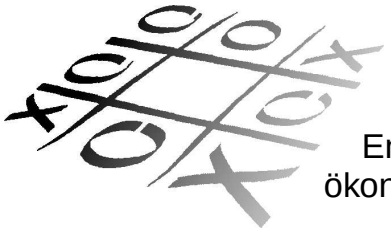

SONDERFORSCHUNGSBEREICH 504

Rationalitätskonzepte,
Entscheidungsverhalten und
ökonomische Modellierung

No. 04-40

On Estimates of Split-Ticket Voting: EI and EMax

Ron Johnston*
and Thomas Gschwend**
and Charles Pattie***

November 2004

Financial support from the Deutsche Forschungsgemeinschaft, SFB 504, at the University of Mannheim, is gratefully acknowledged.

*University of Bristol, email: r.johnston@bristol.ac.uk

**Sonderforschungsbereich 504, email: Thomas.Gschwend@mzes.uni-mannheim.de

***University of Sheffield, email: C.Pattie@sheffield.ac.uk

Universität Mannheim
L 13,15
68131 Mannheim

On Estimates of Split-Ticket Voting: EI and EMax

ABSTRACT: Cho and Gaines have recently criticized work by Burden and Kimball on split-ticket voting in the USA, suggesting that their estimates of the volume of such voting (derived using King's EI method) across Congressional Districts and States are unreliable. Using part of the Burden-Kimball data set, we report on a parallel set of estimates generated by a different procedure (EMax), which employs three rather than two sets of bounds. The results are extremely similar to Burden and Kimball's, providing strong circumstantial evidence for their conclusions regarding the impact of campaign spending and other influences on the volume of split-ticket voting.

The use of King's (1997) method of ecological inference to generate valid estimates of various statistics of interest has not gone uncriticized. Cho and Gaines (2004a) recently condemned Burden and Kimball's (1998) paper which used EI to estimate split-ticket voting and, in a review of Burden and Kimball's (2002) book on the same subject, questioned the validity of the estimates because of the 'shaky statistical foundations of EI' (Cho and Gaines, 2004b, p.243). EI does not meet their criteria for an acceptable procedure for decomposing aggregate data.

We present alternative estimates of the same patterns which sustain Burden and Kimball's, partly countering Cho and Gaines' (2004b, p. 242) claim that those estimates are 'fortuitously, roughly correct'. While accepting the latter's statement that 'an example is never tantamount to proof', our additional evidence – using an alternative estimating procedure that has been available for some two decades, including a paper in *Political Analysis* (Johnston and Pattie, 2000)[1] – provides strong support for Burden and Kimball's estimates. The volume of ticket-splitting estimated by the two procedures is highly correlated.

EI and EMax

EI addresses a long-established problem of ecological inference, generating estimates of the unknown cell values in a 2 x 2 x *n* matrix. On the assumption that no variables other than the row and column sums of each separate 2 x 2 sub-matrix influence the cell sizes, EI produces what King shows, with numerous examples, are accurate estimates of the unknown values for both the sub-matrices and the summed matrix.

EMax, also based on the method of bounds, estimates the sub-matrix cell values, given that cell values for the total matrix (i.e. summed across the *n* sub-matrices) are also known. (The first papers, with applications to electoral data, were Johnston and Hay, 1982, 1983.) It deploys 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, with the latter usually estimated from a national survey. As with EI, EMax assumes that there are no influences on the estimated cell values other than the three sets of constraints; it generates the maximum likelihood estimates of the sub-matrix cell values. The outcome is consistent with fitting a log-linear model without any three-way interaction (Johnston, 1995).

EMax is a more restrictive method for addressing ecological inference problems than EI because three sets of constraints require more prior information than two. Moreover, EI has the added bonus that every estimate comes with a measure of uncertainty.[2] Although a third set of constraints allowing EMax to be deployed might not be available for many applications, in the case of ticket-splitting patterns in the US an estimate of the national split-ticket matrix can be derived from surveys containing reports of retrospective voting behavior. Why not use more information for the estimation problem if it is readily available, therefore? As is generally accepted, better data always trump better methods. But if the better data – in this case, the third set of constraints – are not available, how good are the estimates derived from EI?

EMax Estimates of Split-Ticket Voting in the USA: 1988

We report a comparison of the EMax and EI estimates of split-ticket voting at the 1988 US elections, using the same data set as Burden and Kimball.[3] If the two procedures result in very similar estimates, then some of the force of Cho and Gaines' criticism is removed. If two sets of estimates – derived from different methodologies – result in the same pattern of spatial variation, then the results gained from analyses based on one of the outcomes are not peculiar to that estimating procedure.

The third set of constraints – the national split-ticket voting matrices – were obtained from the 1988 American National Election Studies (NES). For the contests for President and Representative we used data for all respondents living in the 354 Congressional Districts 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 in Table 1.[4]

Figures 1-4 plot Burden and Kimball's EI estimates for selected cells in each of the two matrices against the EMax estimates, indicating very considerable agreement. Figure 1 shows that – save for a small number of outliers – there was a very close fit between the EI and EMax estimates for Democrat loyalty (DD: those who voted for Dukakis for President and for a Democrat candidate for the House). Figure 2 shows a similar very close fit (with only two substantial outliers) between the two sets of estimates for Democrat splitters (DR – those who voted for Dukakis for President and a Republican candidate for the House).

There were only 33 cases in the contests for the Presidency and the Senate. For these, too, Figures 3 and 4 show close agreements between the EI and EMax estimates for Republican loyalists and splitters (RR and RD, respectively).

Regression analyses between the two sets of estimates are in Table 2. All four comparisons of the EI (the independent variable) and EMax (dependent variable) estimates for Presidential and House voting have correlations very close to 1.00 – 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 Democrat loyalists and splitters: fitting curvilinear rather than linear relationships increased the r^2 values to 0.83 and 0.85 respectively. (None of the straight-line regressions gave a 45 degree slope with a constant of zero and regression coefficient of 1.0, however. For this data set, in general EMax produces somewhat higher levels of split-ticket voting at the lower end of the range than EI, with the two converging at the higher end of the range.)

Conclusion

The EI and EMax estimation procedures produced very similar outcomes for the two sets of matrices.[5] In relative terms, the pattern of ticket-splitting across the Congressional Districts and States varies only slightly between the two procedures. Burden and Kimball used their estimates to analyze the impact of campaigning on the volume of ticket-splitting: the greater the intensity of the campaign, the better the party's performance. That their EI estimates of the volume of split-ticket voting by

State and Congressional District correlate so highly with the EMax estimates offers strong circumstantial evidence sustaining their findings regarding such spatial variations and their causes

Much quantitative social science involves estimating unknown variables of interest from known values. Burden and Kimball claimed (1998, 542) their use of EI produced 'more accurate and efficient estimates of the direction and magnitude of ticket splitting in different states and congressional districts', and were superior to those generated by previous research dependent 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 EMax, we have combined aggregate and individual level data and produced estimates 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, should not be dismissed.

Cho and Gaines (2004a, 169) state that 'EI generally does not outperform OLS'. It is difficult to accept the logic of this statement, since EI incorporates information from the bounds; methods which do that – including EMax – are surely superior to any which do not. It is at the core of statistical inference that the more information you have, the better. EI deploys what is available in many cases – i.e. the bounds for the 2 x 2 data matrix for each of the areal units studied. EMax uses a further set of information where that is also available – a third set of bounds – and our reworking of Burden and Kimball's data shows very similar estimates generated by EI and EMax.

Our results suggest a tempering of Cho and Gaines' skepticism regarding the validity of EI-derived estimates: an alternative procedure has produced very similar estimates of the unknown variables, suggesting either that the Burden/Kimball estimates are better than Cho/Gaines argue, or that both of the procedures suffer from fatal flaws and yet 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 at least a stay of execution on Burden and Kimball's estimates let alone the wider criticism of King's EI method implied by their stance

Notes

[1] That technique was used in an early evaluation of spatial variations in ticket-splitting in the US – Johnston and Hay, 1984: Burden and Kimball did not develop 'the first-ever accurate district-level estimates of vote-splitting' (Cho and Gaines, 2004a, p. 169).
[2] As a statistical procedure, EI has standard errors for all of its estimates: developed as a mathematical procedure, EMax does not although we are currently addressing this issue. In addition, EI can use covariates which, as yet, EMax cannot.
[3] We are grateful to Profs Burden and Kimball for providing us with the relevant data files.
[4] These matrices are, of course, estimates from survey data only, and introduce some error to the entropy-maximizing procedure – an issue which we are currently addressing.
[5] Given Cho and Gaines argument of instability in the results from applying EI, it may be that Burden and Kimball serendipitously produced the same pattern of results as we did – an issue for further research on EI.

References

Burden, B. C. and Kimball, D. C. (1998) A New Approach to the Study of Ticket-Splitting. *American Political Science Review,* 92: 533-544.

Burden, B. C. and Kimball, D. C. (2002) *Why Americans Split their Tickets: Campaigns, Competition and Divided Government*. Ann Arbor MI: University of Michigan Press.

Cho, W. K. T. and Gaines, B. J. (2004a) The Limits of Ecological Inference: The Case of Split-Ticket Voting. *American Journal of Political Science,* 48: 152-171.

Cho, W. K. T. and Gaines, B. J. (2004b) Review of Burden and Kimball (2002). *Party Politics*, 10:241-243.

Johnston, R. J. (1995) *The Geography of English Politics*. London: Croom Helm.

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. (2000) Ecological Inference and Entropy-Maximizing: an Alternative Estimation Procedure for Split-Ticket Voting. *Political Analysis*, 8: 333-345.

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 (percentages of row totals)

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 EMax on EI estimates of split-ticket voting, 1998 (standard errors in parentheses)

President	House	a	b	r^2
Bush	Republican	16.8	0.76	0.96
			(0.008)	
Bush	Democrat	6.8	0.76	0.96
			(0.008)	
Dukakis	Democrat	12.0	0.84	0.97
			(0.007)	
Dukakis	Republican	4.1	0.84	0.97
			(0.007)	
President	Senate	a	b	r^2
Bush	Republican	15.2	0.71	0.95
			(0.029)	
Bush	Democrat	13.9	0.71	0.95
			(0.029)	
Dukakis	Democrat	9.4	0.81	0.76
			(0.080)	
Dukakis	Republican	9.9	0.83	0.76
			(0.080)	

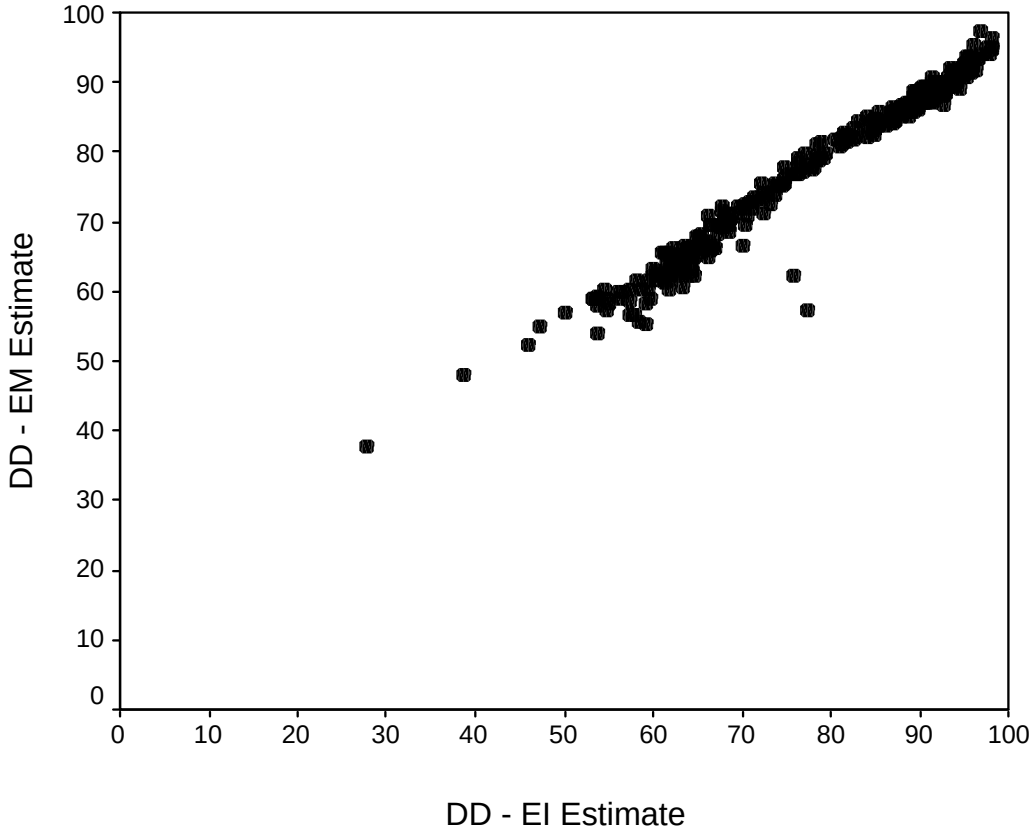

Figure 1. EI and EMax estimates of Democrat loyalty – 1988 Presidential and House elections

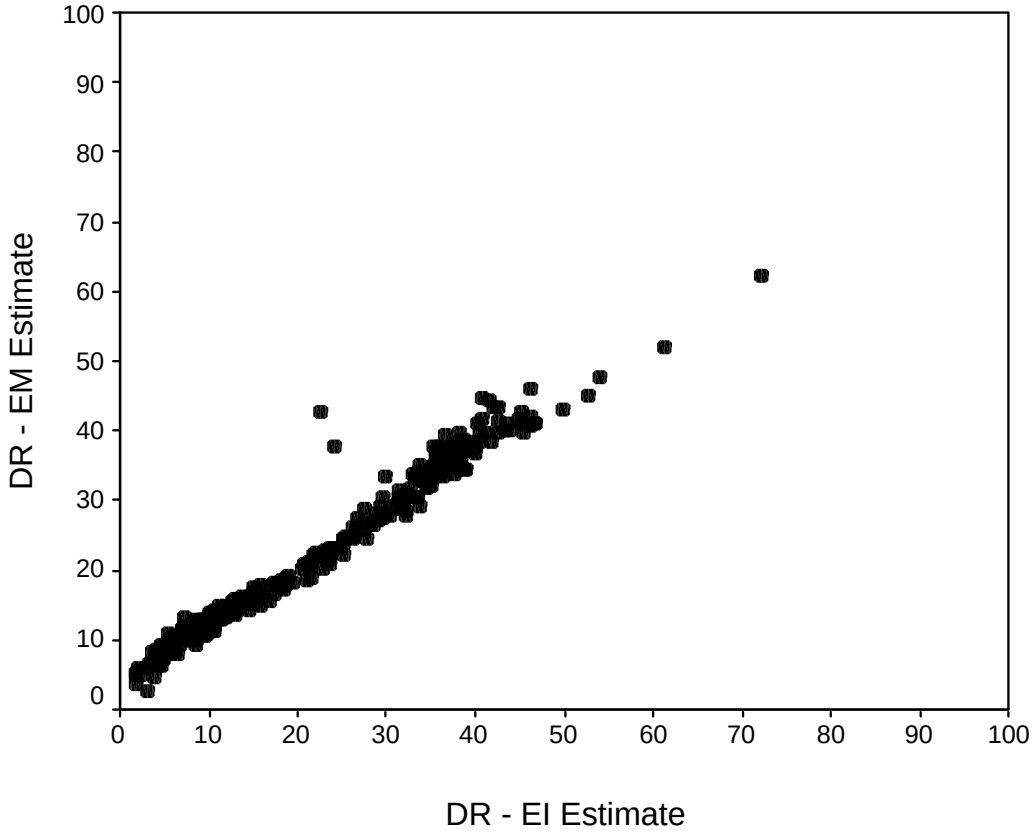

Figure 2. EI and EMax estimates of Democrat splitters – voted Democrat for President and Republican for House in 1988

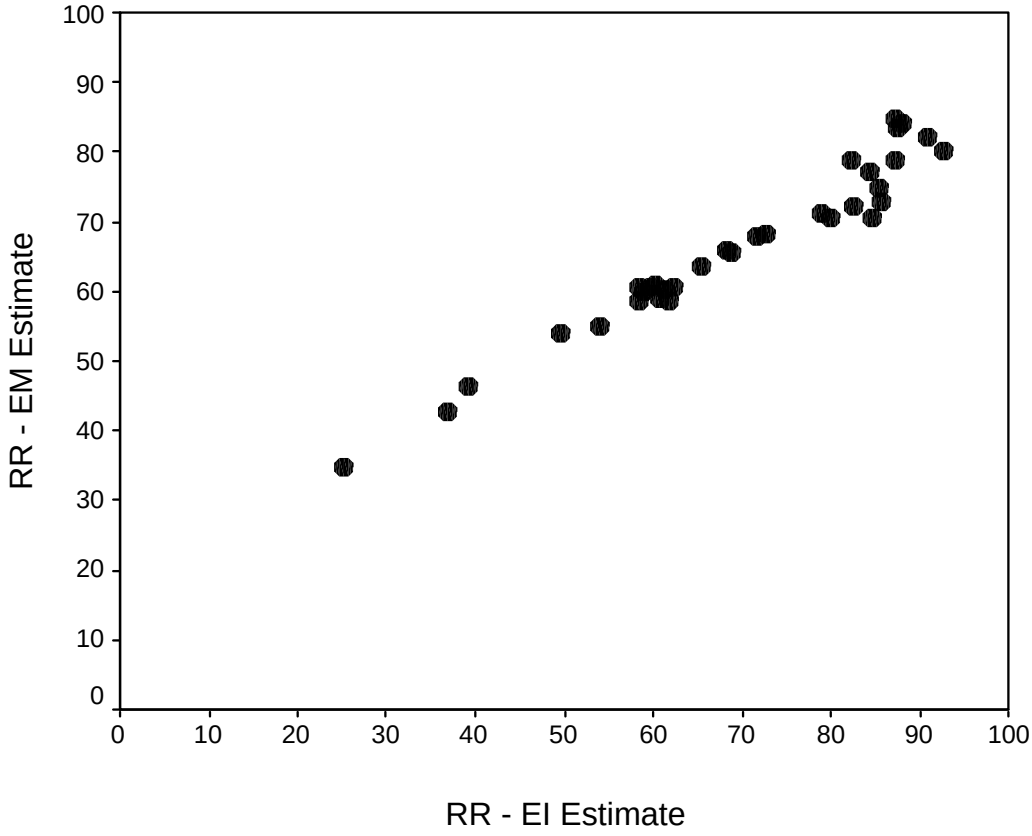

Figure 3. EI and EMax estimates of Republican loyalty – 1988 Presidential and Senate elections

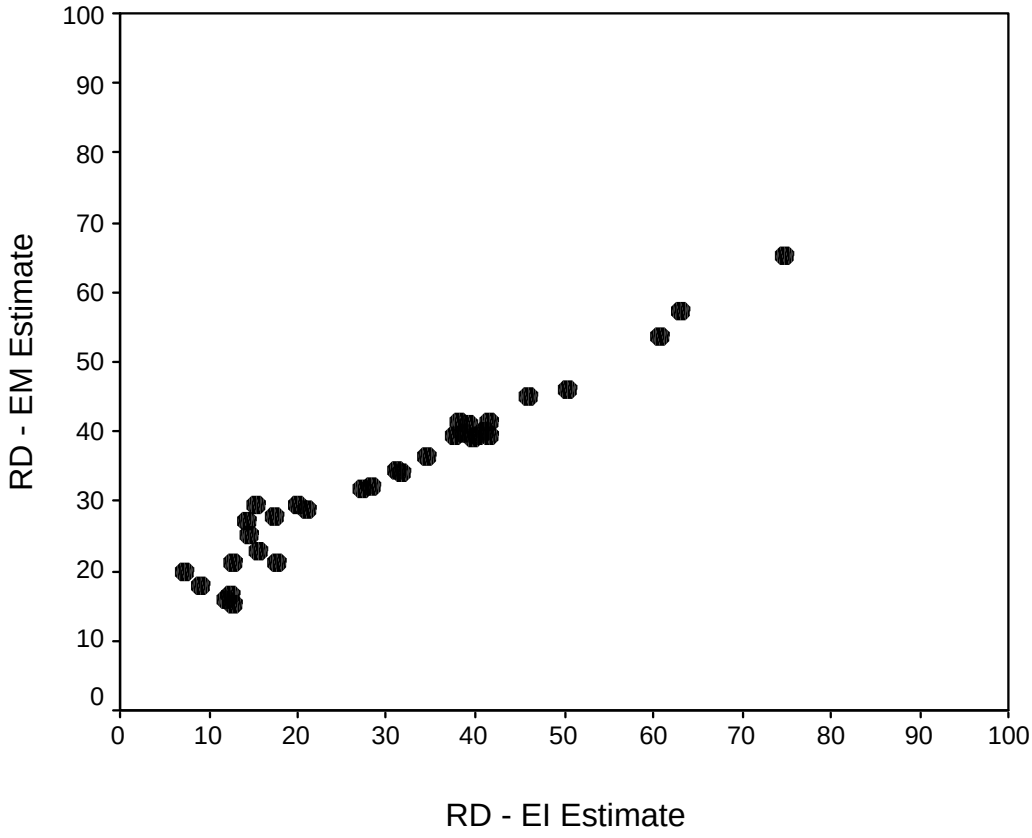

Figure 4. EI and EMax estimates of Republican splitters – voted Republican for President and Democrat for Senate in 1988

Nr.	Author	Title
04-41	Thomas Gschwend	Comparative Politics of Strategic Voting: A Hierarchy of Electoral Systems
04-40	Ron Johnston Thomas Gschwend Charles Pattie	On Estimates of Split-Ticket Voting: EI and EMax
04-39	Volker Stocké	Determinants and Consequences of Survey Respondentsí Social Desirability Beliefs about Racial Attitudes
04-38	Siegfried K. Berninghaus Marion Ott Bodo Vogt	Restricting the benefit flow from neighbors: Experiments on network formation
04-37	Christopher Koch	Behavioral Economics und die Unabhängigkeit des Wirtschaftsprüfers - Ein Forschungsüberblick
04-36	Christopher Koch	Behavioral Economics und das Entscheidungsverhalten des Wirtschaftsprüfers - Ein Forschungsüberblick
04-35	Christina Reifschneider	Behavioral Law and Economics: Überlegungen zu den Konsequenzen moderner Rationalitätskonzepte für die Gestaltung informationellen Kapitalmarktrechts
04-34	Siegfried K. Berninghaus Karl-Martin Ehrhart Marion Ott Bodo Vogt	Searching for ”Stars” - Recent Experimental Results on Network Formation -
04-33	Christopher Koch	Haftungserleichterungen bei der Offenlegung von Zukunftsinformationen in den USA
04-32	Oliver Kirchkamp J. Philipp Reiß	The overbidding-myth and the underbidding-bias in first-price auctions
04-31	Alexander Ludwig Alexander Zimper	Investment Behavior under Ambiguity: The Case of Pessimistic Decision Makers

Nr.	Author	Title
04-30	Volker Stocké	Attitudes Toward Surveys, Attitude Accessibility and the Effect on Respondentsí Susceptibility to Nonresponse
04-29	Alexander Ludwig	Improving Tatonnement Methods for Solving Heterogeneous Agent Models
04-28	Marc Oliver Rieger Mei Wang	Cumulative Prospect Theory and the St.Petersburg Paradox
04-27	Michele Bernasconi Oliver Kirchkamp Paolo Paruolo	Do fiscal variables affect fiscal expectations? Experiments with real world and lab data
04-26	Daniel Schunk Cornelia Betsch	Explaining heterogeneity in utility functions by individual differences in preferred decision modes
04-25	Martin Weber Jens Wuestemann	Bedeutung des Börsenkurses im Rahmen der Unternehmensbewertung
04-24	Hannah Hörisch	Does foreign aid delay stabilization
04-23	Daniel Schunk Joachim Winter	The Relationship Between Risk Attitudes and Heuristics in Search Tasks: A Laboratory Experiment
04-22	Martin Hellwig	Risk Aversion in the Small and in the Large When Outcomes Are Multidimensional
04-21	Oliver Kirchkamp Eva Poen J. Philipp Reiß	Bidding with Outside Options
04-20	Jens Wüstemann	Evaluation and Response to Risk in International Accounting and Audit Systems: Framework and German Experiences
04-19	Cornelia Betsch	Präferenz für Intuition und Deliberation (PID): Inventar zur Erfassung von affekt- und kognitionsbasiertem Entscheiden
04-18	Alexander Zimper	Dominance-Solvable Lattice Games