Coder Reliability and Misclassification in the Human Coding of Party Manifestos: Supplementary Materials

Slava Mikhaylov
University College London
v.mikhaylov@ucl.ac.uk

Michael Laver
New York University
michael.laver@nyu.edu

Kenneth Benoit
London School of Economics
kbenoit@lse.ac.uk

September 9, 2011

Complete Category Listing.

Detailed information on categories, reliability, and classification probabilities from tests.

Code	Description	Overall %	RILE	Fleiss κ All	Fleiss κ RILE	Pr(A*|A) = Left	Right	Other
103	Anti-Imperialism: Positive	0.38	L					
105	Military: Negative	0.77	L					
106	Peace: Positive	0.82	L					
107	Internationalism: Positive	2.79	L					
202	Democracy: Positive	3.55	L	0.18	0.07	**0.50**	0.03	0.47
403	Market Regulation: Positive	2.04	L	0.29	-0.03	**0.75**	0.12	0.14
404	Economic Planning: Positive	0.97	L	0.05	-0.05	**0.18**	0.35	0.47
406	Protectionism: Positive	0.26	L					
412	Controlled Economy: Positive	0.71	L					
413	Nationalization: Positive	0.41	L					
504	Welfare State Expansion: Positive	7.19	L	0.50	0.10	**0.68**	0.03	0.29
506	Education Expansion: Positive	4.44	L	0.46	n/a	**0.78**	0.00	0.22
701	Labour Groups: Positive	2.51	L	0.14	0.05	**0.45**	0.08	0.47
104	Military: Positive	1.32	R					
201	Freedom and Human Rights: Positive	2.56	R					
203	Constitutionalism: Positve	0.59	R					
305	Political Authority: Positive	3.00	R	0.10	0.14	0.24	**0.44**	0.32
401	Free Enterprise: Positive	1.74	R					
402	Incentives: Positive	2.29	R	0.46	0.03	0.20	**0.74**	0.06
407	Protectionism: Negative	0.21	R					
414	Economic Orthodoxy: Positive	1.91	R	0.46	0.16	0.02	**0.77**	0.20
505	Welfare State Limitation: Positive	0.36	R					
601	National Way of Life: Positive	1.03	R					
603	Traditional Morality: Positive	1.41	R					
605	Law and Order: Positive	2.46	R	0.13	n/a	0.00	**0.82**	0.18
606	Social Harmony: Positive	1.44	R	0.44	0.24	0.03	**0.71**	0.26
101	Foreign Special relationships: Positive	0.77	-					
102	Foreign Special relationships: Negative	0.22	-					
108	European Integration: Positive	1.92	-					
109	Internationalism: Negative	0.40	-					
110	European Integration: Negative	0.43	-					
204	Constitutionalism: Negative	0.23	-					
301	Decentralization: Positive	3.19	-					
302	Centalization: Positive	0.16	-					
303	Governmenatal and Administrative Efficiency: Positive	4.60	-	0.02	n/a	0.47	0.00	**0.53**
304	Political Corruption: Negative	0.80	-					
405	Corporatism: Positive	0.27	-	0.21	n/a	0.25	0.00	**0.75**
408	Economic Goals	2.90	-	0.13	0.02	0.16	0.16	**0.68**
409	Keynesian Demand Management: Positive	0.19	-					
410	Productivity: Positive	2.14	-	0.17	0.12	0.01	0.16	**0.83**
411	Technology and Infrastructure: Positive	5.71	-	0.34	0.29	0.41	0.05	**0.54**

Code	Description	Overall %	RILE	Fleiss κ All	Fleiss κ RILE	Pr(A*|A) = Left	Pr(A*|A) = Right	Pr(A*|A) = Other
415	Marxist Analysis: Positive	0.09	-					
416	Anti-Growth Economy: Positive	0.69	-					
501	Environmental Protection: Positive	4.85	-					
502	Culture: Positive	3.04	-					
503	Social Justice: Positive	3.83	-	0.35	0.24	0.12	0.10	**0.78**
507	Education Limitation: Positive	0.04	-					
602	National Way of Life: Negative	0.21	-					
604	Traditional Morality: Negative	0.29	-					
607	Multiculturalism: Positive	0.80	-					
608	Multiculturalism: Negative	0.22	-					
702	Labour Groups: Negative	0.12	-					
703	Farmers: Positive	3.41	-	0.82	0.04	0.03	0.09	**0.88**
704	Middle Class and Professional Groups: Positive	0.86	-					
705	Underprivileged Minority Groups: Positive	1.44	-					
706	Non-economic Demographic Groups: Positive	4.20	-	0.29	0.11	0.17	0.08	**0.75**
000	Uncoded	4.79	-	0.11	0.10	0.41	0.14	**0.45**

Notes

1. Note: The "Overall % column refers to the proportion of coded quasi-sentences assigned to each category from the complete CMP dataset (from *MPP2*).

2. Categories are sorted so that the 13 "Rile-left" categories are listed first, the 13 "Rile-right" categories second, followed by the "Rile-other" categories. The probability that an individual policy category will be classified as belonging to its own "Rile" classification are highlighted in boldface. For quasi-sentences that really belong to "202: Democracy: Positive" for instance—a relative high-frequency category at 3.55% of all CMP quasi-sentences in the combined dataset—the probability is only 0.50 from our tests that it will be assigned a CMP code that is one of the 13 "Rile-left" categories. The probability is almost even (0.47) that it will be coded as a category that is not part of the "Rile" index, and just 0.03 that it will coded as a "Rile-right" category.

3. The limited set of categories in our test documents meant that we could only report misclassification probabilities for the 20 categories identified by the CMP's Master coding.