



INSTITUTO DE SOCIOLOGÍA  
FACULTAD DE CIENCIAS SOCIALES

## Tarea N°2

Inferencia Causal - SOL3063

Estudiante [Andreas Laffert](#)

Profesor Luis Maldonado

Ayudante Gustavo Ahumada

sábado 29, junio 2024

## Pregunta 1

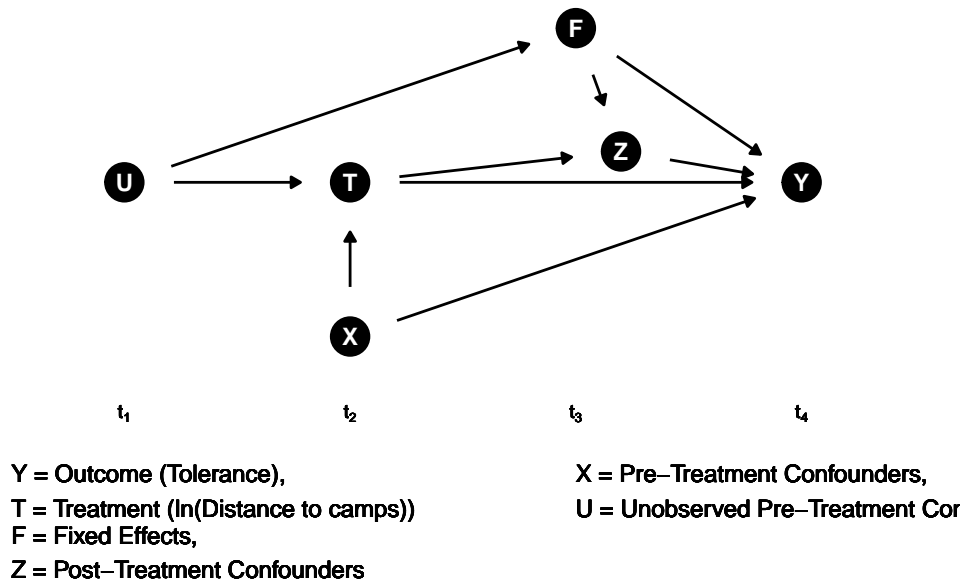
Pepinsky et al. (2024) sostienen que “The causal effect of Distance to nearest camp in Figure 3 is still only identifiable with Länder fixed effects” (p5). En base a este texto, explique esta idea.

Pepinsky et al. (2023) sostienen que los resultados obtenidos por Homola et al. (2020) sobre el efecto de la distancia a los campos de concentración en actitudes políticas contemporáneas son producto de una heterogeneidad espacial no observada, asociada a diferencias entre los Estados-Länder alemanes. Argumentan que la relación causal entre la distancia a los campos y las actitudes hacia exogrupos como la intolerancia está confundida por características económicas y la historia política e institucional de los Estados, que pueden influir en la educación cívica y el currículum escolar, y en consecuencia, en la opinión pública de los individuos (Pepinsky et al., 2023, p. 2). Para abordar este desafío, proponen emplear efectos fijos (EF) para los Estados, ajustando por cualquier factor constante en el tiempo que varía entre ellos y puede influir en la relación causal. El estimador de EF captura las características observables e inobservables constantes en el tiempo dentro de las unidades, pero que varían entre ellas, eliminando así la heterogeneidad no observada y explotando los factores que cambian con el tiempo (Wooldridge, 2009, p. 456). Con esta estrategia, concluyen que el efecto negativo de la distancia a los campos en la intolerancia desaparece al ajustar por EF para los Länder contemporáneos.

Pepinsky et al. (2023) controlan por EF para los Estados-Länder contemporáneos y buscan aclarar en qué medida estos EF son buenos controles en la estimación causal, asegurándose de que no produzcan sesgo de pos-tratamiento. Indican que, para evitar sesgo, deben cumplirse dos supuestos: primero, la variable de confusión  $F$  no debe estar en el camino causal de  $T$  hacia  $Y$ , es decir,  $F$  no debe ser descendiente del tratamiento  $T$  (2023, p. 3). Segundo, aunque se cumpla el primer supuesto, controlar por  $F$  podría generar M-bias, un tipo de sesgo de colisión (Pearl & Mackenzie, 2018). Para evitar este sesgo,  $F$  no debe ser descendiente de una variable  $U1$  de la cual  $T$  también es descendiente, ni de una variable  $U2$  de la cual  $Y$  también sea descendiente (2023, p. 3).

En respuesta a esta revisión crítica, Homola et al. (2024) argumentan que los supuestos planteados por Pepinsky et al. (2023), particularmente el supuesto sobre el no sesgo de pos-tratamiento, son insostenibles. Sostienen que los Estados contemporáneos son consecuencia directa o indirecta de la existencia previa de los campos de concentración, lo que haría insostenible el primer supuesto y sesgaría las estimaciones de Pepinsky et al. (2023). Esto se debe a que los campos de concentración afectan diversas características económicas e institucionales de los Estados, como el desempleo o el currículum escolar. Homola et al. (2024) indican que la proximidad a los campos influye en el currículum escolar, ya que las escuelas cercanas son más propensas a organizar visitas a estos campos, generando variación regional en la educación. Por tanto, controlar por EF generaría estimaciones sesgadas de pos-tratamiento. Proponen utilizar efectos directos controlados mediante la estimación-g secuencial para conocer el efecto causal de la distancia en la intolerancia, considerando las variables pos-tratamiento que median esta relación. Con esta estrategia, sus resultados originales se mantienen.

Figure 1: DAG de la Figura 3 Pepinsky et al. (2024)



Pepinsky et al. (2024) refutan esto, argumentando que Homola et al. (2024) no comprenden los errores de estimación. Sostienen que para que sus estimaciones con EF generen sesgo de pos-tratamiento, debe demostrarse que la distancia a los campos tiene un efecto causal en la formación de los Länder contemporáneos o sus características. Aunque esto no está respaldado por la evidencia presentada, Pepinsky et al. (2024) asumen la estructura causal de Homola et al. (2024), donde la distancia a los campos afecta variables pos-tratamiento  $Z$  (como el currículum escolar) que son descendientes de los EF para Länder (ver Figura 1). Pepinsky et al. (2024) sostienen que, incluso con esta estructura, el efecto causal de  $T$  en  $Y$  es identificable solo con la inclusión de EF para los Estados, ya que esto controla la heterogeneidad no observada que confunde la relación causal. En este sentido, los EF actúan como bloqueadores de confusión del factor  $U$  que afecta tanto a  $T$  como a  $Y$  (Wysocki et al., 2022).

La raíz de esta discusión es si la distancia a los campos es un antecesor causal de los EF para los Länder. Para Pepinsky et al. (2024), la interpretación de Homola et al. (2024) sobre los EF está errada, ya que confunden las características que varían entre Estados con aquellas que varían dentro de ellos. En el caso del currículum escolar, los EF no capturan su variación, ya que esta varía principalmente dentro de los Estados y no entre ellos. “Homola et al. han confundido la observación de que la proximidad a los campamentos afecta a los programas escolares individuales con la observación de que los Estados determinan los programas educativos estatales” (Pepinsky et al., 2024, p. 3). En consecuencia, Homola et al. (2024) atribuyen variables pos-tratamiento, que pueden ser mecanismos explicativos de la relación causal, a atributos que varían entre Estados, fusionando el confundidor  $F$  con la variable pos-tratamiento  $Z$  en la Figura 1.

## Pregunta 2

Tenemos por objetivo evaluar el efecto de la intolerancia en el apoyo a partidos de extrema derecha. Para ello, utilice la distancia a campos de concentración como instrumento. Específicamente, estime los siguientes estimands:

- ITTD sin covariables y ITTD controlando por efectos fijos para Länders.
- ITT sin covariables y ITT controlando por efectos fijos para Länders.
- CACE sin covariables y CACE controlando por efectos fijos para Länders.

Reporte sus estimaciones en dos tablas de calidad: una para estimaciones sin covariables y otra con las estimaciones que controlan por FE para Länders.

Table 1: Modelos de VI sin efectos fijos para apoyo a partidos de extrema derecha en base a Pepinsky et al. (2023)

|                     | Modelos VI          |                    |                  |
|---------------------|---------------------|--------------------|------------------|
|                     | $ITT_D$             | $ITT$              | $CACE$           |
| Distancia al campo  | -0.009**<br>(0.003) | -0.001*<br>(0.001) |                  |
| Intolerancia        |                     |                    | 0.154<br>(0.079) |
| Länder FE           | No                  | No                 | No               |
| R <sup>2</sup>      | 0.005               | 0.002              | -0.206           |
| Adj. R <sup>2</sup> | 0.005               | 0.002              | -0.207           |
| Num. obs.           | 2075                | 2075               | 2075             |
| RMSE                | 0.685               | 0.153              | 0.168            |

Nota: Errores estándares robustos entre paréntesis. \*\* $p < 0.01$ ; \* $p < 0.05$

Table 2: Modelos de VI con efectos fijos para apoyo a partidos de extrema derecha en base a Pepinsky et al. (2023)

|                     | Modelos VI       |                  |                  |
|---------------------|------------------|------------------|------------------|
|                     | $ITT_D$          | $ITT$            | $CACE$           |
| Distancia al campo  | 0.004<br>(0.004) | 0.001<br>(0.001) |                  |
| Intolerancia        |                  |                  | 0.148<br>(0.307) |
| Länder FE           | Sí               | Sí               | Sí               |
| R <sup>2</sup>      | 0.039            | 0.016            | -0.164           |
| Adj. R <sup>2</sup> | 0.031            | 0.009            | -0.173           |
| Num. obs.           | 2075             | 2075             | 2075             |
| RMSE                | 0.676            | 0.153            | 0.166            |

Nota: Errores estándares robustos entre paréntesis. \*\* $p < 0.01$ ; \* $p < 0.05$

## Pregunta 3

En relación con sus estimaciones en el ítem 2:

- a) Interprete sus estimaciones: ITTD , ITT, y CACE. Además, mencione quiénes serían los cumplidores.

El estimador de variable instrumental (VI) es útil para lidiar con posibles sesgos de variable omitida (heterogeneidad no observada) que al estar correlacionada con nuestras covariables introducen fuentes de endogeneidad en la estimación. Este estimador también puede combinarse con métodos de datos de panel, como los efectos fijos, para estimar de forma consistente los parámetros cuando tenemos efectos inobservables y endogeneidad en una o más variables independientes que varían con el tiempo (Wooldridge, 2009, p. 534). Asumiendo una situación de incumplimiento de asignación al tratamiento, se realizó una estimación de VI para conocer la relación causal entre la intolerancia y el apoyo a partidos de extrema derecha, empleando la distancia a los campos de concentración como instrumento con los datos de Pepinsky et al. (2023). En las Tablas 1 y 2 se muestran los modelos de VI para apoyo a partidos de extrema derecha, distinguiendo entre el efecto causal de la asignación del tratamiento (*ITT*) y el efecto causal promedio para los que recibieron el tratamiento (*CACE*) y si se controla (o no) por efectos fijos para Länders.

### ITT\_D e ITT

En cuanto al efecto de asignación del tratamiento ( $ITT_D$  e  $ITT$ ):

- $ITT_D$ : En la Tabla 1, el modelo para  $ITT_D$  indica que el efecto de la distancia sobre la intolerancia es negativo y estadísticamente significativo con un 99% de confianza ( $\beta = -0.009$ ,  $SE = 0.003$ ,  $p < 0.01$ ). Sin embargo, este efecto deja de ser estadísticamente significativo en la Tabla 2 al controlar por EF para Länders, además de que su signo cambia a positivo ( $\beta = 0.004$ ,  $SE = 0.004$ ,  $p > 0.05$ ).
- $ITT$ : Los resultados del modelo para  $ITT$  de la Tabla 1 sugieren que existe una relación negativa de la distancia a los campos de concentración en el apoyo a partidos de extrema derecha ( $\beta = -0.001$ ,  $SE = 0.001$ ), siendo un efecto estadísticamente significativo al 95% de confianza. Al controlar por EF para Länders en la Tabla 2, los resultados del  $ITT$  dejan de ser estadísticamente significativos aunque su dirección se mantiene.

### CACE

Respecto al efecto promedio del tratamiento para aquellos que cumplen con la asignación del tratamiento según el instrumento (*CACE*), los resultados de los modelos de VI muestran que, aunque su efecto es positivo, no es estadísticamente significativo. En la Tabla 1, el modelo *CACE* sin EF sugiere que el efecto promedio de la intolerancia sobre el apoyo a partidos de extrema derecha para los cumplidores es igual 0.15, aunque este efecto no es estadísticamente significativo a un 95% de confianza ( $SE = 0.079$ ,  $p > 0.05$ ). Al controlar por EF para Estados en

modelo *CACE* de la Tabla 2 los resultados siguen siendo los mismos ( $\beta = 0.148$ ,  $SE = 0.307$ ). Con todo, no se encuentra evidencia de que la intolerancia tenga un efecto causal sobre el apoyo a partidos de extrema derecha. En cualquier diseño de variables instrumentales, la subpoblación que toma (o no toma) el tratamiento debido a la variación en el instrumento se conoce como el conjunto de “cumplidores”. En este estudio, los cumplidores serían aquellos cuya proximidad a los campos de concentración determina su nivel de intolerancia y, en consecuencia, su apoyo a partidos de extrema derecha.

## Supuestos de VI y CACE:

Para que una variable  $Z$  sea un instrumento válido para el efecto causal del tratamiento en el resultado debe satisfacer una serie de supuestos, los cuales son similares a la estimación de *CACE* bajo incumplimiento (Wooldridge, 2009):

**Relevancia:**  $Z$  debe estar correlacionada con la variable endógena  $X$ :  $Cor(Z, X) \neq 0$ . Generalmente este supuesto se comprueba con una regresión lineal entre el instrumento y la variable independiente endógena (primera etapa). En este caso, el instrumento  $Z$  distancia a los campos de concentración se asocia significativamente con la variable endógena intolerancia cuando no se incluyen EF para Estados, lo que indica que este efecto depende en gran medida de las diferencias propias de los Estados, que son constantes en el tiempo. Por tanto, los resultados sugieren que la distancia no es un buen instrumento para la relación causal entre intolerancia y apoyo a partidos de extrema derecha.

**Exogeneidad:**  $Z$  no debe correlacionarse con el término de error  $U$  que captura la heterogeneidad no observada:  $Cor(Z, U) = 0$ . Esto generalmente se asume mediante teoría, ya que no es posible de comprobar empíricamente. Sustantivamente, implica asumir que la distancia a los campos no se correlaciona con ningún otro factor no observado que pueda confundir la relación causal de interés.

**Restricción de Exclusión:**  $Z$  debe afectar  $Y$  (apoyo a partidos de extrema derecha) solo a través de  $X$  (intolerancia):  $Cor(Z, Y|X) = 0$ . Esto significa que la distancia a los campos sólo puede afectar al apoyo a partidos de extrema derecha a través de la intolerancia.

- b) Discuta tres supuestos de sus estimaciones de CACEs, ¿son realistas? ¿se podría señalar que los CACEs estimados son creíbles?

Para realizar una estimación *CACE* deben considerarse la evaluación y cumplimiento de, al menos, cinco supuestos: i) SUTVA, ii) restricción de exclusión, iii) asignación aleatoria del instrumento ( $Z$ ), iv) efecto promedio del instrumento ( $Z$ ) sobre el tratamiento ( $D$ ) debe ser diferente de cero, y v) la relación entre el instrumento ( $Z$ ) y la probabilidad de recibir el tratamiento ( $D$ ) debe ser consistente (monotónica) para todas las unidades. En este ejercicio, evaluaré los primeros tres supuestos de *CACE*.

**SUTVA (Stable Unit Treatment Value Assumption):** El cumplimiento de SUTVA implicaría que la distancia a los campos de concentración ( $Z$ ) afecta de manera consistente la intolerancia

( $X$ ) dentro de cada individuo, sin que haya interacciones o efectos cruzados entre individuos debido a variaciones en  $Z$ . Esto significa que  $Z$  influiría de manera uniforme en la intolerancia de todos los individuos sin efectos indirectos entre ellos. Sin embargo, demostrar este supuesto puede ser complicado, ya que, como argumentan Pepinsky et al. (2023), la relación causal entre la distancia a los campos y la intolerancia podría estar confundida por una heterogeneidad espacial no observada asociada a características específicas de los Estados-Länders alemanes. Esto sugiere que las diferencias geográficas podrían interactuar con  $Z$  y  $X$  de maneras que comprometen la uniformidad del efecto de  $Z$  sobre  $X$  entre todos los individuos, lo cual podría plantear dudas sobre el cumplimiento de SUTVA.

**Asignación aleatoria de  $Z$ :** Este supuesto implica que  $Z$  (distancia a los campos de concentración) es asignado de manera aleatoria y no está correlacionado con otras características observadas o no observadas que podrían influir tanto en la intolerancia ( $X$ ) como en el apoyo a partidos de extrema derecha ( $Y$ ). Esto asegura que  $Z$  sea exógeno y no endógeno en relación con los individuos. En nuestro estudio, se garantiza que la distancia a los campos de concentración no está influenciada por características individuales que podrían determinar la presencia o ausencia de campos cercanos a los individuos. Además, es plausible sostener que la distribución de los campos de concentración entre los Länders es potencialmente aleatoria, ya que estos están presentes en algunos Länders y ausentes en otros (Pepinsky et al., 2023)

**Restricción de Exclusión:** Este supuesto establece que la distancia a los campos ( $Z$ ) solo puede afectar el apoyo a partidos de extrema derecha ( $Y$ ) a través del mecanismo de la intolerancia ( $X$ ), sin tener efectos directos no mediados por  $X$  en  $Y$ . Es fundamental para asegurar que las estimaciones de  $CACE$  capturen de manera precisa el efecto causal de interés. Sin embargo, demostrar este supuesto tanto empírica como teóricamente presenta desafíos significativos, dado que pueden existir diversas variables, tanto observadas como no observadas después del tratamiento, que actúen como mediadores o confundidores en la relación causal de interés. Como muestran los resultados, al incluir efectos fijos (EF), la distancia a los campos deja de ser un buen instrumento para la intolerancia, lo cual cuestiona la validez del supuesto de restricción de exclusión.

## Pregunta 4

Utilizando intolerance como variable resultado, estime los modelos 5 y 6 de Table 1 en Pepinsky et al. (2023). Específicamente:

- Reporte sus resultados en una tabla de calidad similar a la Table 1 del artículo bajo replicación. Use las covariables mencionadas arriba en esta pauta<sup>1</sup>.
- En base a la Table 1 de Pepinsky et al. (2023), discuta dos diferencias conceptuales entre el modelo (5) y el modelo (3).
- Utilizando como ejemplo sus estimaciones del modelo (6), explique las nociones de

---

<sup>1</sup>Note que no debe incluir como variable independiente la población en 1925.

Table 3: Replicación de modelos de efectos directos controlados Pepinsky et al. (2023)

|                                   | Modelo 5            | Modelo 6            |
|-----------------------------------|---------------------|---------------------|
| Distancia al campo                | −0.016**<br>(0.003) | 0.006<br>(0.005)    |
| % Judíos (1925)                   | −0.897<br>(1.341)   | 5.924<br>(4.412)    |
| % Desempleo (1933)                | 2.411**<br>(0.589)  | 1.075<br>(0.744)    |
| Participación partido nazi (1933) | −0.156<br>(0.213)   | −0.717**<br>(0.257) |
| Länder FE                         | No                  | Sí                  |
| Método                            | G-est               | G-est               |
| Num. obs.                         | 1376.000            | 1376.000            |

Nota: Errores estándares entre paréntesis. \*\* $p < 0.01$ ; \* $p < 0.05$

efecto directo controlado promedio (ACDE) y efecto directo natural promedio (ANDE).

## Pregunta 5

En nota al pie 2, Pepinsky et al. (2024) sostienen:

Sequential g-estimation, which HPT propose as a solution because they believe that Länder fixed effects are post-treatment variables, is unnecessary. It is also probably biased, because sequential g-estimation also requires different assumptions that they neither acknowledge nor defend, and only identifies a specific causal effect of T under the assumption of sequential unconfoundedness (see Acharya et al., 2016, p. 519), which would require (among other things) that there are no unobserved confounders of the causal relationship between states and contemporary tolerance. Without that assumption, about which HPT are silent, sequential g-estimation does not resolve any identification problem.

Al respecto, explique por qué la estimación-g secuencial que proponen Homola et al. (2024) sería probablemente sesgada.

## Referencias

- Acharya, A., Blackwell, M., & Sen, M. (2016). Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects. *American Political Science Review*, 110(3), 512–529. <https://doi.org/10.1017/S0003055416000216>
- Homola, J., Pereira, M. M., & Tavits, M. (2020). Legacies of the Third Reich: Concentration Camps and Out-group Intolerance. *American Political Science Review*, 114(2), 573–590. <https://doi.org/10.1017/S0003055419000832>
- Homola, J., Pereira, M. M., & Tavits, M. (2024). Fixed Effects and Post-Treatment Bias in Legacy Studies. *American Political Science Review*, 118(1), 537–544. <https://doi.org/10.1017/XPS.2024.1>



017/S0003055423001351

- Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect*. New York: Basic Books.
- Pepinsky, T. B., Goodman, S. W., & Ziller, C. (2023). Modeling Spatial Heterogeneity and Historical Persistence: Nazi Concentration Camps and Contemporary Intolerance. *American Political Science Review*, 118(1), 519–528. <https://doi.org/10.1017/S0003055423000072>
- Pepinsky, T. B., Goodman, S. W., & Ziller, C. (2024). Causation and History in Legacy Studies: A Reply to Homola, Pereira, and Tavits. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4705690>
- Wooldridge, J. M. (2009). *Introductory econometrics: a modern approach* (4th ed). Mason, OH: South Western, Cengage Learning.
- Wysocki, A. C., Lawson, K. M., & Rhemtulla, M. (2022). Statistical Control Requires Causal Justification.

## Código de R

```
knitr::opts_chunk$set(echo = F,
                      warning = F,
                      error = F,
                      message = F)
if (!require("pacman")) install.packages("pacman")

pacman::p_load(tidyverse,
              rio,
              sjmisc,
              AER,
              sandwich,
              DirectEffects,
              estimatr,
              panelr,
              plm,
              clubSandwich,
              texreg,
              ggeffects)

options(scipen=999)
rm(list = ls())
db_or <- rio::import(file = "../input/data/EVS_main.csv") %>%
  as_tibble()

miles <- function(x) {
  format(round(as.numeric(x), 0), big.mark = ".")
}

decimales <- function(x) {
  format(round(as.numeric(x), 2), decimal.mark = ",")
}

custom_extract <- function(model) {
  tr <- extract(model)

  # Identificar índices a conservar (excluyendo "R$^2$", "s_idios" y "s_i
  gof_indices <- which(!(tr@gof.names %in% c("R$^2$", "s_idios", "s_id")))

  # Actualizar gof, gof.names y gof.decimal simultáneamente
```

```

tr@gof.names <- tr@gof.names[gof_indices]
tr@gof <- tr@gof[gof_indices]
tr@gof.decimal <- tr@gof.decimal[gof_indices]

return(tr)
}
# set theme

theme_set(theme_bw())

library(dagitty) # Instalar V8
library(ggdag)

coords <- list(
  x = c(Y=2, T=0, U=-1, F=1, X=0, T1=-1, T2=0, Z=1.2),
  y = c(Y=1, T=1, U=1, F=2, X=0, T1=-1, T2=-1, Z=1.2)
)

HPTbestcase <- dagify(
  Y ~ T + F + X + Z,
  F ~ U,
  Z ~ T + F,
  T ~ U + X,
  exposure = "T",
  outcome = "Y",
  coords = coords
)

tidy_dag <- tidy_dagitty(HPTbestcase)

tidy_dag$data$time <- c("t3", "t3", "t2", "t2", "t1", "t1", "t2", "t2", "

ggdag(tidy_dag, node_size=8, text_size = 4) +
  theme_classic() + remove_axes() + theme(axis.line=element_blank()) +
  ylim(-1.5, 2.5) + xlim(-1.5, 2.5) +
  geom_text(x=-1, y=-.5, label=expression(t[1]), size=3) +
  geom_text(x=0, y=-.5, label=expression(t[2]), size=3) +
  geom_text(x=1, y=-.5, label=expression(t[3]), size=3) +
  geom_text(x=2, y=-.5, label=expression(t[4]), size=3) +
  geom_text(x=-1.5, y=-1, label="Y = Outcome (Tolerance),

```

```

T = Treatment (ln(Distance to camps))", hjust=0) +
  geom_text(x=-1.5, y=-1.4, label="F = Fixed Effects,
Z = Post-Treatment Confounders", hjust=0) +
  geom_text(x=1, y=-1, label="X = Pre-Treatment Confounders,
U = Unobserved Pre-Treatment Confounders", hjust=0)
# Seleccionar: No

db <- db_or %>%
  janitor::clean_names()

# Filtrar: No

# Recodificar
sjmisc::frq(db$state)
db$state <- as.factor(db$state)

# Tratamiento casos perdidos
colSums(is.na(db)) # no NA

# Transformar y derivar: No

# ITT_Y (Z --> Y)

m1_itt <- lm_robust(formula = far_right ~ distance,
  se_type = "HC2",
  data = db)
m2_itt <- lm_robust(formula = far_right ~ distance,
  fixed_effects = ~ state,
  se_type = "HC2",
  data = db)

# IIT_D (Z --> D)

m1_itttd <- lm_robust(formula = intolerance ~ distance,
  se_type = "HC2",
  data = db)
m2_itttd <- lm_robust(formula = intolerance ~ distance,
  fixed_effects = ~ state,
  se_type = "HC2",
  data = db)

```

```

# CACE

m1_iv <- iv_robust(formula = far_right ~ intolerance | distance,
  se_type = "HC2",
  data = db,
  diagnostics = T)
m2_iv <- iv_robust(formula = far_right ~ intolerance | distance,
  fixed_effects = ~ state,
  se_type = "HC2",
  data = db,
  diagnostics = T)

models1 <- list(m1_ittd, m1_itt, m1_iv)
models2 <- list(m2_ittd, m2_itt, m2_iv)

ccoef <- list(
  distance = "Distancia al campo",
  intolerance = "Intolerancia")

texreg::texreg(l = models1,
  include.ci = F,
  caption = paste("(\\#tab:table1)", "Modelos de VI sin efecto"),
  stars = c(0.05, 0.01),
  custom.coef.map = ccoef,
  custom.note = "Nota: Errores estándares robustos entre paréntesis",
  custom.header = list("Modelos VI" = 1:3),
  custom.model.names = c("$ITT_D$", "$ITT$", "$CACE$"),
  leading.zero = T,
  float.pos = "h!",
  use.packages = F,
  booktabs = TRUE,
  scalebox = 1,
  digits = 3,
  custom.gof.rows = list("Länder FE" = c("No", "No", "No")))

ccoef <- list(
  distance = "Distancia al campo",
  intolerance = "Intolerancia")

```



```

        lr +                # M
        immigrants07 +      # M
        unemployment07 +   # M
        unemp +             # M
        educ +              # M
        urban_scale,        # M
data = db)

# extraer para tablas

summary_m1_direct <- summary(m1_direct)
coefnames <- rownames(summary_m1_direct$coefficients)
estimates <- summary_m1_direct$coefficients[, "Estimate"]
stderrors <- summary_m1_direct$coefficients[, "Std. Err."]
pvalues <- summary_m1_direct$coefficients[, "Pr(>|t|)"]
n_obs <- as.integer(length(m1_direct$fitted.values))

m1_acde <- createTexreg(
  coef.names = coefnames,
  coef = estimates,
  se = stderrors,
  pvalues = pvalues,
  gof.names = c("Num. obs."),
  gof = as.integer(c(n_obs))
)

summary_m2_direct <- summary(m2_direct)
coefnames <- rownames(summary_m2_direct$coefficients)
estimates <- summary_m2_direct$coefficients[, "Estimate"]
stderrors <- summary_m2_direct$coefficients[, "Std. Err."]
pvalues <- summary_m2_direct$coefficients[, "Pr(>|t|)"]
n_obs <- as.integer(length(m2_direct$fitted.values))

m2_acde <- createTexreg(
  coef.names = coefnames,
  coef = estimates,
  se = stderrors,
  pvalues = pvalues,
  gof.names = c("Num. obs."),
  gof = as.integer(c(n_obs))
)

```

```

)

ccoef <- list(
  distance = "Distancia al campo",
  prop_jewish25 = "% Judíos (1925)",
  unemployment33 = "% Desempleo (1933)",
  nazishare33 = "Participación partido nazi (1933)"
)

texreg::texreg(l = list(m1_acde, m2_acde),
  caption = paste("(\\#tab:table3)", "Replicación de modelos"),
  stars = c(0.05, 0.01),
  custom.note = "Nota: Errores estándares entre paréntesis.",
  custom.model.names = c("Modelo 5", "Modelo 6"),
  custom.coef.map = ccoef,
  leading.zero = T,
  float.pos = "h!",
  use.packages = F,
  booktabs = TRUE,
  scalebox = 0.90,
  digits = 3,
  custom.gof.rows = list("Länder FE" = c("No", "Sí"),
    "Método" = rep("G-est", 2)))

```