

```

1  /*****
2
3  COLLECTIVE MODEL -- EU-SILC DATA
4  Fall 2022
5
6  Filename:   01_Basics.do
7  Authors:    Jorge Velilla, José Alberto Molina, José Ignacio Giménez-Nadal
8              Department of Economic Analysis
9              MSc in Economics 2022, University of Zaragoza (Spain)
10
11
12  This file provides instructions to compute baseline sample preparation, summary
13  statistics of main variables, t-type tests and tabulates, baseline estimates,
14  and export estimation results.
15
16
17  Licencia Creative Commons "http://creativecommons.org/licenses/by-nc-sa/4.0/"
18  Materiales de software en abierto para estudiantes del Master de Economía:
19  Manejo de datos y obtención de evidencia empírica socioeconómica a través de
20  STATA por José Alberto Molina, José Ignacio Giménez-Nadal y Jorge Velilla se
21  distribuye bajo una Licencia Creative Commons Atribución-NoComercial-
22  CompartirIgual 4.0 Internacional.
23
24  *****/
25
26  *   Clear memory
27  clear all
28
29  *   Set working directory
30  cd "C:\Users\Jorge
31  Velilla\Desktop\UNIVERSIDAD\OTRO\2023\ClaseMaster_2022-23\Master_StataLatex\Stata"
32
33  /*****
34      Request data and make some adjustments:
35          - Keep a given year or a given country
36          - Identify outliers
37          - Delete observations (outliers)
38          - Delete variables
39          - Redefine variables
40          - Generate variables
41          - Label variables
42          - Order variables
43          - Rename variables
44  *****/
45
46  *   Request data
47  use "Data/Sample.dta", clear
48
49  *   Keep Portugal, Italy, Greece, Spain
50  keep if country=="Portugal" | country=="Italy" | country=="Greece" | country==
51  "Spain"
52
53  *   Keep data from 2010 (try both commands)
54  *drop if year<2010
55  keep if year>=2010

```

```

56 * Identify outliers in multivariate data
57 *help bacon
58 bacon age workhours wage nlabinc w_age w_workhours w_wage w_nlabinc numfu numch
hh_nlabinc, p(0.05) gen(todrop)
59
60 * Delete outlier observations
61 drop if todrop==1
62
63 * Delete 'todrop' variable
64 drop todrop
65
66 * Redefine occupation
67 replace occupation_ISC008 = occupation_ISC008/10
68 replace occupation_ISC008 = floor(occupation_ISC008)
69 replace w_occupation_ISC008 = w_occupation_ISC008/10
70 replace w_occupation_ISC008 = floor(w_occupation_ISC008)
71
72 * Generate household earnings
73 generate hh_inc = 52*workhours*wage + 52*w_workhours*w_wage
74
75 * Label household earnings
76 label variable hh_inc "Household labor income"
77
78 * Generate and label log wages
79 generate lwage = log(wage)
80 generate lw_wage = log(w_wage)
81 generate cross_lwage = lwage*lw_wage
82 label variable lwage "H: log wage"
83 label variable lw_wage "W: log wage"
84 label variable cross_lwage "Cross log wages"
85
86 * Order variables
87 order id-wage lwage
88 order id-w_wage lw_wage
89 order id-numch hh_inc cross_lwage
90
91 * Rename variables for simplicity
92 rename hh_inc faminc
93 rename hh_nlabinc famnline
94
95 * Summary statistics of all variables
96 summarize
97
98 * Occupation missing for some individuals! Keep non-missing occupations
99 drop if occupation_ISC008==. | w_occupation_ISC008==.
100 *keep if occupation_ISC008!=. & w_occupation_ISC008!=.
101
102 * Generate country, year, occupation dummies
103 tabulate country, gen(cc_)
104 tabulate year, gen(yr_)
105 tabulate occupation_ISC008, gen(oc_)
106 tabulate w_occupation_ISC008, gen(w_oc_)
107
108
109 /*****
110 Compute summary statistics and baseline tests
111 - Compute sum stats table

```

```

112     - Compute additional sum stats (frequencies, tables...)
113     - Compute t-test p-values for differences between groups
114     *****/
115
116 *   Summary statistics of main variables, by country
117 *   - Portugal:
118 summarize   age edu1 edu2 edu3 workhours wage nlabinc w_age w_edu1 w_edu2   ///
119             w_edu3 w_workhours w_wage w_nlabinc married numfu numch faminc   ///
120             famnlinc w_share_nlab share_wparl edu_gap w_managerial sexratio   ///
121             if country=="Portugal"
122 *   - Italy:
123 summarize   age edu1 edu2 edu3 workhours wage nlabinc w_age w_edu1 w_edu2   ///
124             w_edu3 w_workhours w_wage w_nlabinc married numfu numch faminc   ///
125             famnlinc w_share_nlab share_wparl edu_gap w_managerial sexratio   ///
126             if country=="Italy"
127 *   - Greece:
128 summarize   age edu1 edu2 edu3 workhours wage nlabinc w_age w_edu1 w_edu2   ///
129             w_edu3 w_workhours w_wage w_nlabinc married numfu numch faminc   ///
130             famnlinc w_share_nlab share_wparl edu_gap w_managerial sexratio   ///
131             if country=="Greece"
132 *   - Spain:
133 summarize   age edu1 edu2 edu3 workhours wage nlabinc w_age w_edu1 w_edu2   ///
134             w_edu3 w_workhours w_wage w_nlabinc married numfu numch faminc   ///
135             famnlinc w_share_nlab share_wparl edu_gap w_managerial sexratio   ///
136             if country=="Spain"
137 *   - full sample:
138 summarize   age edu1 edu2 edu3 workhours wage nlabinc w_age w_edu1 w_edu2   ///
139             w_edu3 w_workhours w_wage w_nlabinc married numfu numch faminc   ///
140             famnlinc w_share_nlab share_wparl edu_gap w_managerial sexratio
141
142 *   Tabulate observations by year
143 tabulate year
144
145 *   Tabulate observations by country
146 tabulate country
147
148 *   Tabulate observations by year X country
149 tabulate year country
150
151 *   Test for gender inequality in age, Univ education, wages, labor supply
152 ttest age      == w_age
153 ttest edu3    == w_edu3
154 ttest wage    == w_wage
155 ttest workhours == w_workhours
156
157
158 /*****
159     Compute some figures
160 *****/
161
162 *   Density of wages
163 kdensity wage, addplot((kdensity w_wage))
164
165 *   Density of wages but more paper-style
166 kdensity   wage if wage<50, lcolor(black) lpattern(solid)           ///
167             addplot((kdensity w_wage if w_wage<50,                 ///
168                     lcolor(black) lpattern(dash)))                 ///

```

```

169         ytitle(Density) xtitle(Wage) ///
170         ylabel(, angle(zero)) ///
171         legend(order(1 "Male wage" 2 "Female wage") ring(0) pos(2)) ///
172         scheme(s1mono) ///
173         title(,size(zero))
174
175 * Density of work hours
176 kdensity workhours if workhours<80, lcolor(black) lpattern(solid) ///
177         addplot((kdensity w_workhours if w_workhours<80, ///
178                 lcolor(black) lpattern(dash))) ///
179         ytitle(Density) xtitle(Weekly work hours) ///
180         ylabel(, angle(zero)) ///
181         legend(order(1 "Male hours" 2 "Female hours") ring(0) pos(10)) ///
182         scheme(s1mono) ///
183         title(,size(zero))
184
185 * Wages by country
186 graph bar (mean) wage (mean) w_wage, over(country) ///
187         bar(1, fcolor(gs12) lcolor(black)) ///
188         bar(2, fcolor(gs8) lcolor(black)) ///
189         ytitle(Wage) ///
190         ylabel(, angle(zero)) ///
191         legend(order(1 "Male wage" 2 "Femnale wage")) ///
192         scheme(s1mono)
193
194
195 /*****
196     OLS estimates
197     - Compute estimates
198     - Export estimates
199 *****/
200
201 * Install command 'outreg2'
202 *ssc install outreg2
203
204 * No distribution factors
205 * - i) males
206 regress workhours lwage lw_wage cross_lwage famnlinc ///
207         age edu2 edu3 oc_* ///
208         married numfu numch yr_2-yr_10 cc_1 cc_2 cc_3
209 outreg2 using "Output/Baseline.xls", bdec(3) ctitle(Males) drop (oc_* yr_* cc_* o
210 .*) se replace
211 * - ii) females
212 regress w_workhours lwage lw_wage cross_lwage famnlinc ///
213         w_age w_edu2 w_edu3 w_oc_* ///
214         married numfu numch yr_2-yr_10 cc_1 cc_2 cc_3
215
216 outreg2 using "Output/Baseline.xls", bdec(3) ctitle(Females) drop (w_oc_* yr_*
217 cc_* o.*) se append
218
219 * Introducing distribution factors
220 * - i) males
221 regress workhours lwage lw_wage cross_lwage famnlinc ///
222         w_share_nlab sexratio edu_gap share_wpar1 w_managerial ///
223         age edu2 edu3 oc_* ///
224         married numfu numch yr_2-yr_10 cc_1 cc_2 cc_3
225 outreg2 using "Output/DFactors.xls", bdec(3) ctitle(Males) drop (oc_* yr_* cc_* o

```

```

.*) se replace
223 * - ii) females
224 regress w_workhours lwage lw_wage cross_lwage famnlinc ///
225 w_share_nlab sexratio edu_gap share_wpar1 w_managerial ///
226 w_age w_edu2 w_edu3 w_oc_* ///
227 married numfu numch yr_2-yr_10 cc_1 cc_2 cc_3

228 outreg2 using "Output/DFactors.xls", bdec(3) ctitle(Females) drop (w_oc_* yr_*
cc_* o.*) se append

229
230 * Estimates by country
231 * - Portugal
232 regress workhours lwage lw_wage cross_lwage famnlinc ///
233 w_share_nlab sexratio edu_gap share_wpar1 w_managerial ///
234 age edu2 edu3 oc_* ///
235 married numfu numch yr_2-yr_10 if country=="Portugal"
236 outreg2 using "Output/ByCountry.xls", bdec(3) ctitle(Portugal - Males) drop (oc_*
yr_* o.*) se replace
237 regress w_workhours lwage lw_wage cross_lwage famnlinc ///
238 w_share_nlab sexratio edu_gap share_wpar1 w_managerial ///
239 w_age w_edu2 w_edu3 w_oc_* ///
240 married numfu numch yr_2-yr_10 if country=="Portugal"

241 outreg2 using "Output/ByCountry.xls", bdec(3) ctitle(Portugal - Females) drop (
w_oc_* yr_* o.*) se append
242 * - Italy
243 regress workhours lwage lw_wage cross_lwage famnlinc ///
244 w_share_nlab sexratio edu_gap share_wpar1 w_managerial ///
245 age edu2 edu3 oc_* ///
246 married numfu numch yr_2-yr_10 if country=="Italy"
247 outreg2 using "Output/ByCountry.xls", bdec(3) ctitle(Italy - Males) drop (oc_*
yr_* o.*) se append
248 regress w_workhours lwage lw_wage cross_lwage famnlinc ///
249 w_share_nlab sexratio edu_gap share_wpar1 w_managerial ///
250 w_age w_edu2 w_edu3 w_oc_* ///
251 married numfu numch yr_2-yr_10 if country=="Italy"

252 outreg2 using "Output/ByCountry.xls", bdec(3) ctitle(Italy - Females) drop (w_oc_*
yr_* o.*) se append
253 * - Greece
254 regress workhours lwage lw_wage cross_lwage famnlinc ///
255 w_share_nlab sexratio edu_gap share_wpar1 w_managerial ///
256 age edu2 edu3 oc_* ///
257 married numfu numch yr_2-yr_10 if country=="Greece"
258 outreg2 using "Output/ByCountry.xls", bdec(3) ctitle(Greece - Males) drop (oc_*
yr_* o.*) se append
259 regress w_workhours lwage lw_wage cross_lwage famnlinc ///
260 w_share_nlab sexratio edu_gap share_wpar1 w_managerial ///
261 w_age w_edu2 w_edu3 w_oc_* ///
262 married numfu numch yr_2-yr_10 if country=="Greece"

263 outreg2 using "Output/ByCountry.xls", bdec(3) ctitle(Greece - Females) drop (
w_oc_* yr_* o.*) se append
264 * - Spain
265 regress workhours lwage lw_wage cross_lwage famnlinc ///
266 w_share_nlab sexratio edu_gap share_wpar1 w_managerial ///
267 age edu2 edu3 oc_* ///

```

```

268             married numfu numch yr_2-yr_10 if country=="Spain"
269   outreg2 using "Output/ByCountry.xls", bdec(3) ctitle(Spain - Males) drop (oc_*
   yr_* o.*) se append
270   regress w_workhours lwage lw_wage cross_lwage famnlinc           ///
271             w_share_nlab sexratio edu_gap share_wpar1 w_managerial  ///
272             w_age w_edu2 w_edu3 w_oc_*                             ///
273             married numfu numch yr_2-yr_10 if country=="Spain"

274   outreg2 using "Output/ByCountry.xls", bdec(3) ctitle(Spain - Females) drop (w_oc_*
   yr_* o.*) se append

275
276
277   /*****
278     Define list of variables to simplify code
279     *****/
280
281   *   Main regressors
282   global depvars lwage lw_wage cross_lwage famnlinc
283
284   *   Distribution factors
285   global dfvars w_share_nlab sexratio edu_gap share_wpar1 w_managerial
286
287   *   Male dependent variables
288   global Hvars age edu2 edu3 oc_* married numfu numch
289
290   *   Female dependent variables
291   global Wvars w_age w_edu2 w_edu3 w_oc_* married numfu numch
292
293   *   Year fixed effecs
294   global yrfe yr_2-yr_10
295
296
297   /*****
298     SURE estimates
299     - Loop over countries of interest to simplify code
300     - Compute estimates
301     - Export estimates
302     *****/
303
304   *   Define local list of countries; start loop
305   local ctrylist "Greece Italy Portugal Spain"
306   foreach cc of local ctrylist {
307     sureg (workhours lwage lw_wage cross_lwage famnlinc           ///
308           w_share_nlab sexratio edu_gap share_wpar1             ///
309           w_managerial                                           ///
310           age edu2 edu3 oc_* married numfu numch               ///
311           yr_2-yr_10 )                                           ///
312     (w_workhours lwage lw_wage cross_lwage famnlinc             ///
313       w_share_nlab sexratio edu_gap share_wpar1                 ///
314       w_managerial                                               ///
315       w_age w_edu2 w_edu3 w_oc_* married numfu numch           ///
316       yr_2-yr_10 )                                              ///
317     if country=="`cc'"
318     outreg2 using "Output/SURE.xls", bdec(3) ctitle(`cc')       ///
319     drop(oc_* w_oc_* yr_* o.*) se append
320   }
321

```

322 *** End of dofile