

Exercício 4 - Solução

Microeconometria - IBM0288

Prof. Raphael Gouvea

2026-04-10

1 Dados em Painel — Exercício E10.2 (Stock & Watson)

Este documento apresenta a solução para o exercício E10.2 de Stock & Watson (2020), que investiga a relação entre renda per capita e democracia utilizando um painel de 195 países para os anos de 1960 a 2000 (em intervalos de 5 anos).

A pergunta central é: **a democracia é um bem normal?** Ou seja, os cidadãos demandam mais liberdade política à medida que sua renda cresce?

Antes de iniciar a atividade, vamos chamar os pacotes que serão utilizados. Nesta resolução são utilizados os seguintes pacotes do R.

```
library(tidyverse)
library(skimr)
library(fixest) # melhor pacote para estimar paineis
library(modelsummary) # para ajudar na elaboração de tabelas a partir de modelos
library(gt) # elaboração de tabelas para publicação
library(gtExtras) # temas e funções adicionais para gt
library(readxl)
```

O primeiro passo é ler os dados no R. Vamos salvar os dados no objeto `df`.

```
# Carregar os dados
# Ajuste o caminho conforme a localização do arquivo na sua máquina
df <- read_xlsx(here::here("labs", "income_democracy", "income_democracy.xlsx"))
```

a. Painel Balanceado?

```
# Verificar se o painel é balanceado
df %>%
  count(country) %>%
  summarise(
    n_países      = n(),
    obs_min       = min(n),
    obs_max       = max(n),
    painel_balanceado = if_else(obs_min == obs_max, "Sim", "Não")
  )
```

```
# A tibble: 1 x 4
  n_países obs_min obs_max painel_balanceado
  <int>    <int>    <int> <chr>
1      195        1        9 Não
```

Um painel é **balanceado** quando todos os países possuem o mesmo número de observações ao longo do tempo. Caso contrário, é **não balanceado**, o que pode ocorrer por dados faltantes em alguns países/anos. Como vemos nos dados acima, há 195 países na base de dados, sendo que existe país com apenas uma observação e países com 9 observações. Portanto, o painel é não balanceado.

Observação: Leia a documentação da função `count` para entender como foi feita a análise.

b. Estatísticas Descritivas de Dem_ind

- i. Poderíamos utilizar a função `skim` para obter as estatísticas descritivas básicas. Contudo, a atividade pede outros percentis, como o percentil 10 e o 90.

```
# Tabela formatada com percentis customizados
df |>
  summarise(
    Variável      = "dem_ind",
    N             = sum(!is.na(dem_ind)),
    Mín          = min(dem_ind, na.rm = TRUE),
    Max          = max(dem_ind, na.rm = TRUE),
    Média        = mean(dem_ind, na.rm = TRUE),
    `Desvio Padrão` = sd(dem_ind, na.rm = TRUE),
    `P10`        = quantile(dem_ind, 0.10, na.rm = TRUE),
    `P25`        = quantile(dem_ind, 0.25, na.rm = TRUE),
    `P50`        = quantile(dem_ind, 0.50, na.rm = TRUE),
    `P75`        = quantile(dem_ind, 0.75, na.rm = TRUE),
    `P90`        = quantile(dem_ind, 0.90, na.rm = TRUE)
  )
```

```
# A tibble: 1 x 11
  Variável      N  Mín  Max Média `Desvio Padrão`  P10  P25  P50  P75  P90
  <chr>      <int> <dbl> <dbl> <dbl>          <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 dem_ind    1266     0     1 0.499          0.371     0 0.167  0.5 0.833     1
```

ii. e iii.

```
# Valores para EUA e Líbia
países_interesse <- df %>%
  filter(country %in% c("United States", "Libya"))
```

```
# Média 2000
países_interesse %>%
  filter(year == 2000) %>%
  select(country, year, dem_ind)
```

```
# A tibble: 2 x 3
  country      year dem_ind
  <chr>        <dbl>  <dbl>
1 Libya        2000     0
2 United States 2000     1
```

```
# Média todos os anos
países_interesse %>%
  select(country, dem_ind) %>%
  group_by(country) %>%
  summarise(
    Média = mean(dem_ind, na.rm = T))
```

```
# A tibble: 2 x 2
  country      Média
  <chr>        <dbl>
1 Libya        0.109
2 United States 0.986
```

iv.

```
# Países por faixa de Dem_ind médio
df |>
  group_by(country) |>
  summarise(media_dem = mean(dem_ind, na.rm = TRUE)) |>
  filter(media_dem > 0.95 | media_dem < 0.10 | between(media_dem, 0.3, 0.7)) |>
  mutate(faixa = case_when(
    media_dem > 0.95 ~ "> 0,95",
    media_dem < 0.10 ~ "< 0,10",
    between(media_dem, 0.3, 0.7) ~ "Entre 0,3 e 0,7")
```

```

)) |>
group_by(faixa) |>
slice_head(n = 5) |>
arrange(faixa, desc(media_dem))

```

```

# A tibble: 15 x 3
# Groups:   faixa [3]
  country      media_dem faixa
  <chr>         <dbl> <chr>
1 Afghanistan  0.0967 < 0,10
2 China        0.0804 < 0,10
3 Angola       0.0667 < 0,10
4 Brunei       0.0556 < 0,10
5 Burundi     0.0476 < 0,10
6 Australia    1      > 0,95
7 Barbados    1      > 0,95
8 Belgium     1      > 0,95
9 Belize      1      > 0,95
10 Austria    0.993 > 0,95
11 Argentina   0.666 Entre 0,3 e 0,7
12 Antigua     0.556 Entre 0,3 e 0,7
13 Bolivia     0.551 Entre 0,3 e 0,7
14 Bangladesh  0.533 Entre 0,3 e 0,7
15 Armenia     0.5    Entre 0,3 e 0,7

```

c. Estimativas de MQO

Para estimar um modelo MQO com erro padrão clusterizado, podemos utilizar o pacote `feols` sem incluir os termos de efeito fixo (veja a documentação do pacote).

```

# (c) OLS simples com erros clusterizados por país
feols(dem_ind ~ log_gdppc,
      data      = df,
      cluster  = ~country)

```

```

OLS estimation, Dep. Var.: dem_ind
Observations: 958
Standard-errors: Clustered (country)
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.354828   0.100421 -13.4914 < 2.2e-16 ***
log_gdppc    0.235673   0.011837  19.9099 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.271616  Adj. R2: 0.437878

```

- i. O coeficiente é 0,24 com erro padrão de 0,01. O intervalo de confiança de 95% vai de 0,21 a 0,26. O coeficiente é grande, conforme descrito abaixo.
- ii. Um aumento de 20% no PIB per capita implica que `log_gdp` aumenta aproximadamente 0,20, de modo que se prevê que `Dem_ind` aumente aproximadamente $0,20 \times 0,24 = 0,048$, ou cerca de 1/10 do desvio padrão no conjunto de dados. O intervalo de confiança de 95% para o efeito é (aproximadamente) $0,20 \times 0,21$ a $0,20 \times 0,26$, ou seja, de 0,042 a 0,052.
- iii. Os erros padrão clusterizados são necessários devido a fatores omitidos específicos de cada país nas regressões. O erro padrão clusterizado para `Dem_ind` é 0,012; o erro padrão não clusterizado é menor (0,007) porque ignora a correlação positiva dos erros dentro de cada país.

d. Estimativas em Painel

i.

Países têm histórias, instituições, estruturas sociais e religiões diferentes. Todos esses fatores afetam suas preferências por democracia e podem também estar correlacionados com o desenvolvimento econômico e, portanto, com a renda per capita.

ii.

```
feols(dem_ind ~ log_gdppc | country,
      data = df,
      cluster = ~country)
```

NOTES: 411 observations removed because of NA values (LHS: 103, RHS: 403).
21 fixed-effect singletons were removed (21 observations).

```
OLS estimation, Dep. Var.: dem_ind
Observations: 937
Fixed-effects: country: 129
Standard-errors: Clustered (country)
      Estimate Std. Error t value Pr(>|t|)
log_gdppc 0.083741  0.031543 2.65478 0.0089436 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.188866      Adj. R2: 0.685534
          Within R2: 0.019681
```

O coeficiente estimado cai por um fator de 3, para 0,083 com erro padrão de 0,032. O efeito estimado, embora significativamente menor, ainda é estatisticamente significativo ao nível de 1%. O efeito estimado em (c.ii) também cai por um fator de 3.

iii.

Os dados do Azerbaijão estão disponíveis apenas para o ano 2000. Esses dados são completamente absorvidos pelo efeito fixo específico do país e, portanto, não têm nenhum efeito sobre o coeficiente estimado de `log_gdppc`. Podemos confirmar esse resultado fazendo:

```
feols(dem_ind ~ log_gdppc | country,
      data     = df |> filter(country != "Azerbaijan"),
      cluster  = ~country)
```

NOTES: 411 observations removed because of NA values (LHS: 103, RHS: 403).
20 fixed-effect singletons were removed (20 observations).

```
OLS estimation, Dep. Var.: dem_ind
Observations: 937
Fixed-effects: country: 129
Standard-errors: Clustered (country)
      Estimate Std. Error t value Pr(>|t|)
log_gdppc 0.083741  0.031543 2.65478 0.0089436 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.188866      Adj. R2: 0.685534
          Within R2: 0.019681
```

iv.

A demanda por democracia é possui um efeito de contágio e se alastra entre países, como na “Primavera Árabe” de 2012. Na medida em que esses eventos coincidem com mudanças nas condições econômicas globais, eles estão correlacionados com `log_gdppc` e, portanto, podem gerar viés de variável omitida.

v.

```
feols(dem_ind ~ log_gdppc | country + year,
      data     = df,
      cluster  = ~country)
```

NOTES: 411 observations removed because of NA values (LHS: 103, RHS: 403).
21/0 fixed-effect singletons were removed (21 observations).

```
OLS estimation, Dep. Var.: dem_ind
Observations: 937
Fixed-effects: country: 129, year: 9
Standard-errors: Clustered (country)
      Estimate Std. Error t value Pr(>|t|)
log_gdppc 0.053588  0.04246 1.26208 0.20921
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.17912      Adj. R2: 0.714321
          Within R2: 0.004645
```

O coeficiente estimado cai ainda mais para 0,05, aproximadamente 1/5 do valor obtido quando se omitem os efeitos fixos de tempo e de país. A estimativa não é estatisticamente significativa ao nível de 10%.

vi.

```
feols(dem_ind ~ log_gdppc + age_1 + age_2 + age_3 + age_4 + educ + age_median |
      country + year,
      data = df,
      cluster = ~country)
```

NOTES: 690 observations removed because of NA values (LHS: 103, RHS: 689).
3/0 fixed-effect singletons were removed (3 observations).

OLS estimation, Dep. Var.: dem_ind

Observations: 676

Fixed-effects: country: 92, year: 8

Standard-errors: Clustered (country)

	Estimate	Std. Error	t value	Pr(> t)
log_gdppc	0.028764	0.053579	0.536857	0.592677
age_1	-1.170661	1.915323	-0.611208	0.542586
age_2	-1.717827	1.533917	-1.119896	0.265705
age_3	-3.507567	1.618274	-2.167475	0.032809 *
age_4	-0.895441	1.641053	-0.545650	0.586642
educ	-0.002721	0.022676	-0.120009	0.904741
age_median	-0.000940	0.026123	-0.035987	0.971372

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RMSE: 0.182175 Adj. R2: 0.682527

Within R2: 0.026989

O resultado acima inclui população, idade e educação. Em conjunto, essas variáveis não são estatisticamente significativas na regressão, embora as variáveis de idade sejam significativas ao nível de 10%. Quando essas variáveis são incluídas, o coeficiente estimado de log_gdppc cai ainda mais para 0,03 com erro padrão de 0,05.

Nota: Será que vocês conseguem colocar os cinco modelos em uma tabela como fizemos em atividades anteriores?

e. Conclusões

Com base nas estimações acima, podemos tirar as seguintes conclusões sobre os efeitos da renda sobre a democracia:

1. **Sem efeitos fixos (Modelo 1):** O coeficiente de log_gdppc é positivo e estatisticamente significativo, sugerindo uma associação positiva entre renda e democracia.

2. **Com efeitos fixos de país (Modelo 2):** Ao controlar por características não observadas e invariantes no tempo de cada país (como história, cultura, instituições), o coeficiente se altera — tipicamente diminuindo —, indicando que parte da correlação anterior era espúria.
3. **Excluindo o Azerbaijão (Modelo 3):** Caso o resultado mude de forma expressiva, isso sugere que o Azerbaijão é um *outlier* influente na estimação.
4. **Com efeitos fixos de país e tempo (Modelo 4):** Controlando também por choques globais comuns a todos os países em cada período (como crises internacionais ou ondas de democratização), a estimativa causal do efeito da renda se torna mais confiável.
5. **Com controles demográficos (Modelo 5):** A inclusão de variáveis demográficas permite isolar o efeito da renda de mudanças estruturais populacionais que podem afetar simultaneamente renda e democracia.

No conjunto, a evidência sugere que **o efeito causal da renda sobre a democracia é modesto e menos robusto** do que a correlação bruta indica, especialmente após a inclusão de efeitos fixos. Isso questiona a hipótese de que a democracia é simplesmente um “bem normal” que cresce automaticamente com a renda.