

Evaluación de Impacto - TP3

Leonardo A. Caravaggio

1. Variables Instrumentales

1.1. Efectos de largo plazo del comercio de esclavos

- Existe una amplia literatura sobre los efectos de la colonización europea. Distintos estudios se han enfocado en distintas aristas de este fenómeno, encontrado que los efectos negativos, en muchos casos, han perdurado por siglos. [Nunn y Wantchekon \(2011\)](#) estudian los efectos del comercio de esclavos en la confianza social o interpersonal en el continente africano. En esta sección vamos a replicar algunos de sus resultados.

a) La tabla 5 presenta la resultados principales para las estimación de variables instrumentales. Repliquen estas estimaciones. Los autores usan otro comando, por lo que los resultados que obtengan no tienen que ser exactamente iguales a los de [Nunn y Wantchekon \(2011\)](#). Usen clustered standard errors por distrito (district). Presenten los resultados para la first stage y second stage en un tabla. Incluyan también evidencia para evaluar si el instrumento es débil o no y comenten. Repitan este proceso pero clustereando los errores estándar al nivel del grupo étnico (murdock_name) en otra tabla. Comparen los resultados ¿Cuál estrategia les parece mejor para lidiar con los errores estándar?

Algunas cosas a tener en cuenta: - Las variables dependientes son las cinco variables con “trust” en el nombre. - En NunnWantchekon.do tienen el código para usar los controles sin tener que incluir todas las variables en cada especificación y evitar problemas de strings. - Además de los controles baseline y colonial incluyan, como otra variable de control, ln_init_pop_density. - Las variables ln_export_area y distsea son la independiente endógena y el instrumento, respectivamente.

El trabajo de [Nunn y Wantchekon \(2011\)](#) examina la relación entre los niveles actuales de confianza en África y el impacto de la trata de esclavos. Utilizando datos contemporáneos de encuestas a nivel individual y datos históricos sobre el envío de esclavos por grupo étnico, los autores encuentran que las personas cuyos ancestros fueron fuertemente afectados por la trata de esclavos tienden a ser menos confiadas en la actualidad. El estudio sugiere que esta relación es causal, y se enfoca en mecanismos internos del individuo, como normas culturales, creencias y valores, como las principales influencias.

El artículo se basa en la idea de que la trata de esclavos creó un entorno de inseguridad generalizada, llevando a individuos a volverse unos contra otros, incluso traicionando a amigos y familiares para venderlos como esclavos. Los autores proponen la hipótesis de que en este contexto, se desarrolló una cultura de desconfianza que perdura hasta hoy. Utilizando datos de encuestas de 2005, los autores encuentran que las personas pertenecientes a grupos étnicos más afectados por la trata de

esclavos tienden a tener niveles más bajos de confianza en sus parientes, vecinos, coetnias y gobierno local en la actualidad, respaldando así su hipótesis.

Utilizando los datos provistos, se intenta a continuación replicar los resultados.

En el Cuadro 1 se presenta la salida de la estimación por variables instrumentales utilizando errores estándar clusterizados a nivel de distrito, y en el Cuadro 2 a nivel de grupo étnico. En el primer caso el estadístico F de Kleibergen-Paap Wald arroja un resultado de validez para el instrumento utilizado con un valor que en todos los casos resulta mayor que el valor de [Stock y Yogo \(2005\)](#) para el 10 %. En el segundo caso los valores que se obtienen son sensiblemente menores y resultan inferiores al valor antedicho. Se observa en cambio que los coeficientes estimados no varían por cambiar el tipo de clusterización. En cambio, sí varían los errores estándar estimados.

Cuadro 1: Regresiones Instrumentales - Distrito

	ln_export_area	ln_init_pop_density	KP_F
trust_relatives	-.0998835	-.0460691	202.3524
R.S.E.	.0475199	.0162344	.
trust_neighbors	-.1775261	-.0039939	201.6459
R.S.E.	.0623044	.0212049	.
intra_group_trust	-.2322123	.0088827	201.3151
R.S.E.	.056466	.0205693	.
inter_group_trust	-.0686271	-.0415706	200.6347
R.S.E.	.0487715	.016986	.
trust_local_council	-.355391	.0226054	204.6662
R.S.E.	.0543039	.0181299	.

Nota: Se presentan los coeficientes de las regresiones instrumentales clusterizadas a nivel de distrito.

Cuadro 2: Regresiones Instrumentales - Grupo Étnico

	ln_export_area	ln_init_pop_density	KP_F
trust_relatives	-.0998835	-.0460691	11.23326
R.S.E.	.1048564	.0285036	.
trust_neighbors	-.1775261	-.0039939	11.18456
R.S.E.	.117386	.0325887	.
intra_group_trust	-.2322123	.0088827	11.18083
R.S.E.	.1196355	.0376776	.
inter_group_trust	-.0686271	-.0415706	11.14317
R.S.E.	.1040159	.0292683	.
trust_local_council	-.355391	.0226054	11.22766
R.S.E.	.1589381	.0474609	.

Nota: Se presentan los coeficientes de las regresiones instrumentales clusterizadas a nivel de grupo étnico.

- b) [Blandhol, Bonney, Mogstad, y Torgovitsky \(2022\)](#) muestran que el estimar de 2SLS puede estar sesgado cuando se incluyen controles. Vuelvan a estimar los resultados del punto anterior (con su especificación preferida para los errores estándar únicamente) sin incluir controles. Presenten los resultados en una tabla y comenten.

En una especificación sin controles (cuya salida se presenta en el Cuadro 3) se observa que no se modifican ni los coeficientes estimados, ni el error estándar ni el estadístico F.

Cuadro 3: Regresiones Instrumentales sin Controles

	ln_export_area	ln_init_pop_density	KP_F
trust_relatives	-.0998835	-.0460691	202.3524
R.S.E.	.0475199	.0162344	.
trust_neighbors	-.1775261	-.0039939	201.6459
R.S.E.	.0623044	.0212049	.
intra_group_trust	-.2322123	.0088827	201.3151
R.S.E.	.056466	.0205693	.
inter_group_trust	-.0686271	-.0415706	200.6347
R.S.E.	.0487715	.016986	.
trust_local_council	-.355391	.0226054	204.6662
R.S.E.	.0543039	.0181299	.

Nota: Se presentan los coeficientes de las regresiones instrumentales clusterizadas a nivel de distrito pero sin controles.

- c) Los autores incluyen muchas variables de control ¿Explican su elección en términos de los controles y por qué estos son necesarios? ¿Qué beneficios en términos de sesgo y/o varianza creen que aportan estos controles?

En relación a los controles utilizados por los autores, se busca representar el ingreso individual. Dado que no cuentan con una medida directa del ingreso, emplearon variables como ocupación, educación y condiciones de vida, que son indicadores sólidos del mismo.

La inclusión de un amplio conjunto de covariables por parte de los autores tiene como objetivo capturar cualquier efecto potencial de la influencia europea no relacionada con el comercio de esclavos en la confianza a largo plazo. Controlar por variables coloniales a nivel étnico, en particular, fortalece el cumplimiento de la restricción de exclusión por parte de la variable instrumental.

Añadir más variables puede reducir el sesgo pero también aumentar la varianza. Los autores destacan la necesidad de seleccionar cuidadosamente las variables, demostrando la relevancia de los diversos controles incluidos en este estudio.

- d) Supongamos que no tuviéramos datos disponibles para la variable endógena de interés. Planteen un modelo reduced form con el mismo instrumento. Presenten los resultados en una tabla y compárenlos con los resultados anteriores.

El Cuadro 1.1 muestra las salidas para cada una de las variables endógenas consideradas. Para estas especificaciones se utilizan los mismos controles que los utilizados anteriormente. Para la variable de distancia (anteriormente el instrumento) la relación encontrada es positiva y estadísticamente significativa. Esto quiere decir que en primer término a mayor distancia, mayor el nivel de confianza, lo que resulta compatible con la hipótesis sostenida por el trabajo.

Cuadro 4: Regresiones lineales

	ln_init_pop_density	distsea
trust_relatives	-.0657074	.0001214
trust_neighbors	-.0388197	.0002155
intra_group_trust	-.0367483	.0002827
inter_group_trust	-.0550239	.0000836
trust_local_council	-.0468042	.0004374

Nota: Se presentan los coeficientes de la estimación lineal entre las variables seleccionadas y el tratamiento.

1.2. 1.2 Censura y Corrupción

- Stanig (2014) estudia como la regulación estatal puede afectar la libertad de prensa, particularmente en lo que refiere a la cobertura de casos de corrupción. Lean el trabajo y contesten las siguientes preguntas:

a) ¿Qué es el chilling effect?

El “chilling effect” (efecto de enfriamiento o efecto desalentador) es un concepto legal y social que se refiere a la tendencia de las personas a restringir su comportamiento, expresión o actividades debido al temor a posibles represalias, sanciones o persecución, generalmente por parte de autoridades gubernamentales u otras entidades poderosas. Este temor a menudo surge como resultado de leyes, políticas, regulaciones o prácticas que pueden ser interpretadas como amenazantes para la libertad de expresión, la privacidad, la protesta o la participación en actividades legítimas.

El chilling effect puede tener un impacto negativo en la sociedad, ya que puede llevar a la autocensura y al debilitamiento de la participación pública en asuntos importantes. Cuando las personas sienten que sus acciones o palabras podrían llevar a consecuencias negativas, es menos probable que se expresen libremente, se involucren en la protesta pacífica o desafíen a las autoridades. Esto puede socavar la democracia, la diversidad de opiniones y la rendición de cuentas.

El concepto del chilling effect es particularmente relevante en el contexto de la libertad de expresión y la censura. Las leyes represivas o la vigilancia excesiva, por ejemplo, pueden hacer que las personas eviten hablar abierta y libremente sobre ciertos temas, lo que afecta la discusión pública y la diversidad de opiniones. Por lo tanto, el chilling effect es un tema importante en el debate sobre la protección de los derechos civiles y la libertad de expresión.

- b) Expliquen el experimento natural.

El estudio se centra en entender cómo las leyes que limitan la libertad de prensa afectan la cobertura de noticias sobre comportamientos incorrectos en la política, la burocracia y la policía en los medios de comunicación.

Para llevar a cabo este estudio, los investigadores observaron diferentes estados de un país (en este caso, México), que tienen diferentes restricciones legales en cuanto a la libertad de expresión. Estas restricciones varían de un estado a otro debido a las leyes penales estatales que regulan la libertad de expresión en lugar de una única ley federal que sea igual para todo el país.

El hecho de que los estados tengan diferentes grados de restricciones legales en cuanto a la libertad de expresión se asemeja a un “experimento natural”. Esto se

debe a que, en un experimento controlado, los investigadores manipulan una variable para ver cómo afecta a otra. En este caso, los investigadores no están manipulando las leyes, pero están observando cómo las diferencias en las leyes ya existentes afectan la cobertura de noticias sobre comportamientos incorrectos en los medios de comunicación.

Al comparar los estados con leyes más restrictivas en la libertad de expresión con los estados que tienen leyes menos restrictivas, los investigadores pueden analizar si la rigidez de las leyes tiene un impacto en la cantidad de noticias sobre corrupción y comportamientos incorrectos en la política y la administración pública que se informan en los medios. Esto les permite evaluar cómo las restricciones legales en la libertad de prensa pueden influir en la rendición de cuentas de los funcionarios gubernamentales y en la información disponible para los ciudadanos.

- c) ¿Cuál es el outcome de interés, la variable endógena, y el instrumento?

El “outcome de interés” en este estudio es la cobertura de corrupción (malas prácticas en la política, la burocracia y la policía) en los medios de comunicación. Esto significa que los investigadores están interesados en medir cuántas noticias sobre comportamientos incorrectos en el ámbito político, administrativo y policial se informan en los medios de comunicación.

En los modelos 1 a 3 de la Tabla 2 del trabajo la variable utilizada para esto es la cantidad de artículos sobre corrupción. En el modelo 4 es la cantidad de artículos sobre “petty corruption” (La corrupción menor, burocrática o administrativa tiene lugar en el extremo de la implementación de la política, donde el público se encuentra con los funcionarios públicos. La corrupción menor suele distinguirse de la corrupción política y “a gran escala”. Generalmente se trata de cantidades menores de dinero, pero el daño puede ser significativo en términos sociales.¹). El modelo 5 es una regresión logística que testea la presencia de al menos un artículo sobre gran corrupción.

La “variable endógena” es la rigidez de las leyes de libertad de expresión en los estados. Esto se refiere a cuán estrictas o restrictivas son las leyes que regulan la libertad de prensa en cada estado de México. Esta variable endógena es lo que los investigadores quieren analizar en relación con el outcome de interés.

Se utilizan dos instrumentos: *Homicide Law* (Ley de Homicidio): la pena mínima obligatoria para el delito de homicidio en cada estado. Se utiliza como un indicador de la severidad de las leyes penales en cada estado en relación con los homicidios. Y *Prison Escape* (Fuga de Prisioneros): se refiere a la pena máxima para la evasión de prisión, es decir, ayudar o alentar la fuga de un recluso de una prisión. Se utiliza como un indicador de la severidad de las leyes penales en relación con la evasión de prisioneros en cada estado.

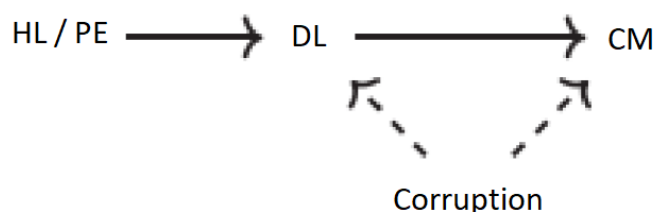
- d) ¿Por qué el autor enfatiza que se incluyen índices de corrupción como controles? Pueden pensarlo como un DAG.

Que los resultados sean robustos a la inclusión de controles a distintos counfoundings sería una prueba de que el chilling effect tiene un efecto causal.

En particular, la inclusión de medidas de corrupción pretende capturar la cantidad de historias reales de corrupción que haya para reportar. Es decir, podría haber menos reportes de historias de corrupción porque hay restricciones a la libertad de expresión o porque verdaderamente hay menos historias de corrupción que contar.

¹<https://etico.iiep.unesco.org/en/petty-corruption>

Figura 1: Directed Acyclic Graph



Nota: DAG de la relación causal entre las restricciones a la libertad de medios y los reportes de corrupción que aparecen en los medios de comunicación. Por HL entiéndase *Homicide Law* por PE *Prision Escape*, por DL *Defamation legislation* y por CM *Corruption reported in Media*.

- e) ¿Qué comparación estamos haciendo al incluir este control? Puede ser útil pensar en este punto como si el índice de corrupción fuese la única variable de control

La inclusión de este control busca garantizar que el efecto de las restricciones sobre la cantidad de noticias publicadas no se vea afectado por los niveles de corrupción.

- f) Muchos estudios sobre los efectos de los medios de comunicación presentan problemas de causalidad reversa (por ejemplo, las personas eligen qué medios consumir) ¿Es esto un problema en este trabajo?

Un efecto de este estilo podría verse si las personas tuvieran la posibilidad de informarse en medios de otro estado.

- g) ¿Cómo se defiende la exclusion restriction? ¿Es satisfactoria esta defensa? ¿Qué ocurre con la first stage?

Se considera extremadamente improbable que las leyes punitivas del sistema penal (como las leyes de homicidio y evasión de prisioneros mencionadas anteriormente) tengan un efecto directo sobre la cobertura de la corrupción en los medios de comunicación. En otras palabras, sería poco probable que las leyes penales de un estado afecten directamente la decisión de los medios de comunicación de informar sobre la corrupción en la política y la burocracia.

Podría, sin embargo, existir una cierta cultura tradicionalista o autoritaria que afecte a todas las variables por igual en un determinado estado. Es decir, generando leyes penales y de libertad de expresión más severas y también mayor interés en los medios de comunicación por poner en evidencia los hechos de corrupción.

Los coeficientes de los dos instrumentos en la *first stage* son positivos. El coeficiente de *Prision Escape* es altamente estadísticamente significativo.

- h) Para realizar tests de placebo, el autor crea una serie de variables. Suponiendo que puedan acceder a los datos, qué otro placebo se podría realizar para aumentar la confianza en este diseño?

El autor genera variables para utilizar como instrumento pero que no tienen correlación con la legislación de libertad de medios. Al fallar la *first stage* debería perderse el efecto encontrado. Esto es lo que sucede de acuerdo al autor. Podría pensarse también en proponer como test de placebo una variable como instrumento que viole la *exclusion restriction*. Utilizar este instrumento debería también tener efectos sobre el coeficiente estimado.

2. Regresión Discontinua

2.1. Head Start

- Ludwig y Miller (2007) analizan los efectos de largo plazo del programa Head Start (HS) a nivel del municipio (county) en USA. Vamos a enfocarnos en los efectos sobre la mortalidad infantil (`mort_age59_related_postHS`).

Las variables en la base de datos `headstart.dta` están descritas en el archivo `LM2007_description.pdf`. En esta vamos a analizar el efecto del programa HS en la mortalidad infantil.

Lean el trabajo de Ludwig y Miller (2007) y contesten las siguientes preguntas:
a) ¿Cuál es la unidad de análisis?

En un primer momento se utiliza el municipio (county) como unidad de análisis. Si bien lo que se busca identificar es el efecto en los chicos, esta información está agregada a nivel de municipio y proviene de la *National Archives and Records Administration (NARA)*, de los Censos y de los *Compressed Mortality Files (CMF)*.

En un segundo momento se utiliza información a nivel individual proveniente del *National Education Longitudinal Study (NELS)*. Para estos datos la unidad de análisis son los estudiantes.

- b) Observando los datos que tenemos, expliquen la temporalidad de los datos ¿Es un panel? ¿Son datos para el mismo año? ¿Por qué tenemos variables medidas en distintos momentos del tiempo?

No, no es un panel.

- c) ¿Cuál es la running variable y cuál el valor del cutoff? ¿Quién determinó este valor? ¿Deberíamos preocuparnos por una posible manipulación a nivel del county?

La *running variable* es la variable sobre la que se determina la asignación al tratamiento. En este caso, la asignación de ayuda para la redacción de subvenciones de Head Start se basó en las tasas de pobreza cinco años antes.

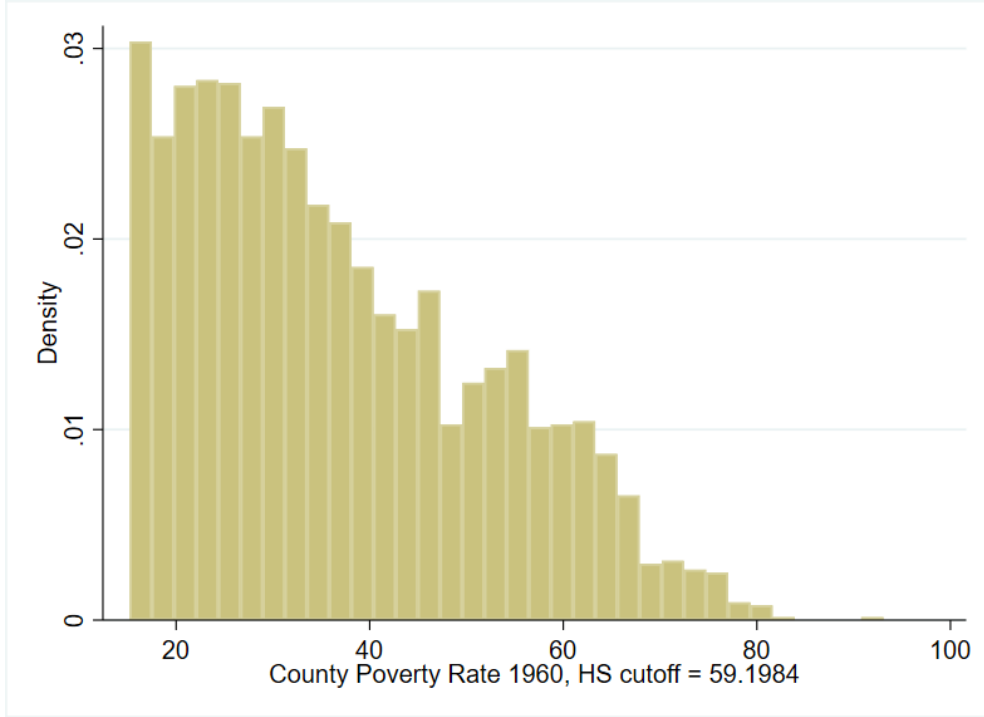
El valor del cutoff es $P_{300} = 59,1984$.

- d) Presentar evidencia descriptiva siempre ayuda a convencer al lector. Muestren un histograma usando como variable en el eje de las x la tasa de pobreza.
- Utilicen los datos de los autores para evaluar el efecto de HS en la mortalidad infantil.

e) Antes de comenzar, propongan una ecuación a estimar para el efecto que quieren medir. No busquen replicar los métodos y resultados de los autores. La idea de este ejercicio es aplicar estimaciones, comandos, y tests que se desarrollaron después de este trabajo. Usen las herramientas que discutimos en la clase, y presenten los resultados en tablas y gráficos. Expliquen los resultados para el efecto de este programa.

Dados los datos disponibles, las particularidades del caso y la pregunta de investigación, realicen todos los checks que crean posibles para validar el diseño de RDD. Expliquen en cada caso qué es lo que hacen y por qué ese check es

Figura 2: Histograma Tasa de Pobreza



Nota: Se presenta el histograma de la variable tasa de pobreza.

importante (por ejemplo, indicando los potenciales problemas si esos checks no se cumplieran). En base a toda esta evidencia, ¿qué pueden concluir sobre los resultados que encontraron para los efectos de HS sobre la mortalidad infantil?

Algunas cosas a tener en cuenta: - Tenemos distintas variables de mortalidad tanto en momentos del momento en el que fueron medidas como en causas de la mortalidad. Piensen cuáles son outcomes adecuados y cuáles pueden servir para otras estimaciones. - Aprovechen las variables de control que tienen. - Para decidir el tipo de errores estándar, piensen en la asignación del tratamiento.

El RDD queda representado por la ecuación (1). Donde el efecto a estimar es la prevalencia de la mortalidad infantil en los márgenes de la aplicación del programa de Head Start (HS).

$$Mort_{i,t+2} = \beta Pob_{i,t} + \gamma(Margin_{i,t}) + \psi X_{i,t} + \phi_s + \epsilon_{i,t} \quad (1)$$

En el Cuadro 5 se presentan las salidas de las estimaciones RDD sin (especificación 1) y con controles (especificación 2). Tal como predice la teoría, la inclusión de controles no cambia mucho los resultados.

En las Figuras 3 y 4 se presentan los Scatter plots de las tasas de mortalidad que buscan identificar el cutoff de la tasa de pobreza.

La Figura 5 presenta el test de Sorting para el cutoff seleccionado.

Referencias

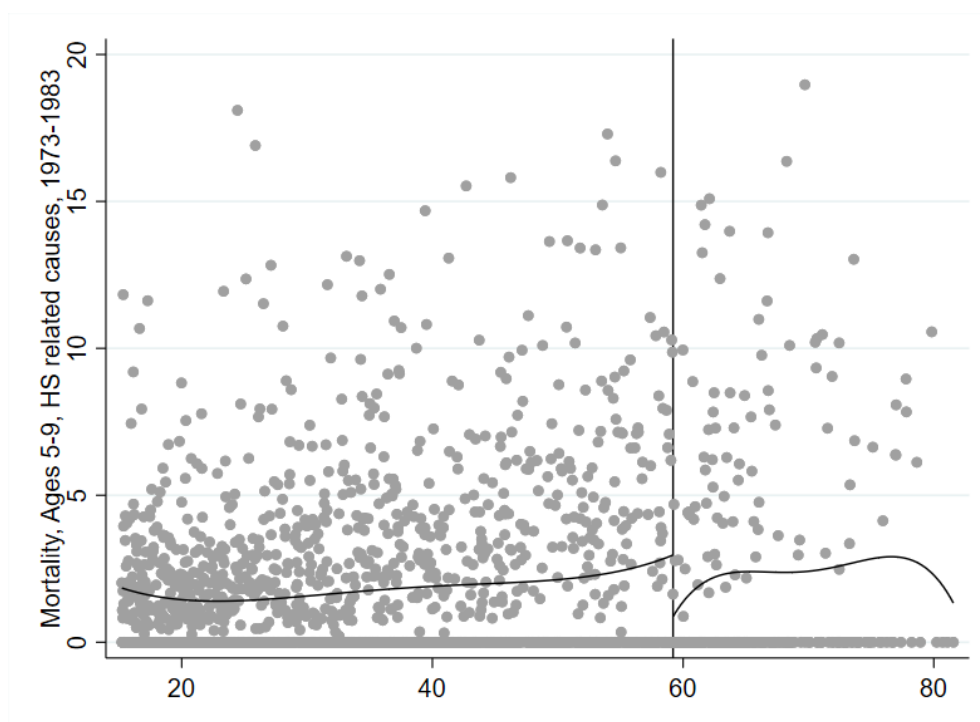
Blandhol, C., Bonney, J., Mogstad, M., y Torgovitsky, A. (2022, January). *When is *tsls* actually late?* (Working Paper n.º 29709). National Bureau of Economic Research. Descargado de <http://www.nber.org/papers/w29709> doi: 10.3386/w29709

Cuadro 5: RDD - Sin Controles

	(1)	(2)
	RDD Simple	RDD con Controles
RD_Estimate	-2.114** (0.990)	-2.314*** (0.898)
Observations	2,783	2,779

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

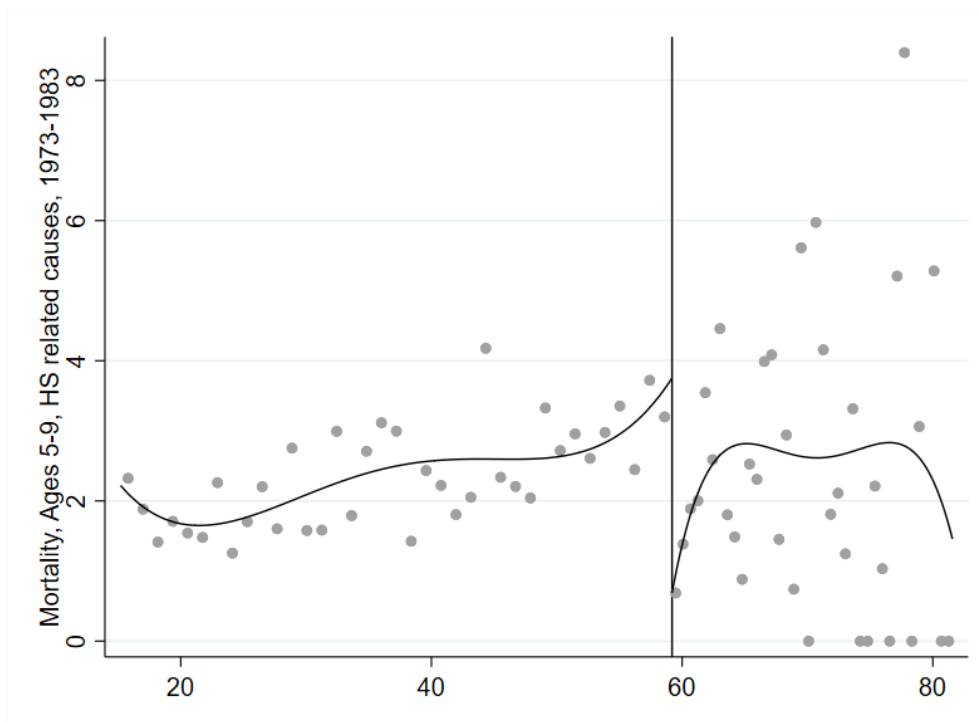
Figura 3: Scatter



Nota: Se observan las tasas de mortalidad en relación a la tasa de pobreza y con la línea vertical el cutoff. Se impone 20 % como límite a la variable de mortalidad de manera de mejorar la visualización.

- Ludwig, J., y Miller, D. L. (2007, 02). Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design*. *The Quarterly Journal of Economics*, 122(1), 159-208. Descargado de <https://doi.org/10.1162/qjec.122.1.159> doi: 10.1162/qjec.122.1.159
- Nunn, N., y Wantchekon, L. (2011). The slave trade and the origins of mistrust in africa. *American Economic Review*, 101(7), 3221-52. Descargado de <https://EconPapers.repec.org/RePEc:aea:aecrev:v:101:y:2011:i:7:p:3221-52>
- Stanig, P. (2014). Regulation of speech and media coverage of corruption: An empirical analysis of the mexican press. *American Journal of Political Science*, 59(1), 175-193. Descargado 2023-10-25, de <http://www.jstor.org/stable/24363603>
- Stock, J., y Yogo, M. (2005). Testing for weak instruments in linear iv regression. En D. W. Andrews (Ed.), *Identification and inference for econo-*

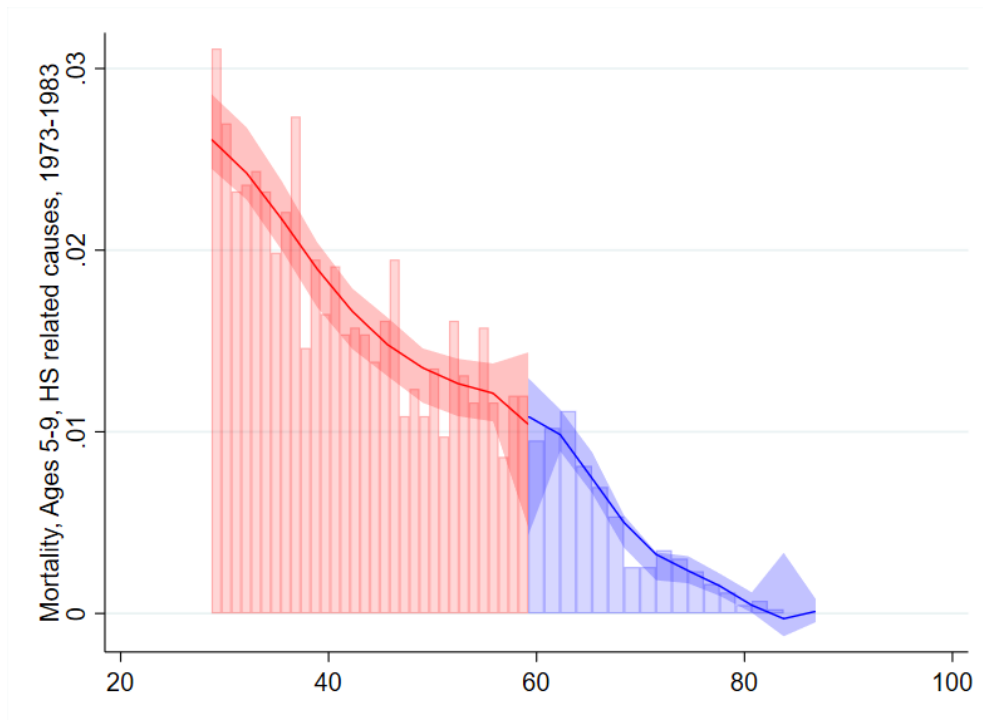
Figura 4: Scatter



Nota: Se observan las tasas de mortalidad en relación a la tasa de pobreza y con la línea vertical el cutoff. Se impone 8 % como límite a la variable de mortalidad de manera de mejorar la visualización.

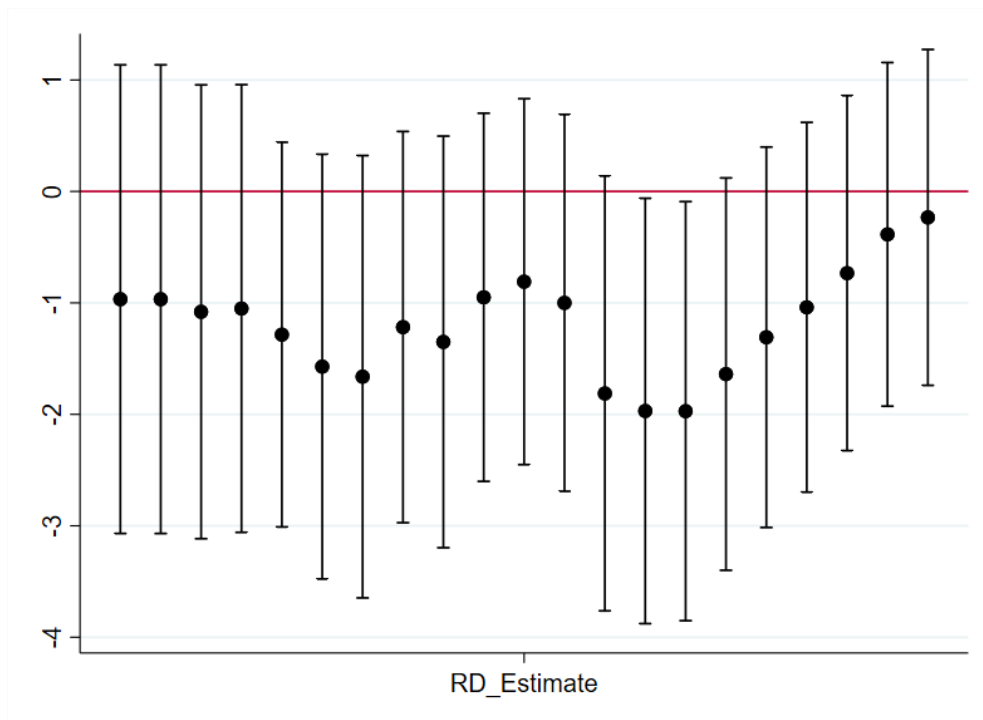
metric models (p. 80-108). New York: Cambridge University Press. Descargado de http://www.economics.harvard.edu/faculty/stock/files/TestingWeakInstr_Stock%2BYogo.pdf

Figura 5: Sorting



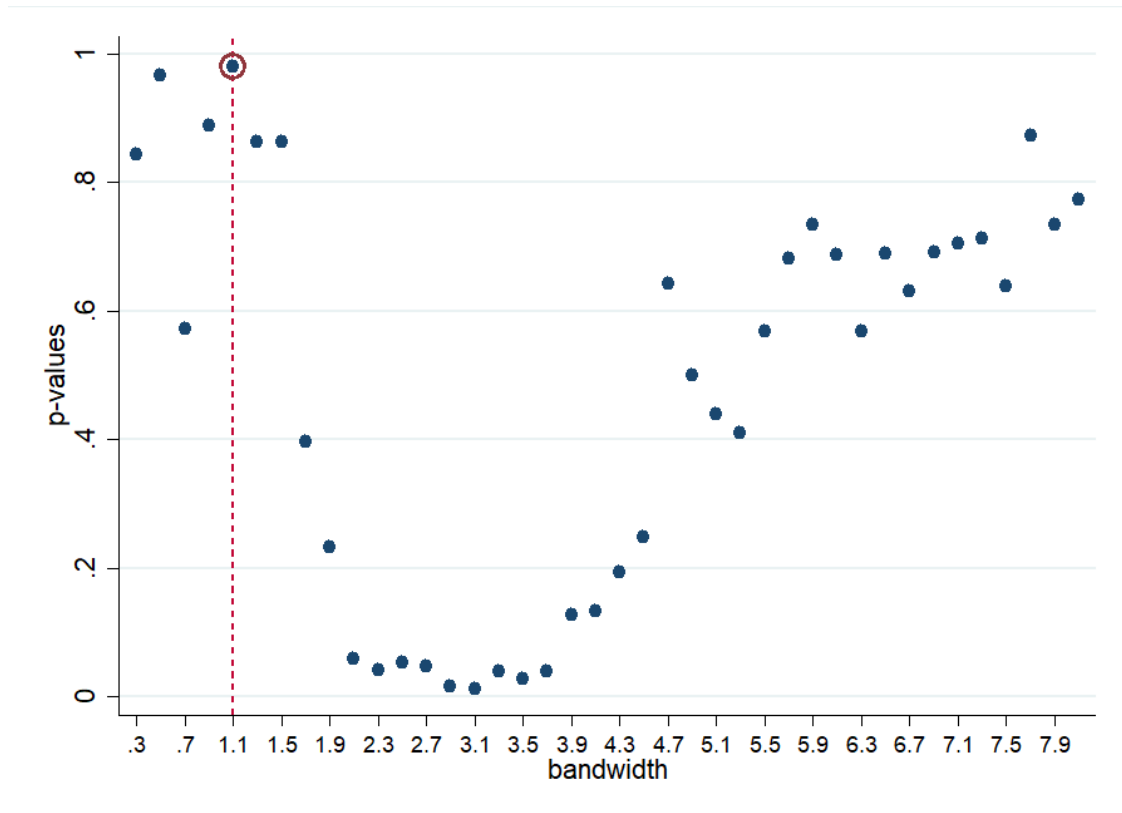
Nota: Se presenta el gráfico de la salida del testeo de sorting.

Figura 6: Cutoffs Irrelevantes



Nota: Se presenta el gráfico de la salida del testeo de Cutoffs Irrelevantes.

Figura 7: Gráfico de P-Values



Nota: Se presenta el gráfico de la salida del testeo de p-values.