



MAESTRÍA EN ECONOMÍA

Economía Aplicada

PROF. MARTIN A. ROSSI

TUTORES: PAOLA LLAMAS Y TOMAS
PACHECO

**Problem Set 2: Multiple
hypothesis testing**

Garcia Ojeda, Juan
Hausvirth, Martina
Hayduk, Gaspar
Salvatierra, Elias Lucas D.

Fecha de entrega: 30 de agosto de 2024

PROBLEM SET 2: TEST DE MÚLTIPLES HIPÓTESIS

GARCIA OJEDA - HAUSVIRTH - HAYDUK - SALVATIERRA

En este problem set se corrige la inferencia del artículo de Attanasio, Cattan, Fitzsimons, Meghir y Rubio-Codina (2020), que investiga los mecanismos por los cuales intervenciones en la niñez tienen efectos en habilidades socioemocionales y cognitivas de los niños.

1. Nuestra Table 1 replica particularmente la Tabla 2 del artículo mencionado, que muestra los resultados de visitas a los ^{typo}hogares en cuatro outcomes: desarrollo cognitivo, desarrollo socioemocional, inversión de los padres en materiales de juego e inversión de los padres en actividades de juego. A continuación se presenten las regresiones en una tabla que contiene los cuatro paneles, donde las columnas refieren a los distintos outcomes y la fila a la variable de tratamiento.

TABLE 1.
Treatment Impacts on Raw Measures and Latent Factors

Panel A: Child's cognitive skills at follow-up						
	(1)	(2)	(3)	(4)	(5)	(6)
	Cognitive	Receptive language	Expressive language	Fine motor	Words	Complex phrases
	b/se	b/se	b/se	b/se	b/se	b/se
Treatment Effect	0.250	0.174	0.029	0.073	0.086	0.057
	(0.063)	(0.063)	(0.062)	(0.059)	(0.064)	(0.056)
Number of Observations	1263	1263	1262	1261	1321	1321
Panel B: Child's socio-emotional skills at follow-up						
	(1)	(2)	(3)	(4)	(5)	
	ICQ: Difficult	ICQ:	ICQ:	ECBQ:	ECBQ:	
	(-)	Unsociable	Unstoppable	Inhibitory control	Attention	
	b/se	(-)	(-)	b/se	b/se	
Treatment Effect	0.073	0.037	0.031	0.000	0.072	
	(0.045)	(0.055)	(0.054)	(0.058)	(0.048)	
Number of Observations	1325	1325	1325	1322	1322	
Panel C: Material investments at follow-up						
	(1)	(2)	(3)	(4)	(5)	
	FCI: Number	FCI: Number	FCI: Number	FCI: Number	FCI: Number	
	of types	of coloring	of toys	of toys	of	
	of play	and drawing	to learn	to learn	shop-bought	
	materials	books	movement	shapes	toys	
	b/se	b/se	b/se	b/se	b/se	
Treatment Effect	0.217	-0.124	-0.049	0.426	0.020	
	(0.064)	(0.056)	(0.065)	(0.088)	(0.061)	
Number of Observations	1325	1325	1325	1325	1325	
Panel D: Time investment at follow-up						
	FCI: Number	FCI: Number	FCI: Number	FCI: Number	FCI: Number	
	of types	of times	of times	of times	of times	
	of play	told a story	read to child	played with	named things	
	activities	to child	in last	toys in last	to child	
	in last 3 days	in last 3 days	3 days	3 days	in last 3 days	
	b/se	b/se	b/se	b/se	b/se	
Treatment Effect	0.280	0.153	0.403	0.183	0.144	
	(0.051)	(0.064)	(0.068)	(0.060)	(0.048)	
Number of Observations	1325	1325	1325	1325	1325	

2. Enfrentamos un problema de múltiples hipótesis cuando lo que queremos testear contiene múltiples outcomes, tratamientos, o subdimensiones. Es necesario corregir la inferencia, dado que los procedimientos estándar que se utilizan habitualmente, dejan de ser confiables cuando se evalúan múltiples hipótesis simultáneamente. El efecto del tratamiento puede dar significativo por azar, y no porque tenga algún efecto real sobre las variables relevantes. Esto es así dado que la probabilidad de cometer al menos un falso positivo crece exponencialmente con el número de hipótesis. Por ejemplo, de testear simultáneamente K hipótesis, es decir, h_1, h_2, \dots, h_K , con un nivel de significatividad de α . Entonces, en cada test, la probabilidad de cometer un falso positivo es α , mientras que la probabilidad de cometer al menos un falso positivo o Family-Wise Error Rate (FWER) crece exponencialmente con el número de hipótesis:

$$\begin{aligned} FWER &= Prob(h_1 \neq FP \vee \dots \vee h_K \neq FP) \\ &= 1 - Prob(h_1 \neq FP \wedge \dots \wedge h_K \neq FP) \\ &= 1 - [Prob(h_1 \neq FP) \times \dots \times Prob(h_K \neq FP)] \\ &= 1 - (1 - \alpha)^K \end{aligned}$$

De esta manera, como mencionamos antes, comparar individualmente los p-values con α puede llevar a conclusiones erróneas.

3. En Table 2 se corrige los p-values utilizando las tres estrategias vistas: Bonferroni; Holm; y Benjamini, Krieger y Yekutieli, y a partir de allí, se exporta una tabla idéntica a la del punto 1 agregando tres filas con los p-values corregidos, que presenta los impactos estimados de recibir visitas domiciliarias en el desarrollo cognitivo, el desarrollo emocional, la inversión de los padres en materiales de juego y la inversión de los padres en actividades de juego.

4. La corrección de **Bonferroni** es la forma mas simple, pero también la mas restrictiva de corregir el problema de múltiples hipótesis. Esto ocurre dado que le quita mucho *power* a la estimación, y la hipótesis nula (H_0) se vuelve muy difícil de rechazar. Lo que hace es ajustar el FWER¹. En una estimación con m hipótesis y un nivel de significatividad α , el ajuste que se hará sobre el nivel de significatividad será α/m .

Por otro lado, el ajuste de **Holm**, si bien es menos restrictiva que Bonferroni, también utiliza el supuesto de independencia para realizar la corrección, lo cual es bastante fuerte, y difícilmente generalizable. Este ajuste lo que hace es ordenar los *p-values* de menor a mayor y comparar, de manera decreciente, cada valor contra el “ranking” del *p-value* $\frac{\alpha}{K+1-k}$. A diferencia del primero, con este método es mas fácil rechazar H_0 . ok queda mas academico si decian que tiene mas poder estadístico.

Finalmente, la corrección de **Benjamini, Krieger & Yekutieli**, es la única de las tres que tiene en consideración la dependencia de los outcomes. Al igual que para Holm, se corre la regresión para cada outcome y se guardan los *p-values* originales de forma decreciente. Sin embargo, lo que hace esta estrategia, a diferencia de las otras dos, es hacer una permutación de la data bajo la hipótesis de no significatividad (es decir, se varía el status de cada individuo aleatoriamente), para luego computar los nuevos *p-values* (P_r^{**}), y ordenarlos de forma decreciente.

¹Family Wise Error Rate

TABLE 2.
Treatment Impacts on Raw Measures and Latent Factors

Panel A: Child's cognitive skills at follow-up						
	(1)	(2)	(3)	(4)	(5)	(6)
	Cognitive	Receptive language	Expressive language	Fine motor	Words	Complex phrases
	b/se	b/se	b/se	b/se	b/se	b/se
Treatment Effect	0.250	0.174	0.029	0.073	0.086	0.057
	(0.063)	(0.063)	(0.062)	(0.059)	(0.064)	(0.056)
P_value	0.000	0.007	0.644	0.224	0.179	0.315
Bonferroni Corrected P_value	0.003	0.147	1.000	1.000	1.000	1.000
Holm Corrected Significance	0.003	0.004	0.017	0.006	0.056	0.007
BKY Corrected P_value	0.001	0.013	0.407	0.215	0.179	0.267
Number of Observations	1263	1263	1262	1261	1321	1321
Panel B: Child's socio-emotional skills at follow-up						
	(1)	(2)	(3)	(4)	(5)	
	ICQ: Difficult	ICQ:	ICQ:	ECBQ:	ECBQ:	
	(-)	Unsociable	Unstoppable	Inhibitory control	Attention	
	b/se	b/se	b/se	b/se	b/se	
Treatment Effect	0.073	0.037	0.031	0.000	0.072	
	(0.045)	(0.055)	(0.054)	(0.058)	(0.048)	
P_value	0.107	0.506	0.569	1.000	0.136	
Bonferroni Corrected P_value	1.000	1.000	1.000	1.000	1.000	
Holm Corrected Significance	0.005	0.010	0.013	0.050	0.005	
BKY Corrected P_value	0.120	0.366	0.398	0.508	0.143	
Number of Observations	1325	1325	1325	1322	1322	
Panel C: Material investments at follow-up						
	(1)	(2)	(3)	(4)	(5)	
	FCI: Number	FCI: Number	FCI: Number	FCI: Number	FCI: Number	
	of types	of coloring	of toys	of toys	of	
	of play	and drawing	to learn	to learn	shop-bought	
	materials	books	movement	shapes	toys	
	b/se	b/se	b/se	b/se	b/se	
Treatment Effect	0.217	-0.124	-0.049	0.426	0.020	
	(0.064)	(0.056)	(0.065)	(0.088)	(0.061)	
P_value	0.001	0.028	0.452	0.000	0.750	
Bonferroni Corrected P_value	0.020	0.587	1.000	0.000	1.000	
Holm Corrected Significance	0.003	0.004	0.008	0.003	0.025	
BKY Corrected P_value	0.004	0.038	0.342	0.001	0.429	
Number of Observations	1325	1325	1325	1325	1325	
Panel D: Time investment at follow-up						
	(1)	(2)	(3)	(4)	(5)	
	FCI: Number	FCI: Number	FCI: Number	FCI: Number	FCI: Number	
	of types	of times	of times	of times	of times	
	of play	told a story	read to child	played with	named things	
	activities	to child	in last	toys in last	to child	
	in last 3 days	in last 3 days	3 days	3 days	in last 3 days	
	b/se	b/se	b/se	b/se	b/se	
Treatment Effect	0.280	0.153	0.403	0.183	0.144	
	(0.051)	(0.064)	(0.068)	(0.060)	(0.048)	
P_value	0.000	0.019	0.000	0.003	0.004	
Bonferroni Corrected P_value	0.000	0.398	0.000	0.060	0.074	
Holm Corrected Significance	0.003	0.004	0.002	0.003	0.003	
BKY Corrected P_value	0.001	0.029	0.001	0.008	0.009	
Number of Observations	1325	1325	1325	1325	1325	

Falta ahondar en detalle en este punto. Cuales dejan de ser significativos?

5. En comparacion a los resultados del articulo, los nuevos resultados incorporan los p values corregidos a partir de la tres estrategias vistas: Bonferroni; Holm; y Benjamini, Krieger y Yekutielle.

A partir de estos nuevos resultados, muchos coeficientes ya no son estadisticamente significativos dado que ahora se rechaza muchas mas veces la hipotesis nula de no significatividad.

Ademas, los p-values corregidos mediante el método de Benjamini, Krieger y Yekutiele tienden a ser más similares a los reportados en el artículo, por la dependencia existente entre los outcomes que hace

que esta metodologia sea hace menos restrictiva en comparación con el resto de los métodos, que asumen independencia entre las hipótesis testadas.

```

1  /*****
2      Semana 3: Problem Set 2
3
4      Universidad de San Andrés
5      Economía Aplicada
6      2024
7  Gaspar Hayduk; Juan Gabriel García Ojeda; Elias Lucas Salvatierra; Martina Hausvirth
8  *****/
9
10
11 * Source: https://www.aeaweb.org/articles?id=10.1257/app.20200204
12
13 /*****
14 Este archivo sigue la siguiente estructura:
15
16 0) Set up environment and globals
17
18 1) Regressions (panel A, B, C y D)
19 2) Correccion de Bonferroni
20 3) Correccion de Holm
21 4) Correccion de Benjamini, Krieger, and Yekutieli (2006)
22
23 *****/
24
25
26
27 * 0) Set up environment
28 *=====
29
30 *gl main "/Users/gasparhayduk/Desktop/Economía Aplicada/aplicadaps2"
31 gl main "C:\Users\Usuario\OneDrive\Juanga\OneDrive\JUANGA\Maestria\Udesa\Economia
32 Aplicada\Problem sets\PS2"
33 gl input "$main/input"
34 gl output "$main/output"
35
36 * Open data set
37
38 use "$input/measures.dta", clear
39
40 * Global with control variables
41
42 global covs_eva "male i.eva_fu"
43 global covs_ent "male i.ent_fu"
44
45 *-- Agregamos etiquetas a la variable de interes y a los outcomes:
46
47 * Outcomes panel A
48 label var b_tot_cog1_st "Cognitive"
49 label var b_tot_lr1_st "Receptive language"
50 label var b_tot_le1_st "Expressive language"
51 label var b_tot_mf1_st "Fine motor"
52 label var mac_words1_st "Words"
53 label var mac_phrases1_st "Complex phrases"
54
55 *Outcomes panel B
56 label var bates_difficult1_st "ICQ: Difficult (-)"
57 label var bates_unsociable1_st "ICQ: Unsociable (-) "
58 label var bates_unstoppable1_st "ICQ: Unstoppable (-)"
59 label var roth_inhibit1_st "ECBQ: Inhibitory control"
60 label var roth_attention1_st "ECBQ: Attention"
61
62 *Outcomes panel C
63 label var fci_play_mat_type1_st "FCI: Number of types of play materials"
64 label var Npaintbooks1_st "FCI: Number of coloring and drawing books"
65 label var Nthingsmove1_st "FCI: Number of toys to learn movement"
66 label var Ntoysshape1_st "FCI: Number of toys to learn shapes"
67 label var Ntoysbought1_st "FCI: Number of shop-bought toys"

```

```

68 *Outcomes panel D
69 label var fci_play_act1_st "FCI: Number of types of play activities in last 3 days"
70 label var home_stories1_st "FCI: Number of times told a story to child in last 3 days"
71 label var home_read1_st "FCI: Number of times read to child in last 3 days"
72 label var home_toys1_st "FCI: Number of times played with toys in last 3 days"
73 label var home_name1_st "FCI: Number of times named things to child in last 3 days"
74
75 *Variable de tratamiento
76 label var treat "Treatment Effect"
77
78
79
80 * 1) Regressions
81 *=====*
82
83
84
85 *****
86 * PANEL A (Child's cognitive skills at follow up)
87 *****
88
89
90 local bayley "b_tot_cog b_tot_lr b_tot_le b_tot_mf"
91 eststo clear
92 foreach y of local bayley{
93 local append append
94 if "`y'"=="b_tot_cog" local append replace
95     cap drop V*
96     reg `y'1_st treat `y'0_st $covs_eva , cluster(cod_dane)
97 }
98 esttab using "$output/Cuadro1_panelA.tex", se replace label ///
99 keep(treat) ///
100 cells(b(fmt(3)) se(par fmt(3))) ///
101
102
103 local macarthur "mac_words mac_phrases"
104 foreach y of local macarthur{
105     cap drop V*
106     reg `y'1_st treat mac_words0_st $covs_ent , cluster(cod_dane)
107 }
108 esttab using "$output/Cuadro1_panelA.tex", se replace label ///
109 keep(treat) ///
110 cells(b(fmt(3)) se(par fmt(3))) ///
111
112
113 *****
114 * PANEL B (Child's socio-emotional skills at follow up)
115 *****
116
117 local bates "bates_difficult bates_unsociable bates_unstoppable"
118 eststo clear
119 foreach y of local bates{
120     cap drop V*
121     reg `y'1_st treat `y'0_st $covs_ent, cl(cod_dane)
122 }
123 esttab using "$output/Cuadro1_panelB.tex", se replace label ///
124 keep(treat) ///
125 cells(b(fmt(3)) se(par fmt(3))) ///
126
127 local roth "roth_inhibit roth_attention"
128 foreach y of local roth{
129     cap drop V*
130     reg `y'1_st treat bates_difficult0_st $covs_ent , cluster(cod_dane)
131 }
132 esttab using "$output/Cuadro1_panelB.tex", se replace label ///
133 keep(treat) ///
134 cells(b(fmt(3)) se(par fmt(3))) ///
135

```

```

136
137 *****
138 * PANEL C (Material investments)
139 *****
140
141 local fciemat "fci_play_mat_type Npaintbooks Nthingsmove Ntoysshape Ntoysbought"
142 eststo clear
143 foreach y of local fciemat{
144     cap drop V*
145     reg `y'1_st treat fci_play_mat_type0_st $covs_ent , cluster(cod_dane)
146 }
147 esttab using "$output/Cuadro1_panelC.tex", se replace label ///
148 keep(treat) ///
149 cells(b(fmt(3)) se(par fmt(3))) ///
150
151 *****
152 * PANEL D (Time investments)
153 *****
154 local fcitime "fci_play_act home_stories home_read home_toys home_name"
155 eststo clear
156 foreach y of local fcitime{
157     cap drop V*
158     reg `y'1_st treat fci_play_act0_st $covs_ent , cluster(cod_dane)
159 }
160
161 esttab using "$output/Cuadro1_panelD.tex", se replace label ///
162 keep(treat) ///
163 cells(b(fmt(3)) se(par fmt(3))) ///
164
165
166
167
168 * 2) Correccion de Bonferroni
169 *=====*
170
171 * Tenemos 21 outcomes de interes, por lo que hay que testear 21 hipotesis. Guardamos ese numero
172 en un escalar:
173 scalar hyp = 21
174
175 * si tengo m hipotesis uso alpha/m como nivel de significancia.
176 * por lo que debemos ver si los p-valores < alpha/m
177 * por comodidad, veremos si m*p-valores < alpha (el codigo hace esto).
178
179 * Hay que correr las regresiones otra vez, almacenar el p-valor y generar el nuevo p-valor
180 corregido. Hacemos esto para cada panel.
181
182 *****
183 * PANEL A (Child's cognitive skills at follow up)
184 *****
185
186 local bayley "b_tot_cog b_tot_lr b_tot_le b_tot_mf"
187 eststo clear
188 foreach y of local bayley{
189     local append append
190     if "`y'"=="b_tot_cog" local append replace
191     cap drop V*
192     reg `y'1_st treat `y'0_st $covs_eva , cluster(cod_dane)
193     eststo: test treat = 0
194     estadd scalar p_value = r(p)
195     estadd scalar corr_p_value = min(1,r(p)*hyp)
196 }
197 esttab using "$output/Cuadro2_panelA.tex", se replace label ///
198 keep(treat) ///
199 cells(b(fmt(3)) se(par fmt(3))) ///
200 stats(p_value corr_p_value blank N, fmt(3 3 0) labels("P_value" "Bonferroni Corrected P_value" " "
    "Number of Observations"))

```



```

201
202 local macarthur "mac_words mac_phrases"
203 foreach y of local macarthur{
204     cap drop V*
205     reg `y'1_st treat mac_words0_st $covs_ent , cluster(cod_dane)
206     eststo: test treat = 0
207     estadd scalar p_value = r(p)
208     estadd scalar corr_p_value = min(1,r(p)*hyp)
209 }
210 esttab using "$output/Cuadro2_panelA.tex", se replace label ///
211 keep(treat) ///
212 cells(b(fmt(3)) se(par fmt(3))) ///
213 stats(p_value corr_p_value blank N, fmt(3 3 0) labels("P_value" "Bonferroni Corrected P_value" " "
    "Number of Observations"))
214
215 *****
216 * PANEL B (Child's socio-emotional skills at follow up)
217 *****
218
219 local bates "bates_difficult bates_unsociable bates_unstoppable"
220 eststo clear
221 foreach y of local bates{
222     cap drop V*
223     reg `y'1_st treat `y'0_st $covs_ent, cl(cod_dane)
224     eststo: test treat = 0
225     estadd scalar p_value = r(p)
226     estadd scalar corr_p_value = min(1,r(p)*hyp)
227 }
228 esttab using "$output/Cuadro2_panelB.tex", se replace label ///
229 keep(treat) ///
230 cells(b(fmt(3)) se(par fmt(3))) ///
231 stats(p_value corr_p_value blank N, fmt(3 3 0) labels("P_value" "Bonferroni Corrected P_value" " "
    "Number of Observations"))
232
233 local roth "roth_inhibit roth_attention"
234 foreach y of local roth{
235     cap drop V*
236     reg `y'1_st treat bates_difficult0_st $covs_ent , cluster(cod_dane)
237     eststo: test treat = 0
238     estadd scalar p_value = r(p)
239     estadd scalar corr_p_value = min(1,r(p)*hyp)
240 }
241 esttab using "$output/Cuadro2_panelB.tex", se replace label ///
242 keep(treat) ///
243 cells(b(fmt(3)) se(par fmt(3))) ///
244 stats(p_value corr_p_value blank N, fmt(3 3 0) labels("P_value" "Bonferroni Corrected P_value" " "
    "Number of Observations"))
245
246 *****
247 * PANEL C (Material investments)
248 *****
249
250 local fcimat "fci_play_mat_type Npaintbooks Nthingsmove Ntoysshape Ntoysbought"
251 eststo clear
252 foreach y of local fcimat{
253     cap drop V*
254     reg `y'1_st treat fci_play_mat_type0_st $covs_ent , cluster(cod_dane)
255     eststo: test treat = 0
256     estadd scalar p_value = r(p)
257     estadd scalar corr_p_value = min(1,r(p)*hyp)
258 }
259 esttab using "$output/Cuadro2_panelC.tex", se replace label ///
260 keep(treat) ///
261 cells(b(fmt(3)) se(par fmt(3))) ///
262 stats(p_value corr_p_value blank N, fmt(3 3 0) labels("P_value" "Bonferroni Corrected P_value" " "
    "Number of Observations"))
263
264 *****

```

```

265 * PANEL D (Time investments)
266 *****
267 local fcitime "fci_play_act home_stories home_read home_toys home_name"
268 eststo clear
269 foreach y of local fcitime{
270     cap drop V*
271     reg `y'1_st treat fci_play_act0_st $covs_ent , cluster(cod_dane)
272     eststo: test treat = 0
273     estadd scalar p_value = r(p)
274     estadd scalar corr_p_value = min(1,r(p)*hyp)
275 }
276 esttab using "$output/Cuadro2_panelD.tex", se replace label ///
277 keep(treat) ///
278 cells(b(fmt(3)) se(par fmt(3))) ///
279 stats(p_value corr_p_value blank N, fmt(3 3 0) labels("P_value" "Bonferroni Corrected P_value" " "
    "Number of Observations"))
280
281
282
283
284 * 3) Correccion de Holm
285 *=====
286
287 * Defino el nivel de significancia:
288 scalar signif = 0.05
289 * Defino la cantidad de hipotesis:
290 scalar hyp = 21
291
292 * Creo una matriz de 21 filas y 1 columna donde almaceno los p-valores:
293 mat p_values = J(21,1,.)
294
295 * Creo la variable para iterar:
296 scalar i = 1
297
298 *-----Ahora debo rellenar la matriz de p-valores. Debo correr todas las regresiones:
299
300 *--- PANEL A (Child's cognitive skills at follow up)
301 local bayley "b_tot_cog b_tot_lr b_tot_le b_tot_mf"
302 foreach y of local bayley{
303     local append append
304     if "`y'"=="b_tot_cog" local append replace
305     cap drop V*
306     reg `y'1_st treat `y'0_st $covs_eva , cluster(cod_dane)
307     eststo: test treat = 0
308     mat p_values[i,1] = r(p)
309     scalar i = i +1
310 }
311
312 local macarthur "mac_words mac_phrases"
313 foreach y of local macarthur{
314     cap drop V*
315     reg `y'1_st treat mac_words0_st $covs_ent , cluster(cod_dane)
316     eststo: test treat = 0
317     mat p_values[i,1] = r(p)
318     scalar i = i +1
319 }
320
321 *--- PANEL B (Child's socio-emotional skills at follow up)
322 local bates "bates_difficult bates_unsociable bates_unstoppable"
323 foreach y of local bates{
324     cap drop V*
325     reg `y'1_st treat `y'0_st $covs_ent, cl(cod_dane)
326     eststo: test treat = 0
327     mat p_values[i,1] = r(p)
328     scalar i = i +1
329 }
330
331 local roth "roth_inhibit roth_attention"

```

```

332   foreach y of local roth{
333       cap drop V*
334       reg `y'1_st treat bates_difficult0_st $covs_ent , cluster(cod_dane)
335       eststo: test treat = 0
336       mat p_values[i,1] = r(p)
337       scalar i = i +1
338   }
339
340   *--- PANEL C (Material investments)
341   local fcimat "fci_play_mat_type Npaintbooks Nthingsmove Ntoysshape Ntoysbought"
342   foreach y of local fcimat{
343       cap drop V*
344       reg `y'1_st treat fci_play_mat_type0_st $covs_ent , cluster(cod_dane)
345       eststo: test treat = 0
346       mat p_values[i,1] = r(p)
347       scalar i = i +1
348   }
349
350   *--- PANEL D (Time investments)
351   local fcitime "fci_play_act home_stories home_read home_toys home_name"
352   foreach y of local fcitime{
353       cap drop V*
354       reg `y'1_st treat fci_play_act0_st $covs_ent , cluster(cod_dane)
355       eststo: test treat = 0
356       mat p_values[i,1] = r(p)
357       scalar i = i +1
358   }
359
360
361   *--- Ahora trabajo con la matriz.
362
363   preserve
364   clear // borro la base que tengo abierta
365   svmat p_values // abro la matriz que guarde antes
366   gen var = _n // identificador de a qué outcome corresponde el pvalor.
367   sort p_values1 // ordeno los p-valores de menor a mayor.
368
369   gen alpha_corr = signif/(hyp+1-_n) // genero el nivel de significancia ajustado segun formula
370
371   gen significant = (p_values1<alpha_corr) // marco con 1 a aquellos p-valores que son menor que el
   nivel de significancia ajustado.
372
373   replace significant = 0 if significant[_n-1]==0 // le doy un cero a los que no son significativos.
374
375
376   *odeno:
377   sort var
378
379
380   *restauro
381   restore
382
383
384   * Almacenos los niveles de significatividad corregidos.
385   * Rechazamos la hipotesis nula de treat = 0 si el p-valor original es menor que el nivel de
   significancia ajustado
386   mat alpha_corr_h = [0.0027778, 0.0035714, 0.016667, 0.00625, 0.055556, 0.0071429, 0.0045455, 0.01,
   0.0125, 0.05, 0.005, 0.0029412, 0.0041667, 0.0083333, 0.0026316, 0.025, 0.0025, 0.0038462,
   0.002381, 0.003125, 0.0033333]
387
388
389   *---Queda agregar los niveles de significancia ajustados a la tabla. Hay que hacerlo para cada
   panel.
390
391   *definimos un iterados para poder ir accediendo a los elementos de la matriz con los niveles de
   significancia ajustados.
392   scalar i = 1
393

```

```

394 *--- PANEL A (Child's cognitive skills at follow up)
395
396 local bayley "b_tot_cog b_tot_lr b_tot_le b_tot_mf"
397 eststo clear
398 foreach y of local bayley{
399 local append append
400 if "`y'"=="b_tot_cog" local append replace
401 cap drop V*
402 reg `y'1_st treat `y'0_st $covs_eva , cluster(cod_dane)
403 eststo: test treat = 0
404 estadd scalar p_value = r(p)
405 estadd scalar corr_p_value = min(1,r(p)*hyp)
406 estadd scalar alpha_corr = alpha_corr_h[1,i]
407 scalar i = i + 1
408 }
409 esttab using "$output/Cuadro2_panelA.tex", p se replace label ///
410 keep(treat) ///
411 cells(b(fmt(3)) se(par fmt(3))) ///
412 stats(p_value corr_p_value alpha_corr blank N, fmt(3 3 3 0) labels("P_value" "Bonferroni
Corrected P_value" "Holm Corrected Significance" " " "Number of Observations"))
413
414
415 local macarthur "mac_words mac_phrases"
416 foreach y of local macarthur{
417 cap drop V*
418 reg `y'1_st treat mac_words0_st $covs_ent , cluster(cod_dane)
419 eststo: test treat = 0
420 estadd scalar p_value = r(p)
421 estadd scalar corr_p_value = min(1,r(p)*hyp)
422 estadd scalar alpha_corr = alpha_corr_h[1,i]
423 scalar i = i + 1
424 }
425 esttab using "$output/Cuadro2_panelA.tex", p se replace label ///
426 keep(treat) ///
427 cells(b(fmt(3)) se(par fmt(3))) ///
428 stats(p_value corr_p_value alpha_corr blank N, fmt(3 3 3 0) labels("P_value" "Bonferroni
Corrected P_value" "Holm Corrected Significance" " " "Number of Observations"))
429
430
431 *--- PANEL B (Child's socio-emotional skills at follow up)
432
433 local bates "bates_difficult bates_unsociable bates_unstoppable"
434 eststo clear
435 foreach y of local bates{
436 cap drop V*
437 reg `y'1_st treat `y'0_st $covs_ent, cl(cod_dane)
438 eststo: test treat = 0
439 estadd scalar p_value = r(p)
440 estadd scalar corr_p_value = min(1,r(p)*hyp)
441 estadd scalar alpha_corr = alpha_corr_h[1,i]
442 scalar i = i + 1
443 }
444 esttab using "$output/Cuadro2_panelB.tex", p se replace label ///
445 keep(treat) ///
446 cells(b(fmt(3)) se(par fmt(3))) ///
447 stats(p_value corr_p_value alpha_corr blank N, fmt(3 3 3 0) labels("P_value" "Bonferroni
Corrected P_value" "Holm Corrected Significance" " " "Number of Observations"))
448
449 local roth "roth_inhibit roth_attention"
450 foreach y of local roth{
451 cap drop V*
452 reg `y'1_st treat bates_difficult0_st $covs_ent , cluster(cod_dane)
453 eststo: test treat = 0
454 estadd scalar p_value = r(p)
455 estadd scalar corr_p_value = min(1,r(p)*hyp)
456 estadd scalar alpha_corr = alpha_corr_h[1,i]
457 scalar i = i + 1
458 }

```

```

459 esttab using "$output/Cuadro2_panelB.tex", p se replace label ///
460 keep(treat) ///
461 cells(b(fmt(3)) se(par fmt(3))) ///
462 stats(p_value corr_p_value alpha_corr blank N, fmt(3 3 3 0) labels("P_value" "Bonferroni
Corrected P_value" "Holm Corrected Significance" " " "Number of Observations"))
463
464
465 *--- PANEL C (Material investments)
466
467 local fcimat "fci_play_mat_type Npaintbooks Nthingsmove Ntoysshape Ntoysbought"
468 eststo clear
469 foreach y of local fcimat{
470     cap drop V*
471     reg `y'1_st treat fci_play_mat_type0_st $covs_ent , cluster(cod_dane)
472     eststo: test treat = 0
473     estadd scalar p_value = r(p)
474     estadd scalar corr_p_value = min(1,r(p)*hyp)
475     estadd scalar alpha_corr = alpha_corr_h[1,i]
476     scalar i = i + 1
477 }
478
479 esttab using "$output/Cuadro2_panelC.tex", p se replace label ///
480 keep(treat) ///
481 cells(b(fmt(3)) se(par fmt(3))) ///
482 stats(p_value corr_p_value alpha_corr blank N, fmt(3 3 3 0) labels("P_value" "Bonferroni
Corrected P_value" "Holm Corrected Significance" " " "Number of Observations"))
483
484
485 *--- PANEL D (Time investments)
486
487 local fcitime "fci_play_act home_stories home_read home_toys home_name"
488 eststo clear
489 foreach y of local fcitime{
490     cap drop V*
491     reg `y'1_st treat fci_play_act0_st $covs_ent , cluster(cod_dane)
492     eststo: test treat = 0
493     estadd scalar p_value = r(p)
494     estadd scalar corr_p_value = min(1,r(p)*hyp)
495     estadd scalar alpha_corr = alpha_corr_h[1,i]
496     scalar i = i + 1
497 }
498
499 esttab using "$output/Cuadro2_panelD.tex", se replace label ///
500 keep(treat) ///
501 cells(b(fmt(3)) se(par fmt(3))) ///
502 stats(p_value corr_p_value alpha_corr blank N, fmt(3 3 3 0) labels("P_value" "Bonferroni
Corrected P_value" "Holm Corrected Significance" " " "Number of Observations"))
503
504
505
506
507
508 * 4) Correccion de Benjamini, Krieger, and Yekutieli (2006)
509 *=====
510
511 *---Ya tenemos los p_valores guardados en la matriz p_values. La abro y la guardo como un .dta:
512 preserve
513 clear
514 * abro la matriz para tratarla como una base de datos.
515 svmat p_values
516 gen outcome = _n
517 rename p_values1 pval
518 save "$output/pvals.dta", replace
519 restore
520
521
522 *--- Ahora toca usar Michael Anderson's code for sharpened q-values.
523 * el input es la base de datos con los p-valores originales (que debo corregirlos)

```

```

524 * el output seran los p-valores corregidos
525 preserve
526
527 use "$output/pvals.dta", clear
528 version 10
529 set more off
530
531 * Collect the total number of p-values tested
532 quietly sum pval
533 local totalpvals = r(N)
534
535 * Sort the p-values in ascending order and generate a variable that codes each p-value's rank
536 quietly gen int original_sorting_order = _n
537 quietly sort pval
538 quietly gen int rank = _n if pval~=.
539
540 * Set the initial counter to 1
541 local qval = 1
542
543 * Generate the variable that will contain the BKY (2006) sharpened q-values
544 gen bky06_qval = 1 if pval~=.
545
546 * Set up a loop that begins by checking which hypotheses are rejected at q = 1.000, then checks
  which hypotheses are rejected at q = 0.999, then checks which hypotheses are rejected at q =
  0.998, etc. The loop ends by checking which hypotheses are rejected at q = 0.001.
547
548 while `qval' > 0 {
549     * First Stage
550     * Generate the adjusted first stage q level we are testing: q' = q/1+q
551     local qval_adj = `qval'/(1+`qval')
552     * Generate value q'*r/M
553     gen fdr_temp1 = `qval_adj'*rank/`totalpvals'
554     * Generate binary variable checking condition p(r) <= q'*r/M
555     gen reject_temp1 = (fdr_temp1>=pval) if pval~=.
556     * Generate variable containing p-value ranks for all p-values that meet above condition
557     gen reject_rank1 = reject_temp1*rank
558     * Record the rank of the largest p-value that meets above condition
559     egen total_rejected1 = max(reject_rank1)
560
561     * Second Stage
562     * Generate the second stage q level that accounts for hypotheses rejected in first stage:
    q_2st = q'*(M/m0)
563     local qval_2st = `qval_adj'*(`totalpvals'/(`totalpvals'-total_rejected1[1]))
564     * Generate value q_2st*r/M
565     gen fdr_temp2 = `qval_2st'*rank/`totalpvals'
566     * Generate binary variable checking condition p(r) <= q_2st*r/M
567     gen reject_temp2 = (fdr_temp2>=pval) if pval~=.
568     * Generate variable containing p-value ranks for all p-values that meet above condition
569     gen reject_rank2 = reject_temp2*rank
570     * Record the rank of the largest p-value that meets above condition
571     egen total_rejected2 = max(reject_rank2)
572
573     * A p-value has been rejected at level q if its rank is less than or equal to the rank of the
    max p-value that meets the above condition
574     replace bky06_qval = `qval' if rank <= total_rejected2 & rank~=.
575     * Reduce q by 0.001 and repeat loop
576     drop fdr_temp* reject_temp* reject_rank* total_rejected*
577     local qval = `qval' - .001
578 }
579
580
581 quietly sort original_sorting_order
582 pause off
583 set more on
584
585 display "Code has completed."
586 display "Benjamini Krieger Yekutieli (2006) sharpened q-vals are in variable 'bky06_qval'"
587 display "Sorting order is the same as the original vector of p-values"

```



```

588
589 keep outcome pval bky06_qval
590 * ACA ESTÁ EL OUTPUT. es una base de datos con los p-valores corregidos.
591 save "$output/sharpenedqvals.dta", replace
592
593 * hay que notar que esta correccion puede reducir los p-valores.
594 * esto puede ocurrir cuando los outcomes estan muy correlacionados entre sí.
595
596 restore
597
598 *-- estas lineas son para ver el output
599 preserve
600 use "$output/sharpenedqvals.dta", clear
601
602 restore
603 *--
604
605 *-- Almaceno los resultados en una matriz:
606
607 mat bky_pval = [0.001, 0.013, 0.407, 0.215, 0.179, 0.267, 0.12, 0.366, 0.398, 0.508, 0.143, 0.004,
608 0.038, 0.342, 0.001, 0.429, 0.001, 0.029, 0.001, 0.008, 0.009]
609
610 *--Queda agregar estos p-valores ajustados a la tabla. Como antes, corro las regresiones y agrego
611 por panel.
612
613 *definimos un iterados para poder ir accediendo a los elementos de la matriz con los niveles de
614 significancia ajustados.
615 scalar i = 1
616
617 *--- PANEL A (Child's cognitive skills at follow up)
618
619 local bayley "b_tot_cog b_tot_lr b_tot_le b_tot_mf"
620 eststo clear
621 foreach y of local bayley{
622   local append append
623   if "`y'"=="b_tot_cog" local append replace
624   cap drop V*
625   reg `y'1_st treat `y'0_st $covs_eva , cluster(cod_dane)
626   eststo: test treat = 0
627   estadd scalar p_value = r(p)
628   estadd scalar corr_p_value = min(1,r(p)*hyp)
629   estadd scalar alpha_corr = alpha_corr_h[1,i]
630   estadd scalar bky_pval = bky_pval[1,i]
631   scalar i = i + 1
632 }
633 esttab using "$output/Cuadro2_panelA.tex", p se replace label ///
634 keep(treat) ///
635 cells(b(fmt(3)) se(par fmt(3))) ///
636 stats(p_value corr_p_value alpha_corr bky_pval blank N, fmt(3 3 3 3 0) labels("P_value"
637 "Bonferroni Corrected P_value" "Holm Corrected Significance" "BKY Corrected P_value" " " "Number
638 of Observations"))
639
640
641 local macarthur "mac_words mac_phrases"
642 foreach y of local macarthur{
643   cap drop V*
644   reg `y'1_st treat mac_words0_st $covs_ent , cluster(cod_dane)
645   eststo: test treat = 0
646   estadd scalar p_value = r(p)
647   estadd scalar corr_p_value = min(1,r(p)*hyp)
648   estadd scalar alpha_corr = alpha_corr_h[1,i]
649   estadd scalar bky_pval = bky_pval[1,i]
650   scalar i = i + 1
651 }
652 esttab using "$output/Cuadro2_panelA.tex", p se replace label ///
653 keep(treat) ///
654 cells(b(fmt(3)) se(par fmt(3))) ///

```

```

651 stats(p_value corr_p_value alpha_corr bky_pval blank N, fmt(3 3 3 3 0) labels("P_value"
"Bonferroni Corrected P_value" "Holm Corrected Significance" "BKY Corrected P_value" " " "Number
of Observations"))
652
653
654 *--- PANEL B (Child's socio-emotional skills at follow up)
655
656 local bates "bates_difficult bates_unsociable bates_unstoppable"
657 eststo clear
658 foreach y of local bates{
659     cap drop V*
660     reg `y'1_st treat `y'0_st $covs_ent, cl(cod_dane)
661     eststo: test treat = 0
662     estadd scalar p_value = r(p)
663     estadd scalar corr_p_value = min(1,r(p)*hyp)
664     estadd scalar alpha_corr = alpha_corr_h[1,i]
665     estadd scalar bky_pval = bky_pval[1,i]
666     scalar i = i + 1
667 }
668 esttab using "$output/Cuadro2_panelB.tex", p se replace label ///
669 keep(treat) ///
670 cells(b(fmt(3)) se(par fmt(3))) ///
671 stats(p_value corr_p_value alpha_corr bky_pval blank N, fmt(3 3 3 3 0) labels("P_value"
"Bonferroni Corrected P_value" "Holm Corrected Significance" "BKY Corrected P_value" " " "Number
of Observations"))
672
673 local roth "roth_inhibit roth_attention"
674 foreach y of local roth{
675     cap drop V*
676     reg `y'1_st treat bates_difficult0_st $covs_ent , cluster(cod_dane)
677     eststo: test treat = 0
678     estadd scalar p_value = r(p)
679     estadd scalar corr_p_value = min(1,r(p)*hyp)
680     estadd scalar alpha_corr = alpha_corr_h[1,i]
681     estadd scalar bky_pval = bky_pval[1,i]
682     scalar i = i + 1
683 }
684 esttab using "$output/Cuadro2_panelB.tex", p se replace label ///
685 keep(treat) ///
686 cells(b(fmt(3)) se(par fmt(3))) ///
687 stats(p_value corr_p_value alpha_corr bky_pval blank N, fmt(3 3 3 3 0) labels("P_value"
"Bonferroni Corrected P_value" "Holm Corrected Significance" "BKY Corrected P_value" " " "Number
of Observations"))
688
689
690 *--- PANEL C (Material investments)
691
692
693 local fcimat "fci_play_mat_type Npaintbooks Nthingsmove Ntoysshape Ntoysbought"
694 eststo clear
695 foreach y of local fcimat{
696     cap drop V*
697     reg `y'1_st treat fci_play_mat_type0_st $covs_ent , cluster(cod_dane)
698     eststo: test treat = 0
699     estadd scalar p_value = r(p)
700     estadd scalar corr_p_value = min(1,r(p)*hyp)
701     estadd scalar alpha_corr = alpha_corr_h[1,i]
702     estadd scalar bky_pval = bky_pval[1,i]
703     scalar i = i + 1
704 }
705
706 esttab using "$output/Cuadro2_panelC.tex", p se replace label ///
707 keep(treat) ///
708 cells(b(fmt(3)) se(par fmt(3))) ///
709 stats(p_value corr_p_value alpha_corr bky_pval blank N, fmt(3 3 3 3 0) labels("P_value"
"Bonferroni Corrected P_value" "Holm Corrected Significance" "BKY Corrected P_value" " " "Number
of Observations"))
710

```



```

711
712 *--- PANEL D (Time investments)
713
714 local fcitime "fci_play_act home_stories home_read home_toys home_name"
715 eststo clear
716 foreach y of local fcitime{
717     cap drop V*
718     reg `y'1_st treat fci_play_act0_st $covs_ent , cluster(cod_dane)
719     eststo: test treat = 0
720     estadd scalar p_value = r(p)
721     estadd scalar corr_p_value = min(1,r(p)*hyp)
722     estadd scalar alpha_corr = alpha_corr_h[1,i]
723     estadd scalar bky_pval = bky_pval[1,i]
724     scalar i = i + 1
725 }
726 esttab using "$output/Cuadro2_panelD.tex", p se replace label ///
727 keep(treat) ///
728 cells(b(fmt(3)) se(par fmt(3))) ///
729 stats(p_value corr_p_value alpha_corr bky_pval blank N, fmt(3 3 3 3 0) labels("P_value"
"Bonferroni Corrected P_value" "Holm Corrected Significance" "BKY Corrected P_value" " " "Number
of Observations"))
730
731
732
733
734
735
736
737
738 *****-----FIN-----*****
*****
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762

```