

CONTROL 1
IN540 “MÉTODOS ESTADÍSTICOS PARA LA ECONOMÍA Y GESTIÓN”

PROFESOR: Cristóbal Humeus
SEMESTRE: Primavera 2006

**NO NECESITA NI CALCULADORA NI FORMULA. SI USTED CONSIDERA
QUE NO TIENE INFORMACION SUFICIENTE PARA CONTESTAR LA
PREGUNTA INDIQUE QUE INFORMACION NECESITARIA.**

1. **Pregunta Empírica (45 puntos)** Un alumno del DII, Pablo, está estudiando el efecto de la educación Pre-Escolar (haber asistido a pre kinder, kinder) en los resultados de la prueba SIMCE de Cuarto Básico. La regresión que está estimando es:

$$\begin{aligned} \text{simce}_{i,j,t} = & \alpha_0 + \gamma \text{Masculino} + \beta_1 \text{Educ}^M + \beta_2 \text{Educ}^P + \theta_2 \text{DKin2} + \theta_3 \text{DKin3} \\ & + \theta_4 \text{DKin4} + \text{DReg} + \text{Dndep} + \epsilon_{i,j,t} \end{aligned} \quad (1)$$

donde $\text{simce}_{i,j,t}$ es el resultado de la prueba simce de matemática del alumno i , en el colegio j en el año t , Masculino es una dummy igual a 1 si es niño (0 si es niña), Educ^M es la educación de la madre, Educ^P es la educación del padre, DKin2 es una dummy si el niño fue sólo a Pre Kinder, DKin3 es una dummy si el niño fue sólo a Kinder, y DKin4 es una dummy si el niño fue a Pre Kinder y Kinder. La variable omitida es que el niño no tiene educación pre-escolar. DReg son dummies de regiones (región omitida es la primera) y Dndep son dummies que indican la dependencia del establecimiento (municipal, subvencionado o particular pagado). La variable omitida es establecimiento municipal, dndep2 es subvencionado y dndep3 es particular pagado.

- (a) **(5 pts)** Los resultados de MCO de esta regresión son los de la tabla 1, columnas (1) a (3). A partir de estos resultados calcule en cuántos puntos subirían los resultados del SIMCE de cuarto básico si todo los niños fueran a pre-kinder y kinder.¹
- (b) **(5 pts)** Por qué disminuyen los coeficientes de las variables DKin2 , DKin3 y DKin4 cuando se incorporan las variables de educación de la madre y educación del padre en la regresión?
- (c) **(5 pts)** El alumno realiza los siguientes test: (a) $\text{dpad4} = \text{dmad4}$, (b) $\text{dpad2} = \text{dmad4}$, donde la hipótesis alternativa es $\text{dpad4} < \text{dmad4}$ y $\text{dpad2} < \text{dmad4}$ respectivamente. Aceptaría o rechazaría estos test a un nivel de significancia de un 5%.
- (d) **(10 pts)** Una amiga de Pablo, Catalina, le dice que tiene un problema de variable omitida y por lo tanto los coeficientes que obtiene para DKin2 , DKin3 y DKin4 son incorrectos. Catalina sostiene que la variable omitida es la habilidad del padre del alumno i , A_i . Si Catalina tuviera razón, como cree que esa variable omitida afectará los coeficientes de DKin2 , DKin3 y DKin4 . Es decir, los coeficientes de la tabla 1 son mayores, menores o iguales a los verdaderos coeficientes. En su respuesta sea **claro** de qué está asumiendo sobre la correlación de A_i y la variable de interés.

¹Para esto necesitará la frecuencia simple de cuántos niños atienden cada una de las categorías de pre-kinder, kinder que se encuentra en la hoja adjunta.

- (e) **(10 ptos)** Catalina le sugiere que use como Instrumento la variable distancia (kilometros) al colegio con pre-kinder y kinder más cercano de donde vivía el niño cuando tenía 4 años. Bajo qué supuesto sería un instrumento válido y bajo qué supuesto sería un instrumento no válido.
- (f) **(10 ptos)** Después de mucho, Pablo se da cuenta que Catalina tiene razón y estima la primera etapa de la variable instrumental. El test F de la variable excluida es 6.5. Qué puede decir sobre los coeficientes y sus errores estándar que obtendrá Pablo en la segunda etapa de la VI?

2. **(20 puntos)** Considere el siguiente modelo de oferta y demanda:²

$$\ln(D) = \alpha_0 + \alpha_1 \ln(P) + \alpha_2 x_2 + \epsilon^D \quad (2)$$

$$\ln(O) = \beta_0 + \beta_1 \ln(P) + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \epsilon^O \quad (3)$$

- (a) Encuentre el precio y la demanda en equilibrio en función de los parámetros del modelo.
- (b) Escriba las formas reducidas de la demanda/oferta y el precio en términos de los π (al igual que en clases).
- (c) Encuentra las soluciones de los α 's y β 's en función de los π 's que usó en la parte (b).
- (d) ¿Qué parámetros del modelo se pueden identificar? Justifique económicamente.
- (e) Explique brevemente (no más de 3 líneas) cómo estimaría esos parámetros. Es decir, ¿qué método de estimación usaría?
3. **Verdadero/Falso/Incierto (35 puntos)** Para cada una de las siguientes afirmaciones indique si son verdaderas, falsas o inciertas. No ocupe más de 4 líneas. Justifique claramente su respuesta.

- (a) Todo estimador que es consistente también es insesgado.
- (b) En la regresión $y = x\beta + \epsilon$, donde ϵ tiene media cero y varianza σ_ϵ^2 . Además se cumple que $E(\epsilon|x) = 0$. Si y es medido con error, es decir sólo observamos $y^* = y + \nu$, donde ν tiene media cero y varianza σ_ν , esto no tiene ningún efecto en la estimación de β ni en los errores estándar de β . Asuma que ϵ y ν son independientes.
- (c) En el trabajo de los mellizos visto en clases, el instrumento de educación usado por los autores fue la educación de cada mellizo reportado por los padres.
- (d) En el método de estimación de Variable Instrumental, siempre es mejor tener mientras más instrumentos que menos.
- (e) Un alumno estimó la siguiente regresión:

$$\ln(y_{i,t}) = \alpha_0 + \beta_1 \ln(x_{i,t}) + \epsilon_{i,t}$$

donde $y_{i,t}$ es la utilidad en miles de millones de pesos antes de impuestos de la empresa i en el año t y $x_{i,t}$ es la cantidad de horas trabajadas por los trabajadores de la empresa i en el año t . Los resultados obtenidos son: $\hat{\alpha}_0 = 5.03$, con error estándar de 2.45 y $\hat{\beta}_1 = 0.4$ con error estándar 0.13. El R^2 de esta regresión fue 0.4. Si en vez de usar los datos de utilidades antes de impuestos se usan los datos después de impuestos $y_{i,t}^* = y_{i,t}(1 - \tau)$ donde $\tau = 0.3$ es la tasa de impuestos. Los coeficientes estimados no cambian porque τ es el mismo para todos.

²Este es el mismo problema de la Tarea 3.

```
. reg xmat4b02 dreg* dsex* dkin* dndep*
```

Source	SS	df	MS	
Model	65961855.8	18	3664547.54	Number of obs = 213796
Residual	549003954213777	2568	11516	F(18,213777) = 1426.94
Total	614965810213795	2876	42747	Prob > F = 0.0000
				R-squared = 0.1073
				Adj R-squared = 0.1072
				Root MSE = 50.677

xmat4b02	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
dreg2	3.5091	.8525343	4.12	0.000	1.838154 5.180046
dreg3	5.527184	1.019783	5.42	0.000	3.528434 7.525935
dreg4	3.435553	.8179976	4.20	0.000	1.832298 5.038808
dreg5	.9251314	.7124026	1.30	0.194	-.47116 2.321423
dreg6	3.41289	.7869126	4.34	0.000	1.870561 4.955219
dreg7	7.297509	.7712758	9.46	0.000	5.785828 8.80919
dreg8	7.42064	.6976042	10.64	0.000	6.053353 8.787926
dreg9	-1.010864	.7805981	-1.29	0.195	-2.540817 .5190886
dreg10	6.339421	.7466242	8.49	0.000	4.876056 7.802785
dreg11	5.749089	1.426267	4.03	0.000	2.953642 8.544536
dreg12	14.8462	1.240365	11.97	0.000	12.41511 17.27728
dreg13	.5015642	.6462237	0.78	0.438	-.7650181 1.768146
dsex1	-3.857605	.2192424	-17.60	0.000	-4.287315 -3.427896
dsex2	(dropped)				
dkin2	11.02521	.4040133	27.29	0.000	10.23336 11.81707
dkin3	16.32761	.5686151	28.71	0.000	15.21314 17.44209
dkin4	18.19609	.4199558	43.33	0.000	17.37299 19.01919
dndep2	18.77228	.238981	78.55	0.000	18.30388 19.24067
dndep3	58.83514	.4653526	126.43	0.000	57.92306 59.74722
_cons	223.0797	.7338181	304.00	0.000	221.6414 224.5179

. reg xmat4b02 dreg* dsex* dkin* dndep* dmad* dpad*

Source	SS	df	MS	
Model	96369131.2	24	4015380.47	Number of obs = 213796
Residual	518596679213771	2425	94495	F(24,213771) = 1655.18
Total	614965810213795	2876	42747	Prob > F = 0.0000
				R-squared = 0.1567
				Adj R-squared = 0.1566
				Root MSE = 49.254

xmat4b02	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
dreg2	1.710415	.8288174	2.06	0.039	.0859538	3.334876
dreg3	5.533527	.9912863	5.58	0.000	3.590631	7.476424
dreg4	5.657683	.7955222	7.11	0.000	4.09848	7.216887
dreg5	1.850768	.6926307	2.67	0.008	.4932289	3.208307
dreg6	7.808336	.766229	10.19	0.000	6.306547	9.310126
dreg7	12.23953	.7513878	16.29	0.000	10.76683	13.71223
dreg8	10.24991	.6786741	15.10	0.000	8.919728	11.5801
dreg9	3.902521	.7602574	5.13	0.000	2.412435	5.392606
dreg10	10.67606	.7270684	14.68	0.000	9.251021	12.10109
dreg11	9.457017	1.386872	6.82	0.000	6.738782	12.17525
dreg12	14.18527	1.205608	11.77	0.000	11.82231	16.54823
dreg13	2.312853	.6283813	3.68	0.000	1.081242	3.544465
dsex1	-3.763406	.2130955	-17.66	0.000	-4.181068	-3.345745
dsex2	(dropped)					
dkin2	7.439764	.3950298	18.83	0.000	6.665515	8.214012
dkin3	10.5099	.5554968	18.92	0.000	9.421141	11.59866
dkin4	12.25363	.4120744	29.74	0.000	11.44597	13.06128
dndep2	12.0345	.2402808	50.09	0.000	11.56355	12.50544
dndep3	43.58998	.5029409	86.67	0.000	42.60423	44.57573
dmad2	16.23781	.2768419	58.65	0.000	15.6952	16.78041
dmad3	23.82915	.4289107	55.56	0.000	22.98849	24.6698
dmad4	19.84821	.4138276	47.96	0.000	19.03712	20.65931
dpad2	12.08282	.2834499	42.63	0.000	11.52727	12.63838
dpad3	17.36256	.4611127	37.65	0.000	16.45879	18.26633
dpad4	13.89024	.3711543	37.42	0.000	13.16279	14.61769
_cons	207.2561	.7297822	284.00	0.000	205.8258	208.6865

Frecuencia Simple de Kinder

kinder	Freq.	Percent	Cum.
1	19,325	9.04	9.04
2	99,781	46.67	55.71
3	14,260	6.67	62.38
4	80,430	37.62	100.00
Total	213,796	100.00	