Pergamon

Electoral Studies 21 (2002) 283–295

Electoral Studies

www.elsevier.com/locate/electstud

The rolling cross-section design

Richard Johnston [a,*], Henry E. Brady [b]

[a] *Department of Political Science, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1*

[b] *Department of Political Science, University of California, Berkeley, CA 94720-1950, USA*

Abstract

This article describes the 'rolling cross-section', a design well-adapted to telephone surveys and to capturing real-time effects in campaigns. In one sense, the design is just a standard cross-section, but the day on which a respondent is interviewed is chosen randomly. As a result, analysis of longitudinal factors is possible with only modest controls. The design necessitates an estimation strategy that distinguishes time-series from cross-sectional effects. We outline alternative strategies and show that the design is especially powerful if it is wedded to a post-election panel wave. We also show how graphical analysis enhances its power. Illustrative examples are drawn from the 1993 Canadian Election Study. We compare the design to some obvious alternatives and argue that, for reasons of cost and simplicity, any national election study based on telephone interviewing is best conducted this way.

The 'rolling cross-section' (RCS) is a design that facilitates detailed exploration of campaign dynamics. Its essence is to take a one-shot cross-section and distribute interviewing in a controlled way over time. Properly done, the date on which a respondent is interviewed is as much a product of random selection as the initial inclusion of that respondent in the sample. Because observations are temporally distributed yet closely spaced, the design moves survey research close to true causal inference. It enables links to debates, news coverage, and campaign advertising, as well as identification of the social and psychological mechanisms that mediate the potential impact from external forces.

The first RCS was an adjunct to the 1984 US National Election Study (ANES); the

* Corresponding author. Tel.: +1-604-822-5456; fax: +1-604-822-5540.

E-mail addresses: rjohnstn@interchange.ubc.ca (R. Johnston), hbrady@bravo.berkeley.edu (H.E. Brady).

PII: S0261-3794(01)00022-1

second was the 1987–88 US 'super Tuesday' primary study.[1] The design's coming of age, however, was the 1988 Canadian Election Study (CES). This study and its successors in 1992–3 and 1997, have been intensely mined for their dynamic properties, and campaign effects have become, arguably, the CES's abiding theme.[2] The Canadian example was followed in the 1996 New Zealand Election Study,[3] in the 1998 ANES pilot study, and the massive 'Year 2000' study at the Annenberg School for Communication, University of Pennsylvania.[4]

New forms of data demand new forms of analysis. Parameters appear that hitherto were not identifiable. RCS data require — and also repay — intense graphical treatment. Indeed, informal, nonparametric, visual analysis is often necessary prior to formal, statistical effort. This paper illustrates these propositions by showing how vote intentions evolve over a campaign, by presenting alternative statistical models to help explain vote dynamics, and by showing how graphical analysis helps to resolve the remaining explanatory uncertainties. First, however, what *is* a rolling cross-section?

1. Basic elements of the design

In itself, conducting a telephone survey as a rolling cross-section is unremarkable. All polls with stretched-out fieldwork harbor temporal heterogeneity. If these properties are acknowledged, it is mainly as a problem, because those interviewed later usually differ in important ways from those interviewed earlier. For the 1988 Canadian team, the conceptual breakthrough was to see the need for protracted sample clearance not as a problem, but as an opportunity. By being self-conscious about release and clearance of the sample, we could convert temporal heterogeneity into an object of study.

A rolling cross-section design using a telephone survey requires, first, a body of telephone numbers sufficiently large to yield a target number of interviews. This body is then broken up into *replicates*. In the CES case, one replicate is generated for each day of projected interviewing.[5] Each replicate is a miniature of the total, in that assignment to a replicate is essentially random, just as initial selection for the total sample is random. The CES goes into the field within days of the start of the election campaign, but, in principle, fieldwork can start at any time.

Although each day has its replicate, the replicate itself is not the proper unit for

[1] The 1984 data were aired in Brady and Johnston (1987) and Bartels (1987, 1988). The 1988 Super Tuesday data, however, seem to lie fallow.

[2] See, in particular, Johnston et al. (1992, 1994, 1996) and Nevitte et al. (2000).

[3] See Vowles et al. (1998), especially ch. 5 (Johnston) and ch. 8 (Miller).

[4] Co-investigators on the Year 2000 Study are Johnston and Kathleen Hall Jamieson. The first publication based on this study is Hagen et al. (2000).

[5] Nothing requires that only one replicate is issued per day. The Annenberg Year 2000 study, for instance, varied the overall intensity and the geographic focus of fieldwork according to the year's electoral rhythms. The RCS response rate model differs subtly from other designs, as shown below.

representing campaign time. This would imply that all numbers in a given replicate remain open for contact by interviewers on one day only. This is obviously bad survey practice, as only the easiest to reach would be contacted. Fig. 1(A) profiles the lag, for the 1993 Canadian data, between the release of a telephone number to sample and actual completion of an interview at that number. From the typical replicate, almost half the interviews are ultimately completed on the day of release. Another one-sixth are completed the next day, one-twelfth, the day after that, and so on. The median lag is 1 day and the mean lag is 2.1 days. Five-sixths of the interviews that will ever be completed from a replicate are completed in 5 days, including the release day. The remaining one-sixth is distributed thinly over succeeding days, and virtually all are completed inside 2 weeks.

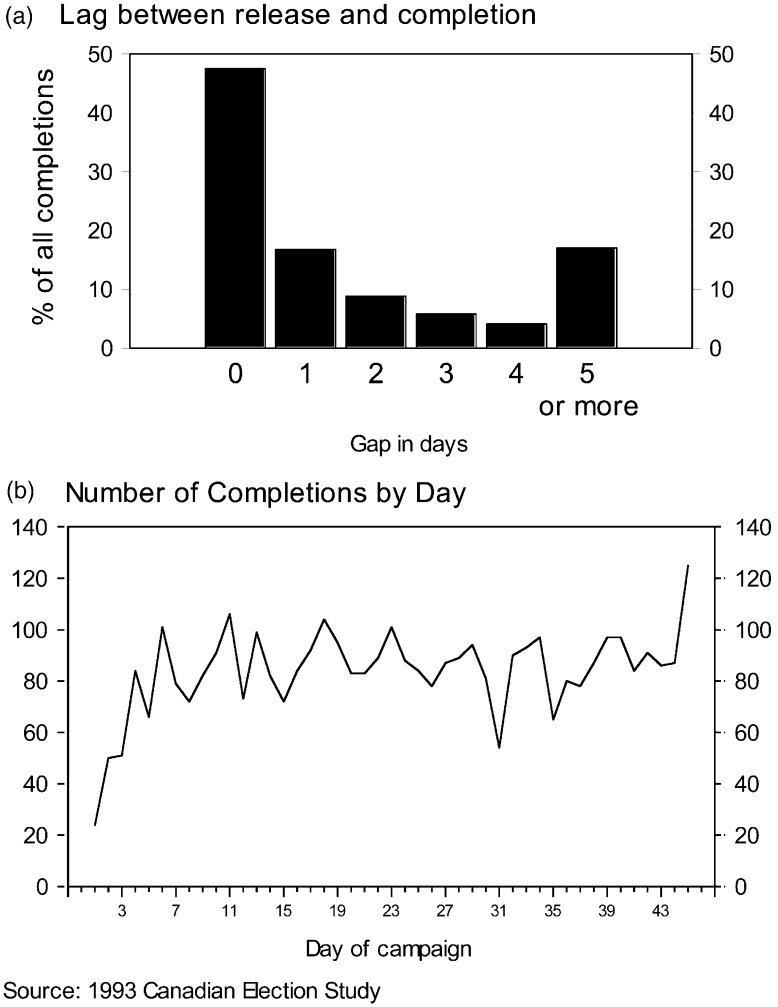

Fig. 1. Distribution of interviews.

The distribution of completed interviews by fieldwork day in the 1993 study is shown in Fig. 1(B). Days 1–3 are qualitatively different from all that follow, in that there were fewer than 50 completions. By day 4, completions clear 80 per day, roughly the daily average from then on. It is at this point that the day of interview becomes, effectively, a random event. Only at the very end does anything beyond stochastic error appear. The last day's take exceeds 120, over one-third greater than usual. This reflects prospective respondents' realization that there is no tomorrow; an interview cannot be scheduled for another day.

Critically, no intensification of effort occurs in this period, and the integrity of the design requires that there be no such intensification. Otherwise, the date of interview would not be a random event (and, frankly, it is not at the beginning and not quite on the last day). This does mean, however, that the overall response rate for an RCS will be lower than for a less time-conscious design. Stable clearance effort necessarily implies that late replicates will not be exhausted as completely as earlier ones. Total completions do not flag, as Fig. 1(B) makes clear, but only because earlier replicates are finally seeing their tardy members cleared. Completions from late replicates begin to drop with 2 weeks to go, as we would expect from discussion, above, of Fig. 1(A). In the second last week, the drop is very gradual, from a daily completion rate in the mid-80s to one in the mid-70s. In the last week, the drop is precipitate, to the 60s and then below 50 on the last day. Had we been able to continue clearing replicates after election day, the response rate could have been about three percentage points higher. Had we released telephone numbers to sample as quickly as house capacity allowed and begun aggressive clearance early, we might have gained still more respondents. But then we would not have had a rolling cross-section.

What do we get when we get a rolling cross-section? Fig. 2(A) shows how a dynamic analysis might start. It tracks one of the critical quantities of the 1993 campaign: the share of vote intentions accruing to the governing Conservative Party.[6] The jagged line is the daily share and the smooth line is the 7-day moving average of that share. Early on, the Conservative share is in the mid-30s, very close to the share (not presented) for its chief rival, the Liberal Party. Between fieldwork days 11 and 13, according to the moving average tracking, the Conservative Party share dropped nearly ten points, and by day 15 there could be no doubt that the party had lost at least that much.

The daily tracking is, of course, highly variable day to day, but it hints that this collapse required only 2 days. It plunged a party then in power into a fight for its very life. Subsequent events admit contrary readings. On one reading, the balance of the campaign is just the gradual playing out of the initial shock. On another, that shock plays out within days, such that the Conservative share now oscillates at a

[6] Quebec respondents are excluded, as campaign dynamics in that province were weaker and subtly different from elsewhere. In the typical day, Quebec dwellers constituted about 25% of the total sample, so the daily readings in Fig. 2 rest on just over 60 respondents.

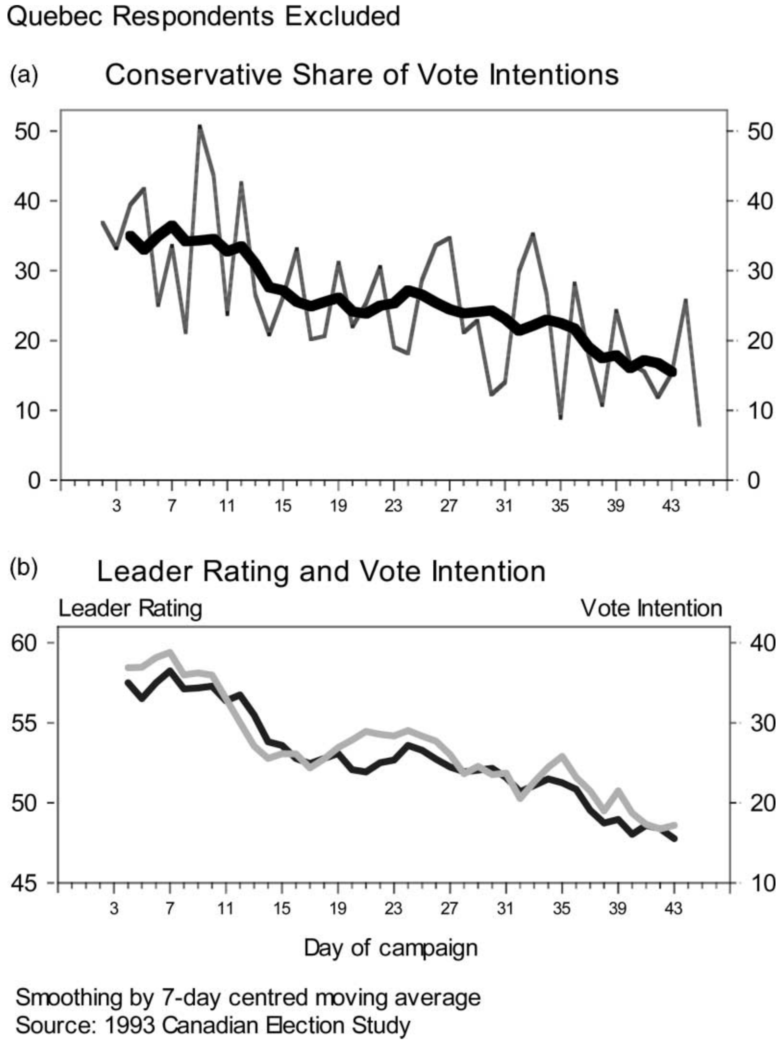

Fig. 2. Conservative vote, Conservative leader. Quebec respondents excluded.

new equilibrium, in the mid-20s. Only after day 34 does the share take its final dive, also of 10 points.[7]

[7] We lean to the second reading. Between days 15 and 34, virtually no trend is visible. The eye may be led to see one, but that, we submit, is a visual trick played by the last 10 days. A sign test on the middle 20 days yields as many positives as negatives on consecutive days. After day 34 virtually every pair of days yields a negative change. The drop over this 10-day period is roughly 10 points, half the total slide.

2. Estimating longitudinal effects

Clearly, something happened early in the campaign to update voters' beliefs about the Conservative Party. Something else might have happened late in the campaign, but this is more contestable. How do we represent the effect of updated beliefs?

Assume for the sake of argument that we have only one relevant belief, be it about an issue, a leader, or a strategic contingency. Consider, then, some belief B_{it} at time t for person i.[8] For the effect of B_{it} on Y_{it} to be related to the campaign proper, the pre-campaign likelihood of voting for the party must be controlled. The simplest way to do this is with the following conditional change model:

$$Y_{it}=\alpha_1+\beta_1 B_{it}+\delta_1 Y^0_{it}+\varepsilon_{it} \tag{1}$$

where Y^0_{it} is the likelihood of person i who is eventually interviewed at time t voting for the party at the beginning of the campaign (t=0).[9] The coefficient β_1 represents the impact of B_{it}. Eq. (1) is essentially the standard model in the campaign-effects literature (Finkel, 1993, p. 6; 1995, pp. 6–7), and it implies that the likelihood at time t of person i voting for the leader is a function of the likelihood at the beginning of the campaign plus the effect of the current belief B_{it}.

The likelihood of voting for the party based upon the person's belief B^0_{it} at t=0, before the campaign, is:

$$Y^0_{it}=\alpha_0+\beta_0 B^0_{it}+\varepsilon^0_{it}. \tag{2}$$

Substituting Eq. (2) into Eq. (1) and collecting terms yields the following:

$$\begin{aligned} Y_{it}&=\alpha_1+\beta_1 B_{it}+\delta_1(\alpha_0+\beta_0 B^0_{it}+\varepsilon^0_{it})+\varepsilon_{it} \\ &=(\alpha_1+\delta_1\alpha_0)+\beta_1 B_{it}+\delta_1\beta_0 B^0_{it}+(\delta_1\varepsilon^0_{it}+\varepsilon_{it}), \end{aligned} \tag{3}$$

where the coefficient of interest is β_1. We can simplify this by writing:

$$Y_{it}=\alpha_2+\beta_1 B_{it}+\delta_2 B^0_{it}+\psi_{it} \tag{4}$$

where, obviously:

$$\alpha_2=(\alpha_1+\delta_1\alpha_0);\ \delta_2=\delta_1\beta_0 \text{ and } \psi_{it}=(\delta_1\varepsilon^0_{it}+\varepsilon_{it}).$$

If B^0_{it} is the person's initial belief, and if B_{it} is the belief at the time of interview, then the coefficient β_1 on B_{it} seems like a reasonable definition of a campaign effect. We can then estimate campaign effects by regressing vote intention at t on beliefs at t as well as on beliefs at the beginning of the campaign.

This brings us to an obvious problem. In the Canadian studies, as in the other

[8] Of course, no person interviewed at one time within the campaign is interviewed at another such time. For notational simplicity, we assume an equal number of respondents I for each time period t. Fig. 1 indicates that this is not quite true, but that it is almost true for all but the very first and the very last days.

[9] The subscript t represents the cross-section in which the respondent is interviewed and i is that respondent's identifier within the cross-section. If the variable is measured at the RCS interview, then we do not add a superscript (hence Y_{it} and not Y^t_{it}), but if it is measured at another time, say, in a pre-election baseline, then we add the superscript for time, in this example, Y^0_{it}.

RCS samples drawn so far, the RCS wave is itself the first one; there is no pre-campaign baseline. It is not even clear that we want one, given that a pre-campaign interview primes respondents in ways that make the campaign wave unrepresentative. Be that as it may, this leaves us lacking direct evidence for B_{it}^{0}. It does not suffice just to enter B_{it} into an estimation of Y_{it}, as doing so would confound longitudinal with cross-sectional impact because of the correlation between B_{it} and the omitted B_{it}^{0}. As a general proposition, there is no way to assess the impact of omitting B_{it}^{0} because there is no reason to assume any particular covariance between time-series and cross-section. This is a point made forcefully by Kramer (1983), and it applies not just to campaign studies such as the CES but also to any analysis that employs the same measures in repeated samples. Just as Kramer's problem has been addressed elsewhere (Markus, 1988), so it can be in the RCS design.

This requires us to consider other starting places for estimation. One place to look is *after* election day, if the design includes a post-election wave. The other place is inside the RCS itself, especially if post-election information is not available.

If key measures from the campaign-wave RCS are repeated post-election, then B_{it}^{T+1} is observed. If we allow ourselves a strong assumption about how beliefs change, we can substitute B_{it}^{T+1} for B_{it}^{0}. Let us assume that B_{it} changes as follows:

$$B_{it}=B_{it}^{0}+D_{t}. \qquad (5)$$

where B_{it}^{0} is the cross-sectional state of belief at time zero and D_t is the time-series variation.[10] This assumes that today's belief equals that at the beginning of the campaign plus some time-series effect that is common across all people. Everyone's opinion, then, is affected by the same amount, D_t, at each time period t.[11] We wish to estimate Eq. (4), and it can be estimated if we have values for B_{it} and B_{it}^{0}. We observe B_{it}, but what can we use for B_{it}^{0}? If Eq. (5) is correct, then B_{it}^{T+1}, a person's belief after the end of the campaign as reported on the post-election interview, will be:

$$B_{it}^{T+1}=B_{it}^{0}+D_{T+1}, \qquad (6)$$

so that B_{it}^{T+1} differs from B_{it}^{0} only in the constant D_{T+1} and it will be perfectly correlated with B_{it}^{0}. Consequently, we can substitute B_{it}^{T+1} into Eq. (4) to represent B_{it}^{0} and the only change in (4) will be in the intercept $\alpha_2^{*}=\alpha_2-\delta_2 D_{T+1}$:

$$Y_{it}=\alpha_2^{*}+\beta_1 B_{it}+\delta_2 B_{it}^{T+1}+\psi_{it} \qquad (7)$$

In reality, B_{it}^{T+1} and B_{it}^{0} will not be perfectly correlated. We can imagine some obvious possibilities: changes that are random, uncorrelated with initial beliefs, as suggested by the literatures on non-attitudes and on measurement error; or change that is systematic, but correlated with original beliefs, as in opinion crystallization. These take us beyond the scope of this paper. We note, however, that where such change

[10] Note: *D* stands for 'diachronic', meaning variation over time.

[11] This assumption is not warranted by the facts but it is a useful starting point, and space does not permit further consideration here. Brady and Johnston (1996) discuss the consequences of relaxing it.

occurs — and thus Eq. (5) is unwarranted — then B_{it}^{0} is just as inadequate as a starting point as B_{it}^{T+1}.

The fact still remains that all variance in the post-election measure B_{it}^{T+1} is, for our purposes, cross-sectional: the election is over and the furor has died immediately; clearance of the post-election sample is relatively rapid; and initiation of contact is orthogonal to the date of pre-election interview. Thus, the post-election reading is a reasonable starting point for separating cross-sectional from longitudinal effects. When taken together, Eqs. (4) and (7) show that the coefficient (β_1) on the RCS wave measure (B_{it}) indicates the campaign effect, while the coefficient (δ_2) on the post-election measure of belief (B_{it}^{T+1}) measures the residual impact of the baseline belief. The rest of the impact is in β_1, and so the full cross-sectional effect is indicated by (β_1+δ_2).[12]

Post-election indicators are not always ready to hand. Neither the 1984 nor the 1988 US RCS has a post-election wave, and not even in the CES is every serious campaign-period factor measured twice. Commonly, then, we are forced to condition on day-to-day information. An obvious starting point is to average across the B_{it} in each daily sample:

$$B_t^{*}=\sum_{i=1,I} B_{it}/I, \tag{8}$$

to get a quantity to represent the time-series effect and to use ($B_{it}-B_t^{*}$) to represent the cross-sectional effect B_{it}^{0}. To see that these indicators will do the trick, go back to the logic in Eq. (5). Assume first, for mathematical simplicity, that the population mean of B_{it}^{0} is set to zero, so that we can scale B_{it}.[13] Then if Eq. (5) is true we can take averages on both sides over each daily sample to get:

$$B_t^{*}=\sum_{i=1,I} B_{it}/I=\sum_{i=1,I} B_{it}^{0}/I+\sum_{i=1,I} D_t/I=D_t, \tag{9}$$

where we have assumed that the sample mean of B_{it}^{0} is zero because the population mean is zero. (That is, we assume that the sample quantity B_t^{*} is a good approximation of D_t, which will be true for large daily samples.) With this result, we can rearrange Eq. (5) to get an expression for B_{it}^{0} in terms of observable quantities:

$$B_{it}^{0}=B_{it}-D_t=B_{it}-B_t^{*}. \tag{10}$$

If we have large daily samples, then B_t^{*} should be close to the true D_t and the estimate of B_{it}^{0} in Eq. (10) should be close to the actual B_{it}^{0}. Hence, we can regress Y_{it} on B_{it} and on ($B_{it}-B_t^{*}$) from (10). This procedure should yield a consistent estimator of

[12] An intuitive statement of these relationships can be found in Johnston et al. (1992). A somewhat formal elaboration can be found in Brady and Johnston (1996), where we weaken the assumptions in Eq. (5).

[13] This is a harmless assumption that, operationally, only requires that we calculate the mean, B_0^{*} of the B_{it}^{0} and subtract this quantity from all B_{it} including B_{it}^{0}. This amounts, then, to setting B_0^{*} to zero.

campaign effects. To see this, go back again to Eq. (4) and substitute $(B_{it}-B_t^*)$ for B_{it}^0:

$$Y_{it}=\alpha_2+\beta_1 B_{it}+\delta_2(B_{it}-B_t^*)+\psi_{it}. \tag{11}$$

Coefficients have the same interpretation as in Eq. (7).

In principle, this is a very attractive strategy. The key phrase in the foregoing, however, is 'large daily sample'. The statistical consistency that made the estimator so attractive accrues as the size of the daily samples increases without limit. Two related consequences of small daily samples occur to us. First, the sampling variance associated with 80+ completions per day introduces extra error into Eq. (8) as an indicator of the true daily mean. This presents us with a classic errors-in-variables problem, which could bias δ_2 towards zero and, possibly, β_1 in the other direction. Second, to the extent that B_t^* is mismeasured, variation in B_{it} will leak into $(B_{it}-B_t^*)$, such that the two will be collinear, inflating standard errors on both β_1 and δ_2.

3. An example

Table 1 gives a practical example, by showing the impact of leader ratings on 1993 Conservative vote intention, estimated each way. Coefficients are extracted from a larger estimation, which also includes policy and expectations variables.[14]

Table 1
The impact of leader ratings on Conservative vote[a]

	Conditioning on:	
Leader	Individual info — post	Daily mean
Conservative		
β_1	0.95*** (0.06)	1.05*** (0.27)
δ_2	0.17** (0.06)	−0.08 (0.27)
Liberal		
β_1	−0.60*** (0.06)	−0.76* (0.38)
δ_2	−0.25*** (0.07)	0.09 (0.38)
Reform		
β_1	−0.43*** (0.06)	−0.71** (0.23)
δ_2	−0.02 (0.06)	0.30 (0.24)
R^2-adj	0.37	0.30
SEE	0.36	0.37
N	1501	1984

[a] Estimation by OLS. Standard errors in parentheses. *$p<0.05$; **$p<0.01$; ***$p<0.001$.

[14] Leader ratings are 100-point thermometer scales compressed to the 0–1 interval. A complete description of the variables can be found in Brady and Johnston (1996). Note, however, that the estimation here varies in some particulars from the one reported there.

The alternative setups provide highly similar estimations of total time-series plus cross-sectional effect, but the daily-mean version assigns more of the total to the time series. Note also that the daily-mean time-series coefficients, although large, are unstable. We suspect this reflects the artifacts just discussed. For all that, however, the alternatives provide broadly similar readings. Our sense is that the pre–post design dominates the daily-mean design. However, if conditioning on post-election information is not possible, daily-mean estimates should not lead us far astray, especially if daily samples are large. For 1993, both treatments suggest that judgments on Kim Campbell, the Conservative leader, powerfully influenced her party's chances.

But did leader judgments, in fact, have this effect? Fig. 2(B) confirms that her ratings fell over the campaign, but perhaps she just fell along with her party, and judgments on her were as much consequence as cause. It is true that other key factors are controlled in Table 1, but controls do not dispose of the causal issue.

This is where visual analysis comes back into play, as judicious manipulation of graphs helps address the issue of causal priority. If re-evaluation of Kim Campbell was critical to, say, the early precipitate drop in Conservative vote intentions, then that re-evaluation had better precede, not follow, the vote shift. Fig. 2(B) attempts to sort out temporal priority. The visual challenge is to get the vertical scale of vote intention and leader evaluation into the same range, so that the temporal priority can be established by horizontal comparison. This is accomplished by:

- assigning leader evaluation to one *Y*-axis and vote intention to the other; and
- setting the vertical range separately on each axis. Mean leader ratings span larger values but a shorter range than percentages of vote intention.

With this done, Fig. 2(B) makes a strong prima facie case that re-evaluation of Campbell was indeed critical to the Conservatives' sharp drop. It also makes clear that, thereafter, she ceased to play an independent dynamic role. Her ratings dropped further, but not with any obvious temporal lead over her party.

4. Discussion

There is no compelling reason *not* to conduct an RCS where fieldwork is temporally dispersed in the pre-election period, as it almost always is. Any campaign period is bound to make response vary over time, even if the campaign only demarcates the run-up to a deadline. Shifts can be very rapid, so almost any survey fieldwork is vulnerable to this heterogeneity. We argue that the RCS is the most effective design for dealing with the problem. Indeed the design converts the problem into an opportunity.

As a design to capture campaign effects, does the RCS dominate a panel? We have already admitted that the repeated-measures design is the canonical way to begin thinking about the problem. Moreover, where the same individuals are measured twice, sampling error is not an issue for estimating net change even as individual trajectories can be captured. Thus, especially where the interval is short, panels have an obvious appeal. This theoretical appeal, however, is very hard to realise in practice. In most electoral contexts the obvious time spans are too long for panels to permit fine-grained causal attribution and for easy maintenance of the panel itself. Repeated panels might permit somewhat fine-grained coverage, but at considerable expense. Moreover, any panel wave other than the very first one is no longer a cross-section and thus is no longer representative of the electorate in the conventional sense, as panel mortality bends later waves away from being demographically representative. No less important, respondents are altered by the earlier interview, and so identification of the campaign's intrinsic dynamic effect is problematic. These considerations also tell against mounting a pre-campaign, conventional cross-section and then drawing out the second, panel wave as a kind of campaign RCS. Besides, a stable second-wave clearance strategy — so critical to making inference with the RCS easy — is almost unimaginable.[15]

At the same time, an RCS is next to impossible to construct after the fact. Commercial polls are now ubiquitous, at least in the US, and it is tempting to use them to reconstruct the dynamics of the campaign. For earlier campaigns, we have no choice but to assemble the data this way. Variation from survey house to house in sample frame, screening, weighting, and question wording is remarkable, however, such that house effects are likely to contribute as much error variance as sampling itself.[16] In addition, commercial data are rarely available at the individual level.

The implication seems clear. If one is to be in the field for a large part of a campaign period, thought and effort should be given to controlling the release and clearance of the sample, so as to make the data collection sensitive, with minimal controls, to events in real time. Obviously, the larger the overall sample, the more sensitive can be samples gathered within temporal subdivisions. For many operations, only a small conceptual leap is required to grasp that the effective size of the overall sample may already be quite large. Commercial operations that publish 'tracking polls' are already accumulating massive total samples. Even commercial companies that publish, say, weekly cross-sections would do better to see the separate weekly studies as part of a consolidated design spanning the whole period of interest, and then to control sampling from start to finish. In an ideal world, the academic community should have the resources to do this for itself. But with a modest raising of self-consciousness, commercial polls can also create an enduring research legacy, even as they enhance the short-run proprietary value of their data.

The RCS as we describe it presupposes interviewing by telephone.[17] In-home interviewing makes controlled release and clearance of a sample all but impossible. Only if the pace of events is sufficiently slow that the time scale is conceived in weeks

[15] This discussion does not preclude a stand-alone cross-sectional survey to establish a pre-election baseline, as long as it is understood that that baseline would permit only aggregate comparisons with the later RCS. And the earlier fieldwork starts relative to election day, the more can early interviews in the RCS itself serve as an aggregate baseline.

[16] See Erikson and Wlezien (1999). The discussion on p. 171 is particularly relevant.

[17] It is conceivable that an RCS could be conducted over the Internet.

rather than days could the design be wedded to a face-to-face survey. That said, the estimation strategy we outline is not peculiar to the RCS. Although it was necessitated by our self-conscious attention to processes operating in real time, its logic is general to *any* analysis that involves survey data collected at different times. What distinguishes the RCS from such merged files is not temporal heterogeneity as such, which may be present in both situations, but the fact that the heterogeneity in the RCS can involve days so closely spaced that clearance from a sample released on one of the days may not have been properly completed before the next relevant day arrives. It is the interleaving of sample replicates, day by day, that distinguishes the design.

If the design opens up new possibilities for analysis, its full exploitation requires new estimators and new ways of thinking. Once in the RCS framework, estimation should take the obvious step beyond the assumptions underlying dynamic analysis with panels. Although we find it useful to start by conceiving of RCS estimation as if it were a panel in which the second wave is released gradually rather than at one shot, we shy away from recommending a pre-campaign baseline wave. However, we must look for a baseline somewhere. One compelling strategy is to find it in a post-campaign wave.[18] We argue that a later wave is probably no worse a representation of the baseline than a pre-campaign wave would have been. Sometimes not even a post-campaign wave is available, and so we also need tools for extracting a notional baseline from inside the RCS proper.

Perhaps the biggest change forced on us by the RCS is to take visual presentation seriously, not just as an expository device but as an original research tool. RCS data commonly do not begin to speak until they are arrayed visually, with the daily noise of sampling smoothed away. Of course, pictures can mislead as readily as inform. But in tandem with formalized, parametric techniques, they supply evidence which is no less powerful for being circumstantial.

Acknowledgements

Data for this paper are drawn from the 1992–93 Canadian Referendum and Election Study, for which Johnston was principal investigator and Brady was a co-investigator. Other co-investigators were André Blais, Elisabeth Gidengil, and Neil Nevitte. The study was supported by the Social Sciences and Humanities Research Council of Canada and by the investigators' universities. Fieldwork was conducted by the Institute for Social Research, York University, under the direction of David Northrup. We thank Mark Franklin, Chris Wlezien, and Charles Franklin for comments on earlier drafts. None of the foregoing are responsible for any errors of analysis or interpretation.

[18] To serve as a baseline in the strict sense we intend here, the post-election wave must be a panel. Where resources permit, it is also useful to collect a fresh cross-section after the election, but the uses of the latter lie elsewhere than as a baseline for the estimations discussed earlier in this paper.

References

Bartels, L.M., 1987. Candidate choice and the dynamics of the presidential nominating process. American Journal of Political Science 31, 1–30.

Bartels, L.M., 1988. Presidential Primaries and the Dynamics of Public Choice. Princeton University Press, Princeton.

Brady, H.E., Johnston, R., 1987. What's the primary message: horse race or issue journalism? In: Orren, G.R., Polsby, N.W. (Eds.) Media and Momentum: the New Hampshire Primary and Nomination Politics. Chatham House, Chatham, NJ.

Brady, H.E., Johnston, R., 1996. Statistical methods for analyzing rolling cross-sections with examples from the 1988 and 1993 Canadian Election Studies. Paper presented at the 1996 Annual Meeting of the Midwest Political Science Association.

Erikson, R.S., Wlezien, C., 1999. Presidential polls as a time series: the case of 1996. Public Opinion Quarterly 63, 163–177.

Finkel, S., 1993. Reexamining the 'minimal effects' model in recent Presidential campaigns. Journal of Politics 55, 1–21.

Finkel, S., 1995. Causal analysis with panel data. In: Sage university papers series. Quantitative Applications in the Social Sciences, 07-105. Sage Publications, Thousand Oaks, CA.

Hagen, M.G., Johnston, R., Jamieson, K.H., Dutwin, D., Kenski, K., 2000. Dynamics of the 2000 Republican primaries. Annals of the American Association for Political and Social Sciences 572, 33–49.

Johnston, R., Blais, A., Brady, H.E., Crête, J., 1992. Letting the People Decide: Dynamics of a Canadian Election. McGill-Queen's University Press, Montreal.

Johnston, R., Blais, A., Brady, H.E., Gidengil, E., Nevitte, N., 1994. The collapse of a party system? The 1993 Canadian General Election. Paper presented to the 1994 Annual Meeting of the American Political Science Association.

Johnston, R., Blais, A., Gidengil, E., Nevitte, N., 1996. The Challenge of Direct Democracy: The 1992 Canadian Referendum. McGill-Queen's University Press, Montreal.

Kramer, G., 1983. Ecological fallacy revisited: aggregate- vs. individual-level findings on economics and elections, and sociotropic voting. American Political Science Review 77, 92–111.

Markus, G.B., 1988. The impact of personal and national economic conditions on the presidential vote: a pooled cross-sectional analysis. American Journal of Political Science 32, 137–154.

Nevitte, N., Blais, A., Gidengil, E., Nadeau, R., 2000. Unsteady State: The 1997 Canadian Federal Election. Oxford University Press, Don Mills, Ontario.

Vowles, J., Aimer, P., Banducci, S., Karp, J. (Eds.), 1998. Voters' Victory? New Zealand's First Election under Proportional Representation. Auckland University Press, Auckland.