

# El Efecto de la Educación en la Pobreza

Vargas, Matías Tomás

Febrero, 2022

**Abstracto:** El objetivo principal de éste trabajo es estimar el efecto de la educación sobre la pobreza en Argentina. La base de datos utilizada es la encuesta permanente de hogares (EPH) del segundo trimestre de 2021 brindada por INDEC. Se utiliza una regresión logística con la probabilidad de ser pobre como variable dependiente y educación (representada por variables dummies), edad y cantidad de horas trabajadas en el oficio principal como variables independientes. Para corregir la selección de muestra se utiliza el método de 2 pasos de Heckman. Los resultados muestran que existe una relación negativa entre educación, edad y cantidad de horas trabajadas con la pobreza. También se concluye que la educación disminuye en mayor medida las probabilidades de ser pobre en una mujer que en un hombre.

## **I. INTRODUCCIÓN:**

La pobreza es una cuestión socio-económica de importancia crítica en el estudio del desarrollo económico de un país. La literatura empírica identifica varios factores que dan lugar a la pobreza como puede ser la urbanización, el nivel de desarrollo de un país, el régimen político, la salud, el sexo y situación civil de una persona entre otros. Si bien todos estos factores son importantes, la educación suele ser el factor más determinante de la pobreza. Dicho componente se encuentra inversamente relacionado con la pobreza, es por ello que a menudo se asocia a la última con personas de bajo nivel educativo. Cuando una persona accede al sistema educativo, la misma adquiere más conocimientos y habilidades que le permiten generar mejores oportunidades de trabajo, y en consecuencia recibir salarios más altos que los sitúa por encima de la línea de pobreza teniendo efectos positivos en el bienestar general de un país. Dicho esto, podemos apreciar porqué es tan importante la inversión en educación y formación de capital humano para el crecimiento de un país. Además, la pobreza puede ser una limitación importante en la calidad de la enseñanza de una institución, ya que puede deteriorar los estándares educativos como así también puede afectar la mentalidad de aquellas personas de bajos recursos en el proceso de aprendizaje por diversas índoles.

La pobreza es un gran problema en Argentina. En el primer semestre de 2021, alrededor del 40.6% de la población estaban por debajo de la línea de pobreza en Argentina. La gran mayoría de las personas que viven por debajo de la línea de pobreza en la Argentina no tienen educación y están desempleadas. Como los recursos económicos de cualquier gobierno son escasos, es importante la correcta asignación de los mismos. Si bien se puede observar un gran porcentaje del presupuesto 2021 de Argentina en “Educación y Cultura”<sup>1 2</sup> en relación a otras funciones, es importante saber cuál nivel educativo tiene mayor impacto reductor de pobreza para la correcta sub-asignación de recursos y de éste modo, reducir la tasa de pobreza que se encuentra a niveles históricos muy altos y a nivel regional por encima del promedio.

Varios estudios analizan los efectos de la educación en la pobreza a nivel individual, mientras que otros lo hacen a nivel agregado (nacional). En éste trabajo se abordará el problema desde el primer punto de vista (microeconómico) mediante el uso de datos transversales para crear modelos de regresión logística.

Mi hipótesis es que a mayor nivel educativo alcanzado por una persona, menor es la probabilidad de la misma de ser pobre. También vamos a estimar los impactos por

---

<sup>1</sup>

[https://www.economia.gob.ar/onp/documentos/presutexto/proy2021/ley/pdf/planillas\\_anexas/capitulo1/anexa103.pdf](https://www.economia.gob.ar/onp/documentos/presutexto/proy2021/ley/pdf/planillas_anexas/capitulo1/anexa103.pdf)

<sup>2</sup>

[https://www.economia.gob.ar/onp/documentos/presutexto/proy2021/ley/pdf/anexo\\_estadistico/capitulo1/cap1pla7.pdf](https://www.economia.gob.ar/onp/documentos/presutexto/proy2021/ley/pdf/anexo_estadistico/capitulo1/cap1pla7.pdf)

género. Esto nos indicará si existen disparidades en los retornos en la educación entre mujeres y hombres en diferentes niveles educativos en el mercado laboral Argentino<sup>3</sup>.

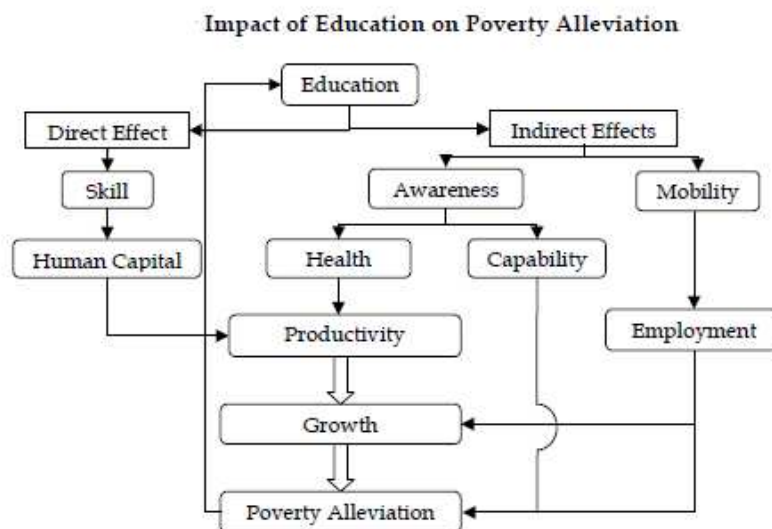
Otros factores que hemos incluido en nuestros datos de muestra que podrían afectar la pobreza incluyen la cantidad de horas trabajadas por semana de un individuo, la edad y estado civil.

## II. REVISIÓN DE LITERATURA:

La relación entre la pobreza y educación ha sido estudiada en gran medida utilizando distintas métricas o variables para las mismas. Los siguientes papers han sido revisados ya que han sido de suma importancia para éste proyecto, cada uno aportando un aspecto único en cuanto a la relación que estamos investigando.

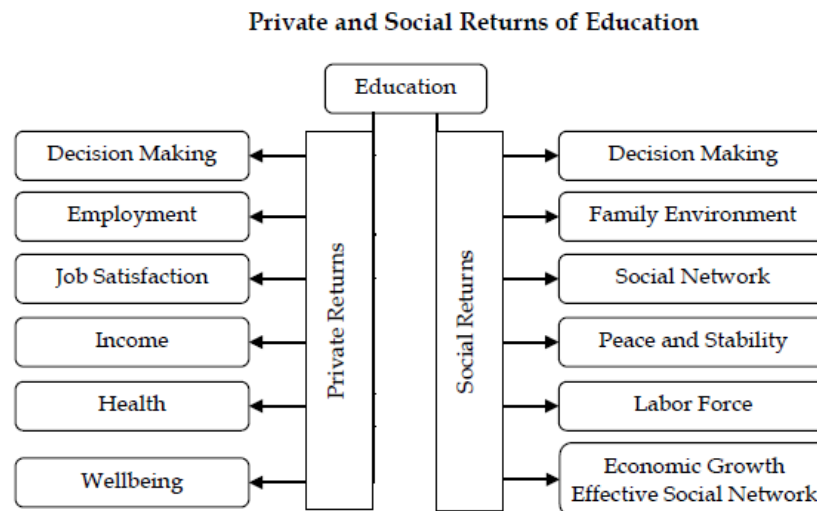
Janjua y Kamal (2011) usan datos de panel para 40 países en desarrollo en el período desde 1999 hasta 2007. Utilizan la técnica de mínimos cuadrados generalizados (GLS) para estimar los coeficientes. En éste paper ambos autores se enfocan en los efectos de la educación, tanto directos como indirectos, en la pobreza. Janjua y Kamal citan a Berg (2008) remarcando que la relación entre la pobreza y educación se puede explicar a través de 3 mecanismos:

- 1- Las personas más educadas ganan más.
- 2- Más (y especialmente mejor calidad) de educación mejora el crecimiento económico y en consecuencia oportunidades económicas e ingresos.
- 3- La educación trae asociada beneficios sociales que mejoran el desarrollo económico, y especialmente la situación de los más carenciados, como menor tasa de fertilidad, más salud a los niños y mayor participación de las mujeres en la fuerza laboral.



<sup>3</sup> <https://www.infobae.com/economia/2021/10/01/pobreza-como-se-encuentra-la-argentina-en-comparacion-con-el-resto-de-los-paises-de-la-region/>

Los autores también remarcan que los retornos a la educación son ambos, privados y sociales. Los privados son aquellos beneficios que reciben los individuos por escolaridad adicional, y los sociales son los beneficios que recibe la sociedad por la escolarización de un individuo.



La conclusión del estudio es que el crecimiento del ingreso per cápita juega un papel moderadamente positivo en la reducción de la pobreza, la reducción de la inequidad en el ingreso juega un papel más importante en la reducción de la pobreza en aquellos países con ingreso per cápita más altos, y por último, la educación secundaria es el principal contribuyente a la reducción de la pobreza.

Ahmad (2019) investiga la relación de largo y corto plazo entre la pobreza y la tasa de alfabetización en Pakistán. La fuente de los datos es el Banco Mundial y utiliza series de tiempo que abarcan el periodo desde 1971 hasta 2013. Un estudio citado por el autor, Arnall (2010), sugiere que la educación es uno de los indicadores más importante de la tasa de pobreza según varios estudios históricos y afirma que la pobreza no se puede erradicar por completo de ningún país a menos que y hasta que se controle la tasa de analfabetismo infantil. El resultado muestra la relación a largo plazo entre la pobreza y la tasa de alfabetización y explica que la pobreza se puede reducir en un 0,759% aumentando la tasa de alfabetización en un 1% en el país. En cuanto la relación a corto plazo entre la pobreza y la tasa de alfabetización, la misma no es significativa.

Njong (2010) evalúa el impacto de los diferentes niveles de educación de las personas empleadas como determinantes de la pobreza en Camerún. Los datos utilizados por el autor provienen de la Encuesta de Hogares de Camerún. Se estima un modelo de regresión logística a partir de los datos transversales, con la probabilidad de que un individuo sea pobre como variable dependiente y un conjunto de niveles educativos y experiencia como variables explicativas. Los resultados muestran a mayor experiencia y logros educativos menor es la probabilidad de ser pobre del individuo empleado. Desde el punto de vista del género, el estudio concluye que el nivel educativo de un hombre reduce más la pobreza que el de una mujer.

De Silva y Sumarto (2015) realizan un estudio a nivel sub-nacional con datos de panel durante el periodo 2002-2012, un modelo de crecimiento neoclásico (o también conocido como modelo de crecimiento de Solow) modificado y un estimador de panel dinámico para investigar el efecto de la salud y educación sobre el crecimiento económico y la pobreza a nivel de distrito en Indonesia. La investigación se hace a nivel sub-nacional ya que hay bastante disparidad en el acceso y calidad de la educación y salud entre las distintas provincias del país. Los hallazgos por los autores son relevantes para comprender el papel de la salud y educación en el crecimiento económico y en la reducción de la pobreza de Indonesia. La conclusión a la que arriban es que la formación literaria tiene un efecto grande y estadísticamente significativo en la tasa de crecimiento del ingreso per cápita. Los distritos con niveles de más baja educación están caracterizados por tener los niveles más altos de pobreza. Hallazgos del estudio revelan que las relaciones entre la salud con la pobreza y el crecimiento económico de largo plazo son más fuertes de lo que se suele dimensionar. Los resultados indican que las disparidades en la salud están estrechamente relacionadas con la ubicación, y que los menos saludables son aquellas personas que viven en los distritos más pobres. Por lo tanto, se ven reducidas en sus capacidades por la mala salud y, además, tienen una habilidad reducida para aprovechar los beneficios que brinda la educación como construir conocimientos y desarrollar habilidades.

Oztunc, Chi Oo y Serin (2015) estudiaron hasta qué punto la educación en las mujeres afectan el crecimiento a largo plazo en la región del pacífico de Asia. El periodo analizado es desde 1990 hasta 2010, e incluye a los siguientes países: Bangladesh, Camboya, China, India, Indonesia, República Democrática Popular Lao, Malasia, Myanmar, Filipinas, Tailandia y Vietnam. Los autores enfatizan el impacto de la educación femenina en el crecimiento económico (medido como PIB), alfabetización, fertilidad y fuerza laboral femenina. Utilizando un análisis de regresión sobre datos panel, encuentran que la tasa de fecundidad, la tasa de participación femenina en la fuerza laboral y la matriculación femenina en la escuela primaria son factores importantes para el crecimiento anual del ingreso per cápita. Los autores al citar Bourguignon y Morrison (2002) remarcen cómo los niveles de educación superior tienden a bajar la tasa de fecundidad al afectar el ingreso per cápita y disminuyendo la tasa de mortalidad infantil. Por último, otro autor que citan y cuyo hallazgo me pareció importante para el desarrollo del trabajo fue Oxaal (1997) que intentó encontrar una relación entre educación y pobreza a través de un análisis de género. Descubrió que las mujeres en los países en desarrollo generalmente reciben menos educación que los hombres. En los hogares pobres, el costo de oportunidad de enviar a las niñas al colegio es más alto que el de los hombres, ya que el tiempo que se debiera asignar a la escolarización de las niñas se utiliza para reemplazar el de sus madres, por ejemplo, en el cuidado de sus hermanos menores. Por tanto, la pérdida de estudio de las niñas durante el horario escolar repercute en la capacidad de las mujeres para aumentar los ingresos del hogar mediante el trabajo asalariado.

### **III. METODOLOGÍA Y DATOS:**

#### **I. Datos:**

La base de datos utilizada en éste estudio es la Encuesta Permanente de Hogares (EPH) relevada por el Instituto Nacional de Estadística y Censos (INDEC) correspondiente al segundo trimestre del 2021<sup>4</sup>. La encuesta contiene información sobre hogares y características de los individuos, nosotros utilizaremos los últimos. La base cuenta con 177 variables y 47,085 observaciones. Las variables del data set tienen relación con las principales características demográficas y socioeconómicas de la población de Argentina. Muchas de las observaciones corresponden a personas menores de 18 años por lo que las excluirémos de nuestro análisis ya que nuestro foco está en personas mayores de edad que se encuentran activos en el mercado laboral. Otra variable que usaremos para aplicar un filtro es el nivel de educación, eliminaremos de la base de datos aquellas personas que no tienen ningún tipo de instrucción (debido a que representan menos del 1% de las observaciones en ésta categoría) y las que no han dado respuesta alguna sobre sus logros en los distintos niveles educativos. Por último, prescindiremos en el análisis de aquellas personas que no indican cuánto es su ingreso mensual ni tampoco la cantidad de horas trabajadas por semana en su ocupación principal. Esto reduce nuestras observaciones a 12,093.

La variable dependiente de nuestro modelo serán los ingresos mensuales provenientes de la ocupación principal de las personas encuestadas. Las variables independientes serán los diferentes niveles educativos a través de dummies, sexo de los encuestados, estado civil y cantidad de horas trabajadas por semana en la actividad principal.

#### **II. Metodología:**

Específicamente se investigará los efectos de los distintos niveles de educación, cantidad de horas trabajadas la última semana en la actividad principal y edad sobre la probabilidad de ser pobre de las personas empleadas mayores de 17 años de edad.

Para llevar a cabo el estudio utilizare una regresión logística. A fin de poder modelar la probabilidad de que una persona se encuentre por debajo de la línea de pobreza en función de las variables explicativas binarizare los ingresos en base al valor de la canasta básica publicada por INDEC (\$ 23,922.00)<sup>5</sup> y de este modo convertirla en una variable dicotómica/binaria. Los ingresos mensuales por debajo del valor de la canasta básica tendrán valor 1 ( $P(Y=1)$ ) indicando que las personas están por debajo de la línea de pobreza, caso contrario el valor es 0 ( $P(Y=0)$ ).

A razón de que hemos aplicado filtros sobre las características de la muestra para obtener los salarios de las personas empleadas, ahora contamos con un sub conjunto de la misma. Es decir, tenemos una selección de muestra no aleatoria porque sólo hay personas asalariadas y, en consecuencia, los estimadores de la regresión pueden llegar a estar sesgados. Sumado a ello, hemos eliminado aquellas personas que se han negado a responder preguntas que brindan información sobre las variables explicativas,

---

<sup>4</sup> <https://www.indec.gob.ar/indec/web/Institucional-Indec-BasesDeDatos>

<sup>5</sup> [https://www.indec.gob.ar/uploads/informesdeprensa/canasta\\_12\\_2134BB3F19C7.pdf](https://www.indec.gob.ar/uploads/informesdeprensa/canasta_12_2134BB3F19C7.pdf)

pudiendo agravar el sesgo. Para enfrentar el problema de selección de muestra utilizaremos el método Heckit. La corrección de Heckman consta de 2 pasos para poder calcular  $\beta$  (estimadores) insesgados. El primero es: utilizando todas las  $n$  observaciones, correr un modelo Probit de  $S_i$  sobre  $Z_i$  y así obtener estimaciones de  $\gamma_h$ . Luego se calcula la razón inversa de Mills  $\lambda_i(Z_i \gamma)$  para cada observación (se necesita para aquellas  $i$  donde  $S_i = 1$ , es decir cuando la variable dependiente de interés “ $y$ ” es observada). El segundo paso es, utilizando la muestra seleccionada (donde las observaciones  $S_i = 1$ ), correr la regresión  $y_i$  sobre  $x_i$ ,  $\lambda_i$ . Python será el lenguaje utilizado para ajustar los modelos estadísticos sobre nuestra base de datos.

#### IV. RESULTADOS:

##### I. Estadísticas descriptivas:

**Tabla N° 1:** Estadísticas descriptivas de las variables cualitativas utilizadas

Atributo		Frecuencia	%
<b>Sexo</b>	Hombre	<b>6477 *</b>	<b>53,56%</b>
	Mujer	5616	46,44%
<b>Nivel educación</b>	Primaria incompl.	311	2,57%
	Primaria compl.	1264	10,45%
	Secundaria incompl.	1948	16,11%
	Secundaria compl.	<b>3768 *</b>	<b>31,16%</b>
	Univ. incompl.	1774	14,67%
	Univ. compl.	3028	25,04%
<b>Estado civil</b>	Casado/a	3264	26,99%
	Sep. o div.	1131	9,35%
	Soltero/a	<b>4028 *</b>	<b>33,31%</b>
	Unido/a	3457	28,59%
	Viudo/a	213	1,76%

\* Moda de la categoría

**Tabla N° 2:** Estadísticas descriptivas de las variables cuantitativas utilizadas.

	Mínimo	25%	50%	75%	Máximo
<b>Edad</b>	17	30	39	49	83
<b>Hs. Trabajadas últ. Sem.</b>	0	24	35	45	140
	Media	Rango	Desv. Estándar	Asimetría	Curtosis
<b>Edad</b>	40,006	66	12,137	0,266	-0,713
<b>Hs. Trabajadas últ. Sem.</b>	33,524	140	16,830	0,030	0,566

La sub-muestra utilizada tiene un tamaño de 12.093 observaciones, 6.477 y 5.616 son hombres y mujeres respectivamente. El mayor nivel educativo alcanzado por la sub-muestra seleccionada es “secundaria completa”. En cuanto a la variable estado civil, la moda es “soltero/a”. La edad media es de 40 años y la desviación estándar respecto la misma es de 12,137 años. La cantidad de horas promedio trabajadas en el oficio principal durante la última semana por los encuestados es de 33,524 hs con una

desviación estándar de 16,83 hs. En cuanto a las distribuciones de probabilidad, ambas variables cuantitativas presentan distribuciones bastantes simétricas con un pequeño sesgo positivo. La curtosis de “Edad” es platicúrtica mientras que la de “Hs. Trabajadas últ. Sem.” es leptocúrtica, indicando que en la primera variable hay una menor cantidad de datos concentrados alrededor de la media en comparación a la última.

## II. Resultados de regresiones:

En éste trabajo, el primer paso del método Heckit para corregir la selección de muestra es estimar la probabilidad de ser una persona empleada, es decir la probabilidad de estar o no incluido en la muestra seleccionada como consecuencia de formar parte de la fuerza laboral. La variable binaria dependiente del modelo es 1 si la observación es una persona empleada asalariada y 0 en caso contrario. Los regresores o variables explicativas son el estado civil con 6 categorías (soltero/a, casado/a, no contesta, separado/divorciado, viudo/a, unido/a), la edad, edad<sup>2</sup>, horas trabajadas en la última semana, y educación que se encuentra conformada por 7 categorías (secundaria incompleta, s/ instrucción, univ. completo, univ. incompleto, primario completo, primario incompleto y secundaria completa).

Tabla N° 3: Modelo probit con respecto la probabilidad de ser empleado/a asalariado (de ser observado o no).

Variables	Mujeres	Hombres
	Coeficientes	Coeficientes
const (Primario incompleto y soltero/a)	-4,1252 (0,101) ***	-3,4573 (0,082) ***
Casado/a	-0,3504 (0,035) ***	0,1380 (0,036) ***
No contesta	-9,7581 (3,62e+11)	-5,0925 (1,55e+04)
Sep./ div.	0,0943 (0,043) **	0,0342 (0,055)
Unido/a	-0,2120 (0,032) ***	0,1240 (0,031) ***
Viudo/a	-0,0729 (0,064)	0,1559 (0,109)
Secundario incompleto	0,1749 (0,064) ***	0,2989 (0,054) ***
Sin instrucción	-0,2709 (0,201)	0,0790 (0,142)
Superior universitario completo	0,7512 (0,063) ***	0,6647 (0,058) ***
Superior universitario incompleto	0,3739 (0,064) ***	0,4221 (0,057) ***
Primario completo	0,1842 (0,067) ***	0,3318 (0,058) ***
Secundario completo	0,3737 (0,062) ***	0,5403 (0,054) ***
Edad	0,1567 (0,005) ***	0,1230 (0,004) ***
Hs. Trabajadas últ. semana	0,0354 (0,001) ***	0,0286 (0,000) ***
Edad <sup>2</sup>	-0,0018 (5,48e-05) ***	-0,0015 (4,92e-05) ***
Pseudo R-squ.	0,4414	0,4147
LLR p-value	0,000	0,000

\*\*\* p<0,01, \*\*p<0,05, \*p<0,1

Los resultados del modelo Probit son mostrados en la tabla N° 3. Los signos de los coeficientes del modelo tienen sentido común, al menos en su mayoría. En cuanto a las dummies que se relacionan con educación todas tienen pendiente positiva y las magnitudes de las mismas están relacionadas con la jerarquía de los distintos niveles educativos, es decir que a mayor nivel educativo, mayores son las probabilidades de tener un trabajo pago. A partir de los signos de los coeficientes edad y edad<sup>2</sup> se



interpreta que a medida que aumentan los años aumentan las probabilidades de ser un empleado asalariado hasta un punto óptimo en el que a una determinada edad comienza a tener una relación lineal negativa con encontrarse dentro de la fuerza laboral. También se aprecia que cuánto más horas se trabaja durante la semana, más son las probabilidades de que una persona tenga un trabajo pago. Por último, se pueden observar diferencias por género en los signos de los coeficientes de las variables dummies correspondientes al estado civil. Mientras que estar casado afecta positivamente al hombre en la probabilidad de tener un empleo pago, el efecto de la misma variable es contrario en la mujer. Éste hecho ocurre también con las variables “unido/a” y “viudo/a”. Cuando las personas que integran ambos sexos se encuentran separados o divorciados como así también solteros, el signo del efecto es el mismo tanto para las mujeres como para los hombres, las probabilidades de ser una persona empleada asalariada aumentan.

**Tabla N° 4:**

Variables	Mujeres	Hombres
	Coeficientes	Coeficientes
const (Primario incompleto)	-0.6804 (0,260) ***	-1.8277 (0,173) ***
Primaria completo	0.0507 (0,268)	-0.2492 (0,172)
Secundaria completo	-0.8164 (0,252) ***	-0.5738 (0,163) ***
Secundaria incompleto	-0.3483 (0,259)	-0.1665 (0,164)
Univ. completo	-1.9924 (0,256) ***	-1.7029 (0,191) ***
Univ. incompleto	-0.9275 (0,255) ***	-0.5587 (0,175) ***
Inverse Mills ratio ( $\lambda$ )	1,7753 (0,089) ***	1,7071 (0,091) ***
Pseudo R-squ.	0,1890	0,1046
LLR p-value	0,000	0,000

\*\*\* p<0,01, \*\*p<0,05,\* p<0,1

El test de log-likelihood ratio nos permite determinar si los modelos son estadísticamente significativos o no, si es probable que los coeficientes de las variables explicativas sean todos iguales a 0. Ya que el LLR p-value < 0.05 rechazamos la hipótesis nula de que todos los  $\beta_k = 0$ , y en consecuencia los modelos estimados para la mujer y el hombre son estadísticamente significativos.

El segundo paso es estimar una regresión logística para calcular la probabilidad de ser pobre ( $y = 1$ ) sobre las variables explicativas seleccionadas y la razón inversa de Mills proveniente del modelo Probit del paso anterior que sirve para corregir el sesgo de selección de muestra. Se corrieron 2 regresiones logísticas. Una para estimar los efectos solamente de la educación (representada a través de variables dummies) sobre las probabilidades de ser pobre, y otra en la cual se agregaron características personales de los encuestados como la edad y cantidad de horas trabajadas por semana para corroborar en primer lugar si éste modelo es más explicativo de la variable dependiente que el anterior, y segundo, observar si los efectos de la educación varían y cómo influyen las nuevas 2 variables sobre las probabilidades de ser pobre.

Los resultados de la primera regresión logística se muestran en la tabla N° 4. En relación a la variable base “primario incompleto”, a nivel general se observa que en ambos sexos

casi todos los coeficientes son negativos y decrecientes a medida que aumenta la instancia educativa alcanzada por una persona. Esto quiere decir que cuando se alcanzan mayores niveles de escolarización menor es la probabilidad de ser pobre. A nivel género hay ciertas particularidades, si bien la variable “primaria compl.” es estadísticamente no significativa en ambos sexos, aumenta levemente la probabilidad ser pobre para las mujeres. El signo de ésta última variable pareciera ser incoherente ya que no coincide con la literatura estudiada previamente aunque para los hombres el signo del parámetro si concuerda con la literatura. Con respecto al sexo masculino, se observa que completar un ciclo educativo es más reductor en la probabilidad de ser pobre que comenzar la instancia educativa siguiente y no finalizarla, mientras que para la mujer éste efecto no se observa ya que los ingresos de las mismas siguen la misma suerte que el nivel alcanzado dentro la jerarquía educativa. En ejemplo: una mujer con la universidad incompleta, en promedio, tiende a tener menos probabilidades de ser pobre en comparación a otra mujer que sólo haya terminado la instancia educativa previa (secundaria completa); para el hombre es al revés, terminar la secundaria reduce aún más las probabilidades de ser pobre que arrancar la instancia siguiente y no terminarla. La magnitud de los coeficientes de todas las variables dummies son menores en las mujeres, con excepción “Primaria completo”.

Tabla N° 5:

Variables	Mujeres	Hombres
	Coeficientes	Coeficientes
const (Primario incompleto)	12,8323 (1,119) ***	0.4865 (1.670)
Primaria completo	-0,4375 (0,272)	-0.1152 (0,215)
Secundaria completo	-2,1517 (0,266) ***	-0.9235 (0,243) ***
Secundaria incompleto	-1,2168 (0,265) ***	-0.6590 (0,203) ***
Univ. completo	-4,0158 (0,296) ***	-1.7316 (0,290) ***
Univ. incompleto	-2,5228 (0,273) ***	-1.0650 (0,230) ***
Edad	-0,3242 (0,034) ***	-0.0465 (0,048)
Edad <sup>2</sup>	0,0033 (0,000) ***	-0.0003 (0,001)
Hs. Trabajadas últ. semana	-0,0856 (0,007) ***	-0.0029 (0,009)
Inverse Mills ratio ( $\lambda$ )	-1,6913 (0,305) ***	2,0843 (0.548) ***
Pseudo R-squ.	0,2501	0,1997
LLR p-value	0,000	0,000

\*\*\* p<0,01, \*\*p<0,05, \*p<0,1

Los LLR p-values < 0.05 para las regresiones de ambos sexos, por lo que rechazamos la hipótesis nula de que todos los  $\beta_k = 0$ , y en consecuencia los modelos estimados para la mujer y el hombre son estadísticamente significativos.

Queda analizar la segunda regresión logística cuyos resultados se observan en la tabla N° 5. En éste modelo se puede observar que todas las dummies que representan la educación tienen coeficientes negativos y decrecientes en relación al nivel educativo más bajo tomado como base en la regresión (primario incompleto). La interpretación es similar al modelo anterior, a mayor nivel educativo alcanzado por una persona, mayores

son sus ingresos disminuyendo las probabilidades de ser pobre. A diferencia del modelo anterior, las probabilidades de ser pobre de los hombres respetan los escalafones de la jerarquía de los distintos niveles educativos, y el signo del parámetro de la variable “Primaria completo” para las mujeres es negativo coincidiendo así con la literatura. La relación del regresor “Hs. Trabajadas últ. Semana” con la variable dependiente es negativa en ambos sexos indicando que a más horas trabajadas, menos son las probabilidades de ser pobre. El p-value de ésta última variable es estadísticamente significativa y no significativa para las mujeres y hombres respectivamente. Los signos de los parámetros de edad son negativos en los modelos de ambos sexos, indicando que a mayor edad menos son las probabilidades de ser pobre. El signo y significatividad de la Edad<sup>2</sup> difiere por sexo, para los hombres es negativo y no significativo mientras que para las mujeres es positivo y significativo. En el segundo caso las mujeres llegan a una cierta edad en la que a partir de la misma las probabilidades de ser pobre aumentan, en el caso de los hombres éste suceso no ocurre, al menos con ésta muestra de datos. La magnitud de los coeficientes de todas las variables dummies relacionados con educación son menores en las mujeres.

En relación a las lambdas ( $\lambda$ ) de la variable de Mills en ambos modelos, los p-values para ambos sexos son estadísticamente significativos. Esto quiere decir que existe un problema de selección de muestra y que si no hubiéramos incluido ésta variable en la regresión, los estimadores estarían sesgados. También indica que existe la posibilidad de que los errores estándares de cada coeficiente sean incorrectos, afectando la inferencia.

Para finalizar, los Pseudo R<sup>2</sup> del segundo modelo son mayores que los del primero en ambos sexos, indicando un mejor ajuste del segundo modelo a nuestros datos. Los LLR p-values < 0.05 para las regresiones de ambos sexos, por lo que rechazamos la hipótesis nula de que todos los  $\beta_k = 0$ , y en consecuencia los modelos estimados para la mujer y el hombre son estadísticamente significativos. En consecuencia, nos quedamos con el segundo modelo tanto para realizar inferencias sobre la muestra como para predecir probabilidades de nuevas observaciones.

## **V. CONCLUSIÓN:**

El objetivo principal de éste trabajo es estimar el efecto de la educación sobre la pobreza en Argentina. La base de datos utilizada es la encuesta permanente de hogares (EPH) del segundo trimestre de 2021 brindada por INDEC. Los resultados de las regresiones logísticas indican que a medida que una persona alcanza instancias educativas más altas, menores son las probabilidades de que la misma sea pobre. Es decir que existe una relación negativa entre educación y pobreza. Estos efectos son coherentes con los que suele indicar la literatura que aborda el tema en cuestión. Una de las razones de éste suceso es que los ingresos, en promedio, de las personas con niveles educativos altos suelen ser mayores que el de aquellas con niveles bajos de educación. La variable edad y cantidad de horas trabajadas la última semana en el oficio principal también tienen

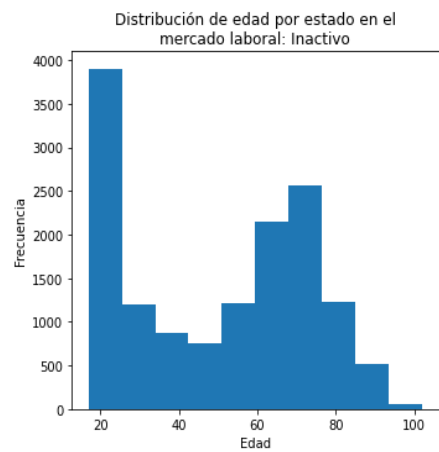
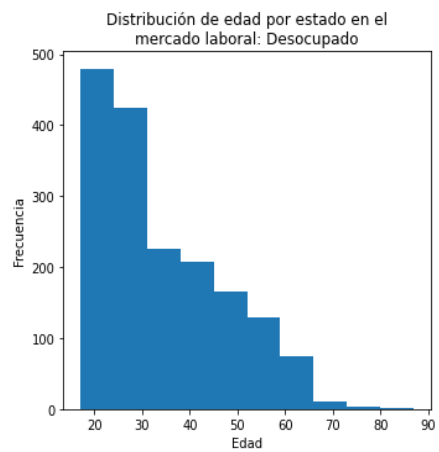
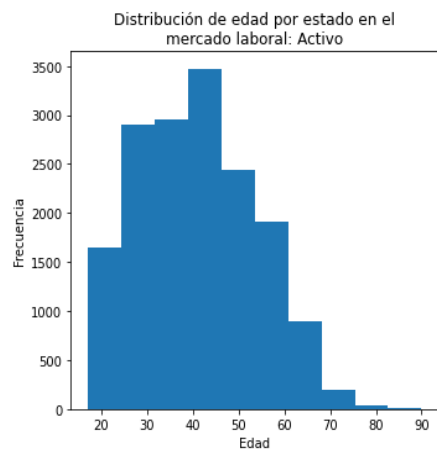
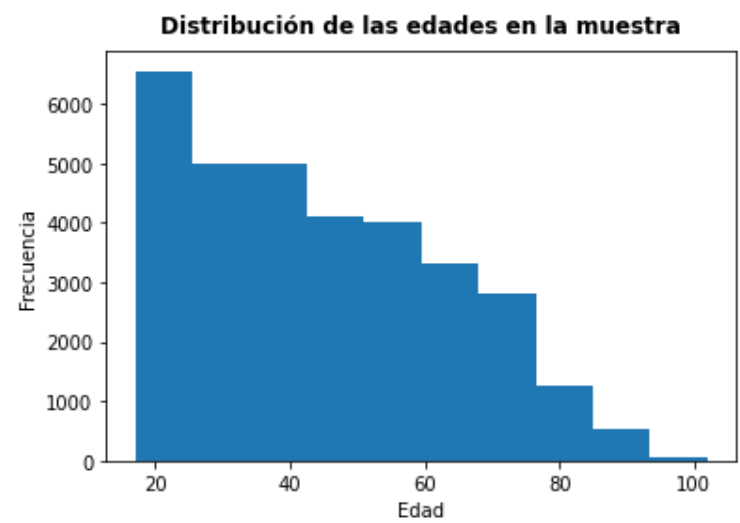
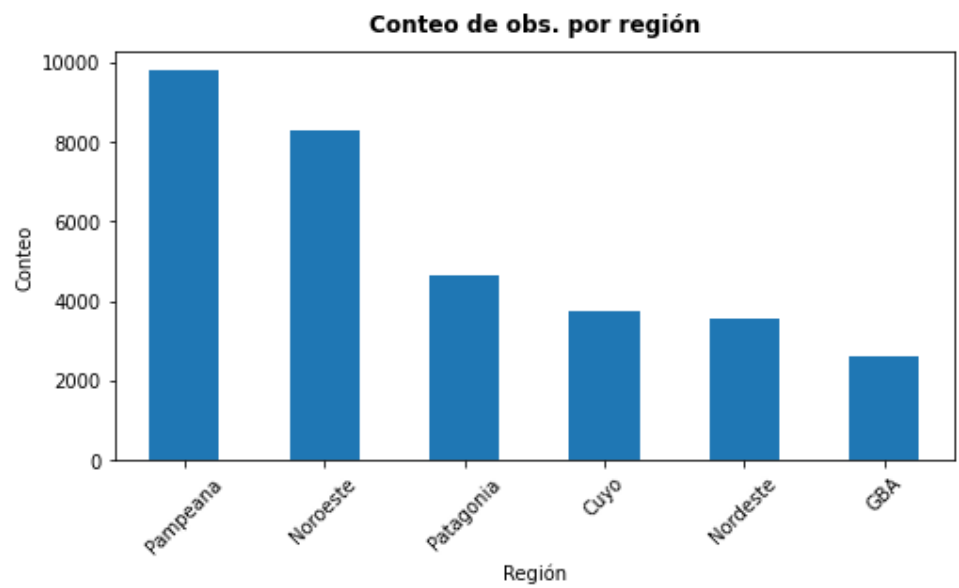
una relación inversa con pobreza. En cuanto a la edad, esto es posible ya que puede estar relacionada con factores como la experiencia en algún trabajo en particular ya que el tiempo dedicado al mismo otorga al individuo cierta “especialización” en el desarrollo de tareas de dicha índole y ese tiempo de especialización puede ser breve si la persona es joven, de éste modo empeorando su sueldo. Por otro lado, la relación inversa de pobreza con cantidad de horas trabajadas en el oficio podría explicarse si se aborda el tema midiendo cada hora de trabajo en una determinada cantidad de dinero, y de este modo si una persona trabaja más horas, disminuye sus probabilidades de ser pobre porque aumentan sus ingresos. Desde el punto de vista de género, los resultados de las regresiones indican que la educación tiene mayor impacto en la reducción de probabilidades de ser pobre en las mujeres que en los hombres en todos los niveles educativos así afirmando que hay sesgo por género. Entonces, si el efecto negativo de la educación sobre la pobreza es mayor en las mujeres, una de las razones es que sus ingresos son menores que los del género masculino con misma educación, y/o que la escolarización incrementa en mayor proporción los ingresos en las primeras que en los últimos.

#### **REFERENCIAS:**

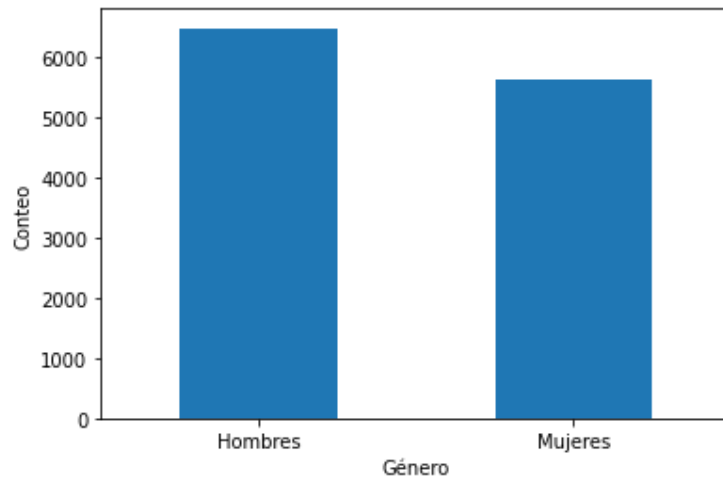
- 1 – Wooldridge, J. M. (2016). Introductory econometrics: A modern approach. 6th ed. Mason, OH: South-Western Cengage Learning.
- 2 – Yousaf Khan, M., Khaliq Alvi, A., Farid Chishti, M. (2019). An investigation on the linkages between poverty and education: a statistical review. Gomal University Journal of Research, Volume 35, Issue 1.
- 3 – Ahmad, N., Batul, E., Saleem, R. (2019). The long run and short run relationship between poverty and literacy rate in Pakistan. Volume 20 Issue 4.
- 4 – Njong, A. M. (2010). The effects of educational attainment on poverty reduction in Cameroon. Faculty of Economics and Management, University of Dschang, BP 285 Dschang, Western Region, Republic of Cameroon.
- 5 – De Silva, I., Sumarto, S. (2015). Dynamics of growth, poverty and human capital: evidence from Indonesian sub-national. Journal of Economic, Volume 40, Number 2.
- 6 – Burke A., Berinhout K., Bonnie P. (2019). Analyzing the Effect of Income Inequality on Poverty. Schools of Economics, Georgia Tech University.
- 7 – Jabbar Abdullah, A., Doucouliagos H., Manning, E. (2011). Education and Income Inequality: A Meta-Regression Analysis. School of Accounting, Economic and Finance, Deakin University.
- 8 – Etang Ndip, A. (2018). Education, Health and Poverty in Sudan. World Bank, Washington, DC. World Bank.
- 9 – Oztunc, H., Chi Oo, Z., Vildan Serin, Z. (2015). Effects of Female Education on Economic Growth: A Cross Country Empirical Study. Fatih University. ISSN 1303-0485 • eISSN 2148-7561.

- 10 – Pervez, S. (2014). Impact of Education on Poverty Reduction: A Co-integration Analysis for Pakistan. *Journal of Research in Economics and International Finance (JREIF)* (ISSN: 2315-5671) Vol. 3(4) pp. 83 – 89.
- 11 – Zamurrad Janjua, P., Ahmed Kamal, U. (2011). The Role of Education and Income in Poverty Alleviation: A Cross-Country. *The Lahore Journal of Economics Analysis. The Lahore Journal of Economics*, pp. 143-172
- 12 – Meghir, C., Rivkin, S. G. (2010). *Econometric Methods for Research in Education*. National Bureau of Economic Research, Cambridge, MA. Working Paper No. 16003.

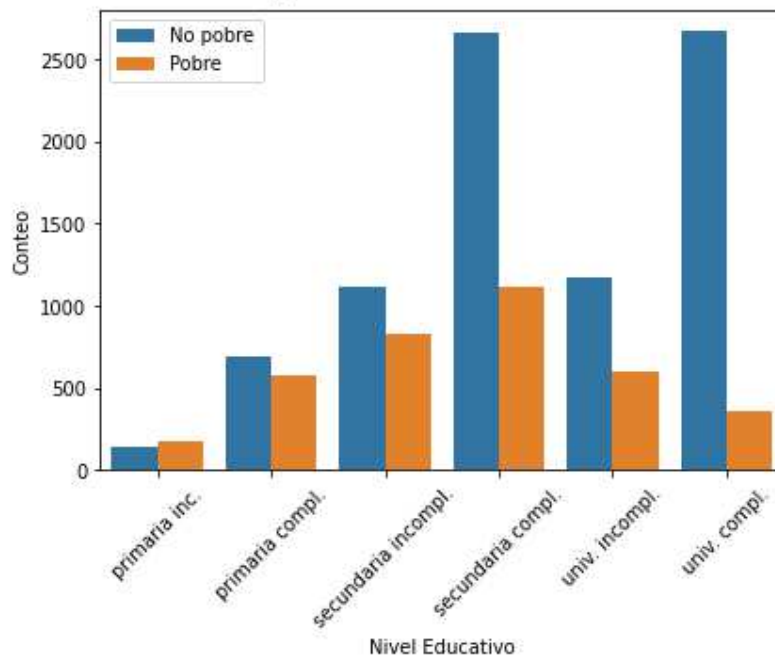
**APENDICE:**

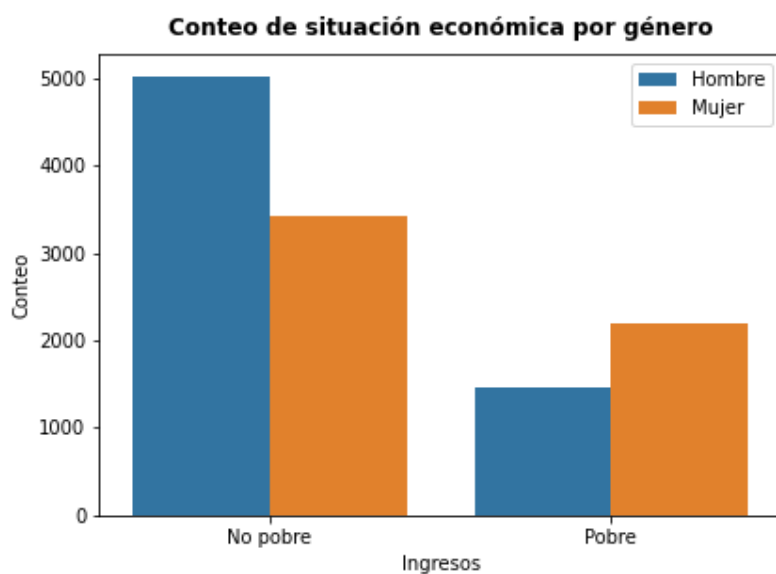
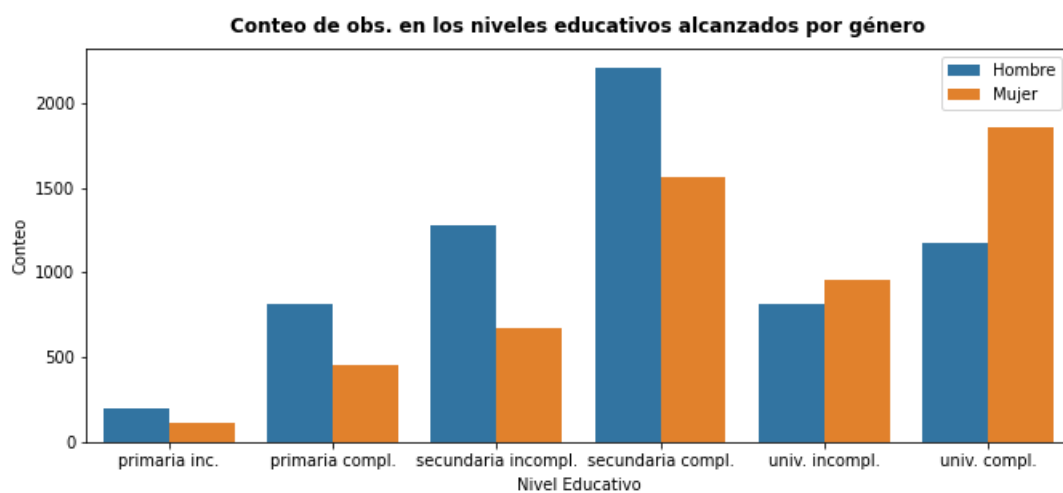
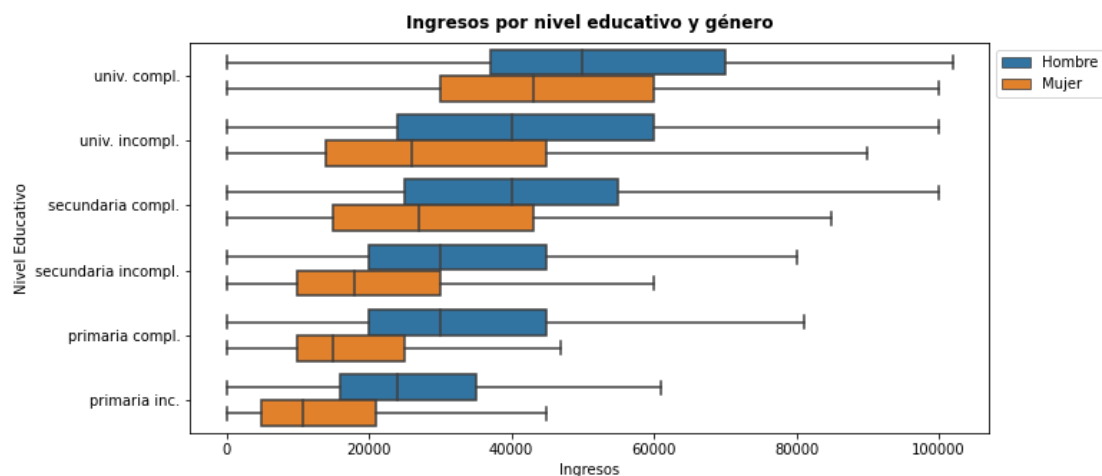


**Conteo de observaciones en la muestra por género**



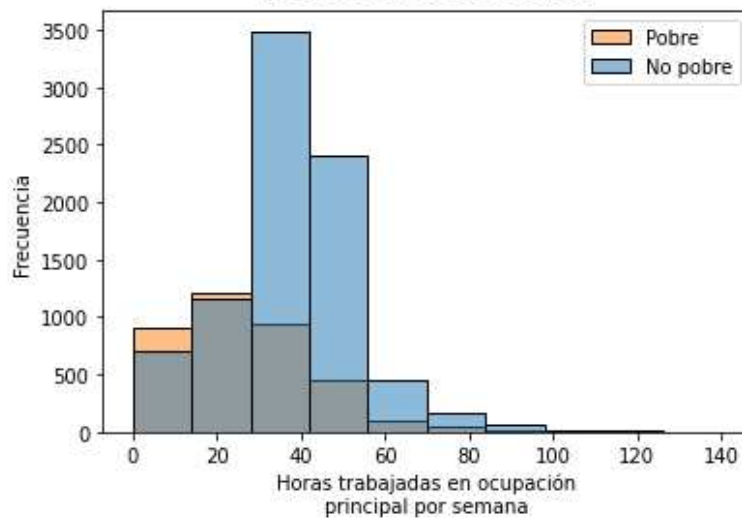
**Conteo en los niveles educativos por situación económica**



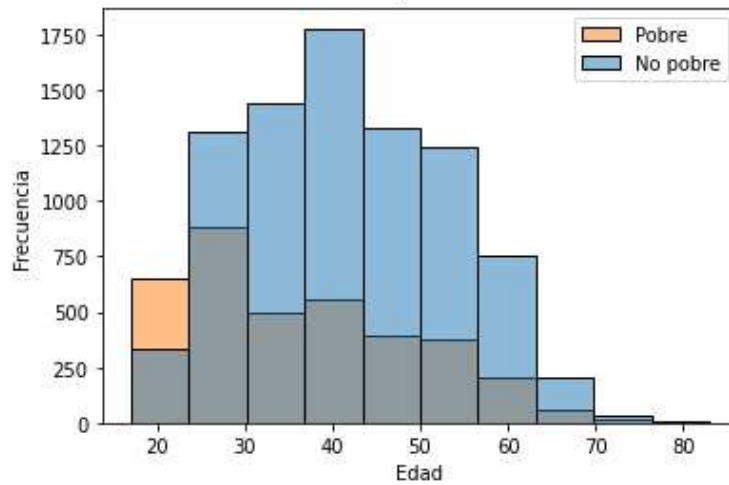




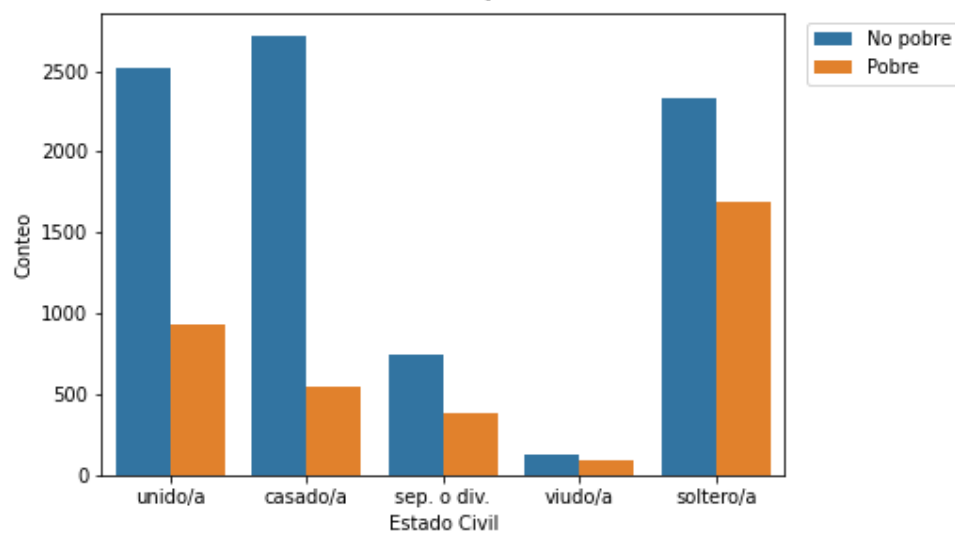
**Distribución de cant. de hs. trabajadas por situación económica**



**Distribución de edad por situación económica**



**Conteo de obs. en Estado Civil por situación económica**





## Hombres:

Probit Regression Results						
Dep. Variable:	observed	No. Observations:	22602			
Model:	Probit	Df Residuals:	22587			
Method:	MLE	Df Model:	14			
Date:	Fri, 18 Feb 2022	Pseudo R-squ.:	0.4147			
Time:	09:33:15	Log-Likelihood:	-8437.5			
converged:	False	LL-Null:	-14415.			
Covariance Type:	nonrobust	LLR p-value:	0.000			
	coef	std err	z	P> z	[0.025	0.975]
const	-3.4573	0.082	-42.319	0.000	-3.617	-3.297
CH07_casado	0.1380	0.036	3.881	0.000	0.068	0.208
CH07_no contesta	-5.0925	1.55e+04	-0.000	1.000	-3.05e+04	3.05e+04
CH07_sep./div.	0.0342	0.055	0.622	0.534	-0.073	0.142
CH07_unido	0.1240	0.031	3.957	0.000	0.063	0.185
CH07_viudo/a	0.1559	0.109	1.426	0.154	-0.058	0.370
NIVEL_ED_Secundario incompleto	0.2989	0.054	5.577	0.000	0.194	0.404
NIVEL_ED_Sin instrucción	0.0790	0.142	0.557	0.578	-0.199	0.357
NIVEL_ED_Superior universitario completo	0.6647	0.058	11.522	0.000	0.552	0.778
NIVEL_ED_Superior universitario incompleto	0.4221	0.057	7.353	0.000	0.310	0.535
NIVEL_ED_primario completo	0.3318	0.058	5.698	0.000	0.218	0.446
NIVEL_ED_secundario completo	0.5403	0.054	10.047	0.000	0.435	0.646
CH06	0.1230	0.004	29.294	0.000	0.115	0.131
PP3E_TOT	0.0286	0.000	68.969	0.000	0.028	0.029
sqr_CH06	-0.0015	4.92e-05	-31.388	0.000	-0.002	-0.001

## Modelo Heckman Paso 2:

### Mod. Hombres N° 1:

Logit Regression Results						
Dep. Variable:	can_basica	No. Observations:	6477			
Model:	Logit	Df Residuals:	6470			
Method:	MLE	Df Model:	6			
Date:	Fri, 18 Feb 2022	Pseudo R-squ.:	0.1046			
Time:	09:33:15	Log-Likelihood:	-3086.3			
converged:	True	LL-Null:	-3446.7			
Covariance Type:	nonrobust	LLR p-value:	1.876e-152			
	coef	std err	z	P> z	[0.025	0.975]
const	-1.8277	0.173	-10.579	0.000	-2.166	-1.489
primaria compl.	-0.2492	0.172	-1.450	0.147	-0.586	0.088
secundaria compl.	-0.5738	0.163	-3.512	0.000	-0.894	-0.254
secundaria inkompl.	-0.1665	0.164	-1.013	0.311	-0.489	0.156
univ. compl.	-1.7029	0.191	-8.910	0.000	-2.077	-1.328
univ. inkompl.	-0.5587	0.175	-3.200	0.001	-0.901	-0.216
mills_ratio	1.7071	0.091	18.786	0.000	1.529	1.885

## Mod. Hombres N° 2:

Dep. Variable:	can_basica	No. Observations:	6477			
Model:	Logit	Df Residuals:	6467			
Method:	MLE	Df Model:	9			
Date:	Fri, 18 Feb 2022	Pseudo R-squ.:	0.1997			
Time:	09:33:15	Log-Likelihood:	-2758.4			
converged:	True	LL-Null:	-3446.7			
Covariance Type:	nonrobust	LLR p-value:	8.770e-291			
	coef	std err	z	P> z	[0.025	0.975]
const	0.4865	1.670	0.291	0.771	-2.786	3.759
primaria compl.	-0.1152	0.215	-0.537	0.591	-0.536	0.306
secundaria compl.	-0.9235	0.243	-3.796	0.000	-1.400	-0.447
secundaria incompl.	-0.6590	0.203	-3.244	0.001	-1.057	-0.261
univ. compl.	-1.7316	0.290	-5.972	0.000	-2.300	-1.163
univ. incompl.	-1.0650	0.230	-4.626	0.000	-1.516	-0.614
CH06	-0.0465	0.048	-0.978	0.328	-0.140	0.047
sqr_CH06	-0.0003	0.001	-0.503	0.615	-0.001	0.001
PP3E_TOT	-0.0029	0.009	-0.311	0.756	-0.022	0.016
mills_ratio	2.0843	0.548	3.802	0.000	1.010	3.159

## Mod. Hombres N° 2 (Efectos Marginales):

Logit Marginal Effects						
=====						
Dep. Variable:	can_basica					
Method:	dydx					
At:	overall					
=====						
	dy/dx	std err	z	P> z	[0.025	0.975]
-----						
primaria compl.	-0.0156	0.029	-0.537	0.591	-0.073	0.041
secundaria compl.	-0.1250	0.033	-3.807	0.000	-0.189	-0.061
secundaria incomp.	-0.0892	0.027	-3.251	0.001	-0.143	-0.035
univ. compl.	-0.2344	0.039	-6.004	0.000	-0.311	-0.158
univ. incomp.	-0.1441	0.031	-4.647	0.000	-0.205	-0.083
CH06	-0.0063	0.006	-0.978	0.328	-0.019	0.006
sqr_CH06	-4.075e-05	8.09e-05	-0.504	0.615	-0.000	0.000
PP3E_TOT	-0.0004	0.001	-0.311	0.756	-0.003	0.002
mills_ratio	0.2821	0.074	3.821	0.000	0.137	0.427

### Mod. Mujeres N° 1:

Logit Regression Results						
Dep. Variable:	can_basica	No. Observations:	5616			
Model:	Logit	Df Residuals:	5609			
Method:	MLE	Df Model:	6			
Date:	Fri, 18 Feb 2022	Pseudo R-squ.:	0.1890			
Time:	09:33:15	Log-Likelihood:	-3047.5			
converged:	True	LL-Null:	-3757.8			
Covariance Type:	nonrobust	LLR p-value:	7.905e-304			
	coef	std err	z	P> z	[0.025	0.975]
const	-0.6804	0.260	-2.614	0.009	-1.191	-0.170
primaria compl.	0.0507	0.268	0.189	0.850	-0.475	0.576
secundaria compl.	-0.8164	0.252	-3.246	0.001	-1.309	-0.323
secundaria incomp.	-0.3483	0.259	-1.346	0.178	-0.855	0.159
univ. compl.	-1.9924	0.256	-7.772	0.000	-2.495	-1.490
univ. incomp.	-0.9275	0.255	-3.642	0.000	-1.427	-0.428
mills_ratio	1.7753	0.089	19.896	0.000	1.600	1.950

### Mod. Mujeres N° 2:

Logit Regression Results						
Dep. Variable:	can_basica	No. Observations:	5616			
Model:	Logit	Df Residuals:	5606			
Method:	MLE	Df Model:	9			
Date:	Fri, 18 Feb 2022	Pseudo R-squ.:	0.2501			
Time:	09:33:15	Log-Likelihood:	-2817.8			
converged:	True	LL-Null:	-3757.8			
Covariance Type:	nonrobust	LLR p-value:	0.000			
	coef	std err	z	P> z	[0.025	0.975]
const	12.8323	1.119	11.466	0.000	10.639	15.026
primaria compl.	-0.4375	0.272	-1.606	0.108	-0.971	0.096
secundaria compl.	-2.1517	0.266	-8.088	0.000	-2.673	-1.630
secundaria incomp.	-1.2168	0.265	-4.589	0.000	-1.736	-0.697
univ. compl.	-4.0158	0.296	-13.585	0.000	-4.595	-3.436
univ. incomp.	-2.5228	0.273	-9.239	0.000	-3.058	-1.988
CH06	-0.3242	0.034	-9.659	0.000	-0.390	-0.258
sqr_CH06	0.0033	0.000	8.184	0.000	0.002	0.004
PP3E_TOT	-0.0856	0.007	-12.571	0.000	-0.099	-0.072
mills_ratio	-1.6913	0.305	-5.539	0.000	-2.290	-1.093

Mod. Mujeres N° 2 (Efectos Marginales):

Logit Marginal Effects						
=====						
Dep. Variable:	can_basica					
Method:	dydx					
At:	overall					
=====						
	dy/dx	std err	z	P> z	[0.025	0.975]
-----						
primaria compl.	-0.0726	0.045	-1.607	0.108	-0.161	0.016
secundaria compl.	-0.3573	0.043	-8.256	0.000	-0.442	-0.272
secundaria incompl.	-0.2020	0.044	-4.618	0.000	-0.288	-0.116
univ. compl.	-0.6668	0.046	-14.388	0.000	-0.758	-0.576
univ. incompl.	-0.4189	0.044	-9.490	0.000	-0.505	-0.332
CH06	-0.0538	0.005	-9.930	0.000	-0.064	-0.043
sqr_CH06	0.0005	6.47e-05	8.344	0.000	0.000	0.001
PP3E_TOT	-0.0142	0.001	-13.173	0.000	-0.016	-0.012
mills_ratio	-0.2808	0.050	-5.582	0.000	-0.379	-0.182
=====						