```
1
2
3   *------------------------------------------------------------------------------
4   *------------------------------------------------------------------------------
5   *
6   *   Title:      Big Data meets Open Political Science:
7   *               An Empirical Assessment of Transparency Standards 2008-2019
8   *
9   *   Authors:    Karin Dyrstad & Jonathon W. Moses
10  *   Journal:    European Political Science
11  *
12  *------------------------------------------------------------------------------
13  *------------------------------------------------------------------------------
14
15  ssc install cem
16  ssc install outreg
17
18  clear all
19  use replicationdata
20
21
22
23  *------------------------------------------------------------------------------
24  *------------------------------------------------------------------------------
25  *
26  *               MAIN ANALYSIS
27  *
28  *------------------------------------------------------------------------------
29  *------------------------------------------------------------------------------
30
31
32
33  ***** Figure 4: BDPS articles published over time, PS and JETS
34  preserve
35  collapse (sum) bd pswos jets if bd ==1, by(year)
36  lab var bd "Total BDPS articles"
37  lab var ps "In PS journal"
38  lab var jets "In JETS journal"
39  tw(connected bd year)(connected ps year)(connected jets year), xlabel(2008(1)2019, angle(45))
    ytitle("Number of BDPS articles") scheme(plotplain) xtitle("")
40  restore
41
42
43
44  ***** Figure 5: Trends in transparency
45  bysort year: egen transp_share_bdps = mean(transp) if bd==1
46  bysort year: egen bdps_n = count(bd) if bd==1
47  bysort year: egen transp_share_ps = mean(transp) if bd==0
48  bysort year: egen transp_share_empirical = mean(transp) if bd==0 & nonempirical==0
49  bysort year: egen transp_share_quant = mean(transp) if bd==0 & analysis==5|bd==0 & analysis==2
50  bysort year: egen transp_share_jets = mean(transp) if bd==0 & jets==1
51  lab var transp_share_bdps "Share of fully transparent BDPS articles, by year"
52  lab var transp_share_ps "Share of fully transparent non-BD PS articles, by year"
53  lab var bdps_n "Number of BDPS observations per year"
54
55  tw(connected transp_share_empirical year)(connected transp_share_quant year)(connected
    transp_share_jets year), xline(2010)  xline(2014) scheme(plotplain)  legend(lab(1 "All empirical
    articles") lab(2 "Quant./mixed methods articles") lab(3 "Articles in JETS journals")) xlabel(2008(
    1)2019, angle(45))  xtitle("") name(ps_connected) ylabel(0(.2)1)
56
57  tw(lfitci transp year if nonempirical==0 & bd==0)(lfitci transp year if analysis==5 & bd==0|
    analysis==2 & bd==0)(lfitci transp year if jets==1), xline(2010)  xline(2014) scheme(plotplain)
    legend(lab(2 "All empirical articles") lab(3 "Quant./mixed methods articles") lab(4 "Articles in
    JETS journals")) xlabel(2008(1)2019, angle(45)) ytitle("")  xtitle("") name(ps_lfitci) ylabel(0(.2
    )1)
58
59  graph combine ps_connected ps_lfitci, col(1)
60
```

```
61
62
63  ***** Figure 6: BDPS vs quant. control group; JETS and non-JETS
64  tw (qfitci transp year if bd==1)(qfitci transp year if bd==0 & quant ==1), ylabel(0(.2)1)  xlabel(
    2008(1)2019, angle(45)) ytitle("") legend(lab(2 "BDPS") lab(3 "Non-BDPS (quant.)")) scheme(
    plotplain) xtitle("") name(fig6_a)
65
66  tw (qfitci transp year if bd==1)(qfitci transp year if bd==0 & quant ==1 & jets==0)(qfitci transp
    year if bd==0 & quant ==1 & jets==1), ylabel(0(.2)1)  xlabel(2008(1)2019, angle(45)) ytitle("")
    legend(lab(2 "BDPS") lab(3 "Non-BDPS, non-JETS (quant.)") lab(4 "Non-BDPS, JETS (quant.)")) scheme
    (plotplain) xtitle("") name(fig6_b)
67
68  graph combine fig6_a fig6_b, col(1)
69
70
71
72  **** Table 3
73  logit transp bd year /* Model 1 */
74  outreg using table, replace
75  logit transp bd year , cl(journal) /* clustering on journals; Model 2 */
76  outreg using table, merge
77  logit transp i.bd##c.year , cl(journal) /* checking if difference is equal across time */
78  outreg using table, merge
79  logit transp bd ib2019.year , cl(journal) /* year as dummy set */
80  outreg using table, merge
81  logit transp bd jets year , cl(journal) /* adding control for JETS; Model 3*/
82  outreg using table, merge
83  logit transp bd year if nonempirical==0,  cl(journal) /* empirical articles only; Model 4 */
84  outreg using table, merge
85  logit transp bd year if quant==1, cl(journal) /* pure quantitative articles only; Model 5 */
86  outreg using table, merge
87  cem  journal (scott) year (#12) ra (#2) if quant==1, treatment(bd) /*coarsened exact matching */
88  logit transp bd [iweight=cem_weights]
89  outreg using table, merge /* Model 6*/
90
91
92
93
94  *-------------------------------------------------------------------------------
95  *-------------------------------------------------------------------------------
96  *
97  *               ONLINE SUPPLEMENTAL MATERIAL
98  *
99  *-------------------------------------------------------------------------------
100 *-------------------------------------------------------------------------------
101
102
103 ** D.4 Reliability: Mean transparency scores, by dataset and RAs
104 ttest transp, by(ra) /* overall */
105 ttest transp if bd==1, by(ra) /* BDPS group */
106 ttest transp if bd==0, by(ra) /* Baseline group (non-BD) */
107
108
109 ** E.1 Publication outlets for BDPS articles, by PS – non-PS journal
110 bysort pswos: tab journal
111
112
113 ** E.2 Descriptive statistics
114 sum transp transp_code transp_data ra psw jets
115 bysort bd: sum transp transp_code transp_data ra psw jets
116
117
118
119
```