

Proyecto 2 econometría aplicada y ciencia de datos

Sergio Sarmiento Hernández

2025-10-03

1 Modelo propuesto:

El presente modelo busca estimar la diferencia en la proporción promedio de aprobación entre empleados y no empleados, controlando por X_{it}

$$y_{it} = \alpha_i + \beta_1 empleo_{it} + BX_{it} + \epsilon_{it}$$

Donde, el vector de controles X_{it} incluye variables como, si la persona i en el periodo t está casada y el número de hijos menores a 18 años de la persona i en el periodo t .

El modelo presentado anteriormente será estimado a partir de 3 estimadores diferentes. El primer modelo es el dado por el estimador pooled (o agrupado), dicho estimador nos ayudaría a comparar directamente empleados con no empleados en toda la muestra, como si fuera un solo corte transversal. Por otro lado, el segundo modelo será estimado por efectos aleatorios (RE), usar este estimador nos permitirá aprovechar tanto la variación entre individuos como la variación dentro de los individuos en el tiempo, de esta forma obendríamos la relación entre empleo y aprobación al presidente de manera más eficientes, siempre y cuando se cumplan los supuestos de dicho estimador. Por último, se usará el estimador de efectos fijos (FE), este estimador nos ayudará a observar cómo cambia la aprobación presidencial de una misma persona cuando cambia su estatus laboral en el tiempo, al eliminar el sesgo por características individuales constantes (α_i), lo que permitirá tener un efecto más limpio del empleo sobre la aprobación.

La creación de las variables de interés y_{it} (aprobación del presidente) y $empleo_{it}$ están determinadas de la siguiente manera. Por un lado, la aprobación del presidente fue tomada

como una variable binaria, donde $y_{it} = 1$ si el individuo i en el año t respondió “Strongly Approve” o “Somewhat Approve” en el momento de la encuesta, y_{it} es igual a cero en cualquier otro caso. De igual forma, la variable empleo es una variable binaria, donde $empleo_{it}$ es igual a uno si la persona respondió que estaba empleado a tiempo completo (FT) o a tiempo parcial (PT) en el año t , en cualquier otro caso es cero.

2 Gráficas de estadísticas descriptivas:

Por un lado, se presenta primero la gráfica de la proporción de empleo en los años 2010, 2012 y 2014:

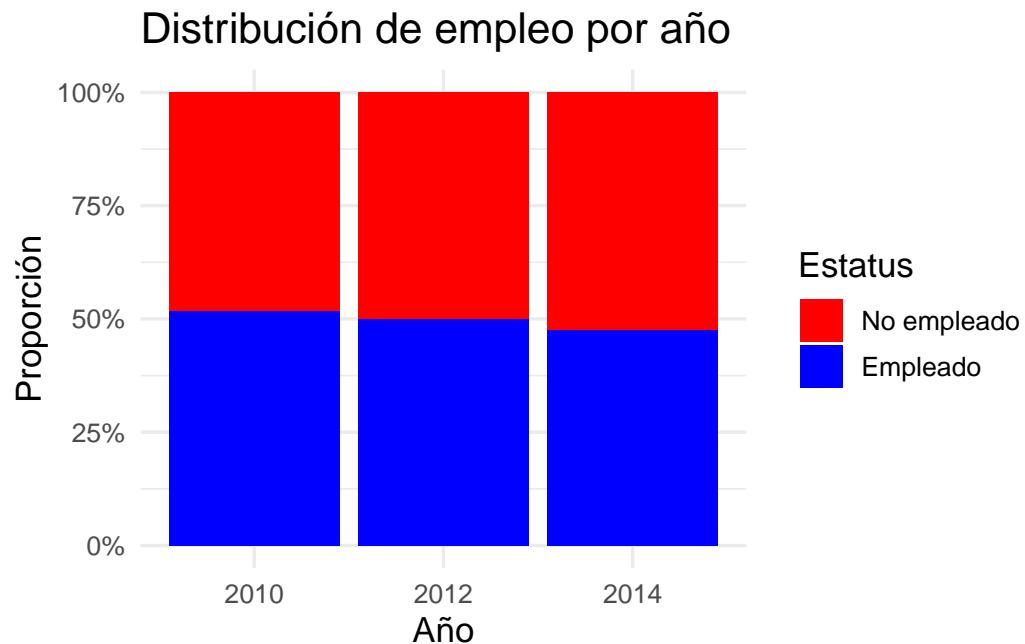


Figura 1: Distribución de empleo por año

Esta primera gráfica muestra que de la población encuestada durante estos tres años, hay una proporción similar de personas empleadas y no empleadas. De igual forma, es visible que a lo largo de este periodo de tres años la población empleada disminuyó ligeramente. Además, la composición es bastante estable: en 2010 aproximadamente 5 de cada 10 reportan empleo de tiempo completo o parcial y para 2014 esa proporción baja ampliando levemente la brecha. En suma, la proporción de empleados y no empleado es bastante pareja, con algunas diferencias durante los tres años.

Porcentaje de aprobación presidencial según estatus laboral

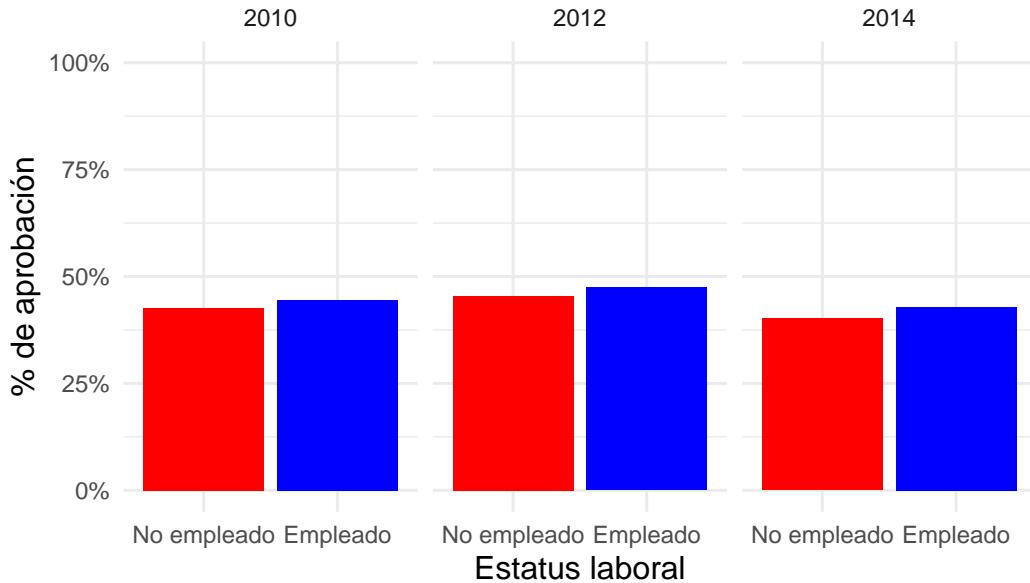


Figura 2: Porcentaje de aprobación presidencial según empleo

A lo largo de los tres años en los que aconteció la encuesta, las personas empleadas muestran una mayor proporción de aprobación que los no empleados, aunque la brecha es pequeña, solo unos puntos porcentuales. La aprobación sube ligeramente de 2010 a 2012 y luego baja en 2014 para ambos grupos. Esto sugiere una asociación positiva pero pequeña entre estar empleado y aprobar al presidente. Aunque en las gráficas podemos ver una relación positiva, es necesario realizar las estimaciones pertinentes para poder obtener conclusiones con fundamentos estadísticos. Más adelante será expuesto si esta relación es positiva o negativa.

3 Estimaciones y resultados

	Pooled OLS	Efectos aleatorios	Efectos fijos
Empleado	-0.023 (0.016)	-0.020 (0.013)	-0.015 (0.015)
Casado	-0.113*** (0.020)	-0.025 (0.019)	0.031 (0.024)
Hijos menores a 18	-0.034*** (0.006)	-0.006 (0.009)	0.008 (0.012)
Num.Obs.	4444	4444	4444
R2	0.015	0.020	0.002

* p < 0.1, ** p < 0.05, *** p < 0.01

3.1 Modelo pooled

A partir del estimador pooled obtenemos un estimador $\hat{\beta} = -0.023$ no significativo estadísticamente. Esto es posible interpretarlo como que estar empleado (FT o PT) se asocia con -2.3 puntos porcentuales en la aprobación del presidente respecto a las personas no empleadas. Dicho estimador es consistente bajo el supuesto de exogeneidad estricta $E[\epsilon_{it} | empleo_{it}, X_{it}]$ y que el efecto individual no esté correlacionado con los regresores. Por último, los errores estándar usados son robustos a heterocedasticidad.

3.2 Efectos aleatorios

Por otro lado, el estimador de efectos aleatorios nos da un estimador $\hat{\beta} = -0.02$ sin significancia estadística. La interpretación de dicho estimador es que estar empleado a tiempo completo o a medio tiempo se asocia, en promedio, con una disminución de -2 p.p. en la aprobación del presidente respecto a las personas no empleadas. Suponemos que tanto la heterogeneidad no observada es una variable aleatoria independiente del $empleo_{it}$, al igual que, el supuesto sobre la heterogeneidad no observada con el error son iid ($\alpha_i \sim (\alpha, \sigma_\alpha^2)$ y $\epsilon_{it} \sim (0, \sigma_\epsilon^2)$). En este caso, los errores estándar usados en esta estimación son robustos a heterocedasticidad y clusterizados a nivel individuo, esto debido a que el error es compuesto $u_{it} = \alpha_i + \epsilon_{it}$, por lo que las observaciones del mismo individuo están correlacionadas a través de α_i , al igual que ϵ_{it} puede ser heterocedástico y tener autocorrelación.

3.3 Efectos Fijos

Por último, el modelo estimado por efectos fijos (FE) nos da una estimador $\hat{\beta} = -0.01$ no significativo. De esta forma, la interpretación de dicho estimador es que estar empleado a tiempo completo o a medio tiempo se asocia, en promedio, con una disminución de -1.5 p.p. en la aprobación del presidente, respecto a las personas no empleadas. Los supuestos claves para FE son, primero que los efectos individuales están potencialmente correlacionados con el $empleo_{it}$. De igual forma, $\hat{\beta}$ es consistente si los α_i son efectos fijos del individuo i y ϵ_{it} es iid. De la misma forma que en la estimación por efectos aleatorios, usamos errores estándar robustos a heterocedasticidad y clusterizados a nivel individuo, debido a que $\hat{\epsilon}_{it}$ pueden seguir mostrando autocorrelación intra-persona, junto con heterocedasticidad entre individuos.