



MAESTRÍA EN ECONOMÍA

Economía Aplicada

PROF. MARTIN A. ROSSI

TUTORES: PAOLA LLAMAS Y TOMAS
PACHECO

**Problem Set 4: Variables
instrumentales**

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

Fecha de entrega: 13 de septiembre de 2024

PROBLEM SET 4: VARIABLES INSTRUMENTALES

GARCIA OJEDA - HAUSVIRTH - HAYDUK - SALVATIERRA

Este trabajo está basado en el paper “Following the Poppy trail: Origins and consequences of Mexican drug cartels” (Murphy y Rossi, 2020). En este artículo usan la migración china a principios del siglo XX como una instrumento de la presencia de un cartel en el presente con el objetivo de estudiar si los carteles mejoraron los outcomes socioeconómicos.

1. La variable “chinese presence” toma valor 1 en todos aquellos municipios en los que se reportó que al menos un chino vivía durante el Censo de 1930.

2. Tal como se realiza en el *paper*, construimos la siguiente tabla que contiene las estadísticas descriptivas de todas las variables disponibles (18 de 24 totales), para los distintos municipios de México, excluyendo a **Distrito Federal** de todas las estimaciones. Esto último ocurre debido a la pérdida de registros de su población en el censo de 1930, lo que afecta significativamente la cobertura de los datos en esa región. En el paper los autores recalcan que no creen que esta pérdida de datos sea sistemática, ya que la mayor parte de los registros originales del resto del país han sobrevivido, por lo que al excluir el Distrito Federal, se aseguran que las estimaciones no se vean sesgadas por la falta de información.

Tabla 1: Resumen de estadísticas descriptivas

Variable	Mean	S.D.	Min	Max
Cartel 2010	0.28	0.45	0	1
Cartel 2005	0.08	0.28	0	1
Chinese presence	0.14	0.347	0	1
Chinese 1930	4.20	35.89	0	1183
IM 2015	0.012	0.99	-2.21	5.02
Impuestos pc mun	141.83	286.32	0.01	6480.58
N index p	2.60	0.35	1	3.65
German presence	0.09	0.29	0	1
Poppy suitability	0.08	0.28	0	1
Distance to U.S. (km)	745.16	260.41	6.68	1349.712
Distance to Mexico City (km)	459.55	375.63	16	2282
Distance to closest port	957.06	264.48	0	1336
Head of state	0.01	0.11	0	1
Population 2015 (000s)	45316.79	129178.82	87	1677678
Surface (km2)	801.12	2110.06	2.21	53256.17
Average temperature	19.82	4.02	10.49	29.08
Average precipitation	89.72	50.84	6.45	338.36
Local population growth (1920-30)	0.22	1.63	-1	54
Population 1930 (000s)	2639.82	7478.08	0	179336
Observations	1982			

Note: All data are at municipality level. Cartel presence 2010 and Cartel presence 2005 are binary variables based on web searches; Chinese population is the number of people born in China and residing in 1930 in a given current municipality; Chinese presence takes the value 1 if Chinese population in 1930 is greater than 0; German presence takes the value 1 if German population in 1930 is greater than 0; Local population growth corresponds to growth in the population in the period 1920 to 1930 in the town that is the head of the municipality today; Head of state is a dummy that takes the value 1 for those municipalities that host the head of state town; Population in 1930 corresponds to the town that is the head of the municipality today; Marginalization is the marginalization index in 2015. Laakso and Taagepera index and Molinar index are from the 2015 election. Per capita tax revenue is for the period 2012–2014. Poppy suitability takes the value of 1 for those municipalities that have climate parameters in the range that FAO considers optimal for poppy (*papaver somniferum*) cultivation. All distances (in Km) are from the centroid of the municipality.

3. La tabla presentada realiza un análisis de robustez en las diferentes formas de medir la presencia de cárteles en dos distintos periodos, 2005 y 2010. Cada columna presenta los coeficientes estimados de la variable "*chinese presence*" sobre la presencia de cárteles en dichos intervalos de tiempo. En todas las especificaciones, los modelos incluyen efectos fijos por estado, y las diferencias entre las columnas radican en la inclusión o no de controles, así como en el periodo de tiempo que se considera para la variable dependiente.

En general, se observa que las especificaciones que incluyen controles (columnas 2 y 4) muestran coeficientes de menor magnitud en comparación con aquellas que no los incluyen (columnas 1 y 3). Los controles usados son los siguientes: presencia de alemanes, la idoneidad de la amapola, la temperatura promedio, la precipitación promedio, la superficie, la población en 1930, la distancia a EE. UU., la distancia a la Ciudad de México, la distancia al puerto más cercano y la variable binaria que toma el valor 1 para aquellos municipios que albergan la localidad que es cabecera de estado. Sin embargo, aún después de incluir los controles, los coeficientes siguen siendo estadísticamente significativos, lo que sugiere que la relación entre la presencia china y la presencia de cárteles es robusta a la inclusión de estas variables adicionales.

Además, se destaca que los coeficientes son más altos para las especificaciones del año 2010 (columnas 1 y 2), en comparación con el otro período. Esto sugiere, al igual que en el *paper* original con una amplitud mayor de fechas, un efecto promedio creciente de la presencia china sobre la presencia de cárteles a lo largo del tiempo.

Tabla 2: Robustness check

	(1)	(2)	(3)	(4)
	Cartel presence	Cartel presence	Cartel presence	Cartel presence
	2010	2010	2005	2005
Chinese presence	0.2158*** (0.0344) [0.000]	0.1011** (0.0363) [0.009]	0.1299*** (0.0242) [0.000]	0.0589** (0.0199) [0.006]
State dummies	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
Clusters	31	30	31	30
Observations	2439	2161	2440	2162

Note: Standard errors clustered at the state level are shown in parentheses. Wild- bootstrapped p-values are shown in brackets (999 replications). The unit of observation is a municipality. The unit of cluster is a state. All models are estimated using Ordinary Least Squares. The set of controls includes German presence, Poppy suitability, Average temperature, Average precipitation, Surface, Population in 1930, Distance to U.S., Distance to Mexico City, Distance to closest port, and Head of state. *Significant at the 10 % level. **Significant at the 5 % level. ***Significant at the 1 % level.

4. Los autores buscan estimar el efecto causal de la presencia de cárteles en variables socioeconómicas. El modelo es:

$$Marginilization_{is} = \alpha CartelPresence_{is} + \pi X_{is} + \mu_s + \epsilon_{is}, \quad (0.1)$$

donde i hace referencia al municipio y s al estado. Marginalization es un índice construido a partir de los siguientes indicadores socioeconómicos: Analfabetismo (porcentaje de analfabetos a los 15 años), Sin primaria (porcentaje de población a los 15 años sin escuela primaria), Ausencia de baño (porcentaje de población sin baño), Ausencia de electricidad (porcentaje de población sin acceso a red eléctrica), Ausencia de agua (porcentaje de población sin acceso a red de agua), Hacinamiento (porcentaje de hogares con algún nivel de hacinamiento), Piso de tierra (porcentaje de ocupantes de viviendas con piso de tierra), Localidades pequeñas (porcentaje de población en localidades con menos de 5000 habitantes), y Salario bajo (porcentaje de fuerza laboral que gana menos de 2 salarios mínimos).

Como la presencia de cárteles es endógena, se instrumenta a la misma con presencia china en 1930 y se estima el modelo 0.1 por 2SLS.

La Tabla 3 es nuestra replicación de la Tabla 7 del paper. La columna 1 no incluye controles. Las columnas 2 a 4 incluyen los siguientes controles: presencia de alemanes, la idoneidad de la amapola, la temperatura promedio, la precipitación promedio, la superficie, la población en 1930, la distancia a EE. UU., la distancia a la Ciudad de México, la distancia al puerto más cercano y la variable binaria que toma el valor 1 para aquellos municipios que albergan la localidad que es cabecera de estado. La columna 3 excluye aquellos municipios con una distancia menor a 100km de la frontera con EE.UU. y la columna 4 excluye al Estado de Sinaloa. Por último, la columna 5 incorpora el crecimiento de la población local.

Podemos notar que los coeficientes estimados indican que la presencia de cárteles en un municipio (instrumentada a partir de la presencia china) disminuye el índice de marginalización aproximadamente 2 desviaciones estándar. Estos resultados son robustos a excluir a los municipios muy cercanos a la frontera con EE.UU., a excluir al Estado de Sinaloa y al controlar por crecimiento poblacional. Todos los coeficientes son significativos al 1% .

Tabla 3: Estimación por IV del efecto de la presencia de cárteles en un índice socioeconómico

	(1)	(2)	(3)	(4)	(5)
	Marginalization	Marginalization	Marginalization	Marginalization	Marginalization
Cartel presence 2010	-2.002*** (0.338) [0.000]	-2.388** (0.779) [0.002]	-2.799** (1.059) [0.008]	-2.420** (0.770) [0.002]	-2.382** (0.788) [0.002]
F-test	39.38	7.74	5.55	8.27	7.64
State dummies	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	Yes
Clusters	31	30	30	29	30
Observations	2439	2161	2107	2143	2159

Note: Standard errors clustered at the state level are shown in parentheses. Wild-bootstrapped p-values are shown in brackets (999 replications). The unit of observation is a municipality. The unit of cluster is a state.

All models are estimated using Two Stages Least Squares. Cartel presence is instrumented using Chinese presence F-test is the F-test of excluded instruments (p-values are shown in parentheses). In Columns (2) to (4) the set of controls includes German presence, Poppy suitability, Average temperature, Average precipitation, Surface, Population in 1930, Distance to U.S., Distance to Mexico City, Distance to closest port, and Head of state. Column (5) further controls for Local population growth. Column (3) excludes municipalities located more than 100 km from U.S. border. Column (4) excludes municipalities located in the state of Sinaloa. insignificant at the 10 % level. **Significant at the 5 % level. *** Significant at the 1 % level.

Como complemento, los autores estiman la ecuación 0.1 pero para cada uno de los 9 componentes del índice de marginalización como variable dependiente. Los resultados de esta estimación se encuentran en la Tabla 4, que es nuestra replicación de la Tabla 8 del paper.

Las estimaciones indican que la presencia de cárteles está asociada negativamente con 8 de 9 componentes del índice de marginalización, y que solamente 4 coeficientes son estadísticamente significativos. En todos los casos, exceptuando el componente de Ausencia de agua, un aumento de la presencia de cárteles implica en promedio una disminución de cada uno de los componentes del índice. Una mayor presencia de cárteles implica en promedio mayores niveles de alfabetismo y mayores porcentajes de la población con escuela primaria, acceso a red eléctrica y en localidades con más de 5000 habitantes.

Tabla 4: Estimación por IV del efecto de la presencia de cárteles en cada uno de los componentes del índice de marginalización

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Illiteracy	Without primary	Without toilet	Without electricity	Without water	Overcrowding	Earthen floor	Small localities	Low salary
Cartel presence 2010	-21.771*	-34.344**	-6.100	-4.595	3.271	-19.994*	-7.041	-100.693*	-27.202*
	(8.941)	(12.729)	(4.193)	(2.657)	(7.258)	(9.295)	(7.003)	(39.096)	(10.825)
	[0.015]	[0.007]	[0.146]	[0.084]	[0.652]	[0.031]	[0.315]	[0.010]	[0.012]
State dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	30	30	30	30	30	30	30	30	30
Observations	2161	2161	2161	2161	2161	2161	2161	2161	2161

Note: Standard errors clustered at the state level are shown in parentheses. Wild-bootstrapped p-values are shown in brackets (999 replications). The unit of observation is a municipality. The unit of cluster is a state. All models are estimated using Two Sta- ges Least Squares. Cartel presence is instrumented using Chinese presence. The set of controls includes German presence, Poppy suitability, Average temperature, Average precipitation, Surface, Population, Distance to U.S., Distance to Mexico City, Dis- tance to closest port, and Head of state. *Significant at the 10 % level. **Significant at the 5 % level. ***Significant at the 1 % level.

5. No se puede testear la exogeneidad del instrumento. Sin embargo, lo que si es posible es aportar distintos tipos de evidencia que ayuden a hacer los supuestos mas creíbles como pruebas de robustes, pruebas historicas, arugmentos teóricos, etc.

Si supieramos que el instrumento es exógeno, se puede asegurar que la variabilidad de la presencia del cartel explicada por la variabilidad del instrumento es exógena. Sin embargo, toda la presencia del cartel puede ser exógena. Para resolver esto, podemos hacer un test de Hausmann bajo el cual se compare los coeficientes de nuestra segunda etapa con los coeficientes de un metodo de minimos cuadrados ordinarios simple que explique marginalización con la presencia de los carteles. La hipotesis nula del test es que ambos coeficientes son iguales, por lo que aceptar la hipotesis nula equivale a afirmar que la variable presencia de cartel no es endógena; por lo que debería usarse OLS en lugar de IV ya que ambos son consistentes pero el estimador de OLS será más eficiente. Para realizar este test, se asume como cierto el supuesto de exogeneidad del instrumento (no podemos testear un supuesto sin hacer otro).

Los resultados del test de Hausman sobre la especificación básica (columna 2 de la Tabla 3) nos arrojan un p-valor de 0.0036, lo que nos permite rechazar la hipótesis nula a cualquier nivel de significatividad relevante. Por tanto, hay evidencia para afirmar que las estimaciones por OLS y IV son estadísticas diferentes y debemos usar el método de IV debido a la endogeneidad de la presencia de cárteles en el modelo para explicar la marginalización.

6. El principal problema del paper es que la elección de la migración china a principios del siglo XX como instrumento para la presencia de cárteles en el presente plantea dudas sobre la exogeneidad del instrumento. Para que un instrumento sea válido, debe cumplir con relevancia y exogeneidad. La exogeneidad exige que el instrumento no esté correlacionado con los errores de la regresión, es decir, no debe afectar los outcomes socioeconómicos a través de otros canales que no sean la presencia de carteles.

Si las razones que motivaron la migración china están relacionadas con los outcomes socioeconómicos que se intentan estudiar como mejores condiciones económicas o sociales que podrían haber atraído a los inmigrantes chinos, se podría violar la restricción de exclusión y esto generaria sesgo en la estimación de interés.

Mas alla de los esfuerzos de los autores por documentar las razones históricas detrás de la migración china y los controles adicionales incluidos en el modelo, siempre persiste vulnerabilidad respecto a la exogeneidad del instrumento. En particular, los autores incluyen la presencia de inmigrantes alemanes como un control para tratar de capturar aspectos de la variabilidad endógena en los patrones de migración. Si bien este control ayuda a argumentar que las características específicas de la migración china no están influyendo directamente en los outcomes socioeconómicos, sigue siendo posible que la inmigración china haya tenido un impacto propio y único sobre las regiones de destino a través de canales que no están incluidos en la estrategia de identificación.


```

1  /*****
   *****/
2          Semana 5: Variables Instrumentales
3          Universidad de San Andrés
4          Economía Aplicada
5  *****/
6  * Gaspar Hayduk; Juan Gabriel García Ojeda; Elías Lucas
   Salvatierra; Martina Hausvirth
7
8  /*****
   *****/
9
10 * 0) Set up environment
11 *=====
   =====*
12 global main "/Users/gasparhayduk/Desktop/Economía Aplicada/PS4"
13 global input "$main/input"
14 global output "$main/output"
15
16 cd "$main"
17
18 * Abrimos la base de datos:
19 use "$input/poppy.dta", clear
20
21
22 * 1) Inciso 1
23 *=====
   =====*
24
25 * Generamos la variable "Chinese presence" como se especifica en
   el paper. Para ello, utilizaremos una dummy que toma valor 0 si
   no existe presencia china dentro del municipio, y 1 en caso
   contrario.
26
27 gen chinesepresence = 0
28 replace chinesepresence =1 if chinos1930hoy > 0
29
30 * Labeledamos algunas variables:
31 label var cartel2010 "Cartel presence 2010"
32 label var cartel2005 "Cartel presence 2005"
33 label var chinesepresence "Chinese presence"
34 label var chinos1930hoy "Chinese population"
35 label var IM_2015 "Marginalization"
36 label var Impuestos_pc_mun "Per capita tax revenue"
37 label var n_index_p "Laakso and Taagepera index"
38 label var dalemanes "German presence"
39 label var tempopium "Poppy suitability"
40 label var distancia_km "Distance to U.S. (km)"
41 label var distkmDF "Distance to Mexico City (km)"
42 label var mindistcosta "Distance to closest port"
43 label var canestado "Head of state"

```

```

43 label var capestado head of state
44 label var POB_TOT_2015 "Population in 2015 (in 000)"
45 label var superficie_km "Surface (000 km2)"
46 label var TempMed_Anual "Average temperature (Celsius)"
47 label var PrecipAnual_med "Average precipitation (mm)"
48 label var growthperc "Local population growth (1920-30)"
49 label var pob1930cabec "Population in 1930 (in 000)"
50
51
52
53
54 * 2) Inciso 2: Estadísticas Descriptivas y dropeo de obs de
    Distrito Federal
55 *=====
    =====*
56 * Dropeamos Distrito Federal
57 drop if estado == "Distrito Federal"
58
59
60 *Luego realizamos la tabla correspondiente.
61
62 ssc install asdoc
63
64 asdoc tabstat cartel2010 cartel2005 chinesepresence chinos1930hoy
    IM_2015 ///
65 Impuestos_pc_mun n_index_p dalemanes tempopium distancia_km
    distkmDF mindistcosta ///
66 capestado POB_TOT_2015 superficie_km TempMed_Anual
    PrecipAnual_med growthperc pob1930cabec, ///
67 stat(mean sd min max) replace save(Cuadro_1.tex)
68
69
70
71 * Ejercicio 3
72 *=====
    =====*
73 *Para este ejercicio replicamos la tabla 5 del paper que
    presenta las 4 formas de medir la presencia de carteles. En
    nuestro caso, contamos únicamente con 2 de las 4 variables
    originales: presencia de carteles en 2010 y en 2005.
74
75 *En primer lugar, realizamos las regresiones hechas por MCO, con
    y sin controles, con errores cluster.
76
77
78 * (1) Cartel presence 2010 – Sin controles
79 eststo clear
80 eststo: reg cartel2010 chinesepresence i.id_estado, cluster(
    id_estado)
81 * Guardamos la información de esta regresión
82 estadd local statedummies "Yes"
83 estadd local controls "No"

```

```

84  estadd local clusters = e(N_clust)
85
86  * (2) Cartel presence 2010 – Con controles
87  eststo: reg cartel2010 chinesepresence dalemanes tempopium
      TempMed_Anual PrecipAnual_med superficie_km pob1930cabec
      distancia_km distkmDF mindistcosta capestado i.id_estado, cluster
      (id_estado)
88  * Guardamos la información de esta regresión
89  estadd local statedummies "Yes"
90  estadd local controls "Yes"
91  estadd local clusters = e(N_clust)
92
93  * (3) Cartel presence 2005 – Sin controles
94  eststo: reg cartel2005 chinesepresence i.id_estado, cluster(
      id_estado)
95  * Guardamos la información de esta regresión
96  estadd local statedummies "Yes"
97  estadd local controls "No"
98  estadd local clusters = e(N_clust)
99
100 * (4) Cartel presence 2005 – Con controles
101 eststo: reg cartel2005 chinesepresence dalemanes tempopium
      TempMed_Anual PrecipAnual_med superficie_km pob1930cabec
      distancia_km distkmDF mindistcosta capestado i.id_estado, cluster
      (id_estado)
102 * Guardamos la información de esta regresión
103 estadd local statedummies "Yes"
104 estadd local controls "Yes"
105 estadd local clusters = e(N_clust)
106
107 * Exportar la tabla a LaTeX
108 esttab using "$output/Cuadro_2.tex", replace ///
109     se label noobs keep(chinesepresence) ///
110     cells(b(star fmt(4)) se(par fmt(4)) p(fmt(3) par("[ ]")) ///
111     stats(statedummies controls clusters N, fmt(0 0 0 0) ///
112     labels("State dummies" "Controls" "Clusters" "Observations"))
    ///
113
114
115
116
117
118 * 4) Inciso 4: replicacion Tabla 7
119 *=====
    =====*
120
121 * La ecuacion es:  $marginalization_{i,s} = \alpha * cartel\_presence + \pi * controles + \mu_s - \epsilon_{i,s}$ . i indexa municipalodad y s
    estado.
122 * se instrumenta cartel_presence con chinese presence in 1930.
123
124 * Especificacion 1: sin controles y state dummies

```

```

124 *--- Especificacion 1: Sin controles y state dummies
125 eststo clear
126 eststo: ivregress 2sls IM_2015 (cartel2010 = chinesepresence) i.
    id_estado, cluster(id_estado)
127 weakivtest
128 estadd scalar ftest = r(F_eff)
129 estadd local statedummies "Yes"
130 estadd local controls "No"
131 estadd local clusters = e(N_clust)
132
133 *--- Especificacion 2: controles y state dummies. Es la basica
134 eststo: ivregress 2sls IM_2015 (cartel2010 = chinesepresence)
    dalemanes tempopium TempMed_Anual PrecipAnual_med superficie_km
    pob1930cabec distancia_km distkmDF mindistcosta capeestado i.
    id_estado, cluster(id_estado)
135 weakivtest
136 estadd scalar ftest = r(F_eff)
137 estadd local statedummies "Yes"
138 estadd local controls "Yes"
139 estadd local clusters = e(N_clust)
140
141 *--- Especificacion 3: especificacion basica solo para
    municipalidades a mas de 100km de la frontera
142 eststo: ivregress 2sls IM_2015 (cartel2010 = chinesepresence)
    dalemanes tempopium TempMed_Anual PrecipAnual_med superficie_km
    pob1930cabec distancia_km distkmDF mindistcosta capeestado i.
    id_estado if distancia_km>=100, cluster(id_estado)
143 weakivtest
144 estadd scalar ftest = r(F_eff)
145 estadd local statedummies "Yes"
146 estadd local controls "Yes"
147 estadd local clusters = e(N_clust)
148
149 *--- Especificacion 4: especificacion basica pero se excluyen
    municipalidades de Sinaloa:
150 preserve
151 drop if estado=="Sinaloa"
152
153 eststo: ivregress 2sls IM_2015 (cartel2010 = chinesepresence)
    dalemanes tempopium TempMed_Anual PrecipAnual_med superficie_km
    pob1930cabec distancia_km distkmDF mindistcosta capeestado i.
    id_estado, cluster(id_estado)
154 weakivtest
155 estadd scalar ftest = r(F_eff)
156 estadd local statedummies "Yes"
157 estadd local controls "Yes"
158 estadd local clusters = e(N_clust)
159
160 restore
161
162
163 *--- Especificacion 5: controles para crecimiento poblacional

```

2021

```

204
205 * Column 2: Without primary
206 eststo: ivregress 2sls SPRIM_2015 (cartel2010 = chinesepresence)
    dalemanes tempopium TempMed_Anual PrecipAnual_med superficie_km
    pob1930cabec distancia_km distkmDF mindistcosta capestado i.
    id_estado, cluster(id_estado)
207 weakivtest
208 estadd scalar ftest = r(F_eff)
209 estadd local statedummies "Yes"
210 estadd local controls "Yes"
211 estadd local clusters = e(N_clust)
212
213
214 * Column 3: Without toilet
215 eststo: ivregress 2sls OVSDE_2015 (cartel2010 = chinesepresence)
    dalemanes tempopium TempMed_Anual PrecipAnual_med superficie_km
    pob1930cabec distancia_km distkmDF mindistcosta capestado i.
    id_estado, cluster(id_estado)
216 weakivtest
217 estadd scalar ftest = r(F_eff)
218 estadd local statedummies "Yes"
219 estadd local controls "Yes"
220 estadd local clusters = e(N_clust)
221
222 * Column 4: without electricity
223 eststo: ivregress 2sls OVSEE_2015 (cartel2010 = chinesepresence)
    dalemanes tempopium TempMed_Anual PrecipAnual_med superficie_km
    pob1930cabec distancia_km distkmDF mindistcosta capestado i.
    id_estado, cluster(id_estado)
224 weakivtest
225 estadd scalar ftest = r(F_eff)
226 estadd local statedummies "Yes"
227 estadd local controls "Yes"
228 estadd local clusters = e(N_clust)
229
230 * Column 5: without water
231 eststo: ivregress 2sls OVSAE_2015 (cartel2010 = chinesepresence)
    dalemanes tempopium TempMed_Anual PrecipAnual_med superficie_km
    pob1930cabec distancia_km distkmDF mindistcosta capestado i.
    id_estado, cluster(id_estado)
232 weakivtest
233 estadd scalar ftest = r(F_eff)
234 estadd local statedummies "Yes"
235 estadd local controls "Yes"
236 estadd local clusters = e(N_clust)
237
238 * Column 6: overcrowding
239 eststo: ivregress 2sls VHAC_2015 (cartel2010 = chinesepresence)
    dalemanes tempopium TempMed_Anual PrecipAnual_med superficie_km
    pob1930cabec distancia_km distkmDF mindistcosta capestado i.
    id_estado, cluster(id_estado)
240 weakivtest

```

```

241 estadd scalar ftest = r(F_eff)
242 estadd local statedummies "Yes"
243 estadd local controls "Yes"
244 estadd local clusters = e(N_clust)
245
246 * Columna 7: earthen floor
247 eststo: ivregress 2sls OVPT_2015 (cartel2010 = chinesepresence)
    dalemanes tempopium TempMed_Anual PrecipAnual_med superficie_km
    pob1930cabec distancia_km distkmDF mindistcosta capestado i.
    id_estado, cluster(id_estado)
248 weakivtest
249 estadd scalar ftest = r(F_eff)
250 estadd local statedummies "Yes"
251 estadd local controls "Yes"
252 estadd local clusters = e(N_clust)
253
254 * Columna 8: small localities
255 eststo: ivregress 2sls PL5000_2015 (cartel2010 = chinesepresence)
    dalemanes tempopium TempMed_Anual PrecipAnual_med superficie_km
    pob1930cabec distancia_km distkmDF mindistcosta capestado i.
    id_estado, cluster(id_estado)
256 weakivtest
257 estadd scalar ftest = r(F_eff)
258 estadd local statedummies "Yes"
259 estadd local controls "Yes"
260 estadd local clusters = e(N_clust)
261
262 * Columna 9: low salary
263 eststo: ivregress 2sls P02SM_2015 (cartel2010 = chinesepresence)
    dalemanes tempopium TempMed_Anual PrecipAnual_med superficie_km
    pob1930cabec distancia_km distkmDF mindistcosta capestado i.
    id_estado, cluster(id_estado)
264 weakivtest
265 estadd scalar ftest = r(F_eff)
266 estadd local statedummies "Yes"
267 estadd local controls "Yes"
268 estadd local clusters = e(N_clust)
269
270 *** Exportamos Resultados
271 esttab using "$output/Cuadro 4.tex", se replace label noobs ///
272 keep(cartel2010) ///
273 cells(b(star fmt(3)) se(par fmt(3)) p(fmt(3) par(`"'["'"`"]'"')) ///
274 stats(statedummies controls clusters N, fmt(0 0 2 0) ///
275 labels("State dummies" "Controls" "Clusters" "Observations"))
276
277
278
279
280 * 4) Inciso 5: ¿Pueden testear la exogeneidad del instrumento?
281 *=====
    =====*

```

282

```
283 * Testeamos la exogeneidad de la presencia del cartel sabiendo
    que el instrumento es exógeno utilizando la especificación
    básica (segunda columna de la Tabla 7):
284 * La hipótesis nula es  $H_0$ : variable exógena.
285 ivregress 2sls IM_2015 (cartel2010 = chinesepresence) dalemanes
    tempopium TempMed_Anuual PrecipAnual_med superficie_km
    pob1930cabec distancia_km distkmDF mindistcosta capestado i.
    id_estado, cluster(id_estado)
286 estat endogenous
287 *el p-valor es 0.0036, menor a cualquier nivel de significancia,
    por lo que rechazamos la hipótesis nula de que la variable es
    exógena.
288
289 *estat endogenous performs tests to determine whether endogenous
    regressors in the model are in fact exogenous (conditional on
    the instrument being exogenous). If the test statistic is
    significant, then the variables being tested must be treated as
    endogenous. This is not the case in our example.
290 * with an unadjusted VCE: the Durbin (1954) and Wu-Hausman
    statistics
291 * with a robust VCE, a robust score test (Wooldridge 1995) and a
    robust regression-based test
292
293
294
295
296
297
298
299
```