Expert-assessed First lag Coef. (s.e.)	Model C4 Expert assessed Second lag Coef. (s.e.)
1) Lagged outcome (ECP)	-1.33 (0.02)	-1.34 (0.02)	-1.32 (0.02)	-1.39 (0.06)
2) **Δ.P**arty l-r position $_{t}$	-0.26 (0.04)	-0.25 (0.04)	-0.06 (0.04)	
3) Party left-right location $_{t-1}$	-0.40 (0.07)	-0.36 (0.07)	-0.06 (0.06)ns	
4) **Δ.**Party l-r location $_{t-1}$				0.03 (0.09)ns
5) Party l-r location $_{t-2}$				0.24 (0.10)
6) Constant	1.29 (0.05)	1.25 (0.05)	1.04 (0.04)	0.87 (0.07)
R-squared	0.66	0.66	0.65	0.63
Observations	4,544	4,544	4,357	2,659
Country-birthyr cohorts	1,971	1,971	1,961	1,729

Notes: All coefficients significant at the p<0.01 level, one-tailed, except as marked "ns".

These effects (if we accept the evidence of Models C1 and C2 that the effects are real)[5] are most likely being suppressed in Model C3 by experts whose judged party locations do not change over time even though the parties locations do in fact change. Such contamination is possible if the same experts are used repeatedly to judge party positions over time (evidently this could not happen with repeated samples of survey respondents but is quite feasible if the same PIs are engaged in consecutive election studies, as already mentioned). Model C4 supports this supposition by showing stronger effects when an additional lag provides more time for PI's judgements to evolve (one would normally expect weaker effects over a longer period).[6]

Considerations developed in this section of Appendix B support my choice of proximity measures to use in different models. However, these considerations might be considered specious in the absence of confirmatory findings from survey experiments that we consider next.

B4 Validating voter awareness of policy shifts

Adams et al. (2018) stress the need to validate evidence for voter awareness of policy shifts by verifying that the evidence remains compelling when cognate variables are employed. Those authors might have added that it would also be important to confirm the findings using alternative data sources. In this part of Appendix B I do both, validating my evidence of voter awareness (as shown by shifts in perceived left-right proximity) by employing an alternative measure of awareness – the reports those voters make of liking a party – and, additionally, referencing findings from an alternative data source.

Table B.4 focuses on responsiveness, with party support as the outcome (successive pairs of rows alternate proximity with liking). Table B5 focuses on representation, with either proximity

[5] Any assimilation effects would cause coefficients in Model C2 to differ from those in Model C1, since expected bias from respondents whose left-right locations echoed the locations of parties they placed should have been eliminated in diff-plugged values.
[6] The changes in sign between effects in Models C3 and C4 are expected theoretically, as explained above.

Table B.4 Robusness of responsiveness across measures of voter awareness and levels of analysis

Outcome: Differenced party support	Party-birthyear level		Party level	
	Model A	Model B	Model C	Model D
Inputs:	Coef (s.e.)	Coef (s.e.)	Coef (s.e.)	Coef (s.e.)
Party support $_{t-1}$	-1.21 (0.01)	-1.22 (0.01)	-0.98 (0.09)	-1.04 (0.08)
Differenced proximity	0.11 (0.00)		0.55 (0.09)	
Proximity $_{t-1}$	0.14 (0.01)		0.60 (0.13)	
Differenced liking		0.24 (0.01)		0.33 (0.04)
Liking $_{t-1}$		0.31 (0.01)		0.38 (0.06)
Constant	-0.06 (0.01)	-0.05 (0.00)	-0.28 (0.08)	-0.10 (0.04)
R-squared	0.64	0.70	0.67	0.73
Observations	13,964	14,911	232	246
Number of groups	6,866	6,987	109	110

Note: All coefficients significant at $p<0.001$.

or liking as the outome. As is clearly evident, in both tables the pattern of findings using party likes/dislikes are very much the same as those using left-right proximity, whether at the birthyear or party levels. Indeed, at the party level, confidence intervals overlap.

Table B.5 Robustness of representation across measures of voter awareness and levels of analysis

Outcome: Proximity/ Liking	Party-birthyear level		Party level	
	Model E	Model F	Model G	Model H
Inputs:	Coef (s.e.)	Coef (s.e.)	Coef (s.e.)	Coef (s.e.)
Proximity $_{t-1}$	-1.44 (0.02)		-1.17 (0.15)	
Liking $_{t-1}$		-1.29 (0.02)		-0.89 (0.12)
Differenced party support $_{t-1}$	-0.08 (0.02)	-0.31 (0.03)	-0.22 (0.13)*	-0.67 (0.22)
Party support $_{t-2}$	-0.26 (0.04)	-0.62 (0.05)	-0.54 (0.18)	-1.01 (0.28)
Constant	1.06 (0.01)	0.63 (0.01)	0.87 (0.11)	0.69 (0.08)
R-squared	0.69	0.65	0.56	0.60
Observations	,13,562	,14,966	130	136
Number of groups	, 6,872	,6, 939	65	65

Notes: Outcome is proximity in Models E and G, liking in Models F and H.
All coefficients significant at $p<0.001$ except where marked * (significant at $p<.01$).

At the birthyear level effects using the likes/dislikes measure are significantly stronger than those using proximity. This could be because left-right proximity focuses on a specific way in which voters might see party policies as being congruent with their own whereas, when voters are asked how much they like a party, this invites them to think of congruence in a more general fashion that would include other considerations than left-right proximity.

So these results suggest that, in adopting a focus on left-right proximity in the main text, my research strategy was conservative. In cognate research, the same strategy yielded substantively identical findings regarding long-term equilibria for party support and left-right proximity based on a completely different data source (European Parliament election studies) and a time-period (1989-2014) that was somewhat longer (Franklin 2015).

C. Error correction models, stationarity and co-integration

The findings of this paper are produced by error correction models (ECMs). For such models to yield valid findings a number of requirements must be met,[7] of which the primary one is also the most difficult to verify: the process under investigation must be in long-term equilibrium (so any short-run disequilibrium will be corrected in due course). This means that, over the long term, all dynamic elements in the model must either be stationary or else co-integrated with the relevant element(s) on the other side of the equals sign (De Boef and Keele 2008). Stationarity means simply a long-run stable mean to which a series reverts after any deviation while co-integration means moving together (thus both non-stationary) in a long-run relationship such that the linear combination of the two series is stationary.

Confirming stationarity or co-integration with regard to my data is not straightforward,

[7] A standard ECM, by construction, meets requirements for balance, a major concern for contributors to a symposium on the topic (Keele, Linn and Webb 2016). There are differences of opinion regarding other requirements, but in this appendix I apply the most stringent of the various possibilities.

however. My time-series for specific country panels are very short: no more than five time-points with an average of 3.5; and I have only 57 parties with data for 4 or more time-points. Given random perturbations in level of support and left-right position of individual parties, a sample of just three or four cases can be expected to show trends that are upward, downward or both, pretty much at random. Examination of my data confirms this expectation at levels of statistical significance appropriate to the small Ns involved. Stationarity is found by an Augmented Dickey-Fuller unit-root test at the 0.05 level for 7 (or 8) out of 69 (or 72) parties (thus random at virtually the 0.1 level of statisticl significance) while a Westerlund test finds co-integration at the 0.05 level in a further 13 out of 57 (thus random at close to the 0.2 level).

I take a two-pronged approach to arguing that my findings are not vitiated by what appear to be short-term anomalies. The first prong is to assert that, from a theoretical standpoint, my data should be in equilibrium and any indication to the contrary is thus spurious – simply capitalizing on chance variations evident in the short-term. This expectation receives face validity for my party-level analyses because the disequilibriating short-term effects in those analyses are not remotely statistically significant. For the birthyear cohort analyses, however, the disequilibriating long-term effects, though small (as expected), are nevertheless statistially significant. The second, more demanding, approach is to take indications of non-cointigration non-stationarity (in the case of any given panel) at face value and demonstrate that when I exclude such panels my findings are substantively unchanged.

The starting point for any assessment of either co-integration or stationarity must be theoretical. I have no basis for supposing that the variables in my models would be co-integrated (which implies non-stationarity) because I do have every reason to suppose that they should be stationary. Political parties come and go, but those that came or went are not part of my sample, which consists only of parties present in the data for at least three time-points. Recent research has established that such parties tend to receive a level of support that is in a long-run stationary equilibrium, (Weber and Franklin 2018). Much the same applies to the left-right locations of such parties (Dalton and McAllister 2015). So I expect stationarity for my primary variables.

Of course such long-term equilibria are quite consistent with the appearance of short-term dis-equilibria in specific panels, simply on the basis of random perturbations (as already mentioned); and my data do fail a test of the "joint requirement" that both of my primary variables (left-right proximity and party support) are co-integrated (using a Westerlund test) over the time-span of any panel that lacks stationarity for one or both of these variables (according to an Augmented Dickey-Fuller unit-root test). However, as already mentioned, the relatively small number of failures to meet the joint requirement is quite consistent with the notion that these failures are just random purturbations, only to be expected when taking short-term "snap shots" of data that, over a longer term, would have proved stationary.

Still, if I proceed with the second prong of my approach by taking at face value the individual failures to meet the joint requirement mentioned above, I can select for analysis only panels that are "clean" in the sense that my primary variables are either stationary or cointegrated. If my results are substantively unchanged when using clean data this will lend support to the idea that, even though the joint requirement cannot be shown to have been met, still my findings are not artifacts of any departure from the clean data requirement.

Table C.1 shows the results.[8] Although significance levels are sometimes low because

[8] At both the party level and party-birthyear-cohort level, the resulting selection of panels is strongly enough balanced to be tested for stationarity of all selected panels, and stationarity cannot be rejecteded at the 0.05

Table C.1 Feedback and representation in data selected for demonstrable stationarity or co-integration at p<0.2

Outcome variable:	Model I Differenced party support (party level)		Model J Differenced left-right proximity (birthyear level)	
Inputs:	Coeff	(s.e.)	Coeff	(s.e.)
1) Party support $_{t-1}$	-1.04	(0.12)**		
2) Differenced Left-right proximity	0.37	(0.17)*		
3) Left-right proximity $_{t-1}$	0.47	$(0.32)^{+}$	-1.54	(0.33)**
5) Differenced party support $_{t-1}$			-0.44	(0.24)*
6) Party support $_{t-2}$			-0.88	(0.28)**
7) Constant	-0.17	(0.22)	1.23	(0.15)**
R-squared	0.72		0.61	
Observations	69		5,448	
Number of parties/birthyear cohorts	23		1,810	

Notes: Fixed effects regression analysis with standard errors in parentheses.
All coefficients significant at 0.05, one-tailed, unless marked "ns" (not significant) or
$^{+}$ (significant at the 0.1 level, one-tailed).

of the small N selected for analysis, Table C1 demonstrates that findings for the primary time-series in this research would have been much the same if focused on panels that were stationary or co-integrated.[9]

Bibliography

Franklin, M. (2015) "Dealignment and Accountability." Paper presented at the ECPR Joint Sessions Workshop on "Accountability without parties" in Warsaw, Poland, April. (Available on researchgate.net by means of a Google Scholar search).

level either for proximity or for party support.

[9] If I compare results from clean data with results from data that includes questionable panels, coefficients from the latter data are generally more highly significant, statistically (coming from larger datasets) but those coefficients are not statistically distinguishable from coefficients deriving from clean data.

D. What is to be done?

This is an unusual appendix; perhaps unique.

The substantive content of the chapter to which it belongs is the result of an accident. While testing for negative feedback in party support I realized that my measure of support at the party level (votes cast for each party as a proportion of electorate size) was indistinguishable, at the party level of analysis, from a measure of turnout. So it occurred to me to wonder what would happen if I substituted an actual measure of turnout (which would be the same for all of the parties competing in each specific election) for my measure of party support in the party-level dataset. The rest, as they say, is history and gave rise to the Rose Festschrift chapter and its online appendices A to C.

The analyses presented in the chapter and its first three appendices are first cuts at an account of why we would expect voter turnout to respond thermostatically to party support and closeness. Although the chapter bravely presents the findings as meeting conventional criteria for statistically significant findings, many alternative model choices would have been possible. And, although the robustness checks in Appendices B and C are supportive, many additional robustness checks are surely called for

More importantly, the paper is unclear as to where it places itself in academic perspective. What we have here might be seen as a proposal for a new subfield for electoral research – a subfield that builds on the original promise of *The American Voter* (1960) to address both voting choice and turnout in concert, using linked theoretical foundations and analytic tools. However, the chapter might equally be seen as belonging with work on thermostatic governance. Knowing where to place the chapter in scholarly terms is critical to deciding how to frame it for maximum impact.

Given this uncertainty, what we have with this chapter should be seen as an academic doodle rather than a serious contribution. It has not been subjected to more than rudimentary peer review. I am fairly sure that submitting the piece, as written, for review by a major journal would result in reviews that would be critical (perhaps outraged) by the extent of the leap that the paper takes, beyond what are the contemporary frontiers on research into negative feedback in policy-making and on research into equilibriating processes in turnout and party support. The fact that, in the process, the paper bight be seen as addressing at least three separate subfields would probably guarantee outrage rather than just criticism.

I seem to be the only person to have addressed this research question as such in over 60 years; and it is clear that I am in way beyond my depth. I have bitten off far more than I can chew. I need help.

I need help just to decide what should be the first step in any serious attempt to bring the ideas presented here into the mainstream of electoral research; and I am very open to the possibility that I have completely missed a major snag hidden somewhere in my own apparent findings. This is why, when asked to organize a Round Table on any topic of my choice, I proposed to organize one on this topic.[10] I hope that, if a sufficient number of scholars who are cleverer than me put their heads together, they may either bring this proto-project to a definitive close or else come up with a viable plan for moving it forward.

[10] See the European Academy of Sciences and Arts Call for Papers: **It's About People** (https://conference.almamater.si/the-2024-call/).