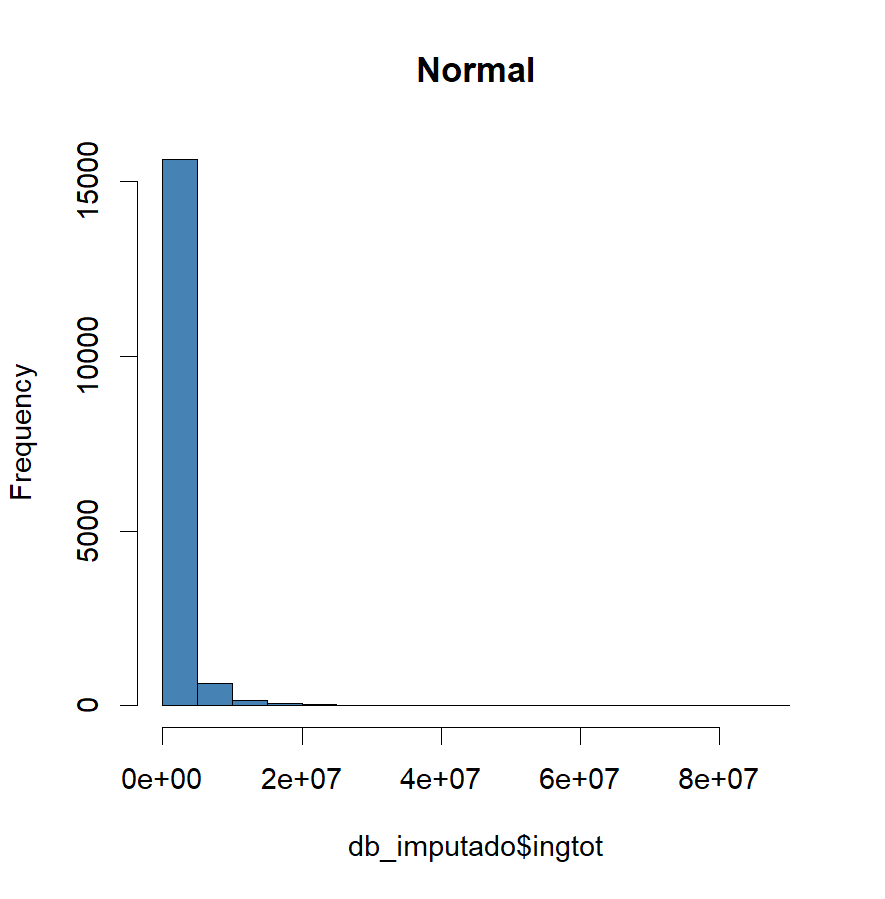
**3. Age-earnings profile.**

A great deal of evidence in Labor economics suggests that the typical worker’s age-earnings profile has a predictable path: Wages tend to be low when the worker is young; they rise as the worker ages, peaking at about age 50; and the wage rate tends to remain stable or decline slightly after age 50.

• **In the data set, multiple variables describe income. Choose one that you believe is the most representative of the workers’ total earnings, justifying your selection.**

La variable escogida para representar los ingresos totales de los trabajadores fue “ingtot” y representa el ingreso total de los mismos. En primer lugar, se escogió esta variable ya que, a comparación de las demás variables relacionadas con los ingresos, representa los ingresos totales de los individuos pues otras variables representan ingresos parciales o fraccionados del ingreso total.

Chart, scatter chart

Description automatically generated

|  |
| --- |
| Min. 1st Qu. Median Mean 3rd Qu. Max.  15.000 800000 1051160 1769379 1723158 85833333 |

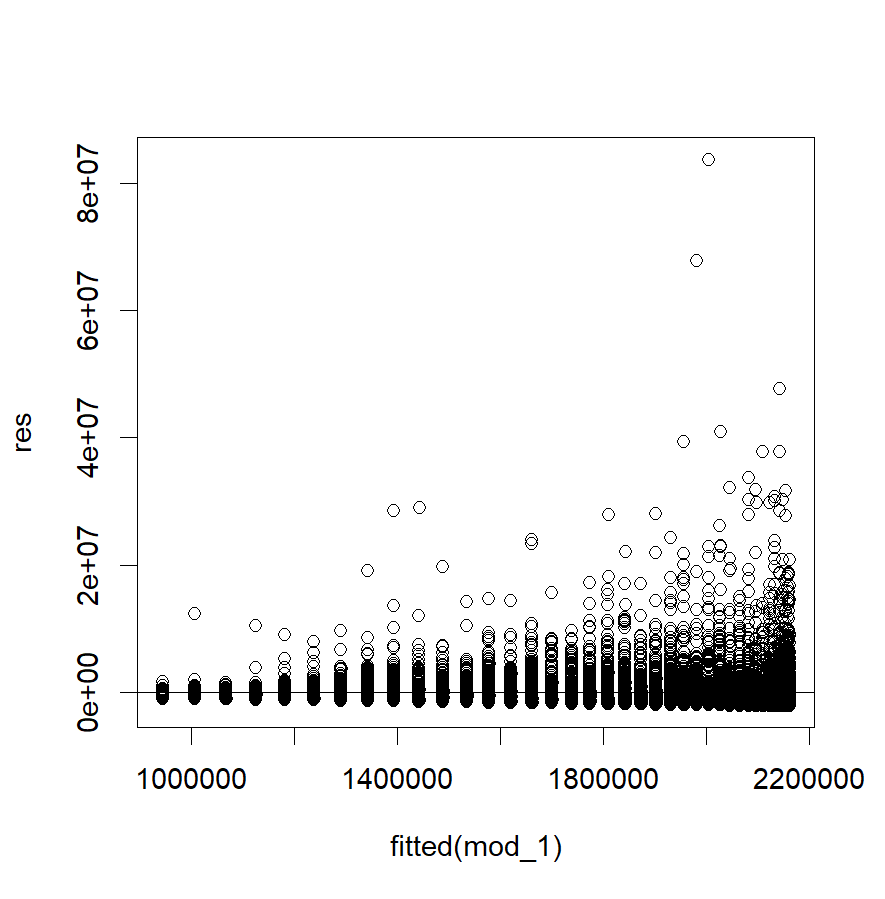
• **Based on this estimate using OLS the age-earnings profile equation:**

**Income = β1 + β2Age+ β3Age2 + u**

|  |
| --- |
| **Resultados regresión ingresos-edad**  Dependent variable: ingtot  **age 91,143.460\*\*\***  **(8,886.416)**  **age2 -799.261\*\*\***  **(102.852)**  **Constant -436,662.900\*\***  **(178,347.200)**  -----------------------------------------------------------------  Observations 16,542  R2 0.017  Adjusted R2 0.017  Residual Std. Error 2,652,732.000 (df = 16539)  F Statistic 144.382\*\*\* (df = 2; 16539) |

**Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01**

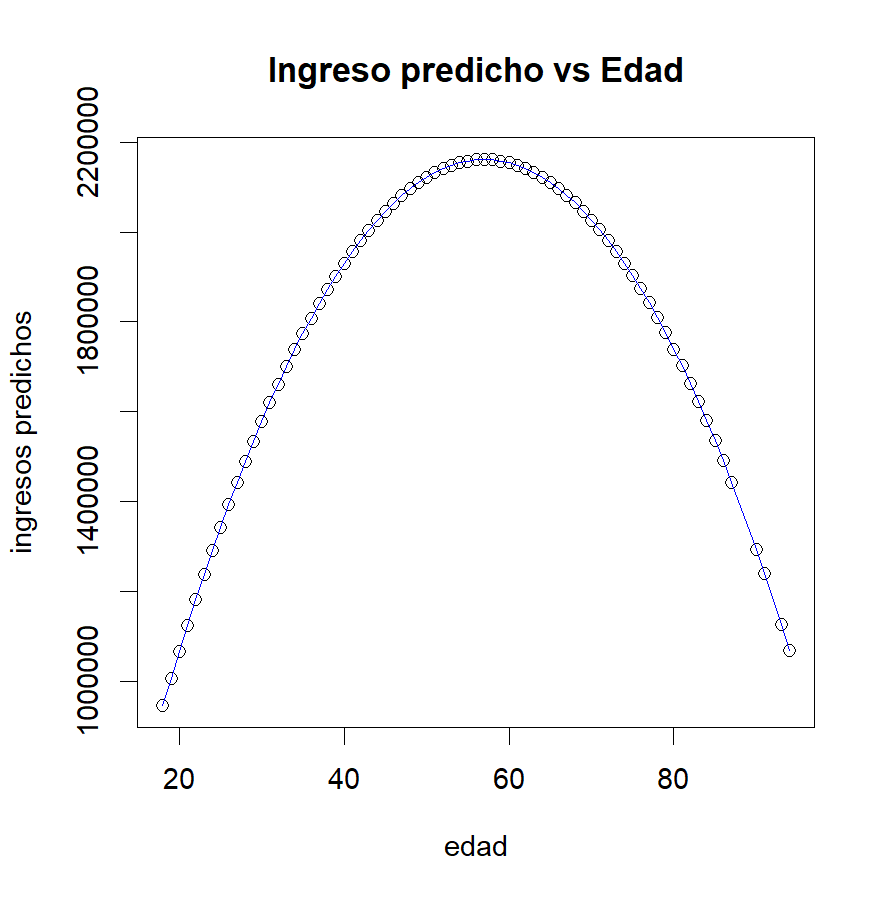
**• How good is this model in sample fit?**

**Chart, histogram

Description automatically generated**

Para poder saber qué tan bueno es este modelo en el ajuste de la muestra, se realizó una gráfica de los residuales frente a los valores predichos del modelo. En esta gráfica se puede observas que el modelo ***Income = β1 + β2Age+ β3Age2 + u*** no se ajusta del todo bien a la muestra puesto que existen múltiples valores

**• Plot the predicted age-earnings profile implied by the above equation.**

****

**• What is the “peak age” suggested by the above equation? Use bootstrap to calculate the standard errors and construct the confidence intervals.**

**4. The earnings GAP**

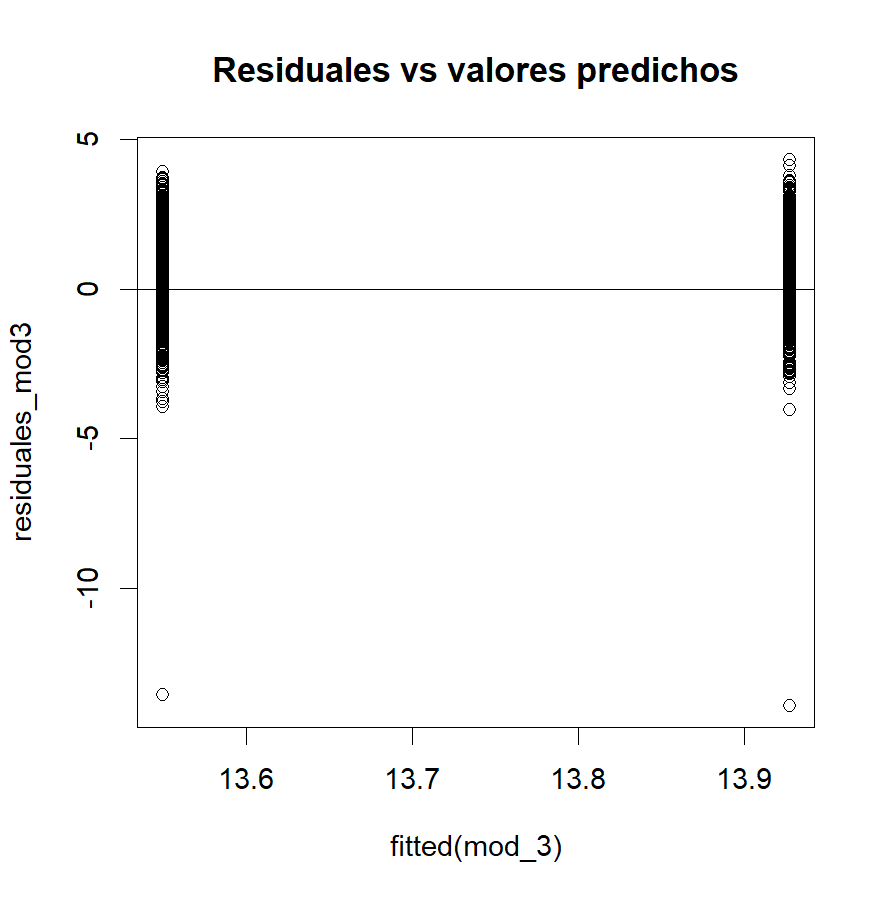
Most empirical economic studies are interested in a single low dimensional parameter, but determining that parameter may require estimating additional “nuisance” parameters to estimate this coefficient consistently and avoid omitted variables bias. Policymakers have long been concerned with the gender earnings gap.

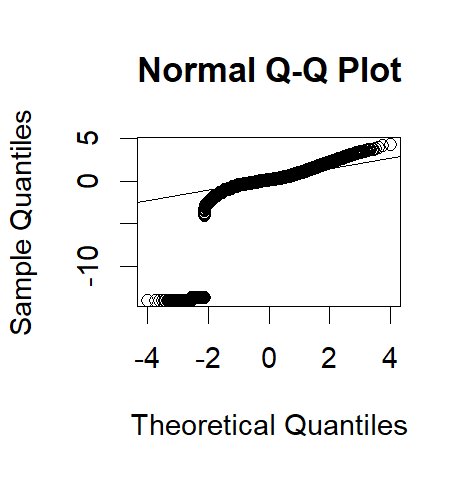
**• Estimate the unconditional earnings gap: log(Income) = β1 + β2Female + u**

|  |
| --- |
| **Regresión Log ingresos-mujer**    Dependent variable: log\_ingreso  -----------------------------------------------  **mujer -0.378\*\*\***  **(0.030)**    **Constant 13.927\*\*\***  **(0.021)**  -----------------------------------------------  Observations 16,542  R2 0.009  Adjusted R2 0.009  Residual Std. Error 1.950 (df = 16540)  F Statistic 154.943\*\*\* (df = 1; 16540) |

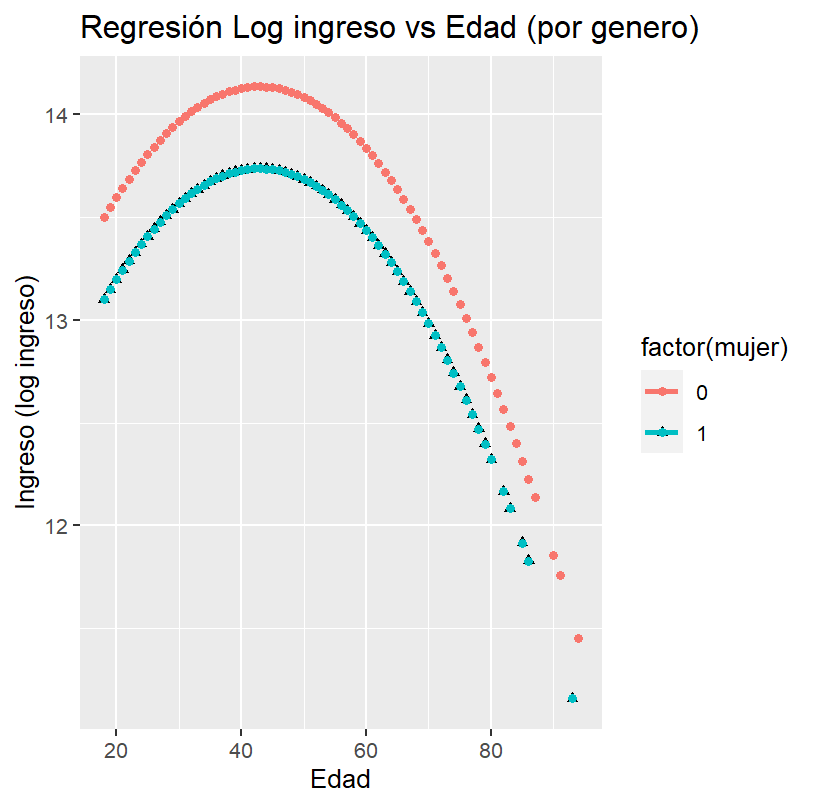
**Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01**

**• How should we interpret the β2 coefficient? How good is this model in sample fit?**

****

****

**• Estimate and plot the predicted age-earnings profile by gender. Do men and women in Bogotá have the same intercept and slopes?**

****

**• What is the implied “peak age” by gender? Use bootstrap to calculate the standard errors and construct the confidence intervals. Do these confidence intervals overlap?**

• Equal Pay for Equal Work? A common slogan is “equal pay for equal work”. One way to interpret this is that for employees with similar worker and job characteristics, no gender earnings gap should exist. Estimate a conditional earnings gap that incorporates control variables such as similar worker and job characteristics (X).

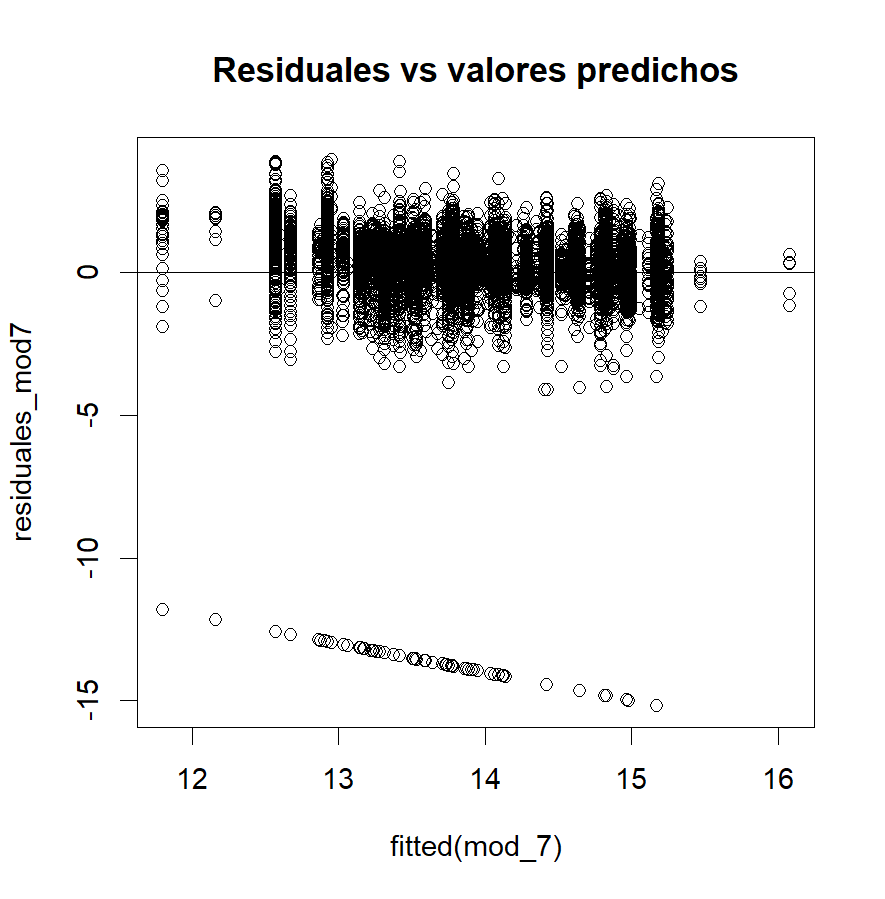
1. **Estimate the conditional earnings gap log(Income) = β1 + β2Female + θX + u**

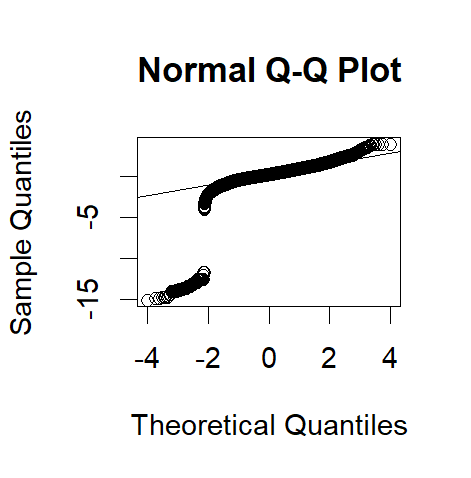
|  |
| --- |
| **Regresión Log ingresos-mujer-oficio**  Dependent variable: log\_ingreso  -----------------------------------------------  **mujer -0.358\*\*\***  **(0.034)**  **.**  **.**  **.**  **.**  **.**  **Constant 15.208\*\*\***  **(0.423)**  -----------------------------------------------  Observations 16,542  R2 0.117  Adjusted R2 0.113  Residual Std. Error 1.844 (df = 16461)  F Statistic 27.391\*\*\* (df = 80; 16461) |

**Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01**

**(b) Use FWL to repeat the above estimation, where the interest lies on β2. Do you obtain the same estimates?**

**(c) How should we interpret the β2 coefficient? How good is this model in sample fit? Is the gap reduced? Is this evidence that the gap is a selection problem and not a “discrimination problem”?**

****

****