

Reporte de Reproducibilidad

Artículo: Política de clases y confianza en los sindicatos en América Latina (Pérez-Ahumada y Carrasco, 2025).

Elisa Calderón Valentina Montecinos Amalia Tsukame
Felipe Vega

2026-05-02

Tabla de contenidos

1	Selección de artículo	2
2	Evaluación de reproducibilidad	3
3	Análisis reproducible	4
3.1	Resultado a reproducir	4
3.2	Proceso de reproducción	6
3.2.1	Procesamiento	6
3.2.2	Reproducción	16
4	Conclusiones	18
5	Recomendaciones	18
6	Referencias	19
7	Apéndice	19
7.1	Material suplementario	19
7.2	Código	19

1. Selección de artículo

■ Referencia del artículo seleccionado

Pérez-Ahumada, P., & Carrasco, K. (2025). Política de clases y confianza en los sindicatos en América Latina. *Latin American Research Review*, 60(2), 410–433. [doi:10.1017/lar.2024.64](https://doi.org/10.1017/lar.2024.64)

■ Resumen del artículo seleccionado

El artículo investiga los factores individuales y contextuales que explican el nivel de confianza de la ciudadanía en los sindicatos latinoamericanos, para esto se utilizan los datos de la encuesta Latinobarómetro de los años 2018 y 2020. Los autores demuestran que, a nivel individual, la confianza sindical es significativamente mayor entre personas pertenecientes a la clase trabajadora y a la clase media asalariada, así como entre quienes se identifican políticamente con la izquierda y confían más en las instituciones políticas. A nivel contextual, el estudio revela que la legitimidad de los sindicatos varía según el modelo de capitalismo de cada país. La confianza tiende a ser mayor en naciones “neo-desarrollistas” (como Brasil y Uruguay) y más baja en países de “capitalismo tercerizado” (como México y Centroamérica). También encontraron que la confianza en los sindicatos en Chile es alta, siendo un país liberal-rentista con sindicatos débiles, y baja en países redistributivos-rentistas y desarrollistas. Finalmente, los investigadores analizan cómo la confianza en los sindicatos depende de variables contextuales como la informalidad laboral, la inflación, el nivel de desempleo y las movilizaciones sociales.

■ Justificación de la selección del artículo

La elección de este artículo se fundamenta por tres motivos principales: su relevancia regional, su aporte teórico e interdisciplinario, y su potencial para el ejercicio de reproducibilidad. En relación a su relevancia regional, este artículo aborda una temática clave para comprender la cohesión social y las dinámicas de la política laboral en América Latina, ya que al analizar cómo los diferentes modelos económicos impactan en la sociedad, permite enriquecer el debate y generar evidencia sobre la percepción que se tiene de las organizaciones de los trabajadores en la región. En segundo lugar, el texto propone una articulación analítica entre las clases sociales, los tipos de capitalismo y la legitimidad social de los sindicatos. Por último, el artículo resulta especialmente adecuado para este ejercicio porque los autores declaran haber liberado sus datos y códigos en un repositorio público, lo que en principio facilita la posibilidad de replicar los análisis y verificar sus resultados.

2. Evaluación de reproducibilidad

- **Disponibilidad de datos y códigos:** El artículo proporciona acceso a los datos y códigos utilizados en el estudio. Ambos se encuentran disponibles de forma abierta en la página oficial de publicación del artículo correspondiente a [Cambridge University Press](#). Para acceder a esta se debe entrar a Supplementary materials, donde se debe descargar el archivo Word adjunto. Tras descargarlo, en el documento se encuentra un enlace que dirige a un [repositorio de libre acceso](#), en el cual se encuentran los archivos mencionados. El repositorio solo dispone de la base de datos procesada (que unifica las bases de ambos años), junto con los códigos de análisis individual y contextual, y las variables contextuales. Notablemente, no se encuentran disponibles las bases de datos originales de los años 2018 y 2020 ni el código empleado en la limpieza de los datos, lo que impide reproducir el flujo completo desde los datos brutos del Latinobarómetro hasta los resultados publicados.
- **Documentación:** Si bien se identifican la fuente de datos, el período de análisis y la estrategia de modelamiento, la documentación del artículo presenta algunas limitaciones que dificultan su reproducibilidad técnica. Aunque se reporta el uso de la encuesta Latinobarómetro y se detalla una muestra final de 25.286 observaciones, el texto omite la justificación de decisiones analíticas fundamentales, como el criterio utilizado para el tratamiento de los valores perdidos (NAs) que permitió llegar a ese tamaño muestral exacto. A nivel de medición, la descripción presenta imprecisiones importantes; por ejemplo, la construcción de la categoría de ‘clase social’ resulta ambigua al no explicitar la distribución empírica de los casos bajo el esquema de Wright, y tampoco se clarifican las variaciones en los códigos de los cuestionarios entre las rondas de 2018 y 2020. Finalmente, y como se mencionó con anterioridad, dentro del repositorio se dispone únicamente de la base de datos ya procesada, y prescinde de scripts o de una explicación exhaustiva sobre el procedimiento de limpieza dentro del artículo.
- **Transparencia:** El artículo es transparente sobre el respaldo académico. Se declara explícitamente que el proyecto fue financiado por ANID/FONDECYT REGULAR 1230056 (“Clases sociales, movimientos obreros y conflicto en tiempos de crisis: Un estudio comparativo de Argentina y Chile”, IP: Pablo Pérez Ahumada) y por COES ANID/FONDAP 1523A0005. Además, los autores mencionan que el texto es un artículo de acceso abierto, distribuido bajo los términos de la licencia Creative Commons Attribution, lo que permite su libre lectura, distribución y reproducción. Por último, se menciona que los autores son parte de instituciones como la Universidad de Chile y el Centro de Estudios de Conflicto y Cohesión Social.

3. Análisis reproducible

Con motivos de evaluar qué tan reproducible es este artículo, se seleccionaron tres gráficos para replicar.

3.1. Resultado a reproducir

La primera imagen que se busca reproducir presenta dos gráficos de líneas (paneles a y b) que ilustran cómo cambia la probabilidad de que las personas confíen en los sindicatos según la clase social y la posición política, comparando los años 2018 (línea gris con círculos) y 2020 (línea negra con triángulos). Ambos gráficos incluyen barras de error que indican la precisión de las estimaciones.

Panel (a): Muestra que la confianza en los sindicatos tiende a aumentar progresivamente desde los empleadores hacia la clase trabajadora y la clase media, los cuales presentan los niveles más altos, superando el 30 %. Además, se observa que en casi todas las clases sociales hubo una leve caída en la confianza entre 2018 y 2020, excepto en la clase media donde se mantuvo estable.

Panel (b): Se observa que la confianza en sindicatos aumenta a medida que la posición se mueve hacia la izquierda. Las personas sin posición política o de derecha muestran menores niveles de confianza (alrededor de 20–25 %), mientras que quienes se identifican con la izquierda presentan los niveles más altos (cerca de 30 %). Al igual que en el panel anterior, la confianza general en los sindicatos fue ligeramente menor en 2020 en comparación con 2018 para casi todos los grupos políticos.

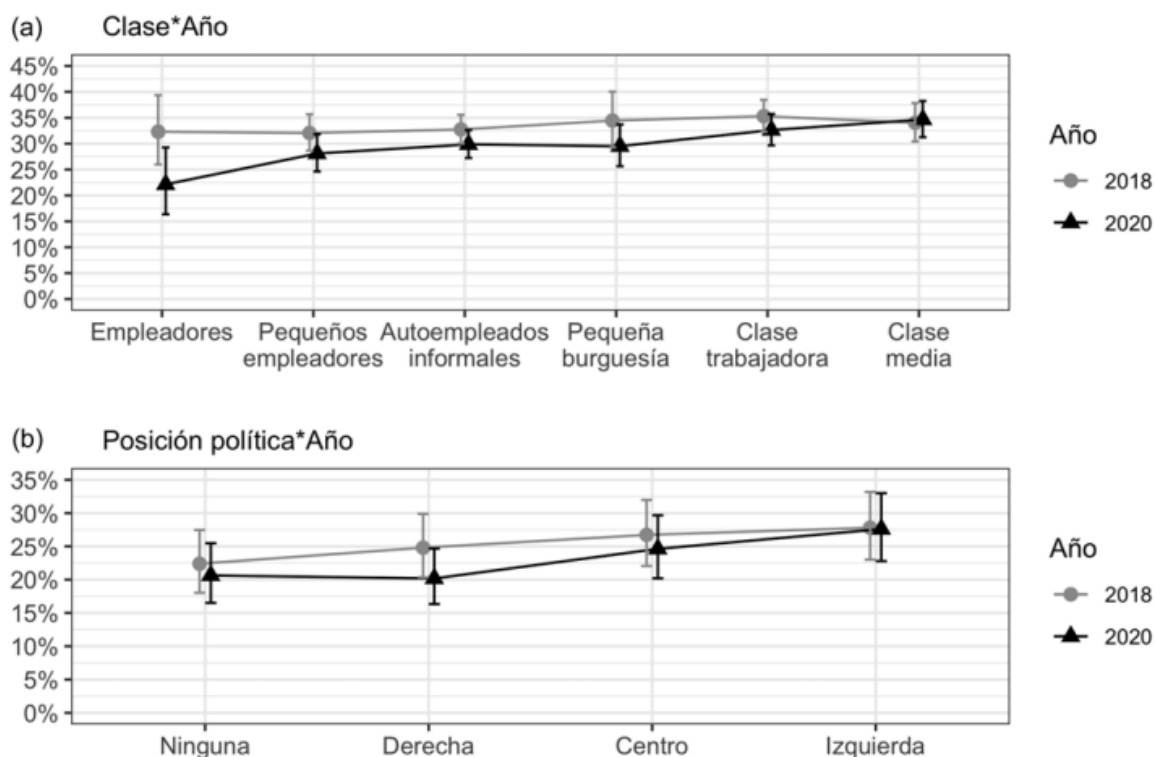


Figura 1. Probabilidad de confiar en los sindicatos (efectos de interacción)

La segunda imagen corresponde a un gráfico de líneas que muestra la probabilidad de confiar en los sindicatos a través de 18 países de América Latina, comparando nuevamente los resultados del año 2018 (línea gris con círculos) y el 2020 (línea negra con triángulos).

Entre los principales hallazgos está que existe una gran diferencia en cómo se percibe a los sindicatos en la región. En el extremo inferior, países como Venezuela (VEN) muestran los niveles más bajos de confianza (alrededor del 10-15%). En el extremo opuesto, Chile (CHL) es el país con mayor probabilidad de confianza en los sindicatos, superando el 40% en ambos años. Además, se observa una caída generalizada de la confianza entre los años 2018 y 2020. En la mayoría de los países, la línea negra (2020) se ubica por debajo de la línea gris (2018).

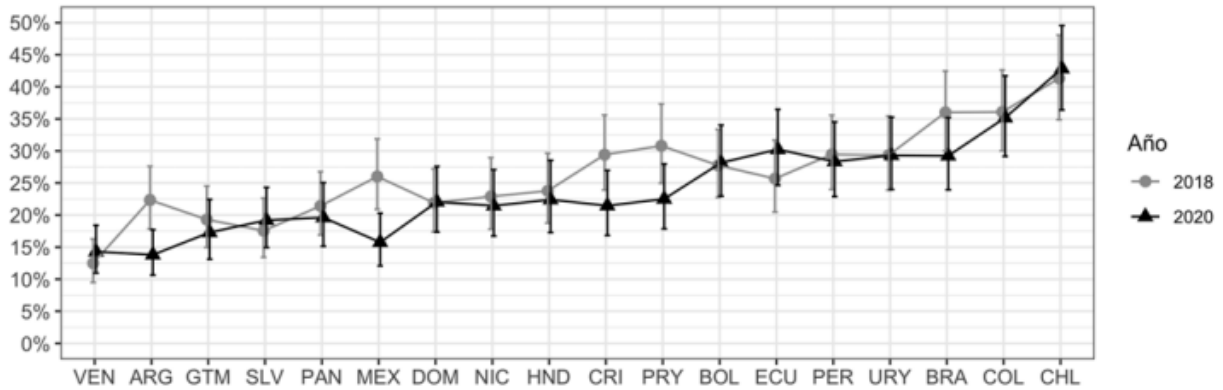


Figura 2. Probabilidad de confiar en los sindicatos en dieciocho países de América Latina (2018 y 2020)

3.2. Proceso de reproducción

3.2.1. Procesamiento

Paso 1: Preparación del entorno y carga de paquetes

En primer lugar, se cargan las librerías necesarias con *pacman*. Luego, se desactiva la notación científica para facilitar la lectura de los datos. Por último, se limpia el entorno de trabajo, borrando cualquier objeto o dato que estuviera previamente cargado.

```
pacman::p_load(dplyr,      # Manipulación de datos
              haven,      # importar datos en .dta o .sav
              car,        # recodificar variables
              sjlabelled, # etiquetado de variables
              sjmisc,     # descriptivos y frecuencias
              sjPlot,     # tablas, plots y descriptivos
              summarytools, # resumen de dataframe
              readxl      # base de datos xlsx
            )

options(scipen = 999)
rm(list = ls())
```

Paso 2: Importar base de datos

Se importan las dos bases de datos a utilizar. La función *readRDS* carga los datos del año 2018 y los asigna al objeto “lb_2018”, y *load* carga los datos del año 2020 que están en formato *.rdata* y lo denomina “lb_2020”.

```
lb_2018 <- readRDS("input/data/original/Latinobarometro_2018_Esp_R_v20190303.rds")
load("input/data/original/Latinobarometro_2020_Esp_Rdata_v1_0.rdata")
lb_2020 <- Latinobarometro_2020_Esp
```

Paso 3: Limpieza de variables latinobarómetro 2018

En primer lugar, se seleccionan las variables necesarias para trabajar.

- **P16NC.C**: Variable que mide la confianza en los sindicatos.
- **S14A, S14B, S15, REEDUC.1, S4**: Variables utilizadas posteriormente para construir la clase social (situación ocupacional, tipo de trabajo, nivel educacional e ingreso subjetivo).
- **P22ST**: Variable que representa la posición política del encuestado.
- **P15STGBSC.D, P15STGBSC.E, P15STGBSC.G**: Variables de confianza en instituciones políticas (Congreso, Gobierno y Partidos políticos, respectivamente).
- **SEXO**: Variable de sexo o género.
- **EDAD**: Variable de edad.
- **IDENPA**: Código identificador del país de la encuesta.

```
#---- 3. Limpieza de variables: latinobarómetro 2018 ----
##---- 3.1. Seleccionar variables ----

datos_2018 <- select(lb_2018,
  P16NC.C, #Conf. en sindicatos
  S14A, S14B, S15, REEDUC.1, S4, #Clase social: sit. ocupacional; t
  P22ST, #Posición política
  P15STGBSC.D, P15STGBSC.E, P15STGBSC.G, #Conf en: congreso, gob. y
  SEXO,
  EDAD,
  IDENPA)
```

Luego, para facilitar la lectura y el manejo de las variables, se renombran a palabras claves que representen una mejor definición.

```
##---- 3.2. Renombrar variables ----

datos_2018 <- rename(datos_2018,
  conf_sindicato = P16NC.C,
  situacion_ocupacional = S14A,
  tipo_trabajo_a = S15,
```

```

    tipo_trabajo_b = S14B,
    nivel_educ = REEDUC.1,
    ing_subj = S4,
    posicion_pol = P22ST,
    conf_congreso = P15STGBSC.D,
    conf_gob = P15STGBSC.E,
    conf_partido = P15STGBSC.G,
    sexo = SEXO,
    edad = EDAD,
    pais = IDENPA)

```

Para la variable **confianza en sindicatos**, se recodifica a una variable dicotómica. Si la respuesta original era 1 o 2 (mucho o algo de confianza), se le asigna el valor 1. Por el contrario, si el valor era 3 o 4 (poco o nada de confianza), se le asigna el valor 0. También se le asigna NA a los valores perdidos.

```

##--- 3.3 Confianza en sindicatos ----

datos_2018 <- datos_2018 %>%
  mutate(conf_sindicato_dic = case_when(
    conf_sindicato %in% c(1, 2) ~ 1, #mucho o algo de confianza
    conf_sindicato %in% c(3, 4) ~ 0, #poca o nada de confianza
    TRUE ~ NA
  ))

```

Para la variable **clase social** se crearon seis categorías basado en el esquema de Erik O. Wright.

- La primera categoría corresponde a **grandes empleadores** y se construye identificando quienes son independientes (valor 1 de la variable situación ocupacional), dueños de empresa (valor 2 de la variable tipo de trabajo), y que además declaran recibir un ingreso subjetivo con el cual “les alcanza bien” y “pueden ahorrar” (valor 1 de la variable ingreso subjetivo).
- La segunda categoría es **pequeños empleadores**, definidos por trabajadores independientes, dueños de empresa y cuyos ingresos percibidos no les alcanza bien ni les permite ahorrar (valores 2,3,4 de ingreso subjetivo).
- La tercera categoría es la **pequeña burguesía** y agrupa a los independientes con ocupaciones de autoempleo (1, 3 y 4 de la variable tipo de trabajo) que cuentan con educación profesional (valor 3 de la variable nivel educativo).

- La cuarta categoría corresponde a **clase media**, que une a los asalariados (valores 2 y 3 de situación ocupacional) en puestos técnicos o de gestión (valores 5, 6, 7 y 8 de la variable tipo de trabajo) y que también poseen formación profesional (valores 3 de nivel educativo).
- La quinta categoría es la **clase trabajadora**, que se define como el grupo de asalariados con puestos técnicos o de gestión pero que no cuentan con educación superior completa (valores 1 y 2 de la variable nivel educativo).
- La sexta categoría corresponde a **autoempleados** informales, compuesta por trabajadores independientes en ocupaciones de autoempleo que poseen niveles educativos bajos o medios (valores 1 y 2). Cualquier combinación de datos que no coincida con estos criterios se clasifica como NA.

```
##--- 3.4 Clase social ----

datos_2018 <- datos_2018 %>%
  mutate(clase_social = case_when(
    situacion_ocupacional %in% c(1, 4, 5) & (tipo_trabajo_a == 2 | tipo_trabajo_b == 2) ~ 1,
    situacion_ocupacional %in% c(1, 4, 5) & (tipo_trabajo_a == 2 | tipo_trabajo_b == 2) ~ 2,
    situacion_ocupacional %in% c(1, 4, 5) & (tipo_trabajo_a %in% c(1, 3, 4) | tipo_trabajo_b %in% c(1, 3, 4)) ~ 3,
    situacion_ocupacional %in% c(2, 3, 4, 5) & (tipo_trabajo_a %in% c(5, 6, 7, 8) | tipo_trabajo_b %in% c(5, 6, 7, 8)) ~ 4,
    situacion_ocupacional %in% c(2, 3, 4, 5) & (tipo_trabajo_a %in% c(5, 6, 7, 8) | tipo_trabajo_b %in% c(5, 6, 7, 8)) ~ 5,
    situacion_ocupacional %in% c(1, 4, 5) & (tipo_trabajo_a %in% c(1, 3, 4) | tipo_trabajo_b %in% c(1, 3, 4)) ~ 6,
    TRUE ~ NA
  ))
```

La variable **posición política** agrupa una escala de 0 a 10 en tres bloques: los valores 0, 1, 2, 3 originales, se les asigna la categoría 1 izquierda. Los valores 4, 5 y 6 la categoría 2 centro. Los valores 7, 8, 9 y 10 la categoría 3 derecha. Por último, se clasifica en la categoría 4 los valores negativos -1, -4, -6 que corresponden a sin identificación.

```
##--- 3.5 Posición política ----

datos_2018 <- datos_2018 %>%
  mutate(posicion_pol_ord = case_when(
    posicion_pol %in% c(0, 1, 2, 3) ~ 1, #Izquierda
    posicion_pol %in% c(4, 5, 6) ~ 2, #centro
    posicion_pol %in% c(7, 8, 9, 10) ~ 3, #derecha
    posicion_pol %in% c(-1, -2, -6) ~ 4, #sin identificación
    TRUE ~ NA
  ))
```

Para crear la variable **confianza en instituciones políticas**, primero se convierten los valores inválidos -1 de las variables confianza en el congreso, en el gobierno y en partidos en NA. Luego, se invierte la escala restando el valor a 4, de manera que, a mayor número hay mayor confianza en las instituciones. Por último, se crea el índice “conf_instituciones”.

```
##--- 3.6 Confianza en instituciones políticas ----

datos_2018 <- datos_2018 %>%
  mutate(
    conf_congreso = case_when(conf_congreso < 1 ~ NA, TRUE ~ conf_congreso),
    conf_gob      = case_when(conf_gob < 1 ~ NA, TRUE ~ conf_gob),
    conf_partido  = case_when(conf_partido < 1 ~ NA, TRUE ~ conf_partido)
  )

datos_2018 <- mutate(datos_2018,
  conf_congreso = 4 - conf_congreso,
  conf_gob      = 4 - conf_gob,
  conf_partido  = 4 - conf_partido
)

datos_2018 <- mutate(datos_2018,
  conf_instituciones = conf_congreso + conf_gob + conf_partido)
```

La variable **sexo** se convierte en una variable dummy, donde 1 es mujer y 0 hombre. Los valores -2 se asignan como NA.

```
##--- 3.7 Sexo (1 = mujer) ----

datos_2018 <- datos_2018 %>%
  mutate(sexo = case_when(
    sexo == 1 ~ 0,
    sexo == 2 ~ 1,
    sexo == -2 ~ NA
  ))
```

La **edad** se mantiene como variable numérica, pero se asigna NA a los valores menores a 1.

```
##--- 3.8 Edad ----
datos_2018 <- datos_2018 %>%
  mutate(edad = case_when(
    edad < 1 ~ NA,
    TRUE ~ edad
  ))
```

PASO 4: Limpieza de variables latinobarómetro 2020

Se realiza el mismo procedimiento y lógica que en el paso 3 pero con ciertas diferencias en el nombre y los valores de las variables en la base de datos.

Primero, se seleccionan las variables.

```
#---- 4. Limpieza de variables: latinobarómetro 2020 ----
##---- 4.1. Seleccionar variables ----
datos_2020 <- select(lb_2020,
  p15st.c, #Conf. en sindicatos
  s24.a, s25, s24.b, reeduc.1, s4, #Clase social: sit. ocupacional;
  p18st, #Posición política
  p13st.d, p13st.e, p13st.g, #Conf en: congreso, gob. y partidos po
  sexo,
  edad,
  idenpa)
```

Luego, se renombran para facilitar su uso.

```
##---- 4.2. Renombrar variables ----
datos_2020 <- rename(datos_2020,
  conf_sindicato = p15st.c,
  situacion_ocupacional = s24.a,
  tipo_trabajo_a = s25,
  tipo_trabajo_b = s24.b,
  nivel_educ = reeduc.1,
  ing_subj = s4,
  posicion_pol = p18st,
  conf_congreso = p13st.d,
  conf_gob = p13st.e,
  conf_partido = p13st.g,
  pais = idenpa)
```

Las variables **confianza en sindicatos** y **clase social** se construyen de la misma manera que en el paso 3.

```
##--- 4.3 Confianza en sindicatos ----
datos_2020 <- datos_2020 %>%
  mutate(conf_sindicato_dic = case_when(
    conf_sindicato %in% c(1, 2) ~ 1,
    conf_sindicato %in% c(3, 4) ~ 0,
    TRUE ~ NA
  ))
```

```
##--- 4.4 Clase social ----
datos_2020 <- datos_2020 %>%
  mutate(clase_social = case_when(
    situacion_ocupacional %in% c(1, 4, 5) & (tipo_trabajo_a == 2 | tipo_trabajo_b == 2) ~ 1,
    situacion_ocupacional %in% c(1, 4, 5) & (tipo_trabajo_a == 2 | tipo_trabajo_b == 3) ~ 2,
    situacion_ocupacional %in% c(1, 4, 5) & (tipo_trabajo_a %in% c(1, 3, 4) | tipo_trabajo_b %in% c(1, 3, 4)) ~ 3,
    situacion_ocupacional %in% c(2, 3, 4, 5) & (tipo_trabajo_a %in% c(5, 6, 7, 8) | tipo_trabajo_b %in% c(5, 6, 7, 8)) ~ 4,
    situacion_ocupacional %in% c(2, 3, 4, 5) & (tipo_trabajo_a %in% c(5, 6, 7, 8) | tipo_trabajo_b %in% c(1, 2, 3, 4)) ~ 5,
    situacion_ocupacional %in% c(1, 4, 5) & (tipo_trabajo_a %in% c(1, 3, 4) | tipo_trabajo_b %in% c(1, 3, 4)) ~ 6,
    TRUE ~ NA
  ))
```

La variable **posición política** se construye igual que en el paso 3, pero la categoría 4 sin identificación contiene los valores 97, 98 y 99 de la base de datos.

```
##--- 4.5 Posición política ----
datos_2020 <- datos_2020 %>%
  mutate(posicion_pol_ord = case_when(
    posicion_pol %in% c(0, 1, 2, 3) ~ 1, #Izquierda
    posicion_pol %in% c(4, 5, 6) ~ 2, #centro
    posicion_pol %in% c(7, 8, 9, 10) ~ 3, #derecha
    posicion_pol %in% c(97, -1, -2) ~ 4, #sin identificación
    TRUE ~ NA
  ))
```

Para la variable **confianza en instituciones políticas** se realiza, en primer lugar, la limpieza de las variables confianza en el congreso, confianza en el gobierno y confianza en los partidos, identificando los valores de respuesta no válidos (-1, -2, -3 y -4)

y reemplazándolos por NA. Luego, el código invierte la escala de las tres variables restando su valor a 4 y, finalmente, suma los valores de estas variables ya transformadas para crear un índice llamado “conf_instituciones”.

```
##--- 4.6 Confianza en instituciones políticas ----

datos_2020 <- datos_2020 %>%
  mutate(
    conf_congreso = case_when(conf_congreso %in% c(-1, -2, -3, -4) ~ NA, TRUE ~ c
    conf_gob      = case_when(conf_gob %in% c(-1, -2, -3, -4) ~ NA, TRUE ~ conf_g
    conf_partido  = case_when(conf_partido %in% c(-1, -2, -3, -4) ~ NA, TRUE ~ c
  ) %>%
  mutate(conf_congreso = 4 - conf_congreso,
         conf_gob      = 4 - conf_gob,
         conf_partido  = 4 - conf_partido,
         conf_instituciones = conf_congreso + conf_gob + conf_partido
  )
```

Las variables **sexo** y **edad** también se construyen de la misma manera que en el paso 3.

```
##--- 4.7 Sexo (1 = mujer) ----

datos_2020 <- datos_2020 %>%
  mutate(sexo = case_when(
    sexo == 1 ~ 0,
    sexo == 2 ~ 1,
    TRUE ~ NA
  ))
```

```
##--- 4.8 Edad ----

datos_2020 <- datos_2020 %>%
  mutate(edad = case_when(
    edad < 1 ~ NA,
    TRUE ~ edad
  ))
```

PASO 5: Unir las bases de datos

Primero, se le agrega a cada base de datos una columna nueva correspondiente al año para poder identificarlas al momento de unir las bases de datos. Luego, si juntan las bases de datos. Por último, se elimina cualquier dato que esté como NA.

```
##--- 5. Unir las bases de datos ----
datos_2018 <- datos_2018 %>% mutate(ano = 2018)
datos_2020 <- datos_2020 %>% mutate(ano = 2020)

datos_lb <- bind_rows(datos_2018, datos_2020)

datos_lb_proc <- na.omit(datos_lb)
```

PASO 6: Etiquetar variables

En este paso, se etiquetan las variables para que sean categóricas. Esto se realizó al final para no tener problemas cuando se unieron ambas bases de datos en el paso 5.

Para la variable **confianza en sindicatos**, se asignan las etiquetas “Mucha o algo de confianza” al valor 1 y “Poca o nada de confianza” al valor 0.

```
#--- 6. Etiquetar variables ----
datos_lb_proc$conf_sindicato_dic <- set_labels(datos_lb_proc$conf_sindicato_dic,1
                                             "Poca o nada de confianza" = 0
                                             ))
```

Para la variable **clase social**, se asignan etiquetas descriptivas a cada una de las seis categorías.

```
datos_lb_proc <- datos_lb_proc %>%
  mutate(clase_social = case_when(
    clase_social == 1 ~ "1. Gran empleador",
    clase_social == 2 ~ "2. Pequeño empleador",
    clase_social == 3 ~ "3. Pequeño burgués formal",
    clase_social == 4 ~ "4. Clase media",
    clase_social == 5 ~ "5. Obrero",
    clase_social == 6 ~ "6. Autoempleado informal"))
```

Para la variable **posición política**, se asignan etiquetas a cada una de las cuatro categorías.

```
datos_lb_proc <- datos_lb_proc %>%
  mutate(posicion_pol_ord = case_when(
    posicion_pol_ord == 1 ~ "Izquierda",
    posicion_pol_ord == 2 ~ "Centro",
    posicion_pol_ord == 3 ~ "Derecha",
    posicion_pol_ord == 4 ~ "Ninguno"))
```

Para la variable **sexo**, se asignan las etiquetas “Hombre” y “Mujer” a los valores 0 y 1 respectivamente.

```
datos_lb_proc <- datos_lb_proc %>%  
  mutate(sexo = case_when(  
    sexo == 0 ~ "Hombre",  
    sexo == 1 ~ "Mujer"))
```

Por último, para la variable **país**, se asignan códigos de texto de tres letras a cada uno de los códigos numéricos originales.

```
datos_lb_proc <- datos_lb_proc %>%  
  mutate(pais = case_when(  
    pais == 32 ~ "ARG",  
    pais == 68 ~ "BOL",  
    pais == 76 ~ "BRA",  
    pais == 152 ~ "CHL",  
    pais == 170 ~ "COL",  
    pais == 188 ~ "CRI",  
    pais == 214 ~ "DOM",  
    pais == 218 ~ "ECU",  
    pais == 222 ~ "SLV",  
    pais == 320 ~ "GTM",  
    pais == 340 ~ "HND",  
    pais == 484 ~ "MEX",  
    pais == 558 ~ "NIC",  
    pais == 591 ~ "PAN",  
    pais == 600 ~ "PRY",  
    pais == 604 ~ "PER",  
    pais == 858 ~ "URY",  
    pais == 862 ~ "VEN"  
  ))
```

PASO 7: Guardar base de datos procesada

Se guarda la base de datos procesada que incluye la combinación de las variables de la base de datos de Latinobarómetro de los años 2018 y 2020.

```
datos_lb_proc <- select(datos_lb_proc,  
  conf_sindicato_dic,  
  clase_social,
```

```

posicion_pol_ord,
conf_instituciones,
sexo,
edad,
pais,
ano)

```

```

saveRDS(datos_lb_proc, file = "input/proc/datos_lb_proc.rds")

```

3.2.2. Reproducción

Figura 1: Probabilidad de confiar en un sindicato por clase social y año, y posición política y año (Reproducción y comparación).

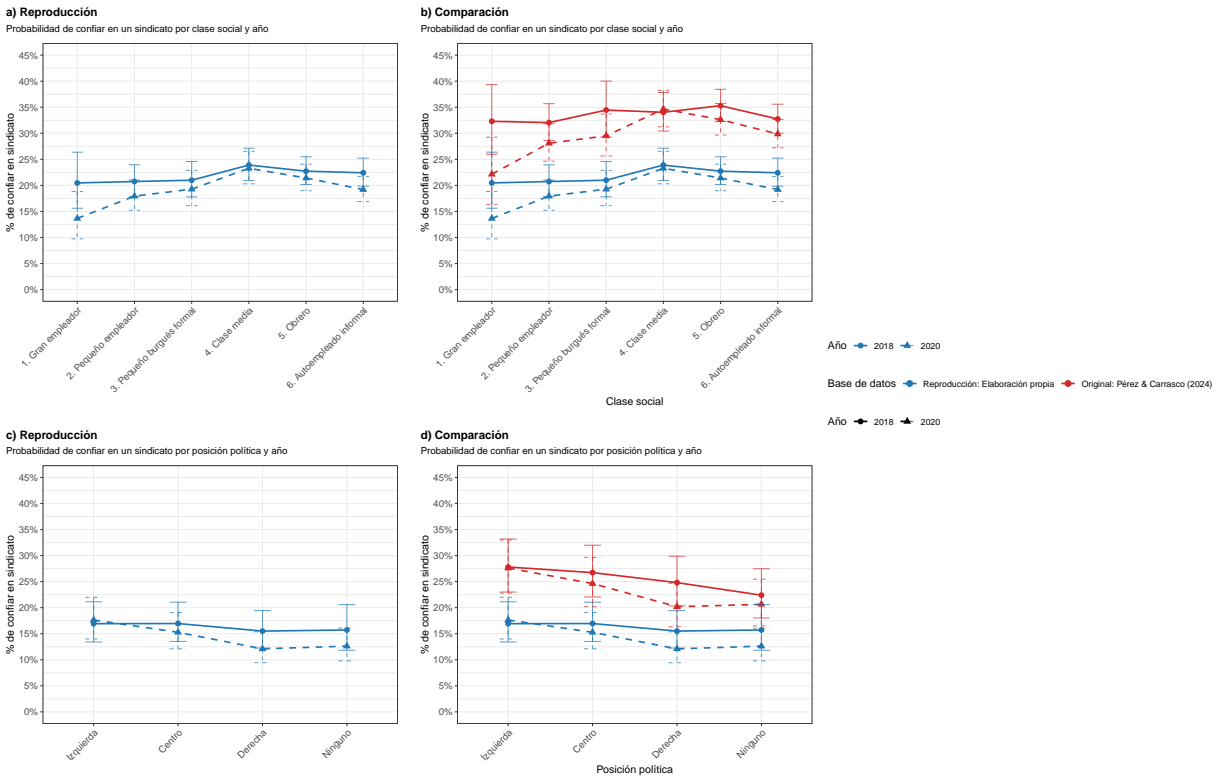


Figura 1: Probabilidad de confiar en un sindicato por clase social y año, y posición política y año (Reproducción y comparación).

Figura 2: Probabilidad de confiar en un sindicato por país y año (Recreación y comparación)

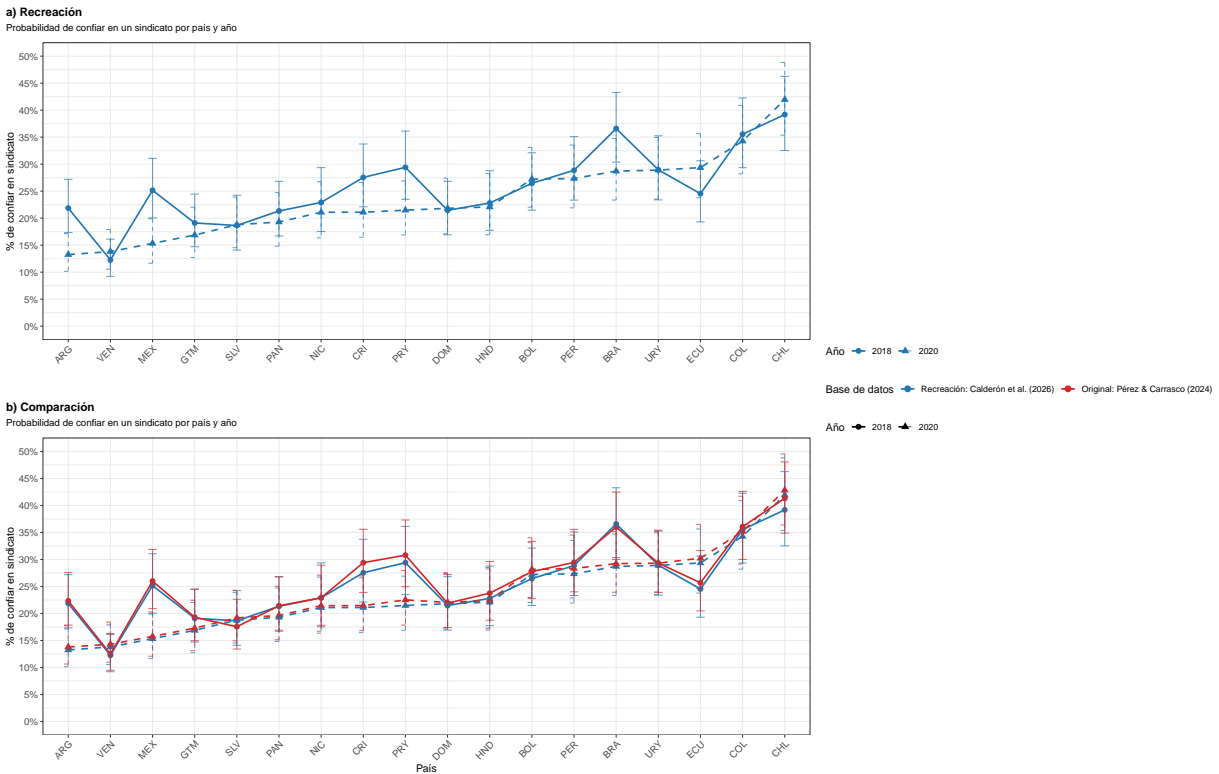


Figura 2: Probabilidad de confiar en un sindicato por país y año (Reproducción y comparación).

Dentro de las principales dificultades que se encontraron para el proceso de reproducción fue la falta de los códigos relacionados a la limpieza de la base de datos. Esto impidió que, a pesar de los diversos intentos realizados, se pudiera reconstruir la muestra final de 25.286 casos señalada en el artículo original.

En una primera instancia, se obtuvo una muestra de 18.000 casos aproximadamente, debido a que no se estaba considerando a las personas que no se encontraban empleadas pero que sí habían otorgado información de su último trabajo en la encuesta. Al solucionar este problema, la muestra aumento a 22.324 datos.

Otra dificultad fue que el libro de códigos del año 2020 identificaba como casos perdidos los valores 97,98 y 99 pero en la encuestas los NA correspondían a los valores -1 y -2.

Sumado a lo anterior, los autores cambian el nombre de las variables, a veces se utiliza “pequeña burguesía” pero en la base limpia se denomina “pequeño burgués formal” y a “clase trabajadora” se nombra “obrero”.

4. Conclusiones

En síntesis, si bien el artículo cumple con distintos criterios que permiten la reproducibilidad (como la disponibilidad de datos y códigos, y la descripción de las hipótesis, variables, métodos y procedimientos), esta resulta limitada. El proceso de reproducción se dificultó por distintas razones: en primer lugar, por la falta del código de limpieza de los datos, por lo cual se debió intentar reconstruir a partir de la información del estudio; en segundo lugar, por la operacionalización compleja de ciertas variables como la clase social, la cual debía hacerse bajo el esquema de Wright; y por último, por la falta de información sobre decisiones metodológicas como el tratamiento de casos perdidos.

En consecuencia, no se pudieron replicar con exactitud las figuras, ya que no se pudo alcanzar la representatividad de la muestra, la cual quedó en 22.384 casos, en comparación con las 25.286 observaciones reportadas. Esta reducción de la muestra afectó también en la reproducción de los gráficos, los cuales difieren a los publicados, debido a que el intercepto es menor, por lo que cae el cálculo del porcentaje de probabilidad, dando resultados diferentes a los establecidos por la investigación. Por estos motivos, se concluye que el nivel de reproducibilidad del artículo seleccionado es restringido, y queda evidenciada la necesidad de una documentación más exhaustiva de las herramientas utilizadas por este.

5. Recomendaciones

En cuanto a las recomendaciones específicas para fortalecer la reproducibilidad de esta investigación, resulta fundamental que los autores incorporen al repositorio la totalidad de los datos y scripts utilizados. Actualmente, el repositorio solo dispone de la base de datos procesada, junto con los códigos de análisis individual y contextual, además de las variables contextuales, sin embargo, faltan las bases de datos originales de los años 2018 y 2020, así como el código empleado en la limpieza de los datos. Complementariamente, el artículo debería integrar una explicación exhaustiva de este procedimiento, de modo que incluso sin disponer del código exacto, el proceso pueda ser replicado.

Asimismo, es necesario que se profundice en la operacionalización de las variables, proporcionando especificaciones para cada año de estudio en particular. Debido a que los códigos de las variables del Latinobarómetro varían entre sus versiones, resulta insuficiente que el artículo solo se mencione la forma en que se operacionalizan para el año 2020. En este sentido, se requiere además una clarificación sobre la variable clase social, puesto que su definición operacional es ambigua y no permite entender con precisión cómo se distribuyeron los casos entre las 6 categorías del esquema de

clases de Wright. Agregando a esto, se debería usar el mismo nombre de las variables en el informe como en las etiquetas de la base de datos, ya que al ser diferentes resulta confuso en el proceso de reproducción.

Por último, se deben explicitar decisiones metodológicas clave, tales como el tratamiento de los casos perdidos (NAs). La omisión de esta información no permite comprender la lógica bajo la cual se filtraron los datos, lo cual es necesario para alcanzar la muestra final de 25.286 observaciones reportadas. Explicitar esto permite que cualquier intento por reproducir los resultados de esta investigación pueda llegar a los mismos hallazgos, incluyendo la replicación de las tablas y gráficos que se muestran a lo largo del texto.

6. Referencias

7. Apéndice

7.1. Material suplementario

- Incluyan aquí detalles metodológicos adicionales, en caso de ser necesario

7.2. Código

- Incluir acá el código original, en caso de estar disponible.