

Entrega Final. Variables Instrumentales

Carlos Bautista

2022-12-03

```
library(tidyverse)
library(stargazer)
library(dagitty)
library(rethinking)
library(AER)
library(haven)
library(kableExtra)
```

a. Artículo de Referencia

Nunn, Nathan, and Leonard Wantchekon. 2011a. “The Slave Trade and the Origins of Mistrust in Africa.” *Am. Econ. Rev.* 101 (7): 3221–52.

b. Descripción General

Este artículo se encuentra dentro de muchos otros que tratan de medir los efectos de largo plazo del esclavismo en África. Su hipótesis principal a demostrar es si las diferencias actuales en los niveles de confianza dentro de África pueden deberse al comercio de esclavos al que fueron sometidas las comunidades en los siglos pasados.

Al vincular información a nivel individual con datos históricos sobre envíos de esclavos por grupo étnico, los autores encuentran que aquellos individuos cuyos ancestros fueron fuertemente atacados durante el comercio de esclavos son menos confiados hoy.

Esta evidencia que obtienen aplicando diferentes estrategias empíricas sugiere que la relación es causal. Además, al analizar los mecanismos causales, muestran que la mayor parte del impacto de la trata de esclavos se produce a través de factores internos del individuo, como las normas, creencias y valores culturales.

c. Análisis

El artículo analiza si las diferencias actuales en los niveles de confianza dentro de África se deben al comercio de esclavos en el transatlántico y el Océano Índico.

Con ayuda de datos de encuestas recientes a nivel individual y con datos históricos sobre el comercio de esclavos por grupo étnico, los autores encuentran que las personas cuyos ancestros fueron fuertemente atacados durante la trata de esclavos son menos confiados hoy. Por ello, la variable de tratamiento es la trata de esclavos (1400-1900) y la resultado los niveles actuales de confianza en las comunidades.

El instrumento elegido para el análisis debe ser una variable que esté correlacionada con el número de esclavos tomados de un grupo étnico pero no correlacionado con ninguna característica del grupo étnico que pueda afectar la confianza de los descendientes. Por ello, eligieron la distancia de la etnia de un individuo hacia la costa durante la trata de esclavos. Este instrumento logra captar la exposición de un grupo étnico a la demanda de esclavos. pero además, la distancia desde la costa no está correlacionada con otros factores que pueda afectar la confianza de sus descendientes.

d. Datos

El conjunto de datos es `slavetrade.dta`. Se puede obtener mayor descripción de los datos e información para replicar el artículo en el sitio web del autor <https://scholar.harvard.edu/nunn/pages/data-0>.

e. Instrucciones

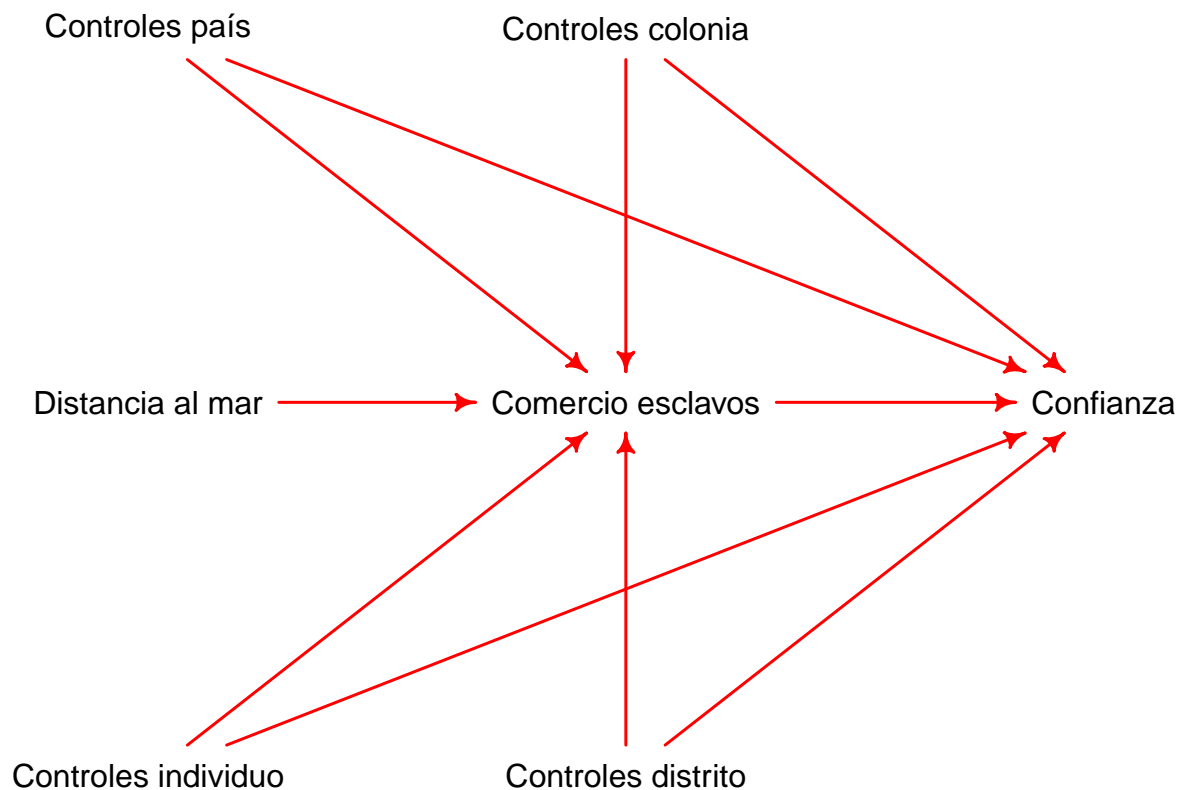
1. Escriba una introducción del análisis que realizará con ayuda de este conjunto de datos de acuerdo al contexto del problema.

COMPLETAR CBP

2. Ilustre correctamente los factores considerados en dicho análisis. Para ello, realice el gráfico directo acíclico (DAG) que muestre la teoría de cambio detrás de la relación causal que busca analizar y cuantificar el artículo.

```
dag_iv <- dagitty('dag{
  "Comercio esclavos" [pos="0.5,0"]
  "Confianza" [pos="1,0"]
  "Distancia al mar" [pos="0,0"]
  "Controles individuo" [pos="0,1"]
  "Controles distrito" [pos="0.5,1"]
  "Controles país" [pos="0,-1"]
  "Controles colonia" [pos="0.5,-1"]
  "Comercio esclavos" -> "Confianza"
  "Distancia al mar" -> "Comercio esclavos"
  "Controles individuo" -> "Comercio esclavos"
  "Controles individuo" -> "Confianza"
  "Controles distrito" -> "Comercio esclavos"
  "Controles distrito" -> "Confianza"
  "Controles país" -> "Comercio esclavos"
  "Controles país" -> "Confianza"
  "Controles colonia" -> "Comercio esclavos"
  "Controles colonia" -> "Confianza"}')
```

```
drawdag(dag_iv, col_arrow="red")
```



3. Luego de leer el conjunto de datos (21,822 observaciones), empiece por realizar el análisis exploratorio de las variables y describa su comportamiento. Tome en cuenta que hay un grupo de variables numéricas y otras categóricas que deben ser descritas de manera particular de acuerdo a su naturaleza.

Se genera resumen estadístico de las variables:

```
dta <- read_dta("data/slavetrade.dta")
attach(dta)

dta %>% select_if(is.numeric) %>% summary()
```

```
## location_id trust_relatives trust_neighbors intra_group_trust
## Min. : 1 Min. :0.000 Min. :0.000 Min. :0.000
## 1st Qu.: 587 1st Qu.:2.000 1st Qu.:1.000 1st Qu.:1.000
## Median :1294 Median :3.000 Median :2.000 Median :2.000
## Mean :1336 Mean :2.189 Mean :1.738 Mean :1.679
## 3rd Qu.:2014 3rd Qu.:3.000 3rd Qu.:3.000 3rd Qu.:3.000
## Max. :2891 Max. :3.000 Max. :3.000 Max. :3.000
## NA's :1084 NA's :1124 NA's :1203
## inter_group_trust trust_local_council ln_export_area export_area
## Min. :0.000 Min. :0.000 Min. :0.00000 Min. : 0.00000
## 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:0.00000 1st Qu.: 0.00000
## Median :1.000 Median :2.000 Median :0.04157 Median : 0.04245
## Mean :1.363 Mean :1.665 Mean :0.53469 Mean : 2.65546
## 3rd Qu.:2.000 3rd Qu.:3.000 3rd Qu.:0.67932 3rd Qu.: 0.97253
## Max. :3.000 Max. :3.000 Max. :3.65603 Max. :37.70739
## NA's :1405 NA's :1499 NA's :120 NA's :120
## export_pop ln_export_pop age age2
## Min. :0.000 Min. :0.000 Min. : 18.00 Min. : 324
```

```

## 1st Qu.:0.000 1st Qu.:0.000 1st Qu.: 25.00 1st Qu.: 625
## Median :0.005 Median :0.005 Median : 33.00 Median : 1089
## Mean :0.113 Mean :0.092 Mean : 36.42 Mean : 1543
## 3rd Qu.:0.107 3rd Qu.:0.102 3rd Qu.: 45.00 3rd Qu.: 2025
## Max. :4.464 Max. :1.698 Max. :130.00 Max. :16900
## NA's :3456 NA's :3456 NA's :266 NA's :266
## male urban_dum occupation religion
## Min. :0.0000 Min. :0.0000 Min. : 0.00 Min. : 0.00
## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.: 2.00 1st Qu.: 2.00
## Median :0.0000 Median :0.0000 Median : 7.00 Median : 4.00
## Mean :0.4997 Mean :0.3661 Mean : 15.79 Mean : 28.53
## 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.: 20.00 3rd Qu.: 10.00
## Max. :1.0000 Max. :1.0000 Max. :995.00 Max. :995.00
## NA's :92 NA's :92
## living_conditions education near_dist distsea
## Min. :1.000 Min. :0.000 Min. : 0.033 Min. : 1.25
## 1st Qu.:2.000 1st Qu.:2.000 1st Qu.: 119.220 1st Qu.: 175.31
## Median :2.000 Median :3.000 Median : 378.142 Median : 393.17
## Mean :2.556 Mean :3.074 Mean : 432.716 Mean : 439.89
## 3rd Qu.:4.000 3rd Qu.:4.000 3rd Qu.: 698.038 3rd Qu.: 675.71
## Max. :5.000 Max. :9.000 Max. :1459.088 Max. :1252.68
## NA's :74 NA's :78 NA's :462 NA's :120
## loc_ln_export_area local_council_performance council_listen
## Min. :0.0000 Min. :1.000 Min. :0.000
## 1st Qu.:0.0000 1st Qu.:2.000 1st Qu.:0.000
## Median :0.0114 Median :3.000 Median :1.000
## Mean :0.4567 Mean :2.512 Mean :1.176
## 3rd Qu.:0.3249 3rd Qu.:3.000 3rd Qu.:2.000
## Max. :3.7388 Max. :4.000 Max. :3.000
## NA's :825 NA's :2074 NA's :1933
## corrupt_local_council school_present electricity_present piped_water_present
## Min. :0.000 Min. :0.0000 Min. :0.0000 Min. :0.0000
## 1st Qu.:1.000 1st Qu.:1.0000 1st Qu.:0.0000 1st Qu.:0.0000
## Median :1.000 Median :1.0000 Median :1.0000 Median :0.0000
## Mean :1.279 Mean :0.7835 Mean :0.5271 Mean :0.4875
## 3rd Qu.:2.000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:1.0000
## Max. :3.000 Max. :1.0000 Max. :1.0000 Max. :1.0000
## NA's :3538 NA's :213 NA's :276 NA's :336
## sewage_present health_clinic_present district_ethnic_frac
## Min. :0.0000 Min. :0.0000 Min. :0.0000
## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.1172
## Median :0.0000 Median :0.0000 Median :0.4062
## Mean :0.2271 Mean :0.4707 Mean :0.4047
## 3rd Qu.:0.0000 3rd Qu.:1.0000 3rd Qu.:0.6760
## Max. :1.0000 Max. :1.0000 Max. :0.9058
## NA's :658 NA's :591
## frac_ethnicity_in_district townvill_nonethnic_mean_exports
## Min. :0.002717 Min. :0.000
## 1st Qu.:0.250000 1st Qu.:0.000
## Median :0.687500 Median :0.041
## Mean :0.599055 Mean :0.386
## 3rd Qu.:0.937500 3rd Qu.:0.345
## Max. :1.000000 Max. :3.656
## NA's :9005

```

```

## district_nonethnic_mean_exports region_nonethnic_mean_exports
## Min. :0.000 Min. :0.0000
## 1st Qu.:0.001 1st Qu.:0.0054
## Median :0.045 Median :0.1016
## Mean :0.365 Mean :0.4264
## 3rd Qu.:0.399 3rd Qu.:0.4760
## Max. :3.656 Max. :3.6559
## NA's :5334 NA's :847
## country_nonethnic_mean_exports centroid_lat centroid_long
## Min. :0.00000 Min. : -32.739 Min. : -16.409
## 1st Qu.:0.01627 1st Qu.: -18.328 1st Qu.: 6.279
## Median :0.11436 Median : -5.147 Median : 28.146
## Mean :0.46895 Mean : -6.867 Mean : 21.581
## 3rd Qu.:0.66756 3rd Qu.: 6.742 3rd Qu.: 33.697
## Max. :2.88457 Max. : 27.817 Max. : 49.246
## NA's :120 NA's :120
## explorer_contact railway_contact dist_Saharan_node dist_Saharan_line
## Min. :0.0000 Min. :0.0000 Min. : 25.42 Min. : 113.9
## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.: 824.34 1st Qu.: 824.3
## Median :0.0000 Median :0.0000 Median :2630.49 Median :2630.5
## Mean :0.4388 Mean :0.4336 Mean :2573.80 Mean :2579.0
## 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:3996.17 3rd Qu.:3996.2
## Max. :1.0000 Max. :1.0000 Max. :5221.35 Max. :5221.3
## NA's :120 NA's :120 NA's :120 NA's :120
## malaria_ecology v30 v33 fishing
## Min. : 0.000 Min. :1.000 Min. :1.000 Min. : 2.500
## 1st Qu.: 3.221 1st Qu.:5.000 1st Qu.:2.000 1st Qu.: 2.500
## Median : 7.991 Median :6.000 Median :3.000 Median :10.000
## Mean :11.506 Mean :6.115 Mean :2.918 Mean : 8.741
## 3rd Qu.:19.774 3rd Qu.:7.000 3rd Qu.:4.000 3rd Qu.:10.000
## Max. :34.640 Max. :8.000 Max. :4.000 Max. :60.000
## NA's :120 NA's :1305 NA's :1283 NA's :120
## exports ln_exports total_missions_area ln_init_pop_density
## Min. : 0.0000 Min. :0.0000 Min. :0.00000 Min. : -4.274
## 1st Qu.: 0.0000 1st Qu.:0.0000 1st Qu.:0.00002 1st Qu.: 1.775
## Median : 0.8472 Median :0.6137 Median :0.00011 Median : 2.503
## Mean : 93.1687 Mean :1.9505 Mean :0.00022 Mean : 2.547
## 3rd Qu.: 29.2104 3rd Qu.:3.4082 3rd Qu.:0.00027 3rd Qu.: 3.494
## Max. :854.9581 Max. :6.7522 Max. :0.00276 Max. : 5.870
## NA's :120 NA's :120 NA's :120 NA's :3456
## cities_1400_dum
## Min. :0.0000
## 1st Qu.:0.0000
## Median :0.0000
## Mean :0.1253
## 3rd Qu.:0.0000
## Max. :1.0000
##

```

4. Realice los gráficos de la distribución de algunas de las variables. Por ejemplo, elija: dos de las variables de confianza como resultados (trust in neighbours, trust in in the local council), la variable de trata- miento slave exports y la variable instrumental distance to the sea. Analice su comportamiento e interprete sus resultados.

Respecto a la variable de confianza en los vecinos, la respuesta con mayor respuesta (moda) fue la de nivel

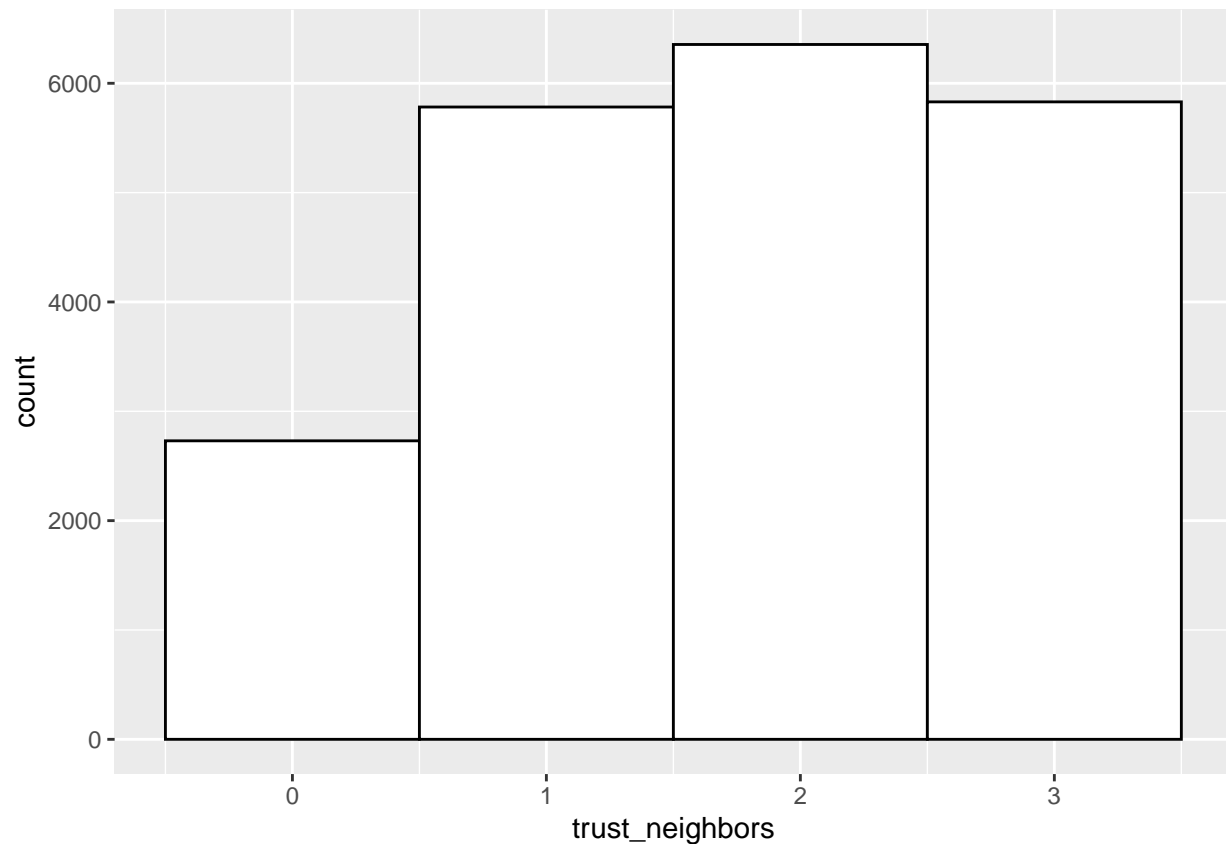
2 (“somewhat”). La respuesta que menos se contestó fue la de confianza nivel 0 (“not at all”).

En el caso de la confianza en el “local council”, la moda fue la confianza nivel 3 (“a lot”), y la menos respondida fue nivel de confianza 0 (“not at all”).

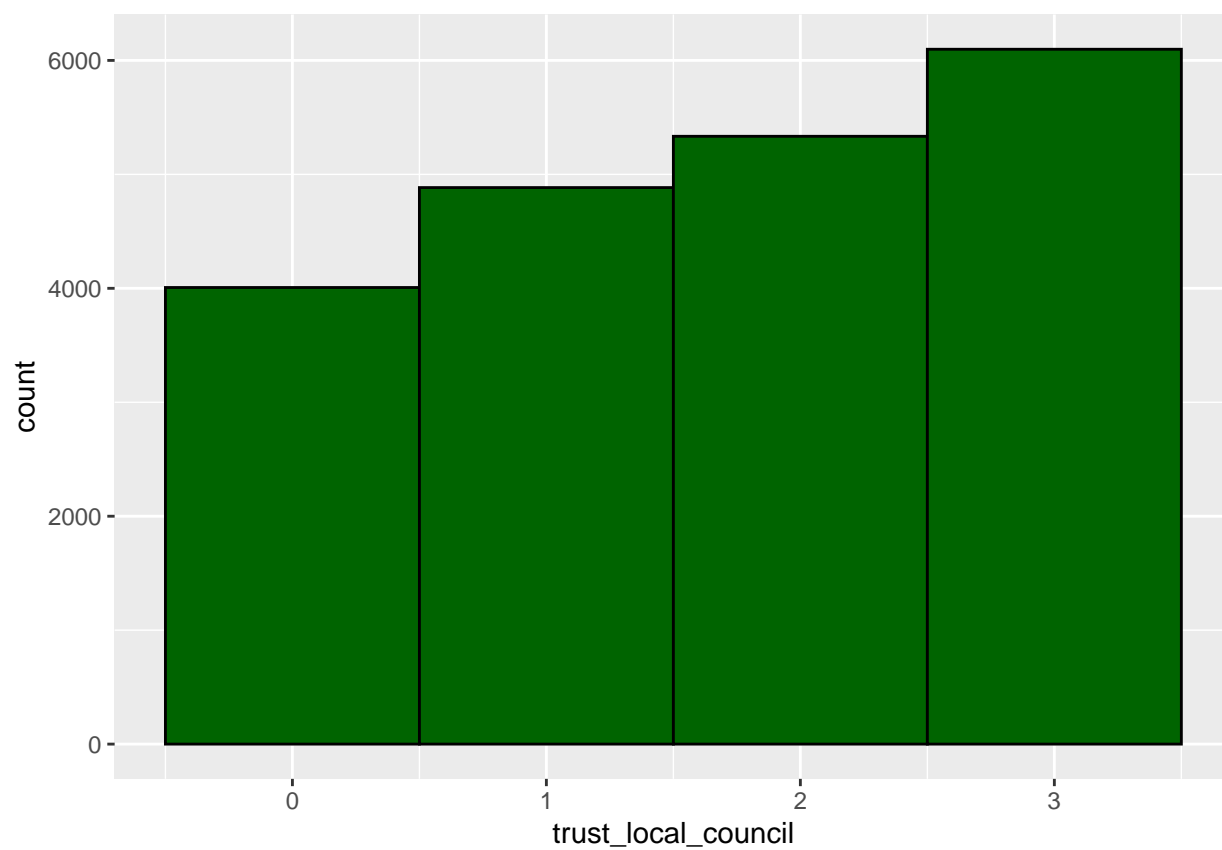
Respecto a la variable del comercio de esclavos, observamos una distribución muy sesgada hacia la izquierda. Podemos observar que la gran mayoría del comercio de esclavos se movía en volúmenes de menos de 100 personas. Esto tiene sentido dado que lo que se transportaba eran personas y las limitaciones de transporte de la época (números de barcos y sus dimensiones determinadas).

Sobre la variable instrumental, también se observa una distribución concentrada en valores bajos (aunque en menor magnitud que el comercio de esclavos), lo cual también tiene sentido dado que mide la distancia al mar de las villas donde habitaban mayoritariamente las etnias africanas, y las características geográficas del continente.

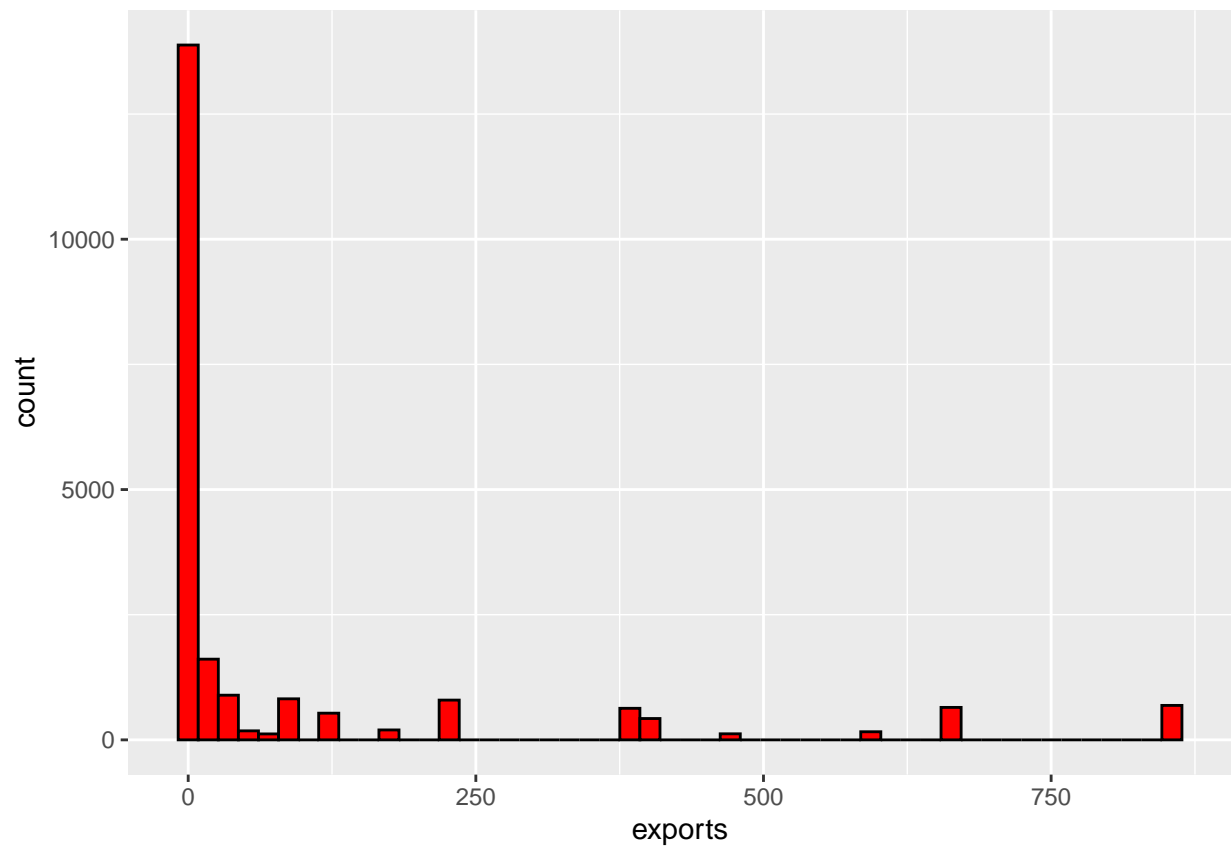
```
g_neig <- ggplot(dta, aes(x = trust_neighbors)) +  
  geom_histogram(binwidth = 1, colour='black', fill='white')  
g_lc <- ggplot(dta, aes(x = trust_local_council)) +  
  geom_histogram(binwidth = 1, colour='black', fill='dark green')  
g_exp <- ggplot(dta, aes(x = exports)) +  
  geom_histogram(bins = 50, colour='black', fill='red')  
g_dist <- ggplot(dta, aes(x = distsea)) +  
  geom_histogram(bins = 20, colour='black', fill='blue')  
  
g_neig
```



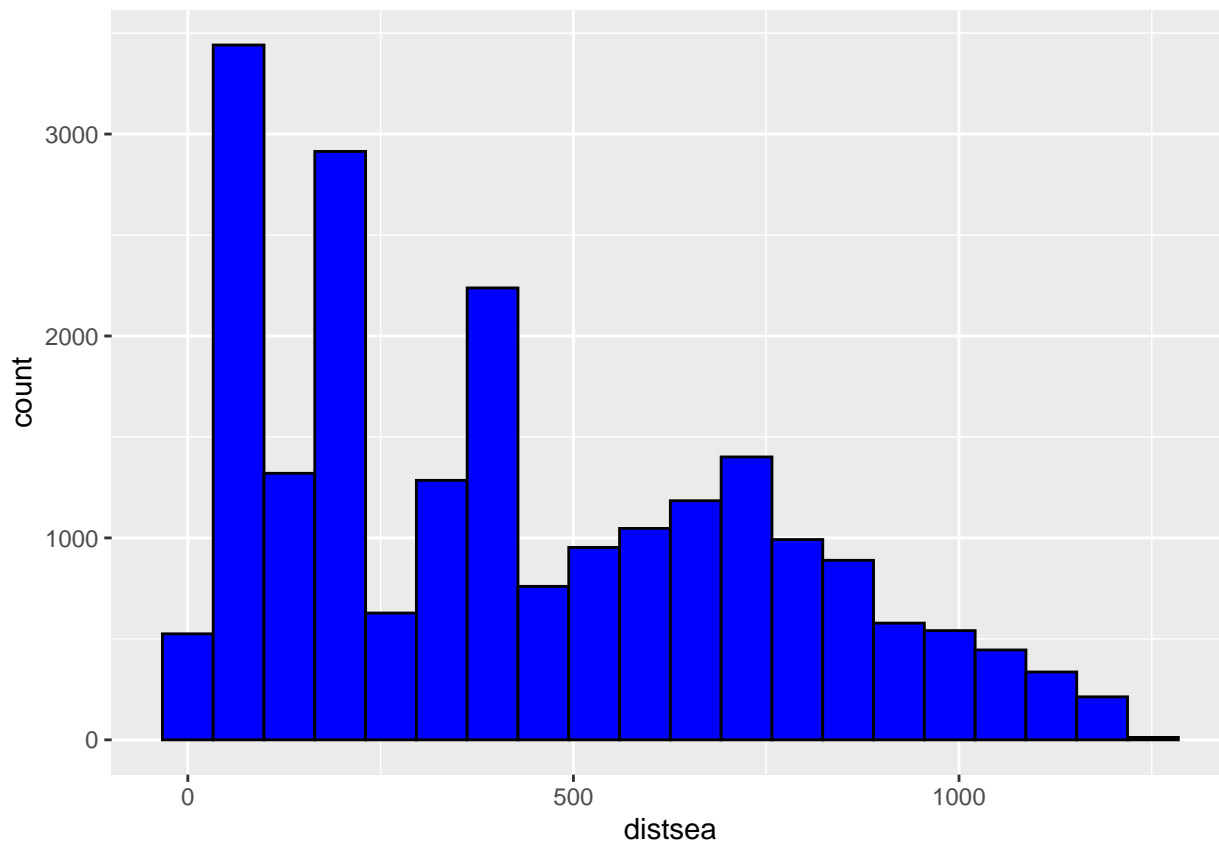
g_lc



g_exp



g_dist



5. Obtenga las estimaciones por mínimos cuadrados ordinarios para los determinantes de la variable Trust of neighbors. Esto es, replique los resultados de la tabla 1 del artículo. Tome en cuenta la corrección que se hace a los errores estándar robustos ante la heterogeneidad. Puede obtener al menos los dos primeros utilizando las funciones en R de Arai (2015) `clx` y `mclx`. Tal como refiere el artículo, los primeros errores estándar robustos consideran sólo heterogeneidad por el nombre de comunidad, los segundos lo hacen por nombre de comunidad y distrito.

Funciones de Arai (2015):

```
# One-way cluster function (Arai 2015)
clx <-
function(fm, dfcw, cluster){
  library(sandwich)
  library(lmtest)
  M <- length(unique(cluster))
  N <- length(cluster)
  dfc <- (M/(M-1))*((N-1)/(N-fm$rank))
  u <- apply(estfun(fm),2,
             function(x) tapply(x, cluster, sum))
  vcovCL <- dfc*sandwich(fm, meat=crossprod(u)/N)*dfcw
  coeftest(fm, vcovCL) }

# Two-way cluster function (Arai 2015)
mclx <-
function(fm, dfcw, cluster1, cluster2){
  library(sandwich)
  library(lmtest)
```

```

cluster12 = paste(cluster1,cluster2, sep="")
M1 <- length(unique(cluster1))
M2 <- length(unique(cluster2))
M12 <- length(unique(cluster12))
N <- length(cluster1)
K <- fm$rank
dfc1 <- (M1/(M1-1))*((N-1)/(N-K))
dfc2 <- (M2/(M2-1))*((N-1)/(N-K))
dfc12 <- (M12/(M12-1))*((N-1)/(N-K))
u1 <- apply(estfun(fm), 2,
            function(x) tapply(x, cluster1, sum))
u2 <- apply(estfun(fm), 2,
            function(x) tapply(x, cluster2, sum))
u12 <- apply(estfun(fm), 2,
            function(x) tapply(x, cluster12, sum))
vc1 <- dfc1*sandwich(fm, meat=crossprod(u1)/N )
vc2 <- dfc2*sandwich(fm, meat=crossprod(u2)/N )
vc12 <- dfc12*sandwich(fm, meat=crossprod(u12)/N)
vcovMCL <- (vc1 + vc2 - vc12)*dfcw
coeftest(fm, vcovMCL)}

```

Seguendo a Nunn & Wantchekon, se calcula el primer modelo de regresión lineal (incorporando los controles específicos), que busca medir el efecto del comercio de esclavos en la confianza de los africanos en sus **vecinos**.

Para replicar la Tabla 1 se estima el mismo modelo pero evaluando con distintas transformaciones de la variable de tratamiento (comercio de esclavos):

El primer modelo toma como variable independiente el volumen total de esclavos y como variable de respuesta el nivel de confianza reportando en sus vecinos, sin embargo por la magnitud de los valores que toma la var. dependiente y dado que está muy fuertemente sesgada la izquierda (como vimos en el histograma anterior), se utilizan diversas transformaciones que buscan estandarizar esta variable. El segundo modelo toma las exportaciones estandarizadas por la extensión del área donde vivían la mayoría de las etnias; el tercer modelo las exportaciones estandarizadas por el la población histórica de cada etnia; el cuarto modelo toma el logaritmo natural de las exportaciones; el quinto, el logaritmo de las exportaciones estandarizado por la extensión territorial; y el sexto, el logaritmo de las exportaciones estandarizado por la población histórica.

Los resultados para todas las variaciones de la exportación de esclavos muestran un efecto causal negativo y significativo en el nivel de confianza en los vecinos. Consistente con la hipótesis de los autores.

Como señalan los autores, dado el potencial de correlación entre los residuales de los grupos, se ajustan los errores estándar robustos y de cluster utilizando las funciones de Arai (2015)

```

# Tabla 1, modelo 1 (M1.1). seleccionamos variables de interés
dta_11 <- dta %>% select(trust_neighbors, exports, #export_area, export_pop,
                        #ln_exports, ln_export_area, ln_export_pop,
                        age, age2, male,
                        urban_dum, education, occupation, religion,
                        living_conditions, district_ethnic_frac,
                        frac_ethnicity_in_district, isocode,
                        murdock_name, district) %>% na.omit()

# MTrust of neighbors
mod_11 <- lm(trust_neighbors ~ exports + age + age2 + male + urban_dum +
             as.factor(living_conditions) + as.factor(education) +
             as.factor(religion) + as.factor(occupation) +

```

```

        district_ethnic_frac + frac_ethnicity_in_district +
        as.factor(isocode),
        data = dta_11)

# Errores estándar usando Arai (2015)
ee1_11 <- clx(mod_11, 1, dta_11$murdock_name)
ee2_11 <- mclx(mod_11, 1, dta_11$murdock_name, dta_11$district)

#summary(mod_11)
#ee1_11

# Tabla 1 modelo 2 (M1.2)
dta_12 <- dta %>% select(trust_neighbors, export_area, #export_pop,
                        #ln_exports, ln_export_area, ln_export_pop,
                        age, age2, male,
                        urban_dum, education, occupation, religion,
                        living_conditions, district_ethnic_frac,
                        frac_ethnicity_in_district, isocode,
                        murdock_name, district) %>% na.omit()
mod_12 <- lm(trust_neighbors ~ export_area + age + age2 + male + urban_dum +
             as.factor(living_conditions) + as.factor(education) +
             as.factor(religion) + as.factor(occupation) +
             district_ethnic_frac + frac_ethnicity_in_district +
             as.factor(isocode), data = dta_12)

ee1_12 <- clx(mod_12, 1, dta_12$murdock_name)
ee2_12 <- mclx(mod_12, 1, dta_12$murdock_name, dta_12$district)

# M1.3
dta_13 <- dta %>% select(trust_neighbors, export_pop,
                        #ln_exports, ln_export_area, ln_export_pop,
                        age, age2, male,
                        urban_dum, education, occupation, religion,
                        living_conditions, district_ethnic_frac,
                        frac_ethnicity_in_district, isocode,
                        murdock_name, district) %>% na.omit()
mod_13 <- lm(trust_neighbors ~ export_pop + age + age2 + male + urban_dum +
             as.factor(living_conditions) + as.factor(education) +
             as.factor(religion) + as.factor(occupation) +
             district_ethnic_frac + frac_ethnicity_in_district +
             as.factor(isocode), data = dta_13)

ee1_13 <- clx(mod_13, 1, dta_13$murdock_name)
ee2_13 <- mclx(mod_13, 1, dta_13$murdock_name, dta_13$district)

# M1.4
dta_14 <- dta %>% select(trust_neighbors, #export_pop,
                        ln_exports, #ln_export_area, ln_export_pop,
                        age, age2, male,
                        urban_dum, education, occupation, religion,
                        living_conditions, district_ethnic_frac,
                        frac_ethnicity_in_district, isocode,

```

```

      murdock_name, district) %>% na.omit()
mod_14 <- lm(trust_neighbors ~ ln_exports + age + age2 + male + urban_dum +
  as.factor(living_conditions) + as.factor(education) +
  as.factor(religion) + as.factor(occupation) +
  district_ethnic_frac + frac_ethnicity_in_district +
  as.factor(isocode), data = dta_14)

ee1_14 <- clx(mod_14, 1, dta_14$murdock_name)
ee2_14 <- mclx(mod_14, 1, dta_14$murdock_name, dta_14$district)

# M1.5
dta_15 <- dta %>% select(trust_neighbors, #export_pop,
  ln_export_area, #ln_export_pop,
  age, age2, male,
  urban_dum, education, occupation, religion,
  living_conditions, district_ethnic_frac,
  frac_ethnicity_in_district, isocode,
  murdock_name, district) %>% na.omit()
mod_15 <- lm(trust_neighbors ~ ln_export_area + age + age2 + male + urban_dum +
  as.factor(living_conditions) + as.factor(education) +
  as.factor(religion) + as.factor(occupation) +
  district_ethnic_frac + frac_ethnicity_in_district +
  as.factor(isocode), data = dta_15)

ee1_15 <- clx(mod_15, 1, dta_15$murdock_name)
ee2_15 <- mclx(mod_15, 1, dta_15$murdock_name, dta_15$district)

# M1.4
dta_16 <- dta %>% select(trust_neighbors, ln_export_pop, age, age2, male,
  urban_dum, education, occupation, religion,
  living_conditions, district_ethnic_frac,
  frac_ethnicity_in_district, isocode,
  murdock_name, district) %>% na.omit()
mod_16 <- lm(trust_neighbors ~ ln_export_pop + age + age2 + male + urban_dum +
  as.factor(living_conditions) + as.factor(education) +
  as.factor(religion) + as.factor(occupation) +
  district_ethnic_frac + frac_ethnicity_in_district +
  as.factor(isocode), data = dta_16)

ee1_16 <- clx(mod_16, 1, dta_16$murdock_name)
ee2_16 <- mclx(mod_16, 1, dta_16$murdock_name, dta_16$district)

```

Se reportan resultados:

```

stargazer(mod_11, mod_12, mod_13, mod_14, mod_15, mod_16, type = "text",
  se = list(c(ee1_11[,2], ee2_11[,2]),
    c(ee1_12[,2], ee2_12[,2]),
    c(ee1_13[,2], ee2_13[,2]),
    c(ee1_14[,2], ee2_14[,2]),
    c(ee1_15[,2], ee2_15[,2]),
    c(ee1_16[,2], ee2_16[,2])),
  digits = 5,
  keep = c("exports", "export_area", "export_pop",

```

```

"ln_exports", "ln_export_area", "ln_export_pop"))

##
## =====
##
## -----
##
## (1) (2) (3)
## -----
## exports -0.00068***
## (0.00014)
##
## export_area -0.01884***
## (0.00509)
##
## export_pop -0.53128***
## (0.14675)
##
## ln_exports
##
## ln_export_area
##
## ln_export_pop
##
## -----
## Observations 20,027 20,027 17,644
## R2 0.15583 0.15633 0.14712
## Adjusted R2 0.15257 0.15307 0.14338
## Residual Std. Error 0.92749 (df = 19949) 0.92722 (df = 19949) 0.91723 (df = 1756)
## F Statistic 47.82569*** (df = 77; 19949) 48.00669*** (df = 77; 19949) 39.35124*** (df = 77; 1756)
## =====
## Note:

```

6. Ahora repita la estimación por mínimos cuadrados ordinarios pero para los determinantes de la variable Trust of others. Esto es, en la medida de lo posible, replique los resultados de la tabla 2. Compare e interprete los resultados que obtuvo en este inciso con los del anterior.

Siguiendo a los autores, se utilizará de aquí en adelante el modelo “base”, que utiliza como variable de tratamiento al logaritmo natural de las exportaciones de esclavos normalizadas por la extensión territorial. Asimismo, para el resto de los ejercicios se utilizan los errores estándar robustos de etnia y distrito.

Con ello, ahora se busca medir el efecto causal del comercio de esclavos sobre la confianza en “otros”, que incluye: [-] Confianza en parientes [-] Confianza en gobierno local [-] Confianza al interior del grupo [-] Confianza entre grupos

Es decir, en términos prácticos, evaluamos el modelo base cambiando la variable de resultado.

```

# filtramos datos para obtener variables de interés. En este caso, podemos generar
# una sola bd dado que utilizaremos solo a ln_export_area como variable de tratamiento
dta_21 <- dta %>% select(trust_relatives,
                        #trust_neighbors,

```

```

        #trust_local_council,
        #intra_group_trust,
        #inter_group_trust,
        ln_export_area,
        age, age2, male, urban_dum, education, occupation, religion,
        living_conditions, district_ethnic_frac,
        frac_ethnicity_in_district, isocode, murdock_name,
        district) %>% na.omit()

dta_22 <- dta %>% select(#trust_relatives,
                        trust_neighbors,
                        #trust_local_council,
                        #intra_group_trust,
                        #inter_group_trust,
                        ln_export_area, age, age2, male, urban_dum, education,
                        occupation, religion, living_conditions,
                        district_ethnic_frac, frac_ethnicity_in_district,
                        isocode, murdock_name,
                        district) %>% na.omit()

dta_23 <- dta %>% select(#trust_relatives,
                        #trust_neighbors,
                        trust_local_council,
                        #intra_group_trust,
                        #inter_group_trust,
                        ln_export_area, age, age2, male, urban_dum, education,
                        occupation, religion, living_conditions,
                        district_ethnic_frac, frac_ethnicity_in_district,
                        isocode, murdock_name,
                        district) %>% na.omit()

dta_24 <- dta %>% select(#trust_relatives,
                        #trust_neighbors,
                        #trust_local_council,
                        intra_group_trust,
                        #inter_group_trust,
                        ln_export_area, age, age2, male, urban_dum, education,
                        occupation, religion, living_conditions,
                        district_ethnic_frac, frac_ethnicity_in_district,
                        isocode, murdock_name,
                        district) %>% na.omit()

dta_25 <- dta %>% select(#trust_relatives,
                        #trust_neighbors,
                        #trust_local_council,
                        #intra_group_trust,
                        inter_group_trust,
                        ln_export_area, age, age2, male, urban_dum, education,
                        occupation, religion, living_conditions,
                        district_ethnic_frac, frac_ethnicity_in_district,
                        isocode, murdock_name,
                        district) %>% na.omit()

```

```

# Tabla 2, modelo 1 (M2.1)
mod_21 <- lm(trust_relatives ~ ln_export_area + age + age2 + male + urban_dum +
  as.factor(living_conditions) + as.factor(education) +
  as.factor(religion) + as.factor(occupation) +
  district_ethnic_frac + frac_ethnicity_in_district +
  as.factor(isocode), data = dta_21)
ee1_21 <- clx(mod_21, 1, dta_21$murdock_name)
ee2_21 <- mclx(mod_21, 1, dta_21$murdock_name, dta_21$district)

# M2.2
mod_22 <- lm(trust_neighbors ~ ln_export_area + age + age2 + male + urban_dum +
  as.factor(living_conditions) + as.factor(education) +
  as.factor(religion) + as.factor(occupation) +
  district_ethnic_frac + frac_ethnicity_in_district +
  as.factor(isocode), data = dta_22)
ee1_22 <- clx(mod_22, 1, dta_22$murdock_name)
ee2_22 <- mclx(mod_22, 1, dta_22$murdock_name, dta_22$district)

# M2.3
mod_23 <- lm(trust_local_council ~ ln_export_area + age + age2 + male + urban_dum +
  as.factor(living_conditions) + as.factor(education) +
  as.factor(religion) + as.factor(occupation) +
  district_ethnic_frac + frac_ethnicity_in_district +
  as.factor(isocode), data = dta_23)
ee1_23 <- clx(mod_23, 1, dta_23$murdock_name)
ee2_23 <- mclx(mod_23, 1, dta_23$murdock_name, dta_23$district)

# M2.4
mod_24 <- lm(intra_group_trust ~ ln_export_area + age + age2 + male + urban_dum +
  as.factor(living_conditions) + as.factor(education) +
  as.factor(religion) + as.factor(occupation) +
  district_ethnic_frac + frac_ethnicity_in_district +
  as.factor(isocode), data = dta_24)
ee1_24 <- clx(mod_24, 1, dta_24$murdock_name)
ee2_24 <- mclx(mod_24, 1, dta_24$murdock_name, dta_24$district)

# M2.5
mod_25 <- lm(inter_group_trust ~ ln_export_area + age + age2 + male + urban_dum +
  as.factor(living_conditions) + as.factor(education) +
  as.factor(religion) + as.factor(occupation) +
  district_ethnic_frac + frac_ethnicity_in_district +
  as.factor(isocode), data = dta_25)
ee1_25 <- clx(mod_25, 1, dta_25$murdock_name)
ee2_25 <- mclx(mod_25, 1, dta_25$murdock_name, dta_25$district)

```

Se reportan resultados y observamos que en el efecto causal sigue siendo negativo y es negativo para todos los distintos “tipos” de confianza evaluados.

```

stargazer(mod_21, mod_22, mod_23, mod_24, mod_25, type = "text",
  se = list(c(ee1_21[,2], ee2_21[,2]),
    c(ee1_22[,2], ee2_22[,2]),
    c(ee1_23[,2], ee2_23[,2]),
    c(ee1_24[,2], ee2_24[,2]),

```

```
c(ee1_25[,2], ee2_25[,2]),
digits = 3,
keep = c("ln_export_area"))
```

```
##
## =====
##                                     Dependent variable:
##                                     -----
##                                trust_relatives    trust_neighbors    trust_local_council
##                                (1)                (2)                (3)
## -----
## ln_export_area            -0.133***            -0.159***            -0.111***
##                            (0.036)              (0.034)              (0.021)
## -----
## Observations                20,062                20,027                19,733
## R2                          0.133                  0.156                  0.196
## Adjusted R2                 0.130                  0.153                  0.193
## Residual Std. Error    0.894 (df = 19984)    0.927 (df = 19949)    0.989 (df = 19654)
## F Statistic            39.830*** (df = 77; 19984) 47.963*** (df = 77; 19949) 61.425*** (df = 78; 19654)
## =====
## Note:
```

7. Analice la variable instrumental a detalle. Utilice regresiones de mínimos cuadrados ordinarios para revisar si el instrumento afecta a la variable de tratamiento slave exports. También revise si el instrumento afecta a cada una de las variables dependientes de confianza. Analice y discuta sus resultados.

Al evaluar mediante MCO la relación entre la variable instrumental y las exportaciones de esclavos encontramos que el efecto sí es significativo y es negativo. Esto está en línea con nuestras expectativas, y nos dice que a mayor distancia del mar, el número de exportaciones de la etnia específica es menor. Es un primer indicio de que podemos usar esta variable como instrumento del comercio de esclavos, además, esperaríamos que no presente problemas de endogeneidad

```
iv_exports <- lm(exports ~ distsea, data = dta, na.action = na.omit)
iv_ln_export_area <- lm(ln_export_area ~ distsea, data = dta, na.action = na.omit)

stargazer(iv_exports, iv_ln_export_area, type = "text")
```

```
##
## =====
##                                     Dependent variable:
##                                     -----
##                                exports    ln_export_area
##                                (1)        (2)
## -----
## distsea            -0.224***    -0.001***
##                    (0.004)      (0.00002)
##
## Constant           191.550***    1.107***
##                    (2.268)      (0.010)
## -----
## Observations                21,702                21,702
```



```
## R2                                0.115          0.184
## Adjusted R2                      0.115          0.184
## Residual Std. Error (df = 21700) 193.105        0.852
## F Statistic (df = 1; 21700)      2,823.994***    4,896.528***
## =====
## Note:                            *p<0.1; **p<0.05; ***p<0.01
```

También evaluamos de manera sencilla el instrumento con las variables de resultado y observamos que sí hay un efecto significativo, aunque de magnitud muy baja y cercana a cero, en todos los casos. Esto también está en línea con lo que esperamos.

Compararemos más adelante vs el modelo mejor especificado, con controles y con uso correcto de instrumento.

```
iv_rel <- lm(trust_relatives ~ distsea, data = dta, na.action = na.omit)
iv_neig <- lm(trust_neighbors ~ distsea, data = dta, na.action = na.omit)
iv_lc <- lm(trust_local_council ~ distsea, data = dta, na.action = na.omit)
iv_intra <- lm(intra_group_trust ~ distsea, data = dta, na.action = na.omit)
iv_inter <- lm(inter_group_trust ~ distsea, data = dta, na.action = na.omit)

stargazer(iv_rel, iv_neig, iv_lc, iv_intra, iv_inter, type = "text")
```

```
##
## =====
##                                     Dependent variable:
## -----
##               trust_relatives      trust_neighbors      trust_local_council
##               (1)                (2)                (3)
## -----
## distsea              0.0001***          0.0002***          0.0004***
##                   (0.00002)          (0.00002)          (0.00002)
##
## Constant             2.145***          1.659***          1.467***
##                   (0.011)          (0.012)          (0.013)
##
## -----
## Observations          20,618            20,580            20,210
## R2                    0.001             0.003             0.016
## Adjusted R2           0.001             0.003             0.016
## Residual Std. Error   0.958 (df = 20616)  1.008 (df = 20578)  1.094 (df = 20208)
## F Statistic           20.625*** (df = 1; 20616) 62.243*** (df = 1; 20578) 326.043*** (df = 1; 20208) 1
## =====
## Note:
```

8. Utilice su instrumento para realizar ahora los modelos para las variables dependientes mediante el uso de la variable instrumental. Ahora deberá replicar los resultados de la tabla 5 del artículo. Tome en cuenta cómo debe incorporar controles para cada especificación.

Siguiendo lo realizado por los autores, calculamos la regresión de dos etapas mediante la función `iv_reg`, y se corrigen los errores estándar robustos y de cluster para comunidad y distrito.

```
dta_51 <- dta %>% select(trust_relatives,
                        #trust_neighbors,
                        #trust_local_council,
                        #intra_group_trust,
```

```

      #inter_group_trust,
      ln_export_area, age, age2, male, urban_dum, education,
      occupation, religion, living_conditions,
      district_ethnic_frac, frac_ethnicity_in_district,
      isocode, murdock_name, district, distsea,
      malaria_ecology, total_missions_area, explorer_contact,
      railway_contact, cities_1400_dum, v30, v33,
      ln_init_pop_density) %>% na.omit()

dta_52 <- dta %>% select(#trust_relatives,
      trust_neighbors,
      #trust_local_council,
      #intra_group_trust,
      #inter_group_trust,
      ln_export_area, age, age2, male, urban_dum, education,
      occupation, religion, living_conditions,
      district_ethnic_frac, frac_ethnicity_in_district,
      isocode, murdock_name, district, distsea,
      malaria_ecology, total_missions_area, explorer_contact,
      railway_contact, cities_1400_dum, v30, v33,
      ln_init_pop_density) %>% na.omit()

dta_53 <- dta %>% select(#trust_relatives,
      #trust_neighbors,
      trust_local_council,
      #intra_group_trust,
      #inter_group_trust,
      ln_export_area, age, age2, male, urban_dum, education,
      occupation, religion, living_conditions,
      district_ethnic_frac, frac_ethnicity_in_district,
      isocode, murdock_name, district, distsea,
      malaria_ecology, total_missions_area, explorer_contact,
      railway_contact, cities_1400_dum, v30, v33,
      ln_init_pop_density) %>% na.omit()

dta_54 <- dta %>% select(#trust_relatives,
      #trust_neighbors,
      #trust_local_council,
      intra_group_trust,
      #inter_group_trust,
      ln_export_area, age, age2, male, urban_dum, education,
      occupation, religion, living_conditions,
      district_ethnic_frac, frac_ethnicity_in_district,
      isocode, murdock_name, district, distsea,
      malaria_ecology, total_missions_area, explorer_contact,
      railway_contact, cities_1400_dum, v30, v33,
      ln_init_pop_density) %>% na.omit()

dta_55 <- dta %>% select(#trust_relatives,
      #trust_neighbors,
      #trust_local_council,
      #intra_group_trust,
      inter_group_trust,

```

```

        ln_export_area, age, age2, male, urban_dum, education,
        occupation, religion, living_conditions,
        district_ethnic_frac, frac_ethnicity_in_district,
        isocode, murdock_name, district, distsea,
        malaria_ecology, total_missions_area, explorer_contact,
        railway_contact, cities_1400_dum, v30, v33,
        ln_init_pop_density) %>% na.omit()

mod_51 <- ivreg(trust_relatives ~ ln_export_area + age + age2 + male + urban_dum +
  as.factor(living_conditions) + as.factor(education) +
  as.factor(religion) + as.factor(occupation) +
  district_ethnic_frac + frac_ethnicity_in_district +
  as.factor(isocode) + malaria_ecology + total_missions_area +
  explorer_contact + railway_contact + cities_1400_dum +
  as.factor(v30) + v33 + ln_init_pop_density | age + age2 + male + urban_dum +
  as.factor(living_conditions) + as.factor(education) +
  as.factor(religion) + as.factor(occupation) +
  district_ethnic_frac + frac_ethnicity_in_district +
  as.factor(isocode) + malaria_ecology + total_missions_area +
  explorer_contact + railway_contact + cities_1400_dum +
  as.factor(v30) + v33 + ln_init_pop_density + distsea,
  data = dta_51)

mod_52 <- ivreg(trust_neighbors ~ ln_export_area + age + age2 + male + urban_dum +
  as.factor(living_conditions) + as.factor(education) +
  as.factor(religion) + as.factor(occupation) +
  district_ethnic_frac + frac_ethnicity_in_district +
  as.factor(isocode) + malaria_ecology + total_missions_area +
  explorer_contact + railway_contact + cities_1400_dum +
  as.factor(v30) + v33 + ln_init_pop_density | . - ln_export_area +
  distsea,
  data = dta_52)

mod_53 <- ivreg(trust_local_council ~ ln_export_area + age + age2 + male + urban_dum +
  as.factor(living_conditions) + as.factor(education) +
  as.factor(religion) + as.factor(occupation) +
  district_ethnic_frac + frac_ethnicity_in_district +
  as.factor(isocode) + malaria_ecology + total_missions_area +
  explorer_contact + railway_contact + cities_1400_dum +
  as.factor(v30) + v33 + ln_init_pop_density | . - ln_export_area +
  distsea,
  data = dta_53)

mod_54 <- ivreg(intra_group_trust ~ ln_export_area + age + age2 + male + urban_dum +
  as.factor(living_conditions) + as.factor(education) +
  as.factor(religion) + as.factor(occupation) +
  district_ethnic_frac + frac_ethnicity_in_district +
  as.factor(isocode) + malaria_ecology + total_missions_area +
  explorer_contact + railway_contact + cities_1400_dum +
  as.factor(v30) + v33 + ln_init_pop_density | . - ln_export_area +
  distsea,
  data = dta_54)

mod_55 <- ivreg(inter_group_trust ~ ln_export_area + age + age2 + male + urban_dum +
  as.factor(living_conditions) + as.factor(education) +
  as.factor(religion) + as.factor(occupation) +
  district_ethnic_frac + frac_ethnicity_in_district +

```

```

as.factor(isocode) + malaria_ecology + total_missions_area +
explorer_contact + railway_contact + cities_1400_dum +
as.factor(v30) + v33 + ln_init_pop_density | . - ln_export_area +
distsea,
data = dta_55)

# usamos errores estándar de etnia y distrito
ee2_51 <- mclx(mod_51, 1, dta_51$murdock_name, dta_51$district)
ee2_52 <- mclx(mod_52, 1, dta_52$murdock_name, dta_52$district)
ee2_53 <- mclx(mod_53, 1, dta_53$murdock_name, dta_53$district)
ee2_54 <- mclx(mod_54, 1, dta_54$murdock_name, dta_54$district)
ee2_55 <- mclx(mod_55, 1, dta_55$murdock_name, dta_55$district)

```

Reportamos resultados:

```

stargazer(mod_51, mod_52, mod_53, mod_54, mod_55,
type = "text",
se = list(ee2_51[,2],
ee2_52[,2],
ee2_53[,2],
ee2_54[,2],
ee2_55[,2]),
digits = 3,
keep = c("ln_export_area"))

```

```

##
## =====
##                                     Dependent variable:
##                                     -----
##               trust_relatives    trust_neighbors    trust_local_council    intra_group_trust    intra_group_trust
##               (1)                (2)                (3)                (4)
## -----
## ln_export_area      -0.190***      -0.245***      -0.221***      -0.251***
##                   (0.067)         (0.070)         (0.060)         (0.088)
## -----
## Observations         16,709         16,679         15,905         16,636
## R2                   0.130          0.159          0.204          0.154
## Adjusted R2          0.126          0.155          0.199          0.150
## Residual Std. Error 0.884 (df = 16617) 0.911 (df = 16587) 0.978 (df = 15812) 0.910 (df = 16544) 0.910 (df = 16544)
## =====
## Note:                                                         *p<0.1; **p<0.01; ***p<0.001

```

9. Finalmente, realice las pruebas de diagnóstico convencionales sobre instrumentos débiles. Utilice las funciones disponibles en el paquete ivreg de R. Interprete sus resultados y concluya en el contexto del problema.

Calculamos estadístico F de la regresión de primera etapa y observamos que para ambas variables de interés, `exports` y su transformación `ln_export_area`, el valor es significativamente mayor a 10. Por lo que podemos afirmar que no tenemos problemas de instrumento débil, es un instrumento fuerte.

```

#summary(mod_51, diagnostics = TRUE)
# Regresión de la primera etapa
iv_test1 <- linearHypothesis(iv_exports, "distsea = 0", vcov = vcovHC, type = "HC1")
iv_test2 <- linearHypothesis(iv_ln_export_area, "distsea = 0", vcov = vcovHC, type = "HC1")

# imprime resultados de estadístico F de la 1SLS
print(c(iv_test1$F, iv_test2$F))

```

```
## [1]          NA 2952.351          NA 4987.299
```

10. Escriba la conclusión general del análisis realizado. Interprete la cuantificación del efecto causal y discuta la utilidad y ventajas de aplicar la estrategia empírica seleccionada para la cuantificación del efecto causal.

En la primera parte, al evaluar el efecto causal del comercio de esclavos en el nivel de confianza actual de las etnias africanas mediante MCO encontramos que sí existe un efecto negativo y significativo. Sin embargo, hay sospechas de que el comercio de esclavos y el nivel de confianza pudieron haber estado afectados por una serie de covariables como las características de los individuos, de la región, o de si los países estuvieron colonizados o no, además de variables omitidas, lo cual genera un problema de endogeneidad.

Por ello, para medir de mejor manera este efecto se instrumentó el comercio de esclavos con la distancia a la costa de la villa específica donde habitaban las etnias africanas durante los años de mayor auge del esclavismo. Esta variable reúne las características deseadas para instrumentar el comercio de esclavos: está correlacionado con las exportaciones (de manera negativa), pero no lo está con las características de los grupos étnicos y de la región que podrían afectar el nivel de confianza actual. Esta es la ventaja de utilizar esta técnica cuando nos enfrentamos a problemas de endogeneidad. Para asegurarnos de que el instrumento es válido, se hicieron las pruebas correspondientes.

De esta forma, se calcula el modelo con variable instrumental para medir el efecto de interés de manera más consistente y válida, sin problemas de endogeneidad. Los resultados nos arrojan un efecto negativo y significativo del comercio de esclavos con el nivel de confianza actual, tanto en los vecinos como en “otros”. Esto está en línea con las primeras estimaciones de MCO, más aún, es importante señalar, que incluso los coeficientes que arroja el modelo de variables instrumentales es muy similar que el de MCO.