On estimates of split-ticket voting

Ronald J. Johnston and Charles M. Pattie

In a recent paper in this journal, Cho and Gaines (2001) have presented a critique of Burden and Kimball's (1998) earlier paper on split-ticket voting at the 1998 Presidential, House and Senate elections in the United States. Their criticism has six main components, of which several relate to Burden and Kimball's use of estimates of split-ticket voting derived employing King's (1997) ecological inference (EI) procedure. They conclude that –

> …the biggest flaw lies in the dependent variable. The authors claim as their principal achievement to have developed the first-ever accurate district-level estimates of vote-splitting. However, their faith in the EI procedure is misplaced, and their findings are not truly findings at all.

In this short note, we present alternative evidence which provides circumstantial support for Burden and Kimball's claim.

EI and EM

King's EI method was developed to solve a long-established problem of ecological inference in social science data analysis, regarding estimates of the unknown cell values in a 2 x 2 matrix when the row and column sums are known not only for that matrix itself but also for the universe of sub-matrices into which it can be subdivided (as in the division of a State into Counties or the United States into Congressional Districts). On the assumption that no variables other than the row and column sums influence the cell sizes, EI produces what King claims (and purports to show with numerous examples) are accurate estimates of the unknown values for both the sub-matrices and the summed matrix.

Although King's EI method was the first to tackle the issue of estimating the cell values for the matrix and the sub-matrices, an alternative procedure has been available for at least two decades which estimates the sub-matrix cell values, given that the total matrix is known. (The first papers were Johnston and Hay, 1982, 1983; the method was used to produce estimates of split-ticket voting at the 1976 Presidential, House and Senate elections – Johnston and Hay, 1984.) This entropy maximization procedure (EM) estimates the individual cell values using three sets of constraints rather than two – the row and column totals for each sub-matrix (as in EI) **plus** the cell values for the total matrix, usually derived from a national survey. As with EI, the EM procedure assumes that there are no influences on variation in the cell values (e.g. the percentage of those who voted for the Republican candidate for President who also voted for the Republican candidate for the Senate) other than the three sets of constraints; it results in the maximum likelihood estimates of the sub-matrix cell values.

The EM procedure has been used in a variety of electoral applications over the last two decades – as in studies of spatial variations in voting by class and inter-election flow-of-the-vote in the UK (e.g. Johnston, Pattie and Allsopp, 1998), of tactical (or strategic) voting there (Johnston and Pattie, 1991), and of split-ticket voting at New

Zealand's first MMP election (Johnston and Pattie, 1999). The full procedure has been described recently, along with a test against actual data on split-ticket voting in New Zealand in 1996 which showed a very close fit between the 'real' and estimated values (Johnston and Pattie, 2000).

EM estimates of split-ticket voting in the USA: 1988

In this note, we report on a comparison of the EM and EI estimates of split-ticket voting at the 1988 US elections, using the same data set as Burden and Kimball. Our argument is that if the two procedures result in very similar estimates, then some of the force of Cho and Gaines' criticism is removed. If two estimating procedures – based on very different methodologies – result in the same pattern of spatial variation, this suggests that the results gained from analyses based on one of the outcomes are not peculiar to that estimating procedure.

To provide the third set of constraints – the national split-ticket voting matrices – we have used the 1988 American National Election Studies. For split-ticket voting in the contests for President and Representative we used data from the 1988 NES for all respondents living in a Congressional District contested by both the Republican and the Democrat parties; for split-ticket voting in the Presidential and Senate races we used the parallel Senate Election Study. The relevant matrices – using data for the two main parties only, as in Burden and Kimball's paper – are given in Table 1. Thus, for example, in the comparison of Presidential with House voting, 72.6 per cent of those who voted for Bush in the former contest also voted for a Republican candidate for the House, and 80.1 per cent of those who preferred Dukakis for the Presidency also voted for a Democrat Representative.

Figures 1-8 plot Burden and Kimball's EI estimates for each of the four cells in each of the two matrices against the EM estimates – showing very considerable agreement. Figure 1, for example, shows that – save for a small number of relative outliers – there was a very close fit between the EI and EM estimates for Republican loyalty (RR: those who voted for Bush for President and for a Republican candidate for the House); since the two are complementary, there is a similar pattern for Republican splitters (RD – voted for Bush for President and a Democrat candidate for the House: Figure 2). The fits are even better for the other two cells (DD – loyal Democrats, Figure 3; DR – Democrat splitters – Figure 4), with only two substantial outliers.

Whereas there were 354 observations in the Presidency-House analyses (i.e. 354 Districts where both the Republicans and the Democrats fielded a candidate), there were only 33 in those for the Presidency and the Senate. Figures 5 and 6 show similar close agreements between the EI and EM estimates for Republican loyalists and splitters (RR and RD, respectively). Finally, with the Democrat loyalists and splitters (DD and DR: Figures 7 and 8), however, although there is a close relationship in each case, it is not linear: for reasons that we cannot discern, the slope of the relationship between the two sets of estimates changes at approximately the 80 per cent value for Democrat loyalists.

Regression analyses for these eight relationships are given in Table 2. All four of the comparisons of the EI (the independent variable) and EM (dependent variable) estimates for Presidential and House voting have extremely high r^2 values – the first

block in the table – as do the first two of the four for Presidential and Senate voting. The smaller r^2 values in the last two regressions in the latter group relate to the Democrat loyalists and splitters which, as Figures 7 and 8 show, had stepped relationships between the two sets of estimates (and fitting curves rather than linear relationships increased the r^2 values to 0.83 and 0.85 respectively).

In sum, therefore, the EI and EM estimation procedures produced very similar outcomes for the two sets of matrices. In relative terms, the pattern of ticket-splitting across the Congressional Districts and States varies only slightly between the two procedures, which offers strong circumstantial evidence sustaining Burden and Kimball's findings regarding such spatial variations and their causes; which ever estimation procedure is used, the relative volume of ticket-splitting is very consistent across the constituent areas.

Conclusion

Much quantitative social science involves estimating unknown variables of interest from known values – because the costs of collection are too great for the needed information to be collected from primary sources (i.e. survey measures of split-ticket voting in every US Congressional District and States). All estimation procedures are based on assumptions regarding the transition from the known to the unknown, and are therefore subject to a certain amount of scepticism unless there is confirmatory evidence which validates the estimation procedure. Cho and Gaines have criticised Burden and Kimball's attempt to derive estimates of split-ticket voting in the USA because of scepticism over the method employed. We have suggested (at least) tempering that criticism, however, in that an alternative estimating procedure – based on a very different approach – has produced very similar estimates of the unknown variables, thereby suggesting either that the Burden/Kimball estimates are better than Cho/Gaines argue…or that both of the procedures suffer from fatal flaws, which produce the same outcomes. Only further methodological evaluations will determine which is the correct conclusion, but we have indicated sufficient doubt regarding Cho and Gaines outright condemnation to suggest a stay of execution on Burden and Kimball's findings.

Burden and Kimball claimed (1998, 542) that they had produced 'more accurate and efficient estimates of the direction and magnitude of ticket splitting in different states and congressional districts', which were superior to those generated by previous research which depended on either aggregate data (which led to the classic ecological inference problem) or individual-level data that 'could not be used to estimate district level parameters'. Using the EM method, we have combined aggregate and individual level data and produced estimates which are very similar to Burden and Kimball's. Given this stability of evidence, we suggest that Burden and Kimball's second claim, regarding the forces which generate these variations across districts and states, is valid – thereby implying that Cho and Gaines' conclusion regarding those findings should themselves be subject to more scepticism than they themselves imply.

Burden, B. C. and Kimball, D. C. (1998) A new approach to the study of ticket-splitting. *American Political Science Review*, 92: 533-544.

Cho, W. K. T. and Gaines, B. J. (2001) Reassessing the study of split-ticket voting. *American Political Science Review*, 95:

Johnston, R. J. and Hay, A. M. (1982) On the parameters of uniform swing in single-member constituency electoral systems. *Environment and Planning A* 14: 61-74.

Johnston, R. J. and Hay, A. M. (1983) Voter transition probability estimates: an entropy-maximizing approach. *European Journal of Political Research* 11: 93-98.

Johnston, R. J. and Hay, A. M. (1984) The geography of ticket-splitting: a preliminary study of the 1976 elections using entropy-maximizing methods. *The Professional Geographer* 36: 291-296.

Johnston, R. J. and Pattie, C. J. (1991) Tactical voting in Great Britain in 1983 and 1987: an alternative approach. *British Journal of Political Science*, 21: 95-108.

Johnston, R. J. and Pattie, C. J. (1999) Constituency campaign intensity and split-ticket voting: New Zealand's first election under MMP, 1996. *Political Science* 51: 164-181.

Johnston, R. J. and Pattie, C. J. (2000) Ecological inference and entropy-maximizing: an alternative estimation procedure for split-ticket voting. *Political Analysis*, 8: 333-345.

Johnston, R. J., Pattie, C. J. and Allsopp, J. G. (1988) *A Nation Dividing? Britain's Changing Electoral Map, 1979-1987*. London: Longman.

King, G. (1997) A Solution to the Ecological Inference Problem – Reconstructing Individual Behavior from Aggregate Data. Princeton: Princeton University Press.

Table 1. The estimated national pattern of split-ticket voting at the 1988 Presidential, House and Senate elections

President	House		Senate	
	R	D	R	D
Bush	72.6	27.4	75.0	25.0
Dukakis	19.9	80.1	33.9	66.1

Table 2. Linear regressions of EM on EI estimates of split-ticket voting, 1998

President	House	a	b	r^2
Bush	Republican	16.8	0.76	0.96
Bush	Democrat 6.8	0.76	0.96	
Dukakis	Democrat	12.0	0.84	0.97
Dukakis	Republican	4.1	0.84	0.97

President	Senate	a	b	r^2
Bush	Republican	15.2	0.71	0.95
Bush	Democrat 13.9	0.71	0.95	
Dukakis	Democrat	9.4	0.81	0.76
Dukakis	Republican	9.9	0.83	0.76

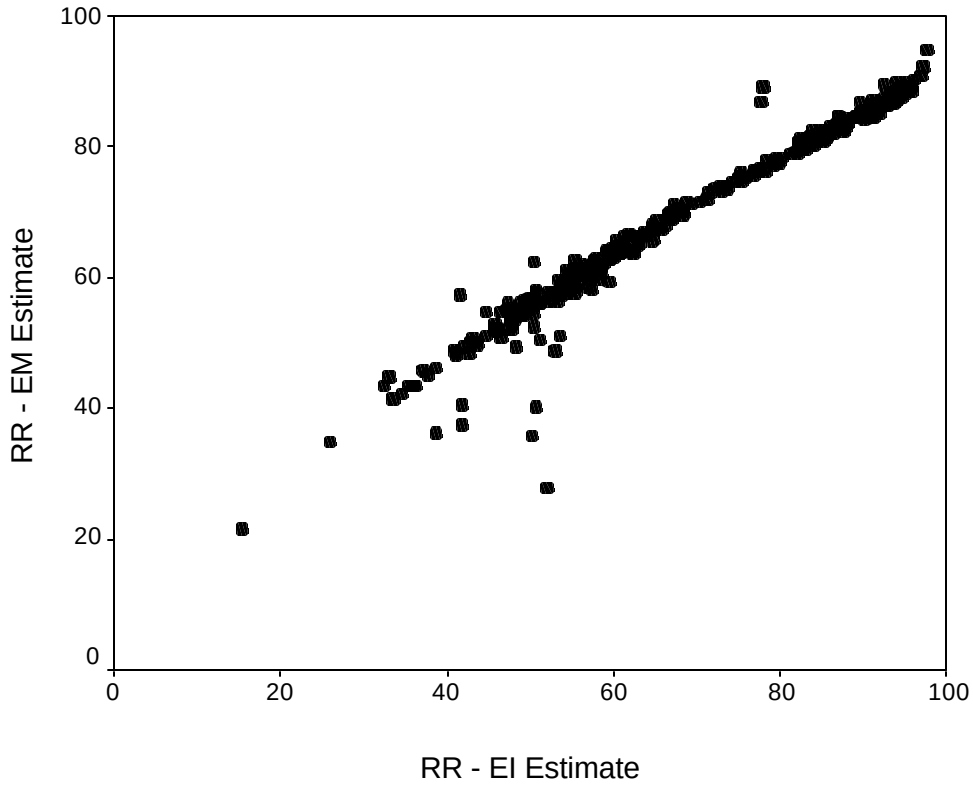

Figure 1. EI and EM estimates of Republican loyalty – 1988 Presidential and House elections

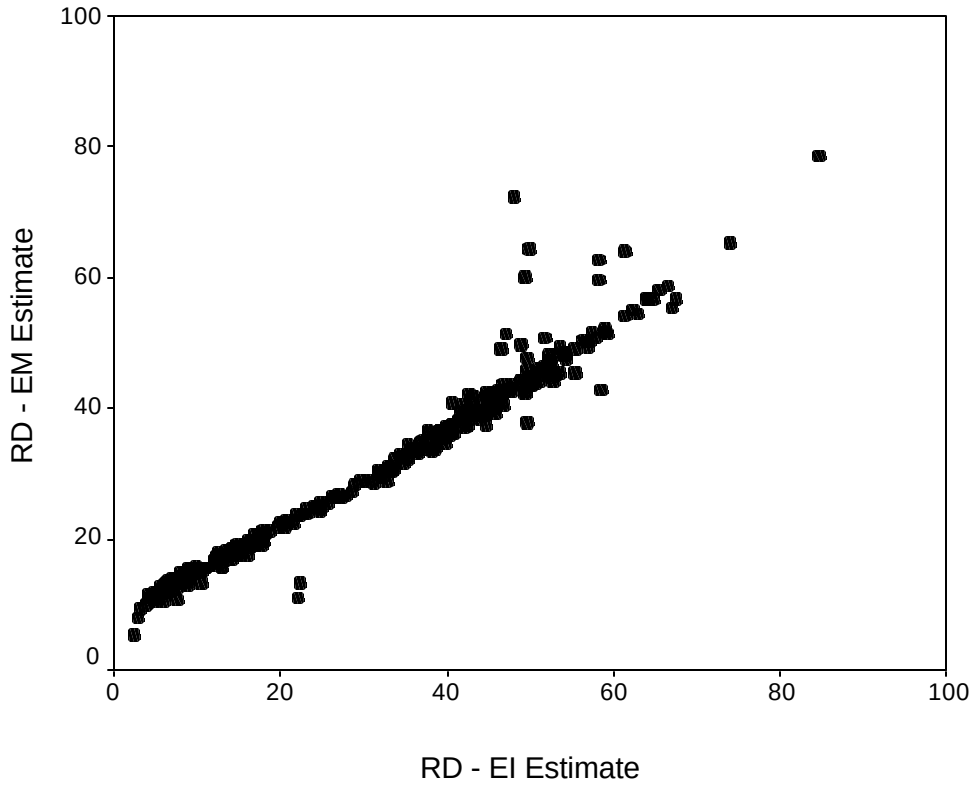

Figure 2. EI and EM estimates of Republican splitters – voted Republican for President and Democrat for House in 1988

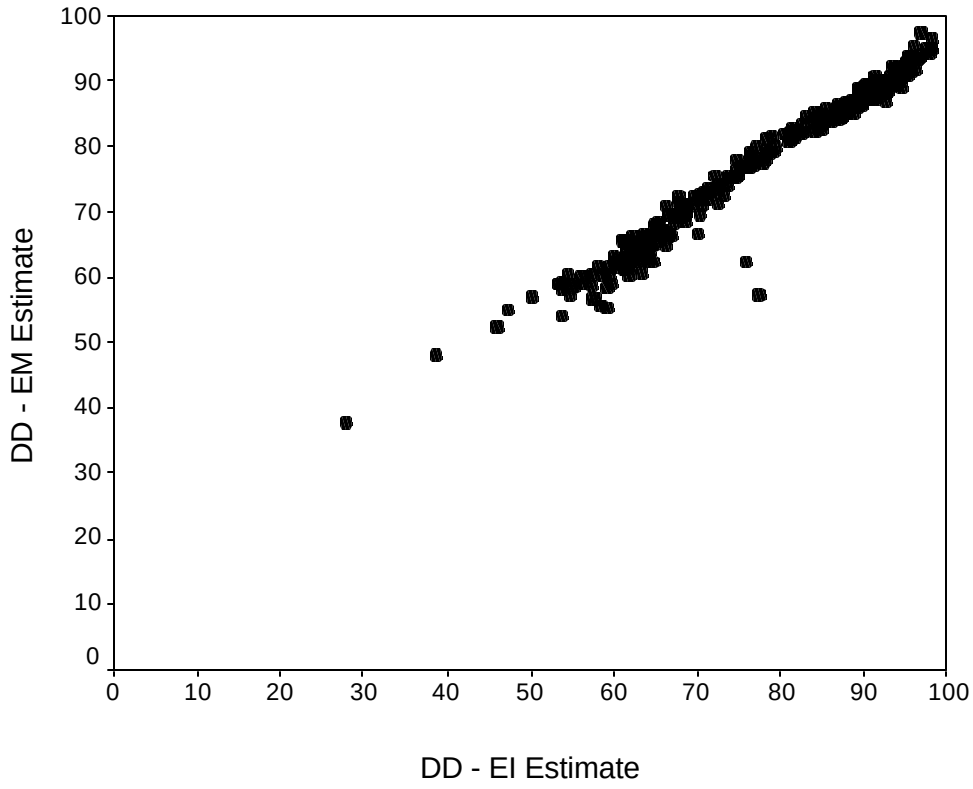

Figure 3. EI and EM estimates of Democrat loyalty – 1988 Presidential and House elections

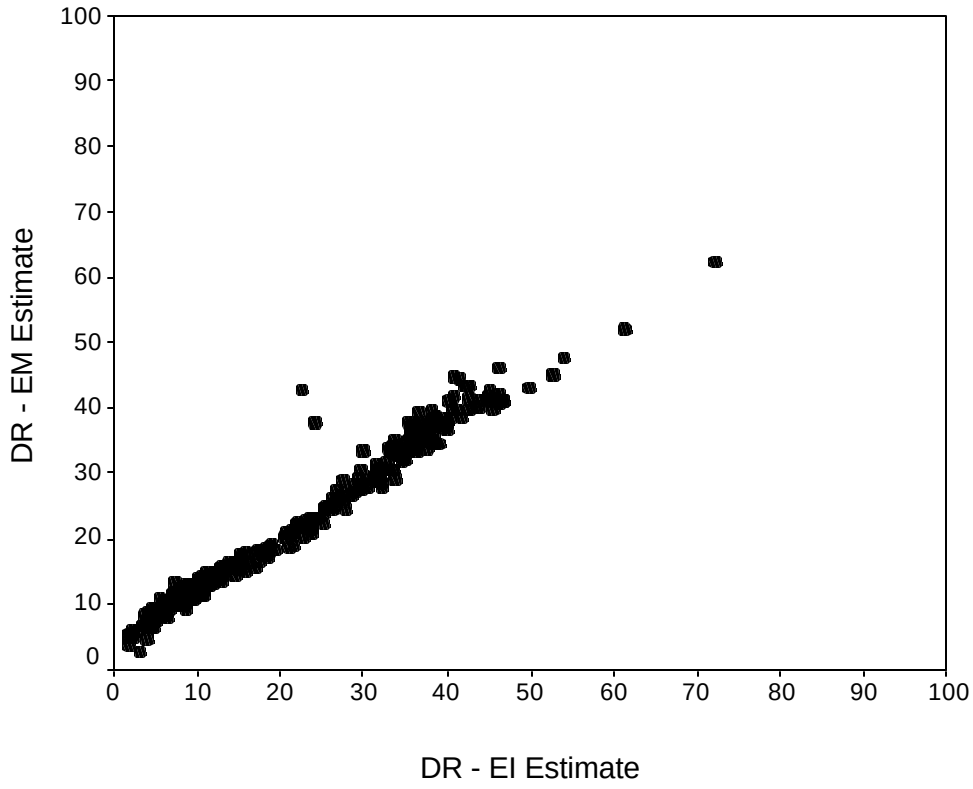

Figure 4. EI and EM estimates of Democrat splitters – voted Democrat for President and Republican for House in 1988

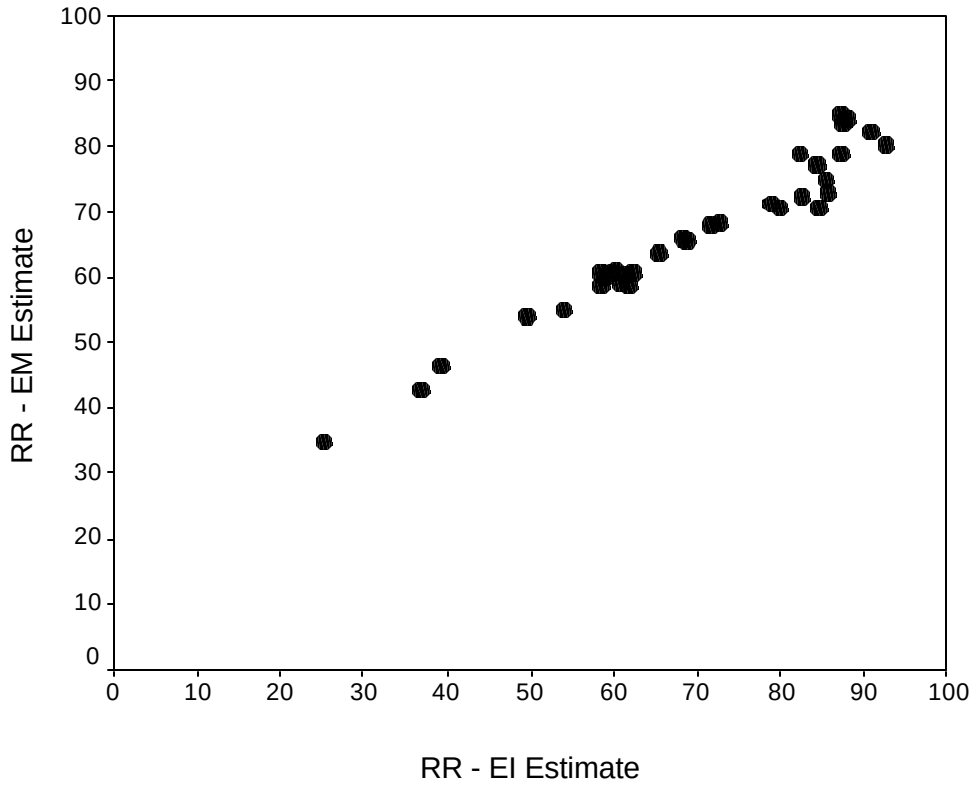

Figure 5. EI and EM estimates of Republican loyalty – 1988 Presidential and Senate elections

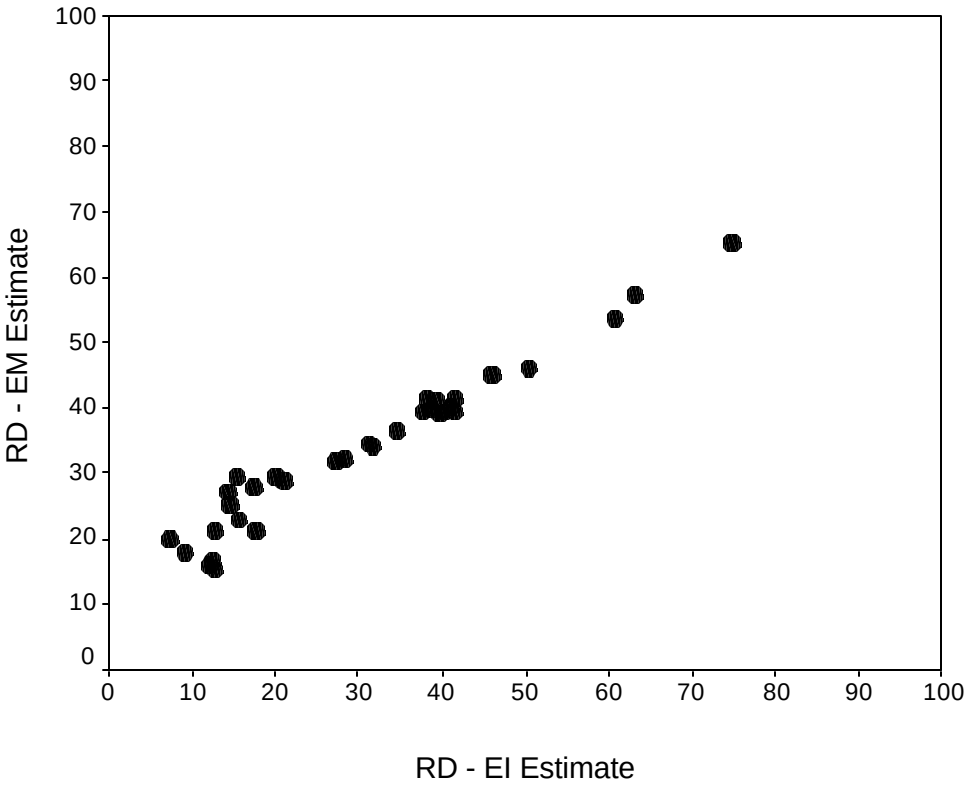

Figure 6. EI and EM estimates of Republican splitters – voted Republican for President and Democrat for Senate in 1988

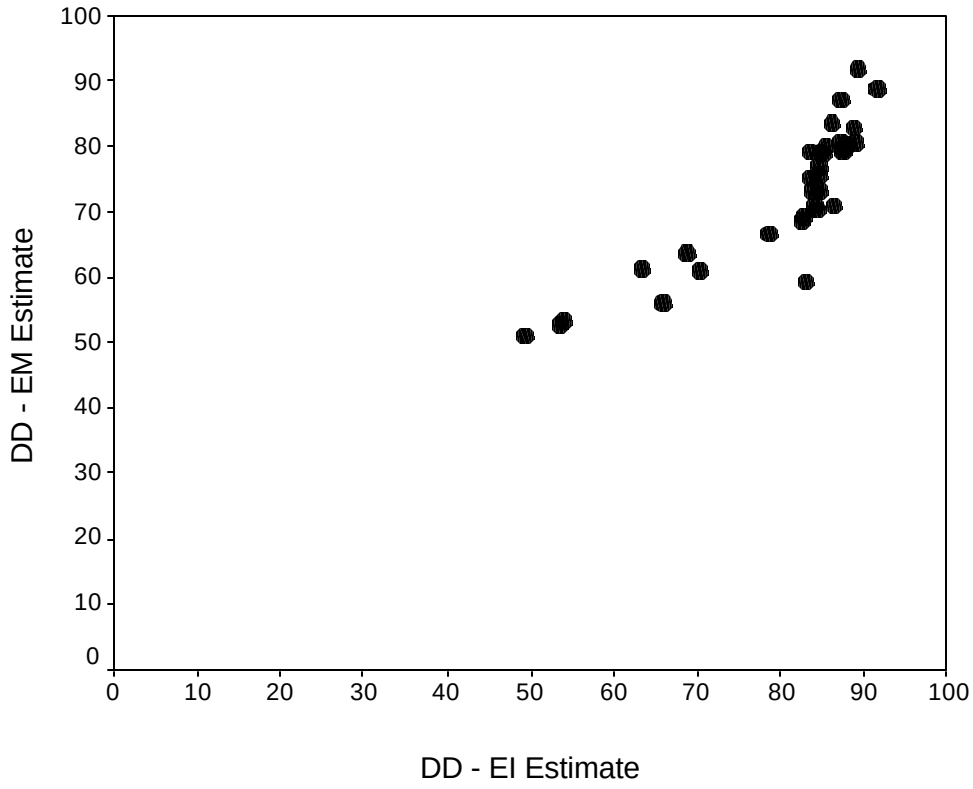

Figure 7. EI and EM estimates of Democrat loyalty – 1988 Presidential and Senate elections

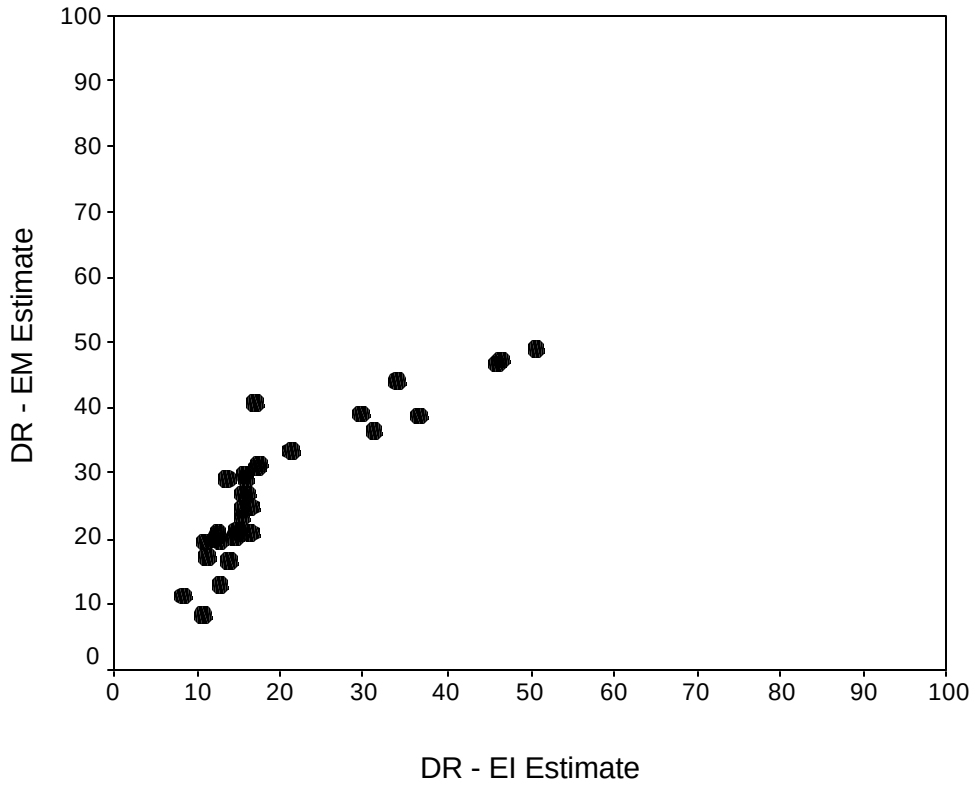

Figure 8. EI and EM estimates of Democrat splitters – voted Democrat for President and Republican for Senate in 1988