

## IV. Resultados

### 1. Modelos lineales

A continuación se muestran los resultados de las regresiones con el modelo lineal. Las primeras cinco corresponden a los modos de transporte urbano, la sexta se refiere a la utilización del automóvil como medio de transporte, y la última es una regresión general de todos los transportes públicos<sup>1</sup>.

**Tabla 21. Modelo de regresiones lineales**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<b>Variables dependientes</b>	<b>gmetro</b>	<b>gbus</b>	<b>gtrole</b>	<b>gpes</b>	<b>gtaxi</b>	<b>ggas</b>	<b>gto_total</b>
<b>ingtri</b>	-0.0003** (0.0001)	-0.0004 (0.0004)	-0.0000172 (0.0000151)	-0.0049** (0.0008)	0.0005 (0.0005)	0.0307** (0.0029)	-0.0051** (0.0009)
<b>sexo</b>	-5.5412 (5.2284)	-10.5388 (21.3776)	0.4948 (1.0297)	-2.0835 (42.5429)	-52.3769* (22.1401)	67.975 (53.6459)	-70.0455 (44.9826)
<b>edad</b>	0.2089 (0.1688)	-0.3017 (0.6184)	-0.0152 (0.0231)	-1.4796 (1.1735)	0.0437 (0.7014)	1.1695 (1.5738)	-1.5439 (1.2621)
<b>ed_formal</b>	-0.0867 (0.6867)	9.8946** (3.2369)	0.2474* (0.1132)	22.9213** (6.5292)	0.4833 (3.0108)	19.8426** (7.0028)	33.4599** (6.8874)
<b>trabajo</b>	6.7947 (6.0398)	-14.7963 (22.6834)	-0.2501 (1.1375)	-23.4388 (49.2468)	35.2977 (24.8329)	-50.5791 (60.9868)	3.6072 (51.0138)
<b>estrato</b>	43.7496** (4.1354)	-275.5904** (36.1444)	1.4189 (0.8909)	137.6471** (45.1839)	20.0613 (19.6513)	118.8409** (39.4417)	-72.7135 (48.917)
<b>vehiculos</b>	-33.251** (4.9365)	-85.8042** (23.4193)	-2.7127** (0.9865)	-524.2834** (44.9517)	29.9603 (21.4171)	1,065.90** (52.1412)	-616.091** (48.0653)
<b>Constante</b>	13.692* (6.3083)	317.8594** (31.7310)	0.1448 (1.0436)	607.9657** (59.3370)	65.8848 (33.6852)	-528.5734** (79.4190)	1,005.55** (63.0665)
<b>Observaciones</b>	2200	2200	2200	2200	2200	2200	2200
<b>R-cuadrada</b>	0.05	0.06	0.01	0.08	0.005	0.36	0.10

Errores estándar robusto entre paréntesis

\* significativa a 5%; \*\* significativa a 1%

<sup>1</sup> Para todas las regresiones del trabajo se utilizó el programa de cómputo Intercooled Stata 7.

- En la tabla 21 podemos ver que el ingreso trimestral es sólo significativo para los modos de Metro, colectivo y para el automóvil, aunque en la regresión del transporte urbano en general también es significativa. Vemos que los coeficientes son negativos para casi todos los modos, excepto para taxi y automóvil. Esto nos dice que al aumentar el ingreso se dejan de utilizar modos como el Metro, colectivo, trolebús y autobús para utilizar más el taxi y el auto particular, es decir, se dejan de utilizar medios de alta capacidad para utilizar medios de baja capacidad. También podemos decir entonces que existe entonces cierto grado de sustitución entre el transporte público y el privado.
- Vemos que la variable *sexo* es sólo significativa para el modo de taxi. Su coeficiente nos indica que al ser hombre el gasto trimestral en promedio disminuye en \$ 52 pesos. Es decir, una mujer gasta en promedio más en taxi que el hombre. Una posible explicación de esto es que para una mujer puede ser más riesgoso viajar en otro tipo de transporte público.
- La *edad* no es una variable que podamos considerar significativa, además esta variable puede presentar cierta correlación con el ingreso, por lo que en estas estimaciones no tiene mucha relevancia. Podemos decir entonces que la gente se traslada, sin considerar cómo lo haga, no importando su edad.
- Los años de educación se presentan como una variable que es significativa en casi todas las regresiones, sólo para Metro y taxi no lo es. Sus coeficientes nos indican que el tener un año más de educación aumentará el uso del transporte público y también el uso de transporte privado. Este resultado es un poco ambiguo ya que nos dice que se utilizará más el transporte pero no nos dice si hay un cambio en el comportamiento del individuo.
- La variable dummy *trabajo* tampoco es significativa y también esperamos cierta correlación con el ingreso, por lo que no es muy relevante. Aquí podemos ver que,

como ya habíamos dicho antes, los individuos se trasladan de un lugar a otro no sólo por el trabajo sino también con otros fines.

- La densidad poblacional que está contemplada en la variable *estrato* no es significativa sólo para dos modos: trolebús y taxi, aunque tampoco lo es considerando al transporte público en general. En los modos que sí es significativa vemos que el que esté situado el hogar en una zona de alta densidad aumenta el uso de los medios de transporte; excepto para el autobús. Este resultado va en contra de la teoría descrita antes, donde se espera que en las zonas de alta densidad sea más utilizado el transporte público y que en las zonas de baja densidad se utilice más el automóvil. También se esperaba que en las zonas de alta densidad existiera un mayor uso de transportes de alta capacidad. El autobús entra en esta categoría, sin embargo vemos que el estar en una zona de alta densidad disminuye el uso de este medio.
- El tener al menos un automóvil en el hogar, que se refleja en la variable *vehiculos*, es significativo en todas las regresiones excepto para taxi, y vemos que el coeficiente para los modos de transporte público es negativo, es decir que el tener al menos un automóvil representa una disminución en el uso de otros medios de transporte; mientras que para la regresión sobre el uso de automóvil (*ggas*) el coeficiente es, obviamente, positivo.

## **2. Modelo Probit**

Mientras que para el modelo de regresiones lineales no era tan relevante el valor de los coeficientes, para el modelo Probit sí lo es, por lo que se generaron variables que tuvieran términos de interacción para ver si estas nuevas variables tenían un impacto significativo en el modelo con el fin de explicarlo mejor.

Por la extensión del número de variables que se generaron, las tablas de la comparación de los resultados se pueden ver en el Anexo E.

De estos resultados obtuvimos que las variables que contienen términos de interacción que resultaron relevantes para el modelo fueron:

- *sexed*: corresponde a la interacción de la variable dummy para el género y la variable de educación ( $\text{sexo} * \text{ed\_formal}$ ).
- *edadcar*: corresponde a la interacción de la variable de edad y la de posesión de automóvil en el hogar ( $\text{edad} * \text{vehiculos}$ ).

Teniendo un modelo más adecuado es posible ya describir los resultados que presenta y con esto podemos explicar lo que sucede con las decisiones sobre cómo se transportan las personas.

En la tabla 23 se presentan los resultados de las regresiones para la variable dependiente *auto*.

**Tabla 23. Resultados de regresiones Probit para *auto***

<b>Variables</b>	<b>(1)</b>	<b>(2)</b>	<b>(3)</b>
<b>ingtri</b>	0.000003** (0.0000005)	0.000003** (0.0000005)	0.000003** (0.0000005)
<b>sexo</b>	0.071368 (0.038501)	0.071406 (0.038343)	-----
<b>edad</b>	-0.001713 (0.001348)	-0.001710 (0.001327)	-----
<b>ed_formal</b>	0.003872 (0.002652)	0.003880 (0.002545)	-----
<b>trabajo</b>	0.000176 (0.016202)	-----	-----
<b>estrato</b>	-0.023908 (0.017890)	-0.023933 (0.017743)	-----
<b>vehiculos</b>	0.409494** (0.0603764)	0.409524** (0.0603046)	0.508426** (0.0171681)
<b>sexed</b>	-0.006491 (0.003410)	-0.006491 (0.003410)	-----
<b>edadcar</b>	0.002358 (0.001376)	0.002357 (0.001373)	-----
<b>Observaciones</b>	2200	2200	2200
<b>Pseudo R<sup>2</sup></b>	0.4424	0.4424	0.4386

Errores estándar entre paréntesis

\* significativa a 5%; \*\* significativa a 1%

La primera regresión presenta el modelo con los dos términos de interacción que resultaron más significativos. Sin embargo, la variable dummy *trabajo* no resultó ser muy significativa por lo que en la segunda regresión la omitimos, dando como resultado una misma pseudo R<sup>2</sup> de 0.4424, lo que nos indica que era una variable irrelevante en el modelo, además suponíamos que existía cierta correlación entre ésta y el ingreso trimestral, por lo que omitirla nos da mejores estimadores.

En la tercera regresión sólo consideramos a las variables que tuvieron la mayor significancia: ingreso trimestral y la propiedad de al menos un automóvil en el hogar. Vemos que en esta regresión el resultado de la pseudo R<sup>2</sup> es de 0.4386, es decir, que estas dos variables explican casi todo el modelo.

Por lo tanto, del modelo original que teníamos anteriormente ahora tenemos el siguiente:

$$\Pr(\text{auto} = 1 | X) = \Phi(\mathbf{b}_0 + \mathbf{b}_1 \text{ingtri} + \mathbf{b}_2 \text{edad} + \mathbf{b}_3 \text{ed\_formal} + \mathbf{d}_1 \text{sexo} + \mathbf{d}_2 \text{estrato} + \mathbf{d}_3 \text{vehiculo} + \mathbf{b}_4 \text{sexed} + \mathbf{b}_5 \text{edadcar})$$

Dado que no se trata de un modelo de regresión lineal, la interpretación de los coeficientes no es tan sencilla. Para facilitar esto, utilizamos un comando en el paquete de cómputo que presenta el efecto marginal de las variables sobre la probabilidad de que se dé un evento dado utilizando los valores medios de todas las variables<sup>2</sup>. Estos son los resultados utilizando los datos de la segunda regresión:

- El efecto marginal promedio que tiene el ingreso trimestral sobre la probabilidad de utilizar el automóvil como medio de transporte es de apenas 0.0003%. Sin embargo, el suponer un aumento marginal en el ingreso de manera trimestral es un poco irrelevante. En cambio, teniendo una variable que presente el ingreso promedio diario de una persona sería más ilustrativo. Es por esto que generamos una nueva variable a partir del ingreso trimestral que, dividida entre los 90 días que representan un trimestre, representa el ingreso promedio diario de una persona. Como esta nueva variable es una transformación de la anterior, los resultados de las regresiones no cambian, sólo que esta vez tenemos que para el aumento marginal en el ingreso diario la probabilidad de utilizar el automóvil aumenta en 0.03%.
- La variable dummy *vehiculos* nos dice que si existe un automóvil en el hogar entonces la probabilidad de utilizarlo como medio de transporte aumenta en casi 41%.
- Para la variable dummy *sexo*, significativa al 6% y tomando en cuenta los términos de interacción, nos dice que, por ser hombre, la probabilidad aumenta en 6.7%.

---

<sup>2</sup> En el programa Stata existe un comando, *dprobit*, el cual presenta resultados más sencillos de interpretar. Todas las regresiones se hicieron utilizando este comando.

- La variable sobre los años de educación (*ed\_formal*), que es muy poco significativa, apenas lo es al 15%, nos dice que por un año más de educación la probabilidad aumenta en sólo 0.4%.
- La densidad poblacional (*estrato*) es sólo significativa a más del 15%. Sin embargo, ésta nos dice que el estar en una zona de alta densidad disminuye la probabilidad en 2.4%.
- La variable *edad* es la menos significativa de todas, lo que nos indica que la edad del individuo no tiene mucha relación en cuanto al modo de transportarse. Aún así, los resultados tomando en cuenta los términos de interacción nos indican que un año más de edad disminuiría la probabilidad en 0.24%.

En la siguiente tabla se incluyen los resultados de las regresiones utilizando el modelo probit para las variables binarias que generamos para los otros medios de transporte. Las variables explicativas y los comandos de regresión son los mismos.

**Tabla 25. Resultados de regresiones probit para transporte público**

	(1)	(2)	(3)	(4)	(5)
	metro	bus	trole	pes	taxi
<b>ingtri</b>	-1.19E-06* (2.08)	-8.14E-07 (1.17)	-7.61E-08 (0.51)	-6.22E-06** (5.23)	-7.35E-07 (1.54)
<b>sexo</b>	-0.015 (1.28)	-0.014 (1.15)	0.0003989 (0.18)	0.037 (1.65)	-0.007 (0.57)
<b>edad</b>	0.001 (1.93)	-0.001* (2.14)	-0.0000647 (0.77)	-0.001 (0.86)	0.0002684 (0.79)
<b>ed_formal</b>	0.001 (0.32)	0.0002947 (0.17)	0.001* (2.24)	-0.0000165 (0.01)	-0.001 (0.66)
<b>estrato</b>	0.113** (7.88)	-0.197** (11.42)	0.003 (1.21)	0.061* (2.36)	0.037** (2.70)
<b>vehiculos</b>	-0.082** (6.36)	-0.058** (4.30)	-0.009** (2.64)	-0.325** (14.09)	0.015 (1.15)
<b>Observaciones</b>	2200	2200	2200	2200	2200
<b>Pseudo R<sup>2</sup></b>	0.0987	0.123	0.0877	0.115	0.0088

Valor absoluto de estadístico z entre paréntesis

\* significativa a 5%; \*\* significativa at 1%

Vemos que los resultados en estas regresiones son poco significativas, sobre todo para el modo de taxi, donde apenas obtenemos una pseudo  $R^2$  de 0.0088 y sólo la variable *estrato* es significativa. Sin embargo, a continuación tenemos un resumen de los resultados:

- El ingreso tiene un efecto marginal muy bajo en la probabilidad de utilizar alguno de los modos anteriores.
- El género de las personas no es una variable significativa para el uso del transporte público.
- La edad sólo es significativa para el modo de autobús donde el efecto marginal sobre la probabilidad es de un aumento del 0.1%.
- La educación tampoco es una variable significativa, sólo para el uso del trolebús el efecto marginal sobre la probabilidad de utilizarlo es un aumento del 0.1%
- La densidad poblacional, contemplada en la variable *estrato*, sí es una variable significativa en este modelo, sólo para colectivos no lo es. Para el Metro, el vivir en una zona de alta densidad significa un aumento de 11% en la probabilidad de utilizarlo. Contrariamente, el vivir en una zona con la misma característica representa una disminución en la probabilidad de utilizar el autobús. Para colectivos y taxis también representa un aumento en la probabilidad, con un 6 y 4% respectivamente.
- Por último, la variable dummy *vehículos* también es significativa en este modelo, con la excepción de los taxis. Vemos que en todos los casos, el tener un vehículo en el hogar representa una disminución en la probabilidad de utilizar los medios de transporte público. Para el que representa una mayor disminución es para los colectivos en 32%, mientras que la disminución más pequeña se presenta en el trolebús, con apenas un 1%.