Using repeated measures data to analyse reciprocal effects: the case of Economic Perceptions and Economic Values

Patrick Sturgis, Department of Sociology, University of Surrey

Peter Smith, Ann Berrington, Yongjean Hu, Department of Social Statistics, University of Southampton

Disentangling causal relationships between abstract concepts is the 'holy grail' of empirical social science. The enterprise is, however, stymied by the seemingly intractable limitations of observational studies which constitute the data source for the vast majority of social scientific research programmes. Without random allocation to 'treatment' and control groups, we are simply unable to discount alternative models that might account equally well for observed covariances between dependent and independent variables, yet do not posit a causal relationship between them. The primary limitations to valid causal inference with observational data relate to the issues of unobserved variable bias and reciprocal causality. The former pertains when X and Y are conditionally independent, given a third variable (or vector of variables), Z, which has not been included in the model. The latter arises when both X and Y exert a direct causal influence on each other over time.

In this paper, we fit a series of models which attempt to estimate the reciprocal relationship between retrospective perceptions of the macro economy (EP) and 'left-right' economic values (EV). EP is measured by three five point items tapping 'past-year' perceptions of unemployment rate, inflation rate and the general standard of living. EV is measured by the six item scale developed by Heath et al (1993). We hypothesise a possible reciprocal relationship between these variables on the grounds that economic perceptions have been shown to be partially endogenous to political preferences (Evans 1997) and because it is likely that beliefs about how the economy should be managed will be influenced by macro-economic performance.

The models we present are deliberately selected to demonstrate the importance of two issues in model specification: correcting for stochastic and systematic errors in the measurement of concepts, and specifying an appropriate covariance structure between error terms in both the measurement and structural components of a model. Measurement error in independent variables results in attenuated effect sizes (Bollen 1989). Failing to adequately account for covariances between error terms can bias estimates of auto-correlation and cross-lag parameters (Williams and Podsakoff 1989). This is not an esoteric statistical point; failing to account for these errors and their covariance structures can lead to seriously flawed causal inference. Our data source for all analyses is the 1992-1997 British Election Panel Study, with an analytical sample size of 1640. Table 1 presents a summary of the cross-lag and stability parameters estimated in each model.
We begin by fitting a 'naïve' model (model 1), which simply regresses the 1992 cross-sectional summed-scale measure of EP on to the 1992 cross-sectional summed-scale measure of EV and find a significant positive relationship. Although this is clearly not a suitable way to address issues of causal reciprocity, it is a frequently used model amongst substantive social researchers. We see that the standardised estimate of the causal influence of EP on EV is the same as that of EV on EP and that - if a reciprocal relationship exists - both estimates will be biased. Next we employ instrumental variables in a Structural Equation Model (SEM) to estimate the non-recursive relationship between EV and EP (Model 2). While the instruments enable identification of the reciprocal paths – both positive and significant – the model requires strict assumptions about the covariances of the instruments with each other and the reciprocally related variables for these estimates to be unbiased. These assumptions are not met in the current instance and it would be difficult, in most cases, to envisage circumstances in which adequate instruments could be found.

It is clear, then that, for both statistical and theoretical reasons, cross-sectional data do not give much leverage on questions of reciprocal causality and we must be extremely cautious when making causal inferences on the basis of this type of association. However, in the current instance, because we are able to make use of longitudinal measures of both variables, we are able to mitigate some of the problems discussed above. We use a cross-lagged panel data model (Campbell and Kenny 1999; Finkel 1995; Marsh and Yeung 1997). Where there are two variables of interest, $Y_1$ and $Y_2$, each variable at $t_2$ is regressed on both its lagged score and the lagged score of the other variable at $t_1$. Cross-lagged panel models are particularly well-suited to examining reciprocal causality because they provide an estimate of the (lagged) effect of each variable of interest on the other(s), net of autocorrelation of each variable with its lagged measurement. Cross-lagged models, therefore, tell us how much variation in $X$ at time $t_1$ is able to predict change in variable $Y$ between times $t_1$ and $t_2$, net of controls specified in the model. The limitation of such models is that they combine both the time specific and within person variance of the analytical variables and are, thus, not capable of a particularly flexible treatment of time.

Model 3 uses the summed-scale measures and is fitted to the first 2 waves of data. This model again provides support for reciprocal effects model, although the path from EP to EV is much weakened relative to Model 2. Model 4 re-specifies each concept as a latent variable and makes a correction for measurement error on the basis of the factor model. The parameter estimates show a marked increase in magnitude, as would be expected after a correction for random error in the independent variables. Model 5 allows the error terms for the same item to covary over time in order to model any systematic component in the error variance. This has the effect of reducing the magnitude of the estimates of the cross-lag and auto-correlation parameters which, again, would be expected because in the previous model
these systematic errors have been forced into the estimates of the structural parameters. Note, also, now that the path from economic perceptions to economic values becomes non-significant in Model 5. This error specification has, therefore, completely changed the nature of the causal inference we would make.

Model 6 extends Model 5 to incorporate all 5 waves of data. The unstandardised path coefficients for the stability and lagged effects are constrained to be equal across waves. This is based on the rationale that (a) these causal influences are likely to be stable over this relatively short period and because estimating each effect freely would impose a highly unlikely ‘linear annual’ time function on the causal mechanism. There is no reason to assume that knowledge and interest influence one another over neat annual cycles and imposing such a structure would likely lead to biased estimates of any underlying causal processes (Gollob & Reichardt, 1987). Although applying this equality constraint does not provide a solution to the problem of mapping discrete measurement intervals on to continuous processes, it does give a longer and, therefore, more realistic time frame over which to examine the hypothesised relationships. We can think of the cross lagged and stability coefficients from this specification then, as representing ‘average’ effects over the five year duration of the panel. Factor loadings for the same item are constrained to equality across waves to impose meaning invariance on the items as measures of political knowledge (Meredith 1993).

Table 1 Summary of Cross-lagged and Auto-correlated Parameter Estimates

<table>
<thead>
<tr>
<th>Path</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value -&gt; Value</td>
<td>-</td>
<td>-</td>
<td>.68</td>
<td>1.01</td>
<td>.97</td>
<td>.90</td>
<td>.94</td>
</tr>
<tr>
<td>Perceptions -&gt; Perceptions</td>
<td>-</td>
<td>-</td>
<td>.26</td>
<td>.58</td>
<td>.27</td>
<td>.70</td>
<td>.71</td>
</tr>
<tr>
<td>Value -&gt; Perceptions</td>
<td>.38</td>
<td>.50</td>
<td>.26</td>
<td>.53</td>
<td>.48</td>
<td>.3</td>
<td>.18</td>
</tr>
<tr>
<td>Perceptions -&gt; Value</td>
<td>.38</td>
<td>.67</td>
<td>.04</td>
<td>.12</td>
<td>n.s.</td>
<td>.1</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Extending the model over 5 waves of data again changes the nature of the causal inference; we again find support for the reciprocal effects model, with both cross-lagged paths positive and significant at the 95% level of confidence. Model 6, however, constrains all covariance paths between the disturbance terms of the endogenous variables to zero. There are good reasons to assume, though, that in many instances this assumption is unwarranted. If, for example, a third variable, Z, causes both endogenous variables Y₁ and Y₂ but Z is not included in the model, the disturbance terms of Y₁ and Y₂ will necessarily be correlated (Anderson and Williams 1992). The problem is exacerbated with repeated measures data because of the likelihood of auto-correlation between the disturbances of the lagged endogenous variables, resulting from a stable unobserved cause of the variable in question.
over time (Williams and Podsakoff 1989; Anderson and Williams 1992). Model 7 estimates the covariances between adjacent disturbance terms for the same endogenous variable and between the disturbance terms for the two endogenous variables at the same wave of measurement. This has the effect of making the causal path from EP to EV non-significant and indicates that the apparent causal effect for this parameter identified in earlier models likely resulted from variables that have not been included in the model.

In summary then, this paper demonstrates the importance of incorporating estimates of measurement error and of adequately accounting for covariance structures in both measurement and structural disturbances when trying to model a reciprocal causal relationship. Failure to do so can result in seriously biased estimates of both stability and cross-lagged parameters, resulting in flawed causal inference.

References
Meredith, W. 1993. 'Measurement invariance, factor analysis and factorial invariance.' *Psychometrika* 58:525-543.