Journal of Elections, Public Opinion and Parties
Vol. 19, No. 1, 117–124, February 2009

Routledge
Taylor & Francis Group

Long Memory Methods and Structural Breaks in Public Opinion Time Series: A Reply to Pickup

EVERETT YOUNG & MATTHEW J. LEBO
Department of Political Science, Stony Brook University, NY, USA

Downloaded By: [Young, Everett] At: 20:41 15 March 2009

Mark Pickup does not take issue with our substantive conclusions about the relationship of Liberal Democratic support to leadership approval. However, he does argue that our claim that vote intention variables are fractionally integrated is unsubstantiated. Instead, he suggests that time series of electoral support might be stationary autoregressive processes with "equilibrium shifts" or "structural breaks". Or, they could be "combined process", asymptotically integrated of order 1. Either way, Pickup claims, such series would *appear* fractionally integrated in statistical tests.

Pickup contends that if one of these alternate possibilities hold – or, he suggests, even if the series *is truly* fractionally integrated – we should avoid the "complications" associated with fractional integration and model electoral support as an autoregressive-moving-average (ARMA) process. He adds that if one insists on fractional differencing to handle autocorrelation, then structural breaks must be accounted for in calculating the level of fractional integration, *d*.

Due to these objections, our article does not make use of fractionally differenced variables in the main equations, as earlier iterations of the paper did. The substantive conclusions remained unaltered but we remain convinced that the use of fractional differencing is a superior way of modeling electoral support. We appreciate Pickup's thoughtful response and the editors' offer to us to explain why fractional integration methods are appropriate to studies of party support over time.

Pickup's arguments against the use of fractional integration (FI) techniques are based on two premises. First, he asks: "Are the added complications of fractional integration necessary?" That is, he assumes that FI is beset with "complications" from which ARMA models are free.

With much respect, we nonetheless do not accept the premise of Pickup's question. Pickup seems to define "simple" in a way we do not. To Pickup, simple does not

Correspondence Address: Everett Young and Matthew J. Lebo, Department of Political Science, Stony Brook University, Stony Brook, NY 11794-4392, USA. Email: vertie@earthlink.net; matthew.lebo@stonybrook.edu

ISSN 1745-7289 Print/1745-7297 Online/09/010117-08 © 2009 Elections, Public Opinion & Parties
DOI: 10.1080/17457280802592055

mean theoretically pleasing, plausible, or elegant. Nor does it mean using minimal parameters or avoiding well established threats to inference. Rather, "simple" is used to mean the avoidance of computationally-intensive filters, however statistically reliable and well grounded in theory they may be. We hold that approximating an FI process – if it is the true data generating process – with a complex ARMA model is non-simple in a more unacceptable way. It is theoretically implausible and inelegant and builds into the data statistical artifacts that cause a range of problems.

Pickup himself acknowledges some of these problems, stating gravely that "to treat a fractionally integrated series as *I(0)* or *I(1)* could result in spurious findings", Yet, a mere five paragraphs later, he suggests that "even under the expectation of heterogeneous individual level behavior" (i.e., under the expectation of FI), an aggregate time series might be a finite order ARMA process (in which case FI would be the wrong model), *or might* "*at least be well represented by a finite ARMA model*" (emphasis added) in which case the "complexities" of FI would be unnecessary. The suggestion to use a less "complex" filter even where it is likely to be an inaccurate representation of the data-generating process, however, is a contradiction of Pickup's earlier statement, with which we agree entirely: it is best to filter the dependent variable using the most accurate model, not necessarily the simplest.

Trying to approximate FI with more traditional methods like ARMA has several complications. Where a series is FI, leaving it in level form or first-differencing it results in over- or under-differencing, is associated with spurious regressions (which Pickup recognizes), can result in biased standard errors and coefficients, inflated R^2 statistics, and faulty conclusions about cointegration.

The second premise behind Pickup's question is the uncritical acceptance that public opinion series in political science are rife with structural breaks. We do not accept this as a given. A reading of Granger and Hyung (2004) reveals that a structural break is defined rather abstractly as a *periodic change in the mean to which a stationary series reverts*. This is not the same as a simple intervention which can move a series away from a long-term mean for several periods. Essentially, a shock is a move away from a long-term equilibrium while a break is a movement of the equilibrium itself.

Extant tests are not sufficient to distinguish whether a series has undergone a large shock with long memory (a component of fractional integration) or whether a permanent equilibrium shift has occurred. Nor is there a sufficiently rigorous real-world understanding of structural breaks that tells us precisely when they have occurred. In the study of economics where much is written about structural breaks, it is easy to think of examples. Exchange rate regimes and banking systems may be fundamentally changed, currencies may be pegged or un-pegged against each other. New legislation can forever after alter the workings of markets.

Yet it is a major leap to assume that the changing of a party leader, the fallout of a particular election, or a particular economic or political event rises to this level of impact worth calling a "structural break". To Pickup it is obvious that myriad political events represent structural breaks, but we think these are more properly described as interventions and handled appropriately through traditional transfer

Downloaded By: [Young, Everett] At: 20:41 15 March 2009

function analysis. As we outline below, estimating *d* under the supposedly safer assumption that events are structural breaks would be a procedure contaminated with bias.

Monte Carlo Evidence

We turn now to Pickup's simulations, the focus of his response. These address the point that electoral support is not integrated at all, but is instead an *I(0)* ARMA process with structural breaks.

Pickup shows that a semi-parametric estimator such as Robinson's (1995) can reject *I(0)* for a white-noise series with structural breaks. Indeed, Pickup shows that even his preferred maximum-likelihood estimator easily rejects *I(0)* for a strongly autoregressive AR1 series even when *no* structural breaks are present. Granger and Hyung (2004) show that it is extremely difficult to distinguish empirically a strongly autogregressive series with structural breaks from one which is fractionally integrated. It may be, then, that our significance tests for *d* are insufficient, *in the absence of theory*, to establish that our dependent variables are not, in reality, *I(0)*.

But we do not work in the theoretical vacuum we are being asked to assume. A theory of heterogeneous processes behind party support is far stronger than any theory of homogenous voter behavior. Indeed, Pickup allows that homogeniety of voter memory is unrealistic.

To elaborate, when outside shocks cause a change in electoral support for a particular party, surely some voters who have changed their minds almost as quickly change them back. Some regress more slowly to their original equilibrium. Still others surely change their minds, only to have their opinions settle back, at varying rates, *toward* previous equilibria, but not fully. Doubtless still others experience their opinions as "online" running tallies, responding to arising news items by incrementing or decrementing their opinions of the parties in, essentially, an *I(1)* random walk.

This is, of course, a textbook description of a process *likely* to generate fractional integration. On the other hand, a theory of voter *homogeneity* – e.g., "all individuals' level of support for a party is AR1 = 0.8 because…" – has not, to our knowledge, been proposed, nor are we aware of any good reason why unproposed, hypothetical models that assume homogeneity should be privileged. Ostensibly, again, the reason is "simplicity". But this reason carries the tacit admission that multi-parameter ARMA models of aggregated series do not claim to model the data-generating process at all. Rather, they are merely "pre-whitening" strategies – i.e., ways of turning the dependent variable into a "white noise" series devoid of autocorrelation.

In fact, fractional differencing is a better pre-whitening strategy than an AR1 model, even when controlling for events. The errors that result from filtering electoral support through a simple AR1 model – when controlling for leadership changes or not – are still contaminated with autocorrelation and require additional AR or MA parameters at more distant lags to correct the problem.

Downloaded By: [Young, Everett] At: 20:41 15 March 2009

The first panel of Figure 1 shows the residual autocorrelation function (ACF) of the Labour vote intent series (1979–1997) when we control for so-called structural breaks. The second panel shows that series after it has been filtered through a Box–Jenkins model with an AR1 parameter.[1] The result is an *over-differenced* series with the most serious effect being the large negative spike created at the first lag. Over-differencing will at least introduce the need for an additional MA parameter and, at worst, this parameter will be non-invertible and extremely problematic for estimation and interpretation.[2] Having to add more parameters that are not *a priori* thought to be part of the model is thus theoretically problematic as well. Without a

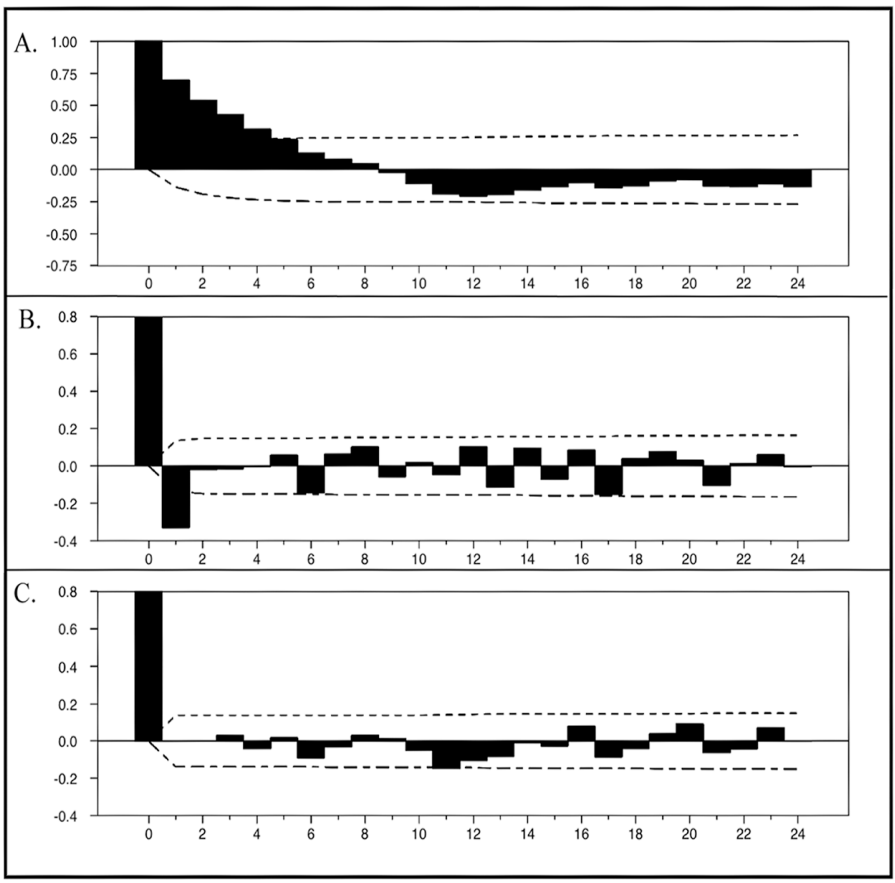

Figure 1. Autocorrelation functions of Labour vote intent from the first period (conservative government), controlling for leadership changes as though they were "structural breaks", in (a) level form, (b) after estimating and controlling for an AR1 parameter, and (c) after fractional differencing. In graph B, effect of overdifferencing – the introduction of a noninvertible MA parameter – by controlling for the AR1 parameter, estimated to be $\phi = 0.99$, is strongly evident.
Note: dotted lines are standard errors * 2 to show statistical significance.

Downloaded By: [Young, Everett] At: 20:41 15 March 2009

belief that they are true components of the data-generating process, why do we follow a procedure that would force us to include them? Again, we do not believe this is a step in the direction of simplicity. The third panel of Figure 1 shows a much cleaner set of residuals using only a single parameter, d, to filter the series. Under a theory of heterogeneous voters, this single parameter fits theoretically as well as statistically.

In fact, not a single one of our electoral support series – Tory, Labour or Liberal Democratic support, in either the Conservative or the Labour government period – can be adequately whitened with a single AR parameter and structural breaks. The reason why a simple AR1 model "overdifferences" the series becomes clear when we address a particular aspect of Pickup's simulations: his choice of an AR1 parameter of 0.85. Pickup selects this AR1 parameter because "this is the autoregressive parameter of the Conservative party popularity series". But this is not true of either the Conservative or Labour government periods. We ran Box–Jenkins models specifying a single AR1 parameter for all six of our electoral-support variables – all three parties in each of two periods – and ran these filters both with and without controls for leadership changes. In all 12 cases the estimate of the AR1 parameter, ϕ, is approximately 1. The six estimates in the Conservative period range from 0.9949 to 1.0005. The six in the Labour period range from 0.9920 to 0.9990. In other words, the series are not autoregressive, they are nonstationary. An AR1 model, designed to capture short-term memory and the basis of Pickup's simulations, is the wrong model.

But could we mistakenly get the same $\phi = 1$ result from a simulated series constructed to have $\phi = 0.85$ with breaks added onto it? If so, our results above would be untrustworthy. But even assuming structural breaks occur and then adding them to our simulated series, we still do not get substantially biased estimates of the autoregressive parameter. We simulated a series with 211 time points, generated from white noise with a standard deviation of 2.5, and constructed to have an AR1 parameter of $\phi = 0.85$. Then we added large breaks of 4, –6, 5, 3, and 8, spaced roughly equally.

And yet a Box–Jenkins model with only a single AR1 parameter and *without controls for the breaks* estimates $\phi = 0.91$. The structural breaks do not fool our estimation by much and it is still clear that the series is not *I(1)*. If we account for the breaks, the estimate of ϕ is a more accurate 0.82, but then again, when we inserted four *randomly generated* interventions to control for breaks that are not even there, the estimate of memory is reduced and the ϕ estimate is a bullseye 0.85.

For all of our electoral support series, resorting to autoregressive integrated moving average (ARIMA) techniques, the AR1 estimates are uniformly $\phi = 1$. Clearly, electoral support is nonstationary. Hence, the more central question is not the one Pickup is most concerned with – i.e., whether $d > 0$ – but whether $d < 1$. Our significance tests still show, with little challenge from Pickup, that in most cases $d < 1$, and significantly so.

Should *d* be Estimated only after Controlling for Structural Breaks?

It is true that estimating d when controlling for supposed structural breaks such as leadership changes, as Pickup suggests, results in new estimates which, while not

Downloaded By: [Young, Everett] At: 20:41 15 March 2009

Table 1. Estimates of *d* under 3 assumptions: (a) leadership changes as just another kind of shock; (b) leadership changes as structural breaks; (c) entirely randomly generated events as structural breaks

Party and period (pd 1 = Conservative gov't, pd 2 = Labour)	Original estimate of *d*, assuming leadership change as shock	Estimate of *d* after controlling for leadership changes as "structural breaks"	Estimate of *d* after controlling for equal number of randomly placed "structural breaks"
Conservative, pd1	0.90	0.85	0.82
Labour, pd 1	0.85	0.67	0.77
Liberals, pd 1	0.79	0.60	0.65
Conservative, pd 2	0.55	0.52	0.55
Labour, pd 2	0.51	0.56	0.57
Liberals, pd 2	0.49	0.41	0.37

Downloaded By: [Young, Everett] At: 20:41 15 March 2009

substantially different than the originals, are predictably lower, as shown in Table 1. Even with these tiny differences, which estimates are appropriate, the old or the new? We make no concession that the new, post-control estimates are superior because we do not concede that large shocks to electoral support should be thought of as equilibrium shifts.

Yet Pickup accepts that structural breaks are ubiquitous in political science. This assertion may at first sound reasonable, but it is theoretically unsupported. For a public opinion variable, the acceptance that what appears to be a memoried shock must actually be a structural break requires quite a gigantic assumption. To illustrate, suppose a new party leader is elected and support for his party soars (or, for another party, drops) by 10%. To call this a "break", we must accept that there exists some natural, inherent, mean level of party support associated with having this particular leader at the party's head, and that this level of support is *for all time* the level of support associated with that leader. It would not be enough to claim that the excitement of electing a new leader is the reason for the ten-point jump. We must assume a permanent shift, flying in the face of a long history of scholarship on the determinants of vote choice. Underlying factors such as party identification, class, socio-economic status, and gender are not so easily and so permanently pushed aside with leadership changes. As our article demonstrates, leadership matters. But to say they are structural breaks requires a much stronger assumption that we are not willing to make.

Further, if events other than leadership changes cause structural shifts – and Pickup accepts that such events are also common – the problem grows. How are we to differentiate an equilibrium-shifting event from a major shock? Is a sudden 2% increase in unemployment equilibrium-shifting? What if it is only 1%?

Pickup's perspective strikes us as open to unscientific decisions about what constitutes an equilibrium shift. The lack of rigor in such an approach is illustrated

in the third column of Table 1. In addition to estimating d under the assumptions that (a) leadership changes are shocks (i.e., not controlling for them first) and (b) that leadership changes are structural breaks, we also estimated d by controlling for imaginary structural breaks. That is, we chose time points randomly and estimated d for our electoral support series as though we suspected structural breaks would be found there. The d estimates were reduced about as much as they were when controlling for "real" events such as leadership changes. Simply adding more parameters reduces estimates of d, but that does not mean that structural breaks are present or that the parameters are appropriate.

Additional Concerns: Maximum-likelihood Estimators and the I(1) Possibility

Entirely aside from structural breaks, Pickup suggests that we would be better off using a maximum-likelihood estimator for d, which can simultaneously estimate an AR1 parameter and d. The suggestion is that the data generating process for party support could be both AR1 and fractionally integrated. If this is the case, our estimate of d is off the mark, because a good deal of the memory accounted for by our d estimates is in fact an AR1 process.

With or without a maximum likelihood estimator (MLE) it is possible to estimate a (p,d,q) model and several researchers using FI have done so (e.g. Box-Steffensmeier & Smith, 1998; Lebo *et al.*, 2000). We attempted this on Tory and Labour electoral support for the first period and found that *any* incorporation of an AR1 parameter in the presence of d was unfounded and that forcing a fractionally differenced series to be filtered by values of ϕ ranging between 0.1 and 0.9 results in an overdifferenced series. That is, after estimating ϕ and d and filtering the original series, still further AR or MA parameters are required to correct the problems of adding the additional parameter. In sum, the only way we were able to achieve a white-noise dependent variable without multiple ARMA parameters was fractional differencing alone.

Finally, we briefly address the possibility less central to Pickup's argument, that party support might be a combined process or some other asymptotically *I(1)* process. Wlezien's paper (2000) on combined series is addressed primarily to series which are a combination of exactly two other series where one of them is *I(1)*. As Wlezien explains, it is difficult in finite series to tell that the combination is *I(1)*, but we know from statistical theory that it is. Yet public opinion is not a combination of merely two series, nor is it the combination of only a few. "Combined series", however, can also result from shocks that are identical across individuals, but decay heterogeneously. This is closer to the data-generating process of public opinion and appears to be a special case of FI. But again, in its homogeneity assumptions, it seems an unrealistic model of public opinion.

Further, we do not think public opinion is asymptotically *I(1)*, even if for certain periods it might appear to wander wildly, because this assumes that party support or opposition can increase without bound. While this is obviously technically impossible for aggregated electoral support series – more than 100% of a polity cannot

Downloaded By: [Young, Everett] At: 20:41 15 March 2009

support a party – we do not think this is merely a case of a measurement ceiling on an unbounded latent variable. Politics is a field of conflict: when one party becomes very popular, other parties adjust their strategies to retake lost ground, and polities are not prone to unbounded agreement. There *are* some mean-reversion-type forces acting on party support, preventing support for any one party from reaching "escape velocity". But this does not mean party support regresses to a *particular* mean. It stays within reasonable bounds, but walks around considerably. It is, in other words, fractionally integrated.

Conclusion

Substantive theory matters. True, AR series with breaks can appear to be FI. But when *theory* suggests FI *and* FI bests ARMA techniques for statistical purposes, resorting to ARMA in the pursuit of "simplicity" is not a retreat to safety – it is just inaccurate modeling. Neither is estimating *d* under the casual assumption of dubious "structural breaks" an act of statistical prudence; rather, it is an invitation to bias. Theory suggests vote intent is FI; statistically, FI appears to be the best strategy. We continue to believe our original strategy, then, would have been a superior approach.

Downloaded By: [Young, Everett] At: 20:41 15 March 2009

Notes

1. That is, we are looking at the ACF of the residuals of a Box–Jenkins model that includes an AR1 and several interventions.
2. See, for example, the online appendix of Lebo and Dickinson (2006), available at <http://ms.cc.sunysb.edu/~mlebo/details.htm>.

References

Box-Steffensmeier, Janet M. & Smith, Renee M. (1996) The dynamics of aggregate partisanship, *American Political Science Review,* 90, pp. 567–580.

Granger, Clive W. J. & Hyung, Namwon (2004) Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns, *Journal of Empirical Finance,* 11, pp. 399–421.

Lebo, Matthew J., Walker, Robert W. & Clarke, Harold D. (2000) You must remember this: dealing with long memory in political analyses, *Electoral Studies,* 19, pp. 31–48.

Robinson, Peter M. (1995) Gaussian semiparametric estimation of long range dependence, *Annals of Statistics,* 23, pp. 1630–1661.

Wlezien, Christopher (2000) An essay on "combined" time series processes, *Electoral Studies,* 19, pp. 77–93.