

# Taller 5 - Regresión Discontinua

Maria Camila Caraballo, Laura Sarif Rivera Sanabria

## Regresión Discontinua

En su artículo “Islamic Rule and the Empowerment of the Poor and Pious”, Meyersson (2014) investiga si la llegada al poder por parte del Partido Islámico tiene algún efecto sobre el empoderamiento de las mujeres en Turquía. Para esto, implementa la metodología de Regresión Discontinua, explotando información de: (1) elecciones locales de alcalde en Turquía del año 1994 y (2) mujeres con educación secundaria completa en el año 2000. Concretamente, estima por MCO la siguiente ecuación

$$y_i = \alpha + \beta m_i + f(x_i) + \varepsilon_i$$

donde  $y_i$  es la proporción de mujeres entre 15 y 20 años con educación secundaria completa en el año 2000,  $x_i$  es el margen de votos con el que ganó o perdió el candidato del partido islámico,  $f(\cdot)$  es un polinomio de grado  $n$  de la variable  $x_i$ ,  $m_i$  es una dicótoma que toma el valor de uno si  $x_i \geq 0$ , es decir, si el alcalde que llegó al poder en 1994 era del partido Islámico y  $\varepsilon_i$  es el término del error. La ecuación es estimada en un vecindario alrededor del corte, el cual, en este caso, es cero.

## 0. Paso inicial

Para esto se deberá eliminar aleatoriamente el 5% de las observaciones y usar la base restante. La semilla que deben usar para que sus resultados sean replicables es su código de estudiante.

```
## hischshr1520m hischshr1520f
## Min. : 0.00 Min. :0.00000
## 1st Qu.:13.77 1st Qu.:0.08513
## Median :18.75 Median :0.15482
## Mean :19.28 Mean :0.16288
## 3rd Qu.:24.07 3rd Qu.:0.22762
## Max. :68.31 Max. :0.68038
```

1) ¿Por qué el autor usa la metodología de Regresión Discontinua para identificar los efectos de interés? ¿Cuál es la intuición detrás? ¿Cuál es el supuesto de identificación?

(Cami)

2) Para cada género, presenten en una tabla los resultados de estimar las siguientes especificaciones

a. Ecuación principal, para toda la muestra, sin incluir controles.

```
##
## Call:
## lm(formula = hischshr1520f ~ i98, data = turquia)
##
## Coefficients:
## (Intercept)          i98
##    0.16581      -0.02413
```

```
##
## t test of coefficients:
##
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.1658128  0.0060063 27.6064 < 2.2e-16 ***
## i98          -0.0241346  0.0058973 -4.0925 4.404e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
## \begin{table}[!htbp] \centering
##   \caption{Efecto de Globo sobre la participación escolar}
##   \label{tab:globo_cluster}
## \begin{tabular}{@{\extracolsep{5pt}}lc}
## \hline
## \hline \hline
## & \multicolumn{1}{c}{\textit{Dependent variable:}} \hline
## \cline{2-2}
## \hline & Participación escolar \hline
## \hline
## Cobertura de TV Globo (1998) &  $-\$0.024^{***}$  \hline
## & ( $-\$0.00001$ ) \hline
## & \hline
## Constant &  $0.166^{***}$  \hline
## & ( $0.00004$ ) \hline
## & \hline
## \hline
## Observations & 2,497 \hline
##  $R^2$  & 0.007 \hline
## Adjusted  $R^2$  & 0.006 \hline
## Residual Std. Error & 0.095 (df = 2495) \hline
## F Statistic &  $17.043^{***}$  (df = 1; 2495) \hline
## \hline
## \hline
## \textit{Note:} & \multicolumn{1}{r}{ $^*p < 0.1$ ;  $^{**}p < 0.05$ ;  $^{***}p < 0.01$ } \hline
## & \multicolumn{1}{r}{Errores estándar agrupados por provincia.} \hline
## \end{tabular}
## \end{table}
```

b. Ecuación principal, para toda la muestra, con controles.

```
##
## \begin{table}[!htbp] \centering
##   \caption{Impacto de Globo sobre escolarización por género}
##   \label{tab:globo_genero}
## \begin{tabular}{@{\extracolsep{5pt}}lcc}
## \hline
## \hline \hline
## & \multicolumn{2}{c}{\textit{Dependent variable:}} \hline
## \cline{2-3}
## \hline & Escolarización femenina & Escolarización masculina \hline
## & Mujeres & Hombres \hline
## \hline & (1) & (2) \hline
## \hline
## Cobertura Globo (i98) &  $0.015^{**}$  &  $0.010^{*}$  \hline
## & ( $0.007$ ) & ( $0.006$ ) \hline
## & & \hline
## X &  $0.028^{*}$  &  $0.034^{**}$  \hline
```

```

## & (0.016) & (0.013) \\
## & & \\
## X_m & $-$0.042 & $-$0.041 \\
## & (0.049) & (0.029) \\
## & & \\
## Islam 1994 & $-$0.001$^{***}$ & $-$0.001$^{***}$ \\
## & (0.0002) & (0.0002) \\
## & & \\
## Número de partidos & $-$0.002 & $-$0.004$^{***}$ \\
## & (0.001) & (0.001) \\
## & & \\
## Población 1994 & 0.006$^{**}$ & 0.005$^{**}$ \\
## & (0.002) & (0.002) \\
## & & \\
## Edad 19 & $-$0.008$^{***}$ & $-$0.006$^{***}$ \\
## & (0.001) & (0.0005) \\
## & & \\
## Edad 60 & $-$0.004$^{***}$ & $-$0.003$^{***}$ \\
## & (0.001) & (0.001) \\
## & & \\
## Sexo & $-$0.0003$^{***}$ & $-$0.0004$^{***}$ \\
## & (0.0001) & (0.0001) \\
## & & \\
## Shhs & 0.005$^{***}$ & 0.004$^{***}$ \\
## & (0.002) & (0.001) \\
## & & \\
## Merkezi & 0.062$^{***}$ & 0.035$^{***}$ \\
## & (0.006) & (0.004) \\
## & & \\
## Merkezp & 0.041$^{***}$ & 0.027$^{***}$ \\
## & (0.009) & (0.009) \\
## & & \\
## Subbuyuk & 0.026$^{**}$ & 0.023$^{***}$ \\
## & (0.011) & (0.008) \\
## & & \\
## Buyuk & 0.023 & 0.033$^{**}$ \\
## & (0.016) & (0.015) \\
## & & \\
## factor(prov\_num)2 & $-$0.033$^{***}$ & 0.033$^{***}$ \\
## & (0.005) & (0.004) \\
## & & \\
## factor(prov\_num)3 & $-$0.075$^{***}$ & $-$0.020$^{***}$ \\
## & (0.003) & (0.002) \\
## & & \\
## factor(prov\_num)4 & $-$0.065$^{***}$ & 0.028$^{***}$ \\
## & (0.008) & (0.006) \\
## & & \\
## factor(prov\_num)5 & $-$0.052$^{***}$ & $-$0.018$^{***}$ \\
## & (0.003) & (0.003) \\
## & & \\
## factor(prov\_num)6 & $-$0.050$^{***}$ & $-$0.015$^{***}$ \\
## & (0.004) & (0.003) \\
## & & \\
## factor(prov\_num)7 & $-$0.004 & 0.001 \\
## & (0.002) & (0.002) \\
## & & \\
## factor(prov\_num)8 & $-$0.079$^{***}$ & $-$0.070$^{***}$ \\
## & (0.006) & (0.005) \\
## & & \\

```

```

## factor(prov\_num)9 & $-$0.022$^{***}$ & $-$0.040$^{***}$ \\
## & (0.006) & (0.004) \\
## & & \\
## factor(prov\_num)10 & $-$0.016$^{***}$ & $-$0.001 \\
## & (0.004) & (0.003) \\
## & & \\
## factor(prov\_num)11 & $-$0.062$^{***}$ & $-$0.085$^{***}$ \\
## & (0.004) & (0.003) \\
## & & \\
## factor(prov\_num)12 & $-$0.078$^{***}$ & $-$0.077$^{***}$ \\
## & (0.006) & (0.005) \\
## & & \\
## factor(prov\_num)13 & $-$0.070$^{***}$ & $-$0.057$^{***}$ \\
## & (0.005) & (0.004) \\
## & & \\
## factor(prov\_num)14 & $-$0.044$^{***}$ & 0.029$^{***}$ \\
## & (0.008) & (0.006) \\
## & & \\
## factor(prov\_num)15 & $-$0.065$^{***}$ & 0.034$^{***}$ \\
## & (0.006) & (0.005) \\
## & & \\
## factor(prov\_num)16 & $-$0.070$^{***}$ & $-$0.035$^{***}$ \\
## & (0.005) & (0.003) \\
## & & \\
## factor(prov\_num)17 & $-$0.097$^{***}$ & $-$0.008 \\
## & (0.008) & (0.005) \\
## & & \\
## factor(prov\_num)18 & $-$0.015 & 0.031$^{***}$ \\
## & (0.010) & (0.007) \\
## & & \\
## factor(prov\_num)19 & $-$0.009$^{*}$ & $-$0.037$^{***}$ \\
## & (0.005) & (0.004) \\
## & & \\
## factor(prov\_num)20 & $-$0.059$^{***}$ & $-$0.054$^{***}$ \\
## & (0.006) & (0.005) \\
## & & \\
## factor(prov\_num)21 & $-$0.099$^{***}$ & $-$0.081$^{***}$ \\
## & (0.004) & (0.003) \\
## & & \\
## factor(prov\_num)22 & $-$0.063$^{***}$ & $-$0.076$^{***}$ \\
## & (0.008) & (0.006) \\
## & & \\
## factor(prov\_num)23 & $-$0.036$^{***}$ & $-$0.033$^{***}$ \\
## & (0.007) & (0.005) \\
## & & \\
## factor(prov\_num)24 & $-$0.080$^{***}$ & $-$0.022$^{***}$ \\
## & (0.005) & (0.004) \\
## & & \\
## factor(prov\_num)25 & $-$0.066$^{***}$ & $-$0.062$^{***}$ \\
## & (0.005) & (0.004) \\
## & & \\
## factor(prov\_num)26 & $-$0.025$^{***}$ & 0.008 \\
## & (0.008) & (0.006) \\
## & & \\
## factor(prov\_num)27 & $-$0.048$^{***}$ & $-$0.016$^{***}$ \\
## & (0.004) & (0.003) \\
## & & \\
## factor(prov\_num)28 & $-$0.105$^{***}$ & $-$0.102$^{***}$ \\
## & (0.007) & (0.006)

```

```

## & & \\
## factor(prov\_num)29 & $-$0.029$^{***}$ & 0.050$^{***}$ \\
## & (0.005) & (0.003) \\
## & & \\
## factor(prov\_num)30 & 0.002 & 0.070$^{***}$ \\
## & (0.005) & (0.003) \\
## & & \\
## factor(prov\_num)31 & $-$0.038$^{***}$ & 0.041$^{***}$ \\
## & (0.006) & (0.004) \\
## & & \\
## factor(prov\_num)32 & $-$0.036$^{***}$ & $-$0.032$^{***}$ \\
## & (0.004) & (0.003) \\
## & & \\
## factor(prov\_num)33 & $-$0.044$^{***}$ & 0.008$^{**}$ \\
## & (0.004) & (0.003) \\
## & & \\
## factor(prov\_num)34 & $-$0.002 & 0.024$^{***}$ \\
## & (0.005) & (0.004) \\
## & & \\
## factor(prov\_num)35 & $-$0.016$^{***}$ & 0.050$^{***}$ \\
## & (0.005) & (0.004) \\
## & & \\
## factor(prov\_num)36 & $-$0.099$^{***}$ & $-$0.014$^{*}$ \\
## & (0.012) & (0.008) \\
## & & \\
## factor(prov\_num)37 & $-$0.031$^{***}$ & $-$0.002$^{*}$ \\
## & (0.002) & (0.001) \\
## & & \\
## factor(prov\_num)38 & $-$0.020$^{***}$ & $-$0.007$^{**}$ \\
## & (0.004) & (0.004) \\
## & & \\
## factor(prov\_num)39 & $-$0.001 & 0.008$^{**}$ \\
## & (0.004) & (0.004) \\
## & & \\
## factor(prov\_num)40 & $-$0.050$^{***}$ & $-$0.042$^{***}$ \\
## & (0.005) & (0.004) \\
## & & \\
## factor(prov\_num)41 & $-$0.084$^{***}$ & $-$0.093$^{***}$ \\
## & (0.005) & (0.004) \\
## & & \\
## factor(prov\_num)42 & $-$0.022$^{***}$ & 0.028$^{***}$ \\
## & (0.003) & (0.003) \\
## & & \\
## factor(prov\_num)43 & $-$0.046$^{***}$ & $-$0.049$^{***}$ \\
## & (0.004) & (0.004) \\
## & & \\
## factor(prov\_num)44 & 0.011$^{**}$ & 0.037$^{***}$ \\
## & (0.005) & (0.004) \\
## & & \\
## factor(prov\_num)45 & $-$0.012$^{**}$ & $-$0.032$^{***}$ \\
## & (0.005) & (0.004) \\
## & & \\
## factor(prov\_num)46 & $-$0.107$^{***}$ & $-$0.082$^{***}$ \\
## & (0.006) & (0.005) \\
## & & \\
## factor(prov\_num)47 & $-$0.011$^{***}$ & 0.009$^{***}$ \\
## & (0.003) & (0.003) \\
## & & \\
## factor(prov\_num)48 & $-$0.081$^{***}$ & $-$0.028$^{***}$

```

```

##      & (0.004) & (0.003) \\
##      & & \\
##      factor(prov\_num)49 & 0.0001 & 0.054$^{***}$ \\
##      & (0.003) & (0.003) \\
##      & & \\
##      factor(prov\_num)50 & $-0.075$^{***}$ & $-0.094$^{***}$ \\
##      & (0.007) & (0.005) \\
##      & & \\
##      factor(prov\_num)51 & $-0.031$^{***}$ & 0.013$^{***}$ \\
##      & (0.003) & (0.003) \\
##      & & \\
##      factor(prov\_num)52 & $-0.028$^{***}$ & $-0.036$^{***}$ \\
##      & (0.005) & (0.004) \\
##      & & \\
##      factor(prov\_num)53 & $-0.052$^{***}$ & $-0.0003$ \\
##      & (0.003) & (0.003) \\
##      & & \\
##      factor(prov\_num)54 & $-0.086$^{***}$ & $-0.034$^{***}$ \\
##      & (0.004) & (0.003) \\
##      & & \\
##      factor(prov\_num)55 & 0.024$^{***}$ & 0.058$^{***}$ \\
##      & (0.005) & (0.004) \\
##      & & \\
##      factor(prov\_num)56 & $-0.113$^{***}$ & $-0.095$^{***}$ \\
##      & (0.004) & (0.003) \\
##      & & \\
##      factor(prov\_num)57 & $-0.029$^{***}$ & 0.019$^{***}$ \\
##      & (0.008) & (0.006) \\
##      & & \\
##      factor(prov\_num)58 & 0.014$^{***}$ & 0.025$^{***}$ \\
##      & (0.002) & (0.002) \\
##      & & \\
##      factor(prov\_num)59 & $-0.052$^{***}$ & $-0.074$^{***}$ \\
##      & (0.006) & (0.005) \\
##      & & \\
##      factor(prov\_num)60 & $-0.021$^{**}$ & 0.015$^{**}$ \\
##      & (0.008) & (0.006) \\
##      & & \\
##      factor(prov\_num)61 & $-0.056$^{***}$ & $-0.033$^{***}$ \\
##      & (0.004) & (0.003) \\
##      & & \\
##      factor(prov\_num)62 & $-0.031$^{***}$ & $-0.012$^{***}$ \\
##      & (0.003) & (0.003) \\
##      & & \\
##      factor(prov\_num)63 & $-0.050$^{***}$ & $-0.029$^{***}$ \\
##      & (0.004) & (0.003) \\
##      & & \\
##      factor(prov\_num)64 & 0.050$^{***}$ & 0.075$^{***}$ \\
##      & (0.002) & (0.002) \\
##      & & \\
##      factor(prov\_num)65 & $-0.045$^{***}$ & $-0.012$^{***}$ \\
##      & (0.005) & (0.004) \\
##      & & \\
##      factor(prov\_num)66 & $-0.067$^{***}$ & $-0.027$^{***}$ \\
##      & (0.004) & (0.003) \\
##      & & \\
##      factor(prov\_num)67 & $-0.022$^{***}$ & $-0.006$^{***}$ \\
##      & (0.002) & (0.002) \\
##      & & \\
##      & & \\

```

```

## factor(prov\_num)68 & $-0.045$^{***}$ & 0.014$^{***}$ \\
## & (0.007) & (0.005) \\
## & & \\
## factor(prov\_num)69 & $-0.065$^{***}$ & 0.019$^{***}$ \\
## & (0.008) & (0.006) \\
## & & \\
## factor(prov\_num)70 & $-0.060$^{***}$ & $-0.019$^{***}$ \\
## & (0.005) & (0.004) \\
## & & \\
## factor(prov\_num)71 & $-0.060$^{***}$ & $-0.001$ \\
## & (0.010) & (0.007) \\
## & & \\
## factor(prov\_num)72 & $-0.00004$ & 0.020$^{***}$ \\
## & (0.005) & (0.004) \\
## & & \\
## factor(prov\_num)73 & $-0.079$^{***}$ & $-0.088$^{***}$ \\
## & (0.005) & (0.004) \\
## & & \\
## factor(prov\_num)74 & $-0.048$^{***}$ & $-0.003$ \\
## & (0.004) & (0.003) \\
## & & \\
## factor(prov\_num)75 & 0.020$^{***}$ & 0.064$^{***}$ \\
## & (0.004) & (0.004) \\
## & & \\
## factor(prov\_num)76 & 0.009 & $-0.025$^{***}$ \\
## & (0.008) & (0.005) \\
## & & \\
## factor(prov\_num)77 & $-0.050$^{***}$ & $-0.028$^{***}$ \\
## & (0.005) & (0.004) \\
## & & \\
## factor(prov\_num)78 & $-0.068$^{***}$ & 0.014$^{**}$ \\
## & (0.010) & (0.006) \\
## & & \\
## factor(prov\_num)79 & 0.013$^{**}$ & $-0.022$^{***}$ \\
## & (0.006) & (0.005) \\
## & & \\
## factor(prov\_num)80 & 0.001 & 0.048$^{***}$ \\
## & (0.003) & (0.002) \\
## & & \\
## factor(prov\_num)81 & $-0.069$^{***}$ & $-0.074$^{***}$ \\
## & (0.004) & (0.003) \\
## & & \\
## Constant & 0.549$^{***}$ & 0.506$^{***}$ \\
## & (0.040) & (0.033) \\
## & & \\
## \hline \\[-1.8ex]
## Observations & 2,497 & 2,497 \\
## R$^{2}$ & 0.554 & 0.411 \\
## Adjusted R$^{2}$ & 0.536 & 0.388 \\
## Residual Std. Error (df = 2402) & 0.065 & 0.061 \\
## F Statistic (df = 94; 2402) & 31.711$^{***}$ & 17.803$^{***}$ \\
## \hline
## \hline \\[-1.8ex]
## \textit{Note:} & \multicolumn{2}{r}{\textit{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01}} \\
## & \multicolumn{2}{r}{\textit{Errores estándar agrupados por provincia.}} \\
## \end{tabular}
## \end{table}

```

c. Ecuación principal, para la submuestra a h unidades alrededor del corte, con controles.

- d. Ecuación principal, para la submuestra a  $h/2$  unidades alrededor del corte, con controles

donde  $h=0.24$  es el ancho de banda óptimo estimado por los autores. Para las especificaciones con controles, supongan que:

$$f(x_i) = \gamma x_i + \delta x_i \times m_i$$

Es decir, un polinomio de grado uno con pendiente distinta a cada lado del corte. Para las especificaciones sin controles, no incluyan ningún polinomio. Todas las especificaciones deben usar errores estándar clúster a nivel de provincia.

### 3) A partir de los resultados encontrados en el anterior punto, respondan:

- ¿Por qué cambian los coeficientes entre especificaciones?
- ¿Cuál parece ser el impacto de la llegada al poder del Partido Islámico para las mujeres? ¿Parece ser este impacto robusto a las especificaciones?

Cami

### 4) Finalmente, presenten evidencia a favor (o en contra) del supuesto de identificación. Para esto,

- Presenten en una gráfica la distribución kernel o el histograma de la variable de asignación  $x_i$ . ¿Parece haber manipulación?

```
“{r}1 # Crear el gráfico de densidad (curva kernel) grafico_kernel <- ggplot(turquia, aes(x = X)) +  
geom_density(fill = “blue”, alpha = 0.5) + # Curva kernel geom_vline(xintercept = 0, color = “red”,  
linetype = “dashed”, linewidth = 1) + labs( title = “Distribución de la variable de asignación (X) - Kernel”, x =  
“Variable de asignación (X)”, y = “Densidad” ) + theme_minimal()
```

## Mostrar en consola

grafico\_kernel

## Guardar el gráfico

```
ggsave(“Tablas/curva_kernel_X.png”, plot = grafico_kernel, width = 8, height = 5, dpi = 300)
```

““

- En una tabla presenten los resultados de estimar la ecuación de interés, sin controles, pero incluyendo  $f(x_i)$ , tomando como variables dependientes: i) la elección de un alcalde del partido Islámico en 1984 ( $i89$ ) y ii) el logaritmo de la población en 1994 ( $lpop1994$ ). Para esto, usen únicamente la muestra de elecciones alrededor del ancho de banda óptimo. Dados sus resultados, ¿parece haber continuidad en estas variables?
- Dados sus resultados en los incisos a y b, ¿es plausible el supuesto de identificación? Expliquen por qué. Cami