

# Apéndice. Código PS4 (STATA)

2023-08-17

```

/*****
Semana 5 y 6: Fuentes de sesgo e imprecisión

Universidad de San Andrés
Economía Aplicada
*****/

Barnes, Fasan, Legaspe y Martin
/*****
Este archivo sigue la siguiente estructura:

*****/

* 0) Set up environment
*=====*/
clear all
global main "C:\Users\Usuario\Desktop\MAESTRIA\Economia Aplicada\TPs\PS4"
global input "$main/input"
global output "$main/output"

cd "$main"

* Open data

use "$input/poppy.dta", clear
*=====
* 1) Generen la variable "Chinese presence" tal como lo especifica el paper.
gen chinese_presence = .
replace chinese_presence = 1 if chinos1930hoy > 0 & chinos1930hoy !=.
replace chinese_presence = 0 if chinos1930hoy == 0

*=====
* 2) Estadística descriptiva de las variables relevantes (como en la Tabla 1)
drop if estado == "México"

* Generamos labels para hacerla mas facil de leer
label var cartel2010 "Cartel presence (2010)"
label var cartel2005 "Cartel presence (2005)"
label var chinese_presence "Chinese presence"
label var chinos1930hoy "Chinese Population"
label var IM_2015 "Marginalization"
label var Impuestos_ "Per capita tax revenue"
label var dalemanes "German presence"

```

```

label var capestado "Head of state"
label var pob1930cabec "Population 1930"
label var cartel2010 "Cartel presence 2010"
label var mindistcosta "Distance to closest port"
label var capestado "Head of state"
label var growthperc "Local population growth"

* Armamos una unica tabla de descriptivas
estpost summarize cartel2010 cartel2005 chinese_p chinos1930 IM_2015 Impuestos_
dalemanes tempopium distancia_km distkmDF mindistcosta capestado POB_TOT_2015
superficie_km TempMed PrecipAnual growthperc pob1930cabec

esttab using "$output/Table 1.tex", cells("mean sd min max") ///
nontitle nonumber replace label

=====
*3) Replicamos las regresiones hechas por OLS de las columnas 3 a 6 de la Tabla5.

* Columna 3
regress cartel2005 chinese_presence i.id_estado, cluster(id_estado)
outreg2 using "$output/Table 2.rtf", replace label

* Columna 4
regress cartel2005 chinese_presence dalemanes tempopium TempMed PrecipAnual
superficie_km pob1930cabec distancia_km distkmDF mindistcosta capestado
i.id_estado, cluster(id_estado)

outreg2 using "$output/Table 2.rtf", append label

* Columna 5
regress cartel2010 chinese_presence i.id_estado, cluster(id_estado)
outreg2 using "$output/Table 2.rtf", append label

* Columna 6
regress cartel2010 chinese_presence dalemanes tempopium TempMed PrecipAnual
superficie_km pob1930cabec distancia_km distkmDF mindistcosta capestado
i.id_estado, cluster(id_estado)

outreg2 using "$output/Table 2.rtf", append label

=====
*4)Reproducimos las Tabla 7 y 8 utilizando la presencia del cartel en 2010"
=====
* TABLA 7
=====

*col 1
ivregress 2sls IM_2015 i.id_estado (cartel2010=chinese_presence), cluster(id_estado)
outreg2 using "$output/Table 3.rtf", replace label
estat firststage

```

```

*col 2
ivregress 2sls IM_2015 dalemanes tempopium TempMed PrecipAnnual superficie_km
pob1930cabec distancia_km distkmDF mindistcosta capestado
i.id_estado (cartel2010=chinese_presence), cluster(id_estado)

outreg2 using "$output/Table 3.rtf", append label
estat firststage

*col 3
preserve
drop if distancia_km < 100
ivregress 2sls IM_2015 dalemanes tempopium TempMed PrecipAnnual superficie_km
pob1930cabec distancia_km distkmDF mindistcosta capestado
i.id_estado (cartel2010=chinese_presence), cluster(id_estado)

outreg2 using "$output/Table 3.rtf", append label
estat firststage

restore

*col 4
preserve
drop if estado == "Sinaloa"
ivregress 2sls IM_2015 dalemanes tempopium TempMed PrecipAnnual superficie_km
pob1930cabec distancia_km distkmDF mindistcosta capestado
i.id_estado (cartel2010=chinese_presence), cluster(id_estado)

outreg2 using "$output/Table 3.rtf", append label
estat firststage

restore

*col 5
ivregress 2sls IM_2015 dalemanes tempopium TempMed PrecipAnnual superficie_km
pob1930cabec distancia_km distkmDF mindistcosta capestado growthperc
i.id_estado (cartel2010=chinese_presence), cluster(id_estado)

outreg2 using "$output/Table 3.rtf", append label
estat firststage

*-----*
* TABAL 8
*-----*

*col 1
ivregress 2sls ANALF_2015 i.id_estado (cartel2010=chinese_presence),
cluster(id_estado)

outreg2 using "$output/Table 4.rtf", replace label

```

```

*col 2
ivregress 2sls SPRIM_2015 dalemanes tempopium TempMed PrecipAnual
superficie_km pob1930cabec distancia_km distkmDF mindistcosta capestado
i.id_estado (cartel2010=chinese_presence), cluster(id_estado)

outreg2 using "$output/Table 4.rtf", append label

*col 3
ivregress 2sls OVSDE_2015 dalemanes tempopium TempMed PrecipAnual superficie_km
pob1930cabec distancia_km distkmDF mindistcosta capestado
i.id_estado (cartel2010=chinese_presence), cluster(id_estado)

outreg2 using "$output/Table 4.rtf", append label

*col 4
ivregress 2sls OVSEE_2015 dalemanes tempopium TempMed PrecipAnual superficie_km
pob1930cabec distancia_km distkmDF mindistcosta capestado
i.id_estado (cartel2010=chinese_presence), cluster(id_estado)

outreg2 using "$output/Table 4.rtf", append label

*col 5
ivregress 2sls OVSAE_2015 dalemanes tempopium TempMed PrecipAnual superficie_km
pob1930cabec distancia_km distkmDF mindistcosta capestado
i.id_estado (cartel2010=chinese_presence), cluster(id_estado)

outreg2 using "$output/Table 4.rtf", append label

*col 6
ivregress 2sls VHAC_2015 dalemanes tempopium TempMed PrecipAnual superficie_km
pob1930cabec distancia_km distkmDF mindistcosta capestado
i.id_estado (cartel2010=chinese_presence), cluster(id_estado)

outreg2 using "$output/Table 4.rtf", append label

*col 7
ivregress 2sls OVPT_2015 dalemanes tempopium TempMed PrecipAnual
pob1930cabec distancia_km distkmDF mindistcosta capestado
i.id_estado (cartel2010=chinese_presence), cluster(id_estado)

outreg2 using "$output/Table 4.rtf", append label

*col 8
ivregress 2sls P02SM_2015 dalemanes tempopium TempMed PrecipAnual superficie_km
pob1930cabec distancia_km distkmDF mindistcosta capestado
i.id_estado (cartel2010=chinese_presence), cluster(id_estado)

outreg2 using "$output/Table 4.rtf", append label

*=====
*5) Testear Exogeneidad

*-----*
* Exogeneidad de la presencia del cartel (Hausman Test)

```

```

*-----*
* regression IV
ivregress 2sls IM_2015 dalemanes tempopium TempMed PrecipAnnual superficie_km
pob1930cabec distancia_km distkmDF mindistcosta capestado
i.id_estado (cartel2010 = chinese_presence)

est store iv

* regression OLS
reg IM_2015 cartel2010 dalemanes tempopium TempMed PrecipAnnual superficie_km
pob1930cabec distancia_km distkmDF mindistcosta capestado i.id_estado

est store ols

* Hausman Test
hausman iv ols

* Alternativa
ivregress 2sls IM_2015 dalemanes tempopium TempMed PrecipAnnual
superficie_km pob1930cabec distancia_km distkmDF mindistcosta
capestado i.id_estado (cartel2010 = chinese_presence), cluster(id_estado)

estat endogenous

/* The null hypothesis is that there is not systematic difference between
the estimates.
If you reject the null hypothesis, the coefficients are different so you should
use IV, education is endogenous. If you don't reject the null hypothesis, you
should use OLS, because being the coefficients consistent, ols is the most
efficient. */

*=====*
*6) Sargan Test

* Estimate by IV and predict the residuals
ivregress 2sls IM_2015 tempopium TempMed PrecipAnnual superficie_km pob1930cabec
distancia_km distkmDF mindistcosta capestado
i.id_estado (cartel2010 = chinese_presence dalemanes), cluster(id_estado)

predict resid, residual

* Regress the residuals on all the exogenous variables (instruments and controls)
reg resid chinese_presence dalemanes tempopium TempMed PrecipAnnual superficie_km
pob1930cabec distancia_km distkmDF mindistcosta capestado
i.id_estado, cluster(id_estado)

* Obtain the R2 and use it to compute the statistic S=nR2
ereturn list
display chi2tail(1,e(N)*e(r2))

```

```
/* The null hypothesis is that all moment conditions are valid. If the test is
rejected, you cannot determine which the invalid moment conditions are.
In this case you cannot reject the null hypothesis: the instruments
are exogenous*/
```