**Tabla 1. Correlaciones entre las variables**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | ingreso | edad | escolaridad | educ\_madre | educ\_padre |
| ingreso | 1 | -0,02 | 0,45 | 0,30 | 0,33 |
| edad | -0,02 | 1 | -0,30 | -0,36 | -0,31 |
| escolaridad | 0,45 | -0,30 | 1 | 0,49 | 0,51 |
| educ\_madre | 0,30 | -0,36 | 0,49 | 1 | 0,73 |
| educ\_padre | 0,33 | -0,31 | 0,51 | 0,74 | 1 |

En primer término, se observan **correlaciones moderadas** entre las variables independientes y la variable dependiente. Esto sería un primer indicador para **repensar la especificación** del modelo de relaciones entre variables. Un segundo elemento a destacar es la r**elación elevada entre tres de las variables independientes**. El nivel de escolaridad de la persona entrevistada está bastante relacionado con el nivel educativo de la madre y del padre; asimismo, la más alta correlación se observa entre las variables que expresan el nivel educativo de la madre y del padre. Ello indica una “medición superpuesta” que tentativamente mermará la capacidad explicativa general del modelo.

**Tabla 2. Comparación de los modelos especificados**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Modelo 1** | **Modelo 2** | **Modelo 3** | **Modelo 4** |
| *Constante* | -449.795,88\*\*\* | -898.512,04\*\*\* | -1.006.107,51\*\*\* | -1.049.197,37\*\*\* |
| (16.023,19) | (28.681,81) | (28.685,04) | (29.061,78) |
| *Escolaridad* | 88.496,02\*\*\* | 95.741,84\*\*\* | 80.568,90\*\*\* | 78.483,45\*\*\* |
| (1.237,38) | (1.286,00) | (1.427,25) | (1.444,85) |
| *Edad* | - | 7.841,78\*\*\* | 9.690,56\*\*\* | 10.350,64\*\*\* |
| - | (417,33) | (419,52) | (425,63) |
| *Educación padre* | - | - | 54.805,77\*\*\* | 38.198,44\*\*\* |
| - | - | (2.357,38) | (3.035,08) |
| *Educación madre* | - | - | - | 28.256,55\*\*\* |
| - | - | - | (3.261,56) |
| *R2* | 0,20 | 0,21 | 0,23 | 0,24 |
| *R2 Ajustado* | 0,20 | 0,21 | 0,23 | 0,24 |
| *Num. obs.* | 20.365 | 20.365 | 20.365 | 20.365 |
| \*\*\*p < 0,001; \*\*p < 0,01; \*p < 0,05 | | | | |

Un primer elemento relevante es la **poca capacidad explicativa** que tienen las diferentes variantes del modelo. Se parte con sólo un 20% de varianza explicada, y se logra solamente un 24% en la versión con la mayor cantidad de variables (R2 ajustado). También destaca la **centralidad de la variable escolaridad**, en tanto explica por sí sola un 20% de la varianza de la variable ingreso. El **resto de las variables**, incluso combinadas, sólo aumentan en un 4% la varianza explicada de la variable independiente: tienen poca relevancia para el modelo, en términos estadísticos.

Un segundo elemento relevante es la **significatividad estadística de todos los coeficientes** estimados, en todos los modelos (p < 0,001 en todos los casos). Es decir, las relaciones encontradas son significativas en términos estadísticos (y no azarosas).

Un tercer elemento relevante es el **formato de la interpretación de los resultados**. Para orientar dejo dos formatos, uno más extenso y otro más sintético ajustado al modelo APA (ambos refieren al modelo 4, que incluye todas las variables).

* Por cada año adicional de escolaridad, el ingreso derivado de la ocupación principal aumenta en promedio en $78.483, *controlando por* la edad de la persona, la educación del padre y la educación de la madre. Este efecto es significativo al 99% de confianza, por lo que podemos inferir que la escolaridad se asocia positivamente a los ingresos por trabajo de la ocupación en la población de chilenos/as.
* La escolaridad de las personas tiene un efecto positivo y significativo en el ingreso derivado de la ocupación principal, controlando por la edad de la persona, la educación del padre y la educación de la madre (β=$78.483,45; p<0,01).

Un tercer elemento relevante es de carácter analítico o teórico. ¿Es la **especificación de las variables indicadas suficiente para explicar la variabilidad del ingreso** en Chile? Parece que no, seguramente estamos dejando elementos afuera del modelo.

Notas:

1. El **estadístico F** que se observa en los otros resultados (en R) indica una prueba de hipótesis sobre la disminución de la suma de los términos del error, busca determinar si esta disminución es o no significativa.
2. El elemento que se encuentra entre paréntesis en la tabla 2 es el **valor de la prueba T, de significación estadística** de cada coeficiente. En general esto no se interpreta, en la medida que interesa su si este resultado es o no significativo, privilegiándose el valor “p”.