

Tarea calificada - Módulo 2

Sonia Rosmery Asto Mercado - 20166142

Claudia Mirela Vivas Alejandro - 20141150

October 2022

1. Variables Instrumentales

- a. **Replique la tabla 2 (Columnas 1-6) del paper y discuta (siguiendo los argumentos de los autores) sobre la interpretación de los resultados en términos de causalidad.**

El cuadro 1 presenta las regresiones por OLS del logaritmo del PBI per cápita sobre la variable de protección contra el riesgo de expropiación utilizando diversas muestras.

En la columna (1), se regresa solo el log del PBI con la variable de protección contra el riesgo de expropiación, es decir, sin utilizar controles. En este caso, se utiliza toda la muestra y se obtiene un efecto de 0.53 significativo al 1 % y positivo (mayor índice de protección implica mejor desempeño económico). Asimismo, se observa que la variable de protección contra el riesgo de expropiación explica el 61 % del desempeño económico para 111 países.

En la columna (2), se regresa el mismo modelo, pero esta vez en la muestra base de 64 países. En este caso, el efecto de las instituciones sobre el PBI per cápita no ha

Cuadro 1: OLS Regressions

Dependent variable: Log GDP per capita in 1995						
	(1)	(2)	(3)	(4)	(5)	(6)
Average Protection Against Expropriation Risk in 1985-1995	0.53*** (0.04)	0.52*** (0.06)	0.46*** (0.06)	0.42*** (0.05)	0.47*** (0.06)	0.40*** (0.06)
Latitude			0.87* (0.49)	0.65 (0.53)	1.58** (0.71)	0.88 (0.63)
Asia dummy				-0.28 (0.18)		-0.58** (0.23)
Africa dummy				-0.97*** (0.17)		-0.88*** (0.17)
"Other" continent dummy				-0.31 (0.26)		0.11 (0.38)
R^2	0.61	0.54	0.62	0.72	0.57	0.71
N	111	64	111	111	64	64

Standard errors in parentheses

* $p < 0,10$, ** $p < 0,05$, *** $p < 0,01$

variado mucho respecto al modelo anterior, pero el porcentaje de explicación de la variable instituciones al desempeño económico se redujó a 54 %. El efecto estimado es de 0.52 y es significativo al 1 %, con un error estándar de 0.06.

En las columnas (3) y (5), se añade la variable "latitud" a los modelos, lo que disminuye un poco el efecto de las instituciones sobre los resultados académicos. Sin embargo, en el modelo (3) se regresa sobre toda la muestra de 111 países, mientras que en el modelo (5) se regresa sobre la muestra base de 64 países. En ambos casos, la latitud es una variable significativa, pero el coeficiente (0.87) del modelo (3) es mucho menor al coeficiente (1.58) del modelo (5). Adicionalmente, el modelo (3) explica el 62 % del resultado económico y el modelo (5) explica el 57 %.

Finalmente, en las columnas (4) y (6), se añaden más controles como las variables dummy por continente. El efecto estimado disminuye a 0.42 y 0.40, respectivamente,

pero no genera un cambio drástico, al menos en comparación a los modelos (3) y (5). Asimismo, vemos que al añadir más variables de control, la variable latitud deja de ser significativa. La variable de África es significativa al 1 %, lo que significa que los países africanos son 97 puntos logarítmicos más pobres, incluso luego de estimar el efecto de las instituciones, lo que podría dar indicio de que los países africanos son pobres no solo por sus instituciones, sino también por otros factores. Algo que resaltar, es que el modelo (4) y (6) explican el desempeño económico en más del 70 %, mayor a los modelos anteriores.

b. Discuta brevemente el(los) instrumento(s) utilizado(s) por los autores, especialmente en relación a las condiciones que este(os) debe(n) cumplir para identificar insesgadamente el parámetro de interés.

El objetivo principal del artículo de Acemoglu et al (2001) es estimar el efecto de las instituciones en el desempeño económico de los países. Entonces, debido a que las instituciones y el desempeño económico tienen causalidad en ambos sentidos, los autores deciden utilizar una variable exógena, como instrumento, para explicar las instituciones actuales, medida por el índice de protección contra el riesgo de expropiación. Este instrumento es la tasa de mortalidad esperada de los primeros colonos europeos en las colonias.

La idea fundamental del autor para escoger este instrumento es el siguiente:

- La tasa de mortalidad potencial determinó los lugares en los que los colonos europeos se asentarían. Es decir, las colonias con bajas tasas de mortalidad fueron elegidos por los colonos europeos para establecerse.
- Aquellas colonias donde se establecieron los colonos europeos desarrollaron instituciones muy similares a las europeas. En cambio, las demás, de mayor mortalidad, desarrollaron peores instituciones.

- Las instituciones que se crearon al principio se mantienen hasta la actualidad, por lo que las instituciones actuales son un reflejo de ello.
- Por último, las instituciones afectan el desempeño económico.

Las condiciones que debe cumplir la tasa de mortalidad potencial para que se pueda identificar el efecto en el PBI per cápita son las siguientes:

- No tiene que estar correlacionado con el término de error, es decir, no debe explicar directamente el PBI per cápita. Por ende, el efecto de la tasa de mortalidad sobre el PBI per cápita es sólo a través del mecanismo de las instituciones.
- Tiene que estar correlacionado con las instituciones, es decir, debe ser una variable explicativa significativa para las instituciones.

La primera condición no es tan fácil de cumplir, ya que Acemoglu y otros (2001) reconocen la preocupación de que las tasas de mortalidad de los colonos afecten el entorno actual de las enfermedades. De esta manera, las tasas de mortalidad tendrían un efecto directo sobre el desempeño económico. Sin embargo, los autores justifican la plausibilidad de la primera condición indicando que es poco probable, ya que las enfermedades tuvieron un efecto limitado sobre los adultos indígenas que habían desarrollado inmunidad en comparación del efecto mortal sobre los europeos. Para sustentar esto, los autores recurren a tasas de mortalidad históricas donde muestran que las tasas de mortalidad de los europeos en estas colonias eran más altas debido a la falta de inmunidad. Finalmente, los autores mediante pruebas de sobreidentificación tratan de detectar si la mortalidad tiene un efecto directo sobre el PBI per capita y no encuentran evidencia de este efecto directo. La segunda condición es probada por los autores mediante la regresión de las instituciones (índice de protección contra el riesgo de expropiación) y la tasa de mortalidad en la primera etapa del modelo 2SLS. En ella, se encuentra que la tasa de mortalidad explica más del 25 % de la

variación de las instituciones actuales.

- c. **Replique la tabla 4 (Columnas 1-8) del paper e interprete los resultados obtenidos con 2SLS vs OLS.**

Cuadro 2: IV Regressions of Log GDP per capita

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Two-Stage Least Squares								
Average Protection Against Expropriation Risk in 1985-1995	0.94*** (0.16)	1.00*** (0.22)	1.28*** (0.36)	1.21*** (0.35)	0.58*** (0.10)	0.58*** (0.12)	0.98*** (0.30)	1.11** (0.46)
Latitude		-0.65 (1.34)		0.94 (1.46)		0.04 (0.84)		-1.18 (1.76)
Asia dummy							-0.92** (0.40)	-1.05* (0.52)
Africa dummy							-0.46 (0.36)	-0.44 (0.42)
"Other" continent dummy							-0.94 (0.85)	-0.99 (1.00)
Panel B: First Stage for Average Protection Against Expropriation Risk in 1985-1995								
Log European Settler Mortality	-0.61*** (0.13)	-0.51*** (0.14)	-0.39*** (0.13)	-0.39*** (0.14)	-1.21*** (0.22)	-1.14*** (0.24)	-0.43** (0.17)	-0.34* (0.18)
Latitude		2.00 (1.34)		-0.11 (1.49)		0.99 (1.43)		2.01 (1.39)
Asia dummy							0.33 (0.50)	0.47 (0.50)
Africa dummy							-0.27 (0.41)	-0.26 (0.41)
"Other" continent dummy							1.24 (0.84)	1.06 (0.84)
R ²	0.27	0.30	0.13	0.13	0.47	0.47	0.30	0.33
Panel C: Ordinary Least Squares								
Average Protection Against Expropriation Risk in 1985-1995	0.52*** (0.06)	0.47*** (0.06)	0.49*** (0.08)	0.47*** (0.07)	0.48*** (0.07)	0.47*** (0.07)	0.42*** (0.06)	0.40*** (0.06)
N	64	64	60	60	37	37	64	64

Standard errors in parentheses

* $p < 0,10$, ** $p < 0,05$, *** $p < 0,01$

Los resultados obtenidos mediante 2SLS, al utilizar el instrumento de mortalidad, logra controlar el efecto endógeno de la variable de instituciones (la protección contra

el riesgo de expropiación), de manera que se elimina la correlación entre la variable explicativa (la protección contra el riesgo de expropiación) y el término de error. Por lo tanto, los resultados obtenidos en el modelo de 2 etapas son consistentes e insesgados respecto al verdadero parámetro poblacional de nuestra variable de interés, en comparación al modelo OLS, ya que cuando hay presencia de endogeneidad, los resultados del modelo OLS no son consistentes. Entonces, al controlar este efecto endógeno se logra capturar el efecto real de las instituciones sobre el desempeño económico, por lo que este efecto es mayor al obtenido en los modelos OLS.

En el cuadro 2 se presenta la estimación del efecto de las instituciones en el desempeño económico mediante la estimación de MCO en 2 etapas. En el Panel A, se muestra la segunda etapa de la regresión, es decir, regresionamos el GDP per cápita respecto al índice de protección contra el riesgo de expropiación. En el Panel B, se muestra la primera etapa de la regresión en el que regresionamos el índice de protección contra el riesgo de expropiación respecto a la tasa de mortalidad de los colonos europeos. Por último, en el Panel C, se muestra los resultados de la regresión de OLS simple del GDP per cápita respecto al índice de protección contra el riesgo de expropiación.

Respecto a la primera etapa (Panel B), podemos resaltar la significancia al 1 % de la mortalidad de los colonos (en log) en los modelos (1-6), así como el efecto negativo de la mortalidad en la calidad de instituciones que se crean. Es decir, a menor tasa de mortalidad de los países, el índice de protección contra el riesgo de expropiación es mayor, es decir, se crean mejores instituciones. Asimismo, los modelos (5) y (6), en los cuales se excluye a África de la regresión, estiman un impacto de -1.21 y -1.14 respectivamente, siendo los efectos negativos de mayor impacto en todos los modelos. Adicionalmente, vemos que el agregar variables de control como latitud y variables dummy por continente no son significativas para explicar el índice de

protección contra el riesgo de expropiación.

En la segunda etapa, en el modelo (1) y (2), se utiliza toda la muestra base de 64 países. En el modelo (1) se estima que el impacto de las instituciones sobre el PBI per cápita es de 0.94, con una significancia al 1 % y un error estándar de 0.16. Esta estimación es casi el doble del efecto encontrado por OLS (0.52), pero igual de significativo. En el modelo (2), se añade como variable de control a la latitud de los países; sin embargo, esta no es significativa para ningún modelo. Asimismo, se encuentra un efecto significativo al 1 % de 1, muy cercano al modelo anterior, con un error estándar de 0.22, mayor al anterior modelo.

En los modelos (3) y (4), se excluye de la muestra a países Neo-Europeos (Canadá, EE.UU, Nueva Zelanda y Australia) y se encuentra que en ambos modelos, el efecto de las instituciones sobre el PBI per cápita es significativo al 1 % con un coeficiente de 1.28 y 1.21 respectivamente, un poco mayor al encontrado en los modelos (1) y (2). De esta manera, podemos concluir que el impacto de las instituciones al desempeño económico no está impulsado por los países Neo-Europeos. Asimismo, vemos que la variable latitud no cambia ni el signo ni la significancia de los resultados. Por último, el efecto encontrado es mucho mayor al que se estima con OLS.

En la columna (5) y (6), se utiliza solo la muestra de países que pertenecen a un continente distinto al de África. En el modelo (5), si bien el impacto de las instituciones al desempeño económico sigue siendo positivo, la magnitud del impacto es mucho menor a lo encontrado en los modelos (1-4). El impacto de 0.58 sigue siendo significativo al 1 % con un error estándar de 0.10 y mayor a lo registrado por OLS (0.48). En el modelo (6), al añadirle la latitud, los resultados no varían (0.58), ya que esta variable no es significativa. Al igual que el modelo (5), la estimación sigue siendo significativa al 1 % y mayor al modelo OLS (0.47).

Por último, en la columna (7) y (8) se utiliza la muestra base de 64 países, pero controlando por variables dummies por continentes, con América como grupo omitido. Vemos que agregar estas variables no modifica el efecto estimado, ya que al menos en el modelo (7) el 0.98 de efecto estimado mantiene su significancia 1 %, aunque en el modelo (8), el resultado (1.11) es significativo al 5 %. Asimismo, se observa que la variable dummy de asia si es significativa tanto al 5 % como al 10 % para los modelos (7) y (8), respectivamente. Estos resultados, al igual que los modelos anteriores son superiores al efecto estimado por OLS.

- d. **Para saber si estamos o no ante la presencia de instrumentos débiles, testee usando el estadístico de Cragg y Donald junto con las tablas de Stock y Yogo con respecto al tamaño o size del test de Wald. ¿Cómo testean los autores la fortaleza del instrumento? Compare ambos resultados.**

Como parte del análisis de robustez, los autores verifican la validez de sus instrumentos usando un test de sobreidentificación, donde usan a las variables relacionadas a las instituciones tempranas y el establecimiento de las colonias europeas como instrumentos adicionales. De esta forma se evaluará si estos instrumentos adicionales y el instrumento de mortalidad son exógenos tomando en cuenta que al menos uno de los instrumentos es exógeno. El test de sobreidentificación será rechazado (es decir el instrumento será débil) si se cumplen tres escenarios (1) la ecuación de interés no tiene constante, (2) los instrumentos adicionales tienen efecto sobre el PBI per capita y (3) la mortalidad de los colonos tenga efecto sobre el PBI per capita a través de otra vía además de la especificada por los autores. Los autores sostienen que los resultados de la tabla 5 respaldan la fortaleza del instrumento "log mortalidad de los colonos", debido a que no se demostró que las variables instrumentales hayan tenido efecto sobre la variable dependiente.

Por otro lado, la presencia de instrumentos débiles indica una débil correlación entre

las variables endógenas y los instrumentos, esto sesga la estimación en la dirección del estimador MCO. Por esta razón, es importante testear si los instrumentos que se emplean en una estimación por variables instrumentales son fuertes, el estadístico más riguroso en la actualidad es el de Cragg y Donalds (1993) con las tablas de Stock y Yogo (2005) con respecto al tamaño del test de Wald.

Cuadro 3: Instrumentos débiles

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estadístico F	22.9468	13.0932	8.64607	7.82624	30.5419	21.6138	6.23297	3.4556
Prob >F	0.0000	0.0006	0.0047	0.0070	0.0000	0.0000	0.0154	0.0681

Acemoglu et al (2001) solo usa un instrumento "tasa de mortalidad" para limpiar de endogeneidad a la variable riesgo de expropiación", de esta forma, nos encontramos en el caso de una variable endógena y un instrumento, A continuación, en la tabla 3 se calcula el estadístico para todas las especificaciones que realizaron los autores y en la tabla 4 mostramos los valores críticos de las tablas de Stock y Yogo. Los valores críticos de esta última tabla ayudan a determinar el nivel de significancia de la inferencia.

Cuadro 4: Critical Values

	n = 1, r =			
K	0.10	0.15	0.20	0.25
1	16.83	8.96	6.66	5.53

- El estadístico F de la primera identificación es de 22.95 el cual es mayor a 16.83, por lo tanto, se puede decir que la significancia del test a la Wald es a lo más 10% cuando teóricamente era de 5%.
- El estadístico F de la segunda identificación es de 13.09 el cual es mayor a 8.96, por lo tanto, se puede decir que la significancia del test a la Wald es a lo más 15% cuando teóricamente era de 5%.
- El estadístico F de la tercera identificación es de 8.65 el cual es mayor a 6.66

por lo tanto, se puede decir que la significancia del test a la Wald es a lo más 20 % cuando teóricamente era de 5 %.

- El estadístico F de la cuarta identificación es de 7.83 el cual es mayor a 6.66 por lo tanto, se puede decir que la significancia del test a la Wald es a lo más 20 % cuando teóricamente era de 5 %.
- El estadístico F de la quinta identificación es de 30.54 el cual es mayor a 16.83 por lo tanto, se puede decir que la significancia del test a la Wald es a lo más 10 % cuando teóricamente era de 5 %.
- El estadístico F de la sexta identificación es de 21.62 el cual es mayor a 16.83 por lo tanto, se puede decir que la significancia del test a la Wald es a los más 10 % cuando teóricamente era de 5 %.
- El estadístico F de la séptima identificación es de 6.23 el cual es mayor a 5.53 por lo tanto, se puede decir que la significancia del test a la Wald es a los más 25 % cuando teóricamente era de 5 %.
- El estadístico F de la octava identificación es de 3.45 el cual no logra superar los valores críticos de la tabla de Stock y Yogo. En ese caso podemos decir, que la significancia empírica de ese modelo tiene un tope mayor al 25 %.

En conclusión, se puede decir que el instrumento log mortalidad de los colonos es fuerte a distintos niveles de significancia, excepto en la octava especificación. Esto se puede condecir con la justificación de la validez de instrumentos que realizan los autores a través de la sobreidentificación.

- e. **Replique la tabla 8 (excepto el panel C) del paper y comente sobre la interpretación que hacen los autores. Comente además acerca de los tests de sobreidentificación realizados aún cuando utilizan un solo instrumento.**

Cuadro 5: Overidentification Test

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Two-Stage Least Squares										
Average protection against expropriation risk, 1985-1995	0.87*** (0.14)	0.92*** (0.20)	0.71*** (0.15)	0.68*** (0.20)	0.72*** (0.14)	0.69*** (0.19)	0.34 (0.22)	0.33 (0.22)	0.42** (0.16)	0.38** (0.17)
Latitude		-0.47 (1.24)		0.34 (1.08)		0.31 (1.05)		1.77 (1.12)		1.55 (0.95)
Panel B: First Stage for Average Protection Against Expropriation Risk										
European settlements in 1900	0.03*** (0.01)	0.03*** (0.01)								
Constraint on executive in 1900			0.32*** (0.08)	0.26*** (0.09)						
Democracy in 1900					0.24*** (0.06)	0.20*** (0.07)				
Constraint on executive in first year of independence							0.18** (0.08)	0.17** (0.07)		
Democracy in first year of independence									0.16*** (0.05)	0.15*** (0.05)
Panel D: Second Stage with Log Mortality as Exogenous Variable										
Average protection against expropriation risk, 1985-1995	0.81*** (0.23)	0.88*** (0.29)	0.45* (0.25)	0.42 (0.30)	0.51** (0.23)	0.48* (0.27)	0.04 (0.34)	0.07 (0.31)	0.10 (0.26)	0.11 (0.25)
Log European settler mortality	-0.07 (0.16)	-0.05 (0.18)	-0.25 (0.16)	-0.26 (0.17)	-0.21 (0.15)	-0.22 (0.16)	-0.50** (0.22)	-0.43** (0.18)	-0.47** (0.18)	-0.41*** (0.15)
Latitude		-0.52 (1.15)		0.38 (0.89)		0.28 (0.86)		1.08 (0.98)		1.01 (0.88)

Standard errors in parentheses

Dependent variable: Log GDP per capita in 1995

* $p < 0,10$, ** $p < 0,05$, *** $p < 0,01$

Los autores construyen la tabla 5 para justificar la validez del instrumento que usan para esto usan un grupo de instrumentos adicionales. En el panel A se observa el resultado del estimador 2LSL para las diferentes especificaciones, estas se diferencian por los instrumentos que usan y en algunos casos se aumenta la variable latitud como variable de control. El panel B, muestra la primera etapa de la estimación, donde se regresa el riesgo de expropiación contra los instrumentos y variables de control. Finalmente, el panel D incluye a la variable log mortalidad de los colonos como variables de control.

Como se habia mencionado en la pregunta anterior, existen tres escenarios en los que

el test de sobreidentificación se rechazaría, por lo tanto, el instrumento no sería válido. Los autores se apoyan de los resultados del Panel A para comprobar, el segundo escenario, los instrumentos adicionales no tienen efecto sobre el PBI per capita, ya que los resultados son cercanos a los de la tabla 2. Esto es favorable, pues quiere decir que los nuevos instrumentos no tienen efecto sobre la variable dependiente de la ecuación estructural. Por ejemplo, en la columna (1) de la tabla 2 se usó a la log mortalidad de los colonos como instrumento de las instituciones, donde el efecto estimado fue de 0.94 con 0.16 de error estandar este resultado es comparable con la columna (1) de la tabla 5, donde se usó el asentamiento de los europeos en 1900 como instrumento de las instituciones, el efecto estimado fue de 0.87 con 0.14 de error estándar.

El no rechazo del test de sobreidentificación del tercer escenario se justifica en el Panel D, ya que al incluir la log mortalidad de los colonos como variable de control se observa que el efecto es negativo, en magnitudes reducidas y no significativos . En otras palabras, no existe efecto directo de la log mortalidad de los colonos sobre la el log PBI per capita que no sea a través del riesgo de expropiación. No obstante, los estimados de la variable log mortalidad de los colonos de las columnas (7) - (10) tienen una mayor magnitud y son significativos, estos resultados podría deberse a alguna falla de base de datos de la replica de esta investigación. Pero de no ser el caso, podríamos decir que los instrumentos no son válidos.

- f. **Utilizando los resultados de replicación de las tablas 2 y 4, realice el test de Hausman e interprete sus resultados. Nota: Esto debe hacerse para aquellas especificaciones que son iguales; es decir: tabla 2 columnas 2, 5 y 6 vs tabla 4 columnas 1, 2 y 8 respectivamente.**

Para verificar la existencia de endogeneidad en los estimadores IV o 2SLS se usa el test de Durbin-Wu-Hausman, este testea exogeneidad débil, donde la hipótesis nula

es que las variables son exógenas. Dicho esto, en la tabla 6 se observa el resultado de este test a las especificaciones (1), (2) y (8) de la tabla 2, pues esta tabla contiene los resultados del estimador 2SLS.

Se observa que en los tres casos el p-value del test es menor a 0.01, por lo que se rechaza la hipótesis de exogeneidad, esto quiere decir que la variable de riesgo de expropiación presenta endogeneidad porque mantiene correlación con el término de perturbación. Por tanto, usar variables instrumentales ayudará a limpiar la variable de efectos endógenos. El beneficio de usar variables instrumentales por 2SLS se observa en la tabla 2, pues las estimaciones son mayores y significativas en comparación a las estimaciones por OLS de la tabla 1, esto se debe a que se limpia la variable de riesgo de expropiación de fuentes exógenas.

Cuadro 6: Test de Durbin Wu Hausman			
Test de Endogeneidad			
Ho: Variables son exógenas			
	(1)	(2)	(8)
Durbin (score) $\chi^2(1)$	18.189	15.2386	9.37062
	(p = 0.0000)	(p = 0.0001)	(p = 0.0022)
Wu-Hausman F()	24.2196	18.7508	9.77725
	(p = 0.0000)	(p = 0.0001)	(p = 0.0028)

- g. Lea el paper **The Colonial Origins of Comparative Development: An Empirical Investigation: Comment de David Albouy (2012)**. ¿Cuál es el argumento central del paper? ¿Qué hace el autor para cuestionar los resultados de Acemoglu et al. (2001)? ¿Qué puede decir respecto a sus resultados en d) luego de leer el paper?

El aporte de Albouy (2012) es poner en tela de juicio la validez del instrumento usado por Acemoglu et al (2001), (AJR en adelante) y, por tanto, la credibilidad de sus resultados, ya que sostiene que la construcción de la base de datos de la log mortalidad de los colonos europeos presenta tres problemas. En primer lugar, de los 64 países usados para la estimación, solo 28 de estos tenían tasa de mortalidad que se

habían generado dentro de sus bordes geográficos, mientras que los otros 36 países tenían tasa de mortalidad imputadas basadas en supuestos que realizaron los autores, por ejemplo, algunos países recibieron la misma tasa de mortalidad de sus países vecinos, esto genera un problema de heterocedasticidad porque los errores estarán correlacionados por efectos de la clusterización. De esta forma, si se quitan los 36 países con tasa de mortalidad imputadas entonces la magnitud del efecto del riesgo de expropiación sobre el log PBI per capita cae y algunas variables de control ganan significancia. En segundo lugar, existe una distinción entre las tasa de mortalidad que afectaron a los soldados que estuvieron en los cuarteles en comparación con los soldados que estuvieron en campaña, pues estos últimos estaban más expuestos a enfermedades infectocontagiosas. Así Albouy resalta que AJR usa tasa de mortalidad altas para países con mayores riesgos de expropiación y con bajo log PBI per capita, de esta forma las tasa de mortalidad sufren de endogeneidad. En tercer lugar, AJR usan las máximas tasas de mortalidad de los obreros africanos importados recogidas de Curtin (1968) y no las tasas de mortalidad promedio.

Para cuestionar los resultados Albouy realiza tres pruebas de robustez:

- Eliminar los 36 con tasa de mortalidad imputadas (tasa de mortalidad originadas fuera de sus fronteras), incluido los datos latinoamericanos de referencia para usar una muestra reducida de 28 países.
- Agregar dos variables de control que indique cuando se toman tasas de mortalidad de soldados en campaña o de obreros africanos importados.
- Agregar los nuevos datos de AJR (2005) donde agregan las tasas de mortalidad reales para países como Australia, Bahamas, Guyana, Hong Kong, Honduras y Singapore.

Luego de leer la argumentación de Albouy, la postura que tomamos es más escépti-

ca ya que los datos de la tasa de mortalidad (instrumento) podrían presentar fallas sistemáticas que beneficiarían la hipótesis de AJR. Ya que los datos no estarían representando el verdadero proceso generador de datos. De esta forma, si los datos presentan sesgos entonces el test de instrumentos débiles de Cragg y Donal de la tabla 3 podría estar sesgado también. En pocas palabras, es necesario cuidar mucho del procesamiento de los datos, ya que datos de baja calidad nos dirigirán a resultados sesgados y lo que es peor, equivocados.

h. Utilizando la base de datos albouy 2012.dta replique la tabla 2 (Columnas 1-4, paneles A-D) de Albouy (2012) y comente sobre los resultados obtenidos.

La tabla 7 describe la primera etapa, donde se regresa el riesgo de expropiación contra la log mortalidad de los colonos. Asimismo, la columna (1) no usa variables de control, la columna (2) incluye a la latitud como una variable de control, la columna (3) excluye a los países "Neo-europeos" como Australia, Canadá, Nueva Zelanda y Estados Unidos y la columna (4) incluye a los continentes de África, Asia y otros ¹.

El panel A, usa la muestra original (64 países), asimismo se considera la existencia de heterocedasticidad de los errores por clusters. Se observa que la log mortalidad es un buen predictor del riesgo de expropiación ya que es significativo en todas las especificaciones, asimismo los errores estándar crecen debido a la heterocedasticidad.

El panel B solo cuenta solo los 28 países que tienen tasas de mortalidad dentro de sus fronteras, es decir se aplica la primera prueba de robustez. Los resultados muestran que los estimados de la variable log mortalidad se reducen en tamaño y se vuelven menos significativos. Además, la variable de control latitud se vuelve significativa, contrariamente al no aplicar esta prueba de robustez esta variable no era significativa en el Panel A.

¹ Australia, Malta y Nueva Zelanda

El panel C usa la muestra original y controla por tasas de mortalidad de soldados en campaña y de obreros africanos importados. Los resultados muestran que los estimados de la log mortalidad se vuelven aún más pequeños y menos significantivos.

Finalmente, el panel D reduce la cantidad de países a 28 e incluye las variables de control del panel C. Se observa que las variables de control se hacen significativas y los estimados de la log mortalidad pierden significancia cuando se aplica la prueba de robustez.

i. **Finalmente lea el paper *The Colonial Origins of Comparative Development: An Empirical Investigation: Reply* de Acemoglu et al. (2012). ¿Cómo contraargumentan los autores los resultados y críticas de Albouy (2012)?**

Las críticas de Albouy (2012) se enfocan principalmente en 2 puntos: (i) la asignación de datos para algunos países latinoamericanos y africanos no son confiables, así como (ii) la mezcla de información de las tasas de mortalidad de soldados en campaña y en cuarteles. De esta manera, Albouy (2012) acusa a Acemoglu et al (2001) de construir una serie de tasas de mortalidad que beneficie sus resultados.

Acemoglu et al (2012) se defienden de las críticas y responden a cada una de ellas. En primer lugar, respecto a la asignación de tasas de mortalidad a países vecinos, los autores defienden la confiabilidad de sus datos tanto para América Latina, como para los demás países restantes. Por un lado, para los países latinoamericanos, Acemoglu et al (2012) ofrecen una forma alternativa de vincular los conjuntos de datos de Curtin y Gutiérrez, de manera que crean nuevas series de tasas de mortalidad usando como benchmark a Jamaica y República Dominicana, en vez de usar a México como benchmark como lo hacían en el paper original. Asimismo, los autores proponen el uso de data de las estaciones navales británicas en tierras sudamericanas. De esta manera, Acemoglu defiende los datos utilizados para Latinoamérica al usar distintas

series de tasas de mortalidad para realizar nuevamente las regresiones y así demostrar la solidez de sus resultados en Acemoglu et al (2012) . Por otro lado, para los países restantes, los autores se basan en literatura de salud pública y seguros de vida para asegurar la existencia de información independiente que respalda las tasas de mortalidad usadas para estos países.

Respecto al segundo problema, Acemoglu et al (2012) recalcan que las tasas de mortalidad entre los soldados en campaña y en cuarteles no son tan distintos sistemáticamente como lo afirma Albouy (2012). Para argumentar ello, los autores se basan en el registro histórico que indica que aveces las campañas podían mantener tasas de mortalidad menores a las tasas de los civiles o de los soldados que se encontraban en los cuarteles. Esto principalmente, porque los soldados de campaña no permanecen mucho tiempo en un solo lugar, por lo que mudarse cada cierto tiempo podía reducir el impacto de una epidemia e incluso prevenir enfermedades que surgan por un saneamiento deficiente.

Acemoglu et al (2012) no solo se defienden, sino también critican la forma en la que Albouy (2012) lidia con los problemas anteriores. En primer lugar, los autores critican a Albouy por descartar los datos de 36 países argumentando que la información de dichos países no es confiable, ya que dicha información no se genera en sus fronteras. En segundo lugar, para ellos, la construcción de la dummy "campaña", utilizada para identificar a los soldados que estuvieron en campaña y en el cuartel, no es posible. Además, de ello señalan que Albouy realizó una clasificación de campaña inconsistente y selectiva, ya que asigna tasas de "no campaña" distintas al registro histórico en al menos diez países, de esta manera sesgando su data. Teniendo en cuenta lo mencionado, Acemoglu et al (2012) concluyen que los resultados que muestra Albouy en su paper no son sólidos ni consistentes.

Cuadro 7: First-Stage Estimates

	(1)	(2)	(3)	(4)
Panel A: Original data				
log settler mortality	-0.61*** (0.17)	-0.51** (0.19)	-0.39** (0.17)	-0.53*** (0.18)
Latitude		2.00 (1.47)		
Panel B: Removing conjectured mortality rates				
log settler mortality	-0.59*** (0.19)	-0.42* (0.22)	-0.32* (0.19)	-0.43** (0.20)
Latitude		3.34* (1.63)		
Panel C: Original Data, adding campaing and laborer indicators				
log settler mortality	-0.43** (0.18)	-0.37* (0.20)	-0.29* (0.17)	-0.24 (0.21)
Original mortality from campaign	-0.74 (0.46)	-0.73 (0.47)	-0.45 (0.42)	-1.32** (0.62)
Data from forced labor	-1.56* (0.88)	-1.35 (0.87)	-1.32 (0.85)	-2.14** (0.95)
Panel D: Removing conjectured mortality rates adding campaing and laborer indicators				
log settler mortality	-0.35 (0.22)	-0.21 (0.25)	-0.18 (0.22)	-0.24 (0.24)
Original mortality from campaign	-1.15* (0.57)	-1.17** (0.52)	-0.81 (0.49)	-1.02 (0.65)
Data from forced labor	-1.88** (0.67)	-1.33* (0.68)	-1.50** (0.64)	-1.64** (0.77)

Standard errors in parentheses

Dependent variable: Expropriation Risk

* $p < 0,10$, ** $p < 0,05$, *** $p < 0,01$