Modeling Dynamics in the Study of Conflict: A Comment on Oneal and Russett

Nathaniel Beck
Department of Political Science
University of California, San Diego
La Jolla, CA 9209
USA[1]

Draft of April 17, 2001.

[1]I wish to thank the always collegial John Oneal and Bruce Russett for providing their data, and, in particular, John Oneal for answering all my queries and meeting all my requests. Chris Zorn also provided very helpful comments on GEE, and my basic approach owes much to my constant coauthors, Jonathan Katz and Richard Tucker, though they bear no blame for anything said here.

Introduction

Oneal and Russett (hereinafter "OR") offer some new analyses on the role of trade in reducing conflict. These new analyses are largely a response to the work of Katherine Barbieri, but there is also some discussion about the best way to model dynamics in the dyad-year analysis of international conflict (that is, with a binary dependent variable). This discussion contrasts the event history methods (PEACEYRS) that I have proposed (Beck, Katz and Tucker, 1998) with the "generalized estimating equation" (GEE) of Liang and Zeger (1986). OR find stronger results for the pacific effect of trade using the GEE correction than using the PEACEYRS correction. The claim of this response is that the PEACEYRS correction is appropriate for the OR data whereas the GEE correction *alone* is not. But it is also the case that we may be able to improve on the PEACEYRS correction by combining it with the GEE correction. The substantive consequence of this is that trade has a pacific effect for all dyads (using imputed data when trade data is missing), but it just misses having a statistically significant pacific effect for "politically relevant dyads" (PRDs), whether we analyze the onset of militarized disputes or all dispute years. More importantly, the substantive effect of trade on peace is cut in half when we include the PEACEYRS correction (whether we use GEE or not).

In this brief reply I discuss why the PEACEYRS event history approach is superior to the GEE approach for the OR dyad year data. The GEE approach has a fine pedigree and is a useful approach for the appropriate data; but, for the OR problem, the GEE approach does little and the event history approach appears reasonable.[1] These issues are orthogonal to the controversy between Barbieri and OR, and so I need not take any position on that controversy. For convenience, I work only with the specifications preferred by OR.

Since this paper focuses on my differences with OR, I should stress our commonalities. We agree that trade does not increase conflict. For all dyads, we agree that it does decrease conflict, though we disagree on the quantitative assessment of that impact. For PRDs, we both find that trade decreases conflict, but my analyses indicate that this result is (barely) not statistically significant.[2] If we ignore issues of statistical significance, the GEE estimates imply that the impact of trade on decreased conflict for PRDs is at least twice as large as implied by the PEACEYRS estimators (for the onset of conflict) and three times larger when all dispute years are analyzed. These differences are not trivial. I stress that OR report results based both on the GEE and PEACEYRS analyses, and do not commit to the former; therefore the differences between OR and myself are that the value of their stronger findings on the impact of trade on conflict.

The issues raised in this reply go beyond the study of conflict and trade. As we as a discipline become more methodologically sophisticated, and as programs such as Stata

[1] There is no claim that it is optimal, only that it provides more accurate results than does the GEE for this data set.

[2] OR and I agree that one should use robust (Huber) standard errors which allow for all observations in the same dyad to be correlated. I only present non-robust standard errors in Table 1 to show the relative impact of GEE versus using robust standard errors. All substantive conclusions in this reply are based on the robust standard errors; OR report only such errors.

make sophisticated analysis easier, we see more attention paid to "details" such as temporal dependence in binary models. As we see in the current case, the particular method chosen to deal with these "details" can make a big difference. It is therefore important that we begin to think about which of these technical treatments is appropriate for any particular data set. In some cases this will be GEE; in other cases it will be PEACEYRS and in other cases neither approach is correct. But we must look at methods as more than black boxes, and think about the models which underlay the various methods. For the OR analysis of conflict and trade, we shall see that the PEACEYRS method is more reasonable than the GEE method.

From a very practical perspective, if the PEACEYRS correction is correct (that is, either the data show duration dependence or that dependence is not completely modeled by the GEE), then omitting the *PEACEYRS* variables leads to omitted variable bias. But it is easy enough to test whether the *PEACEYRS* variables are needed in any specification via a standard hypothesis test (of the null hypothesis that all the coefficients on the *PEACEYRS* variable are zero). This test is trivial to perform (whether or not one does ordinary logit or GEE). If scholars incorrectly omit the *PEACEYRS* variables, results based on such analyses are incorrect (biased). As we shall see, this is the case for the OR GEE analyses.

GEE or event history for conflict data

The event history (PEACEYRS) method I prefer has its underpinnings spelled out in Beck, Katz and Tucker (1998). But the basic idea behind it is extremely simple and intuitive: the information in the dyad year conflict data sets is exactly the time between conflicts. While we often see event history modelers examine the time between events, it is uncontroversial that we can model event history data as a sequence of conditional binary dependent variable (logit or probit) models (Sueyoshi, 1995).

In my work I have chosen to model the event histories of peace as the discrete time (logit) version of the most common model used by event history analysts, the Cox (1975) semi-parametric proportional hazards model. But while this model is flexible, it of course makes assumptions: proportionality of the hazards (Box-Steffensmeier and Zorn, 2001) or homogeneity of the data (Beck and Katz, 2001). Thus it cannot be claimed that the PEACEYRS approach is correct, but it can be claimed that it is derived from a very flexible and very commonly used event history model. My preferred approach requires no new statistical work, and is very intuitive.

This is not to claim that the event history approach is appropriate for all binary longitudinal data. There are surely many data sets where the time between events (ones) is not the best way to conceive of the information in the data set. The prototypical biostatistical data set of Liang and Zeger, based on whether individuals take a pill on a sequence of days, is probably not appropriate for the event history approach, since it is not likely that the best way to think of the information in this data being summarized by the number of consecutive days between taking a pill But the various theories of the cause of peace and conflict, when studied in a binary dyad year context, seem to me to be exactly about why some dyads

remain at peace longer than other dyads.

What of OR's [MS p. 10, cite however fits the volume] claim that "[i]t is no more surprising, then, that the statistical significance of interdependence [trade] declines in the presence of the PEACEYRS variables ... than that a measure of inflation will be less significant in a regression that also includes a control for the money supply[?]" A more correct analogy would be "it is no more surprising that interdependence declines in the presence of the PEACEYRS variable than that controlling for extensive serial correlation of the errors causes many independent variables to lose statistical significance in time series analysis." Omitting the PEACEYRS variable, as was noted in Beck, Katz and Tucker (1998) is equivalent to assuming that data which show duration dependence can be modeled ignoring such dependence. We know that ignoring duration dependence leads to non-optimal estimates and incorrect standard errors. Including PEACEYRS in the logit is like correctly modeling the error term in a time series analysis; such correct modeling is not optional, and is not on the same level as deciding which independent variables to include in a regression. Ignoring duration dependence, when the data indicates it is present, must lead to incorrect estimates. I return to the issue of whether PEACEYRS needs to be in the specific OR specifications in the next section.

Does the Liang and Zeger GEE solve OR's problem of modeling dependent data without introducing the troublesome PEACEYRS variable. The GEE is (or was) new statistical work designed to deal with panel data with a binary (or more complicated) dependent variable (see Zorn, 2001 for a good discussion of GEE relevant to political science). It is a very general fix-up, and sees the dependence of observations as a nuisance which has its primary impact on the estimated standard errors. It is almost certainly the case that using the GEE must be better than ignoring the issue of interdependent observations, but no one here is claiming that ignoring interdependence is a good thing.

The first problem with GEE is that it is a bit of black box. Unlike my preferred event history approach, it is a bit hard to tell exactly what GEE is doing to the OR data. As but one example of this, OR never even tell us what the estimated correlation between the observations is; in these analyses they view the temporal aspects of the data simply as a nuisance which impedes estimation, rather than a feature of the data which is to be modeled. To put this more technically, GEE is a "quasi-maximum likelihood approach" which takes the ordinary logit model for the data but then "fixes" the various equations which are used to estimate the model parameters. Thus, in using the GEE estimates for prediction purposes, one would simply use the standard logit prediction formulae, *ignoring any temporal dependence in the data*. GEE does not provide a model for temporal dependence; such dependence simply enters the black box used in the maximization.

But even if we feel we understand the GEE, are its assumptions relevant to the OR data? The GEE assumption (which only enters the maximization routines) is that the disputes follow a first order autocorrelation (AR1) pattern, that is, the correlation of $y_{i,t}$ and $y_{i,t-1}$ is ρ. The correlation of adjacent observations is identical for all observations. This means that if we observe a dyad in conflict last year, the correlation with this year's conflict is the same as if the dyad was at peace last year; similarly, the correlation between a dyadic observation in year twenty of peace and year twenty-one of peace is the same as that between year one

and year two of peace (and a year of dispute followed by a year of peace). This is very different from the event history findings (Beck, Katz and Tucker, 1998), which show that disputes are more likely to break out after a very short spell of peace than after a long spell of peace.

The simple AR1 pattern is particularly odd for the OR *ONSET1* measure. Here, subsequent years of a dispute are dropped from the data set. Thus the conditional probability of a dispute following a dispute is very low by the very nature of the way in which the data is constructed. Given that disputes are rare, the conditional probability of peace following either a dispute or peace is high. But the GEE assumes that the correlation between adjacent temporal observations is the same, regardless if the prior observation was one of peace of war. AR1 models may be perfectly appropriate for economic time series, where good years follow good and bad years follow bad, but such a model seems completely inadequate to capture the dynamics of the OR dispute data (whether *ONSET1* or *DISPUTE*, which allows all dispute years into the analysis).

How does one tell which approach to use in practice? To see if the PEACEYRS approach is necessary, one simply adds the *PEACEYRS* (spline) variables to the specification and then tests the null hypothesis that those variables do not belong in the specification, using an ordinary χ^2 test? Thus one does not need to resort to a purely theoretical argument as to whether the event history approach is necessary; the data indicates whether it is. If we do not reject the null hypothesis that the *PEACEYRS* variable belongs in the specification, we can then do ordinary logit (or its GEE analogue).

Note also that one can do both GEE and the PEACEYRS correction. There is no theoretical claim that the event history approach completely eliminates any other temporal relations in the data. Thus one can simply add the *PEACEYRS* variables to the GEE specification, of course after testing whether this is needed. I now turn to an analysis of the OR results.

Some Analyses

To focus attention on the difference between the GEE and event history approaches, I work with the exact specifications[3] as used by OR to generate their results in Tables II (all dyads) and III.[4] For reasons of space, I show results only for the critical OR variable, the ratio of

[3]Thus I do not examine other alternatives based on event history ideas, such as modeling second disputes differently from first disputes; nor do I examine alternative choices of independent variables. Given that the OR specification with PEACEYRS for PRDs just misses statistical significance, I assume that some specification would produce marginally statistically significant results. One obviously should not make much fuss over whether P values are .045 or .055.

[4]The analyses based on Table III are limited to PRDs only, rather than all dyads broken down into contiguous dyads, major power dyads and non-relevant dyads. Since OR find no statistically significant difference between major power and contiguous dyads (the two subtypes of PRDs), and no effect of trade in non-PRDs, there is no loss in simply analyzing all PRDs; in any event, my results for PRDs are very similar to their results.

dyadic trade to GDP for the less trade dependent partner (denoted *TRADE* here).[5]

It should be noted that all the OR results use robust corrected standard errors, which allow for dyadic observations to be correlated in some unspecified way. There is no disagreement that this is a useful approach for this type of data. However, to see what GEE and PEACEYRS are doing to the data, I first present the results of an ordinary logit analysis which ignores dynamics, and then show the effect of each of the "fixes" with and without the impact of using robust standard errors. (Each analysis shows the usual maximum likelihood standard errors ("ML") and the robust standard errors ("Rob.") which allows for observations on the same dyad to be dependent.) Results are in Tables 1 corresponding to the three OR analyses.

Table 1: The effect of trade on conflict

	All Dyads			PRDs Only					
	ONSET1			*ONSET1*			*DISPUTE*		
		SE			SE			SE	
Correction	$\widehat{\beta}$	ML	Rob.	$\widehat{\beta}$	ML	Rob.	$\widehat{\beta}$	ML	Rob.
None	−118.9	18.6	45.9	−108.8	17.6	52.2	−168.0	20.2	73.4
GEE	−115.7	21.2	42.9	−105.4	20.9	48.3	−169.4	20.9	73.6
ρ (GEE)	.15			.20			.03		
PEACEYRS	−51.9	14.9	23.4	−28.7	13.0	22.3	−32.6	14.4	28.0
Test of PEACEYRS[a]	337.2			328.3			587.3		
PEACEYRS and GEE	−60.2	16.5	24.9	−42.2	15.9	26.8	−63.3	19.7	39.1
ρ (GEE)	.05			.08			.08		
Test of PEACEYRS[a]	246.6			210.6			420.6		

[a] χ^2_4 statistic on H_0:All PEACEYRS $\beta = 0$, using robust variance estimates

The results are easy to summarize. GEE results (without PEACEYRS) are essentially identical to ordinary logit results, both in estimated coefficients and standard errors. The only difference between the GEE results reported by OR and their earlier results (Oneal and Russett, 1997) based on ordinary logit (beyond some changes in the data and slight changes in the specification) is because the current GEE results also report robust standard errors. GEE by itself has almost no effect on either ordinary logit estimates or standard errors. Thus to believe the GEE results, one would have to believe that there is trivial temporal dependence in the data. With the data displaying long sequences of peace, such a belief is hard to maintain. Note that GEE estimates the correlation of adjacent observations

[5] When comparable, my results are identical to those provided by OR. A complete file of my statistical results is available on my web site (`http:\\weber.ucsd.edu\~nbeck`). The data used way kindly made available by John Oneal and may be obtained from him.

as between .15 and .2 for the *ONSET1* variable and .03 for the *DISPUTE* variable. Such trivial estimated correlation obviously has only a trivial effect on coefficient estimates and standard errors.[6] The largest difference between the two types of analyses is for the *DISPUTE* dependent variable; as argued in Beck, Katz and Tucker (1998), this dependent variable is more vulnerable to duration dependence than is the *ONSET1* variable.[7]

What about the notion that the *PEACEYRS* and *TRADE* variables are highly similar, and so inclusion of both makes it impossible to see the true effect of either. If we simply regress *TRADE* on the four components of the *PEACEYRS* spline, the R^2 for all dyads is .005 and for PRDs .015; looking at all dispute years, this figure increases to an enormous .016. In other words, *TRADE* is not very well predicted by the temporal spline. The large impact of controlling for duration dependence in the *DISPUTE* analysis cannot be because *TRADE* and the *PEACEYRS* spline are essentially the same variable; they are clearly not.

Finally, I note that if we combine the event history and GEE approaches, by adding the *PEACEYRS* variable to the OR GEE specification, the results are slightly more favorable to the liberal peace argument for the onset of disputes, and considerably more favorable for the less preferred *DISPUTES* variable (though the effect of adding the *PEACEYRS* variables still reduces OR's estimated impact of trade by about two thirds). Combining methods leaves the pacific impact of trade for PRDs just barely statistically insignificant, and slightly increases our estimate of the pacific impact of trade for all dyads.

Conclusion

I have no problem with any of the OR analyses using the event history (*PEACEYRS*) method. These indicate that trade has a significant pacific impact if we look at all dyads, but no significant pacific impact if we look only at PRDs (whether all dispute years or only the onset of disputes). There is clearly no evidence in these analyses indicating that trade is conflictual.

But the GEE results, which show a much stronger pacific impact of trade, are not correct. The GEE model of the interdependence of the data is not consistent with the temporal interdependence manifested in the conflict dyad-year data. In essence, the GEE results are identical to the ordinary logit results which ignore temporal dependence. If one accepts the argument made in Beck, Katz and Tucker (1998) that the original findings of Oneal and Russett (1997) that trade has pacific effects are flawed, then the GEE results reported here by OR here must be equally flawed.

[6] The χ^2 tests of whether the *PEACEYRS* variable belong in the specification indicate that we can clearly reject the null hypothesis that they do not belong, whether or not we also use a GEE correction. The reported χ^2 statistics are quite large, indicating that the *PEACEYRS* variables are important predictors of disputes.

[7] Since I wish to focus on the GEE, I only respond briefly to OR's [manuscript, p .12] claim that "researchers should be concerned with all years in which states are involved in conflict." While it is hard to disagree with that sentiment, OR's model of conflict assumes that the process which determines the duration of peace is identical to that which determines the length of war. Dynamic models of transitions (Diggle, Liang and Zeger, 1994), however, always begin with the assumption that these two processes might be different.

This does not mean that GEE is never useful, nor that the event history method preferred here is always right. As we have seen, GEE combines quite easily with the event history approach. There will clearly be cases where the GEE approach is preferred to the event history method. But the OR analysis of trade and conflict is not one of those cases. Their reports, based on only the GEE correction, do not adequately represent the impact of trade on conflict. Their more accurate reports, and related reports presented here, show a much more modest pacific impact of trade.

References

Beck, Nathaniel and Jonathan N. Katz. 2001. "Throwing Out the Baby With the Bath Water: A Comment on Green, Kim and Yoon." *International Organizations*. 55:487–95.

Beck, Nathaniel, Jonathan N. Katz and Richard Tucker. 1998. "Taking Time Seriously: Time-Series–Cross-Section Analysis with a Binary Dependent Variable." *American Journal of Political Science*. 42(4):1260–88.

Box-Steffensmeier, Janet M. and Christopher J. W. Zorn. 2001. "Duration Models and Proportional Hazards in Political Science." *American Journal of Political Science*. 45(3).

Cox, D. R. 1975. "Partial Likelihood." *Biometrika*. 62(2):269–76.

Diggle, Peter J., Kung-Yee Liang and Scott L. Zeger. 1994. *Analysis of Longitudinal Data*. Oxford: Oxford University Press.

Liang, Kung-Yee and Scott L. Zeger. 1986. "Longitudinal Data Analysis Using Generalized Linear Models." *Biometrika*. 73:13–22.

Oneal, John R. and Bruce M. Russett. 1997. "The Classical Liberals Were Right: Democracy, Interdependence, and Conflict, 1950–1985." *International Studies Quarterly*. 41(2):267–94.

Sueyoshi, Glenn T. 1995. "A Class of Binary Response Models for Grouped Duration Data." *Journal of Applied Econometrics*. 10(4):411–31.

Zorn, Christopher J. W. 2001. "Generalized Estimating Equation Models for Correlated Data: A Review With Applications." *American Journal of Political Science*. 45(2):479–90.