

```

1 *****
2 * A Practical Guide to Using Panel Data
3 * Simonetta Longhi and Alita Nandi
4 * ISER, University of Essex
5 * Chapter 8
6 *****
7
8 version 12
9 clear all
10 set more off
11 capture log close
12
13
14 cd "C:\My Documents"
15
16 global datadir "S:\final"
17
18
19 log using Example_Chapter8.log, replace
20
21
22
23 * 8.1. Introduction
24 *-----
25
26 use "$datadir/rindresp", clear
27 rename r* *
28 quietly mvdecode _all, mv(-9/-1)
29
30 * When "quietly" is added before the command, Stata
31 * will run the command without showing the results of the command
32
33 * 8.2. Continuous Variables
34 * 8.2.1. Correlations Coefficients vs. Regression Coefficients
35 *-----
36
37 label var paygu "Usual gross pay per month"
38
39 set more on
40
41 correlate paygu age
42 scatter paygu age, scheme(smanual)
43 more
44 twoway (scatter paygu age) (lfit paygu age), scheme(smanual)
45 more
46
47 generate LnW = ln(paygu)
48
49 kdensity paygu
50 more
51 kdensity LnW
52 more
53 twoway (scatter LnW age) (lfit LnW age), scheme(smanual)
54
55 set more off
56
57 generate age2 = age^2
58
59 generate Female = 1 if sex == 2
60 replace Female = 0 if sex == 1
61 label var Female "Dummy for women"
62
63 generate Married = 1 if mastat == 1 | mastat == 2 | mastat == 7
64 replace Married = 0 if (mastat >= 3 & mastat <= 6) | (mastat >= 8 & mastat < .)
65 label var Married "Whether married or cohabiting"
66
67 * qfachi: highest academic qualification
68 recode qfachi (-9/-1 = .) (1 = 2)
69 tabulate qfachi, gen(Q)
70 label var Q1 "1st degree or higher"
71 label var Q2 "HND,HNC,teaching"
72 label var Q3 "A level"
73 label var Q4 "O level"
74 label var Q5 "CSE"
75 label var Q6 "none of these qualif"

```

```

76
77 * region2: Government Office Regions
78 tabulate region2, gen(R)
79 label var R1 "North East"
80 label var R2 "North West"
81 label var R3 "Yorkshire & Humber"
82 label var R4 "East Midlands"
83 label var R5 "West Midlands"
84 label var R6 "East of England"
85 label var R7 "London"
86 label var R8 "South East"
87 label var R9 "South West"
88 label var R10 "Wales"
89 label var R11 "Scotland"
90 label var R12 "Northern Ireland"
91
92 save DatasetR.dta, replace
93
94 regress LnW age
95 * Compare with:
96 regress LnW age age2 Female Married Q1-Q5 R1-R6 R8-R12
97
98 * 8.2.2. Different Types of Standard Errors
99 *-----
100
101 regress LnW age age2 Female Married Q1-Q5 R1-R6 R8-R12, vce(robust)
102
103 * 8.2.3. Post-estimation Commands
104 *-----
105
106 predict yhat_a, xb
107 label var yhat_a "Predictions (xb) on full sample"
108 predict yhat_a_se, stdp
109
110 predict yhat_b if e(sample), xb
111 label var yhat_a "Predictions (xb) on estimation sample"
112 predict yhat_b_se if e(sample), stdp
113
114 test Female
115 test Q1 Q2 Q3 Q4 Q5
116
117 * Run the model again after including a dummy for part-time workers
118 * jbft: employed full time
119 tabulate jbft, gen(PT)
120 rename PT2 PartTime
121 regress LnW age age2 Female Married PartTime Q1-Q5 R1-R6 R8-R12, vce(robust)
122
123
124 * 8.2.4. Diagnostics
125 *-----
126
127 regress LnW age age2 Female Married Q1-Q5 R1-R6 R8-R12
128
129 * Heteroskedasticity
130 set more on
131 rvfplot, yline(0)
132 more
133 estat hettest
134
135 * Multicollinearity
136 estat vif
137 collin age age2 Female Married Q1-Q5 R1-R6 R8-R12
138
139 * Normality
140 predict Resid, residuals
141 kdensity Resid, normal
142 more
143
144 mvtest norm Resid, stats(all)
145
146 set more off
147
148 * 8.2.5. Interaction Terms
149 *-----
150 generate MarriedWoman = Married * Female

```

```

151 regress LnW age age2 Female Married MarriedWoman Q1-Q5 R1-R6 R8-R12, vce(robust)
152 test Married + MarriedWoman = 0
153
154 * Vs separate estimations: men
155 regress LnW age age2 Married Q1-Q5 R1-R6 R8-R12 if Female == 0, vce(robust)
156 estimates store ResultsMen
157 * and women
158 regress LnW age age2 Married Q1-Q5 R1-R6 R8-R12 if Female == 1, vce(robust)
159 estimates store ResultsWomen
160
161
162
163 * 8.3 Estimating Weighted Regression Models
164 *-----
165
166 * Add in psu and strata variables from rhhsamp using link variable rhid
167 * The variable identifying households in the rhhsamp file is called rhid
168 * while above we have renamed it hid
169 generate double rhid = hid
170 merge m:1 rhid using "$datadir\rhhsamp", keepusing(rhid rpsu rstrata)
171 keep if _merge == 3
172 drop _merge
173 rename rpsu psu
174 rename rstrata strata
175
176 regress LnW age age2 Female Married Q1-Q5
177 estimates store R_Regions
178
179 regress LnW age age2 Female Married Q1-Q5 R1-R6 R8-R12
180 estimates store R_NoWeights
181
182 regress LnW age age2 Female Married Q1-Q5 R1-R6 R8-R12 [pweight = xrwtuk1]
183 estimates store R_Weights
184
185 replace psu = hid if memorig == 7
186 svyset psu [pw=xrwtuk1], strata(strata)
187 svy: regress LnW age age2 Female Married Q1-Q5 R1-R6 R8-R12
188 estimates store R_SampleDesign
189
190 * Save the results in a table *
191 * See Chapter 14 *
192 *****
193
194 * Table 8.1
195 quietly estout R_Regions R_NoWeights R_Weights R_SampleDesign ///
196     using "WageModels.out", ///
197     cells(b(star fmt(%9.3f)) se(par fmt(%9.3f))) ///
198     style(tab) stats(r2 N, fmt(%9.3f %9.0g)) ///
199     labels(R2 Observations) label collabels(, none) ///
200     starlevels(+ 0.05 * 0.01) ///
201     postfoot("St. err. in parenthesis; + Significant 5%, * Significant 1%") ///
202     replace
203
204
205 * 8.4. Binary Variables
206 * 8.4.1. Linear and Non-linear models
207 *-----
208
209 * jbstat: current economic activity
210 recode jbstat (1 2 = 1) (3/10 = 0), gen(Employed)
211 label define Employed 1 "Employed/Self-Employed" 2 "Unemployed or Inactive"
212 label value Employed Employed
213
214 regress Employed age age2 Married Q1-Q5 R1-R6 R8-R12 if Female == 0
215 probit Employed age age2 Married Q1-Q5 R1-R6 R8-R12 if Female == 0
216 logit Employed age age2 Married Q1-Q5 R1-R6 R8-R12 if Female == 0
217
218 * 8.4.2. Coefficients, Marginal Effects, Odd Ratios
219 *-----
220
221 probit Employed age age2 Married Q1-Q5 R1-R6 R8-R12 if Female == 0
222 margins, ///
223     dydx(age age2 Married Q1 Q2 Q3 Q4 Q5 R1 R2 R3 R4 R5 R6 R8 R9 R10 R11 R12)
224 * Compare with:
225 probit Employed age age2 i.Married i.Q1-Q5 i.R1-R6 i.R8-R12 if Female == 0

```

```

226 margins, ///
227     dydx(age age2 Married Q1 Q2 Q3 Q4 Q5 R1 R2 R3 R4 R5 R6 R8 R9 R10 R11 R12)
228
229 logit Employed age age2 i.Married i.Q1-Q5 i.R1-R6 i.R8-R12 if Female == 0
230 margins, ///
231     dydx(age age2 Married Q1 Q2 Q3 Q4 Q5 R1 R2 R3 R4 R5 R6 R8 R9 R10 R11 R12)
232
233 logit Employed age age2 Married Q1-Q5 R1-R6 R8-R12 if Female == 0, or
234
235 * 8.4.3. Different Types of Standard Errors and Other Post-estimation commands
236 *-----
237
238 probit Employed age age2 Married Q1-Q5 R1-R6 R8-R12 if Female == 0, vce(robust)
239 logit Employed age age2 Married Q1-Q5 R1-R6 R8-R12 if Female == 0, vce(robust)
240
241 * 8.5. Multiple Outcomes
242 * 8.5.1. Ordered Outcomes
243 *-----
244
245 generate Unemployed = 1 if jbstat == 3
246 replace Unemployed = 0 if jbstat == 1 | jbstat == 2
247
248 * lfsato: satisfaction with: life overall
249 tab lfsato Unemployed, miss
250
251 oprobit lfsato i.Unemployed i.Female age age2 i.Married i.Q1-Q5 i.R1-R6 i.R8-R12
252 margins, dydx(Unemployed Female Married)
253
254 ologit lfsato i.Unemployed i.Female age age2 i.Married i.Q1-Q5 i.R1-R6 i.R8-R12
255 margins, dydx(Unemployed Female Married)
256
257 ologit lfsato Unemployed Female age age2 Married Q1-Q5 R1-R6 R8-R12, or
258
259 regress lfsato Unemployed Female age age2 Married Q1-Q5 R1-R6 R8-R12
260
261 * 8.5.2. Unordered Outcomes
262 *-----
263
264 recode jbstat (1 2 = 1) (3 = 2) (4/10 = 3), gen(LMS)
265 label var LMS "Labour Market Status"
266 label define LMS 1 "Has Job" 2 "Unemployed" 3 "Inactive"
267 label values LMS LMS
268
269 mprobit LMS i.Female age age2 i.Married i.Q1-Q5 i.R1-R6 i.R8-R12, baseoutcome(1)
270 margins, dydx(Female Married) predict(outcome(1))
271 margins, dydx(Female Married) predict(outcome(2))
272 margins, dydx(Female Married) predict(outcome(3))
273
274 mlogit LMS Female age age2 Married Q1-Q5 R1-R6 R8-R12, baseoutcome(1)
275 margins, dydx(Female Married) predict(outcome(2))
276 margins, dydx(Female Married) predict(outcome(3))
277
278
279 * 8.6 Heckman Selection Models
280 *-----
281 generate Works = (jbhas == 1 | jboff == 1)
282 * JBHAS: Did paid work last week
283 * JBOFF No work last week but has job
284
285 regress LnW age age2 Married PartTime Q1-Q5 R1-R6 R8-R12 ///
286     if Female == 1 & age >= 23 & age < 60, vce(robust)
287
288 heckman LnW age age2 Married PartTime Q1-Q5 R1-R6 R8-R12 ///
289     if Female == 1 & age >= 23 & age < 60, ///
290     select(Works = age age2 Married Q1-Q5 R1-R6 R8-R12 nchild) ///
291     first mills(MR) vce(robust)
292
293 log close
294

```