

Variación en los precios de los carros importados como función de sus atributos

Miguel Armando Paniagua

Instituto Superior de Economía y Administración de Empresas (ISEADE)

Resumen

Se usan datos sobre el parque vehicular de El Salvador para analizar el efecto de los atributos de los autos importados sobre su precio. Las estimaciones se realizan por medio de mínimos cuadrados ponderados, en términos Log - Nivel. Se modela este precio controlando por la marca, años entre fabricación e importación del vehículo (*Tiempo*), color, cilindraje y ser nuevo al momento de importarlo. Un auto pierde 10 % de su valor con cada año adicional en la variable *Tiempo*; se toma como base la marca Toyota, puesto que representa el 21 % de las placas particulares del conjunto de datos analizados. La mayor parte de las marcas (14/16) consideradas se valoran menos que la marca base (-35 %). Se recomienda registrar de forma completa la información sobre asientos y puertas e incluir una variable con el tipo de automotor (sedan, camioneta, etc.) en la base sobre el parque vehicular del país.

Palabras clave: El Salvador, usados, ponderados, precios, modelo.

Introducción

Con base en registros del Ministerio de Obras Públicas (MOP), se constata que el parque vehicular salvadoreño en 2018 ascendió a 1,175,341 unidades activas (67.25 % con placas particulares). Cada año ingresa una alta cantidad de autos importados al país, principalmente desde Estados Unidos de América y la reventa de automóviles (importados o adquiridos en el mercado local) es una importante fuente de adquisición de automotores en el país. Entre 2016 y 2017 se estimó que ingresaron 44,000 carros usados adicionales (Molina, 2017).

Este mercado es importante puesto que grandes segmentos poblaciones podrían decidir invertir en activos privados para su desplazamiento, ante la situación de un sistema de transporte colectivo con múltiples falencias, y diferenciales de precios

importantes, entre los autos transados en el mercado de segunda mano y los nuevos.

De forma adicional, se identifican múltiples actividades conexas con la reventa de autos usados. En 2016 el país importó \$ 63.6 millones en partes y accesorios para vehículos; \$ 697 miles se importaron en concepto de carrocería para vehículos (Molina, 2017). Al momento, no se identifica ninguna investigación que modele de forma empírica los determinantes de los precios en el parque vehicular en El Salvador, y que provea evidencia para medir el impacto de los diversos atributos de los carros sobre su valor.

El análisis sobre los precios del parque vehicular está documentado en la literatura empírica y teórica. Con base en el enfoque de la economía de la información (Akerlof, 1970) indicó que el mercado de autos usados sufre como principal falla la asimetría de información, entre compradores y vendedores, ya que solo estos últimos pueden saber si el carro transado es de buena calidad (durazno) o de mala calidad (limón), lo cual resultaría en disminuir de forma ineficiente la disponibilidad de autos de buena calidad en el mercado.

La literatura sobre la estimación de los determinantes del precio de los autos gira en torno a los modelos hedónicos de precios (Cumhur & Senturk, 2009). Este enfoque usa información sobre las características de los autos y las trata como variables dependientes en el ajuste de modelos que estiman su relación con el precio del vehículo. Entre las variables usualmente incluidas están: el tamaño del automóvil (carrocería y motor), antigüedad, marca, colores y la posesión o no de una serie de atributos auxiliares, pero relevantes, tales como: presentar aire acondicionado en funcionamiento, tipo de transmisión, tipo de combustible, número de asientos y puertas, color, entre otros (Cumhur & Senturk, 2009) (Ramírez, Uriarte, Delbianco, & Larrosa, 2017).

La obtención de los datos ha sido abordada, ya sea por medio de encuestas y levantamiento directo (Cumhur & Senturk, 2009) o por medio de técnicas de *web scraping* (Ramírez, Uriarte, Delbianco, & Larrosa, 2017). Con relación a los métodos cuantitativos, el ajuste de los modelos hedónicos se realiza usando de Mínimos

Cuadrados Ordinarios (MCO), en sus diversas formas funcionales, así como *Random Forest* y Redes Neuronales, para casos de predicción de precios (Pal, Sundararaman, Arora, Kohli, & Palakurthy, 2018).

Este trabajo usa MCO y Mínimos Cuadrados Ponderados (MCP), para el ajuste de los modelos planteados. La validez de las estimaciones por medio de MCO requiere el cumplimiento de varios supuestos sobre los errores. De principal importancia es la ausencia de heterocedasticidad. Una de las alternativas ante violaciones a este requerimiento es el enfoque de Mínimos Cuadrados Ponderados (MCP), los cuales asignan menor peso en la estimación a los *outliers*, para reducir su nivel de influencia en la línea de regresión (Wooldridge, 2010).

Los resultados de este trabajo pueden ser útiles al menos para tres propósitos: (i) contar con información que explique la generación de precios en el mercado de autos de segunda mano en El Salvador; (ii) proveer insumos para la apoyar la estrategia comercial, tanto de los proveedores de autos nuevos, como usados y (iii) reducir la asimetría de información entre compradores y vendedores, y así fortalecer la eficiencia en ese mercado.

Metodología de la investigación

Este trabajo cuantitativo es de corte exploratorio y busca establecer precedentes (Sampieri, 2014) en el estudio del precio del parque vehicular en El Salvador. El análisis econométrico se efectúa por medio de regresión lineal. Se usan dos fuentes: (i) el conjunto de datos abiertos sobre el parque vehicular publicado por el Ministerio de Obras Públicas (Ministerio de Obras Públicas de El Salvador, 2018); y (ii) un conjunto de datos de generación propia, por medio de *web scraping*, con base en los datos del sitio en internet www.encuentra24.com.

Los conjuntos de datos fueron manipulados y analizados usando SPSS 23, Microsoft Excel 19.0 y R 3.5.3. De este último lenguaje, se usaron las librerías *rvest* -para el web scraping-, *gvlma*, *tseries*, *MASS* y *sandwich* (adicionales a las utilidades base), para el ajuste y diagnóstico de los modelos estimados por medio de MCO y MCP.

El conjunto de datos del MOP cuenta con 24 variables, entre ellas el valor del vehículo. Se incluyó información solo sobre automotores con placas particulares y con estados activo de circulación. Se creó la variable *Tiempo* (fecha de importación menos fecha de fabricación) y se seleccionó solo aquellos autos que no excedieran 30 años, ni que presentaran inconsistencias (valores negativos, por ejemplo). Este proceso de filtro dejó 407, 074 observaciones válidas para el análisis.

Se creó una variable dicotómica *para cada marca presente* en la selección final, que representase al menos 1 % en el total de observaciones (19 marcas, que aglutinan 78 % de los autos incluidos), y el resto fue clasificada como *Otras*. La marca Toyota representa el 21 % de los autos considerados, por lo que se tomó como categoría base del análisis. Esto implica que no entra como variable binaria, y que el resto de estimadores categóricos deben compararse contra esta clase. De las marcas iniciales, 2 no presentan efectos estadísticamente significativos¹ (Isuzu y Daewoo), por lo que fueron excluidas del modelo final.

Para controlar por el tamaño del carro se creó la variable *DistCil*, con base en el cálculo de los quintiles de la distribución del cilindraje de los autos de la muestra. Así, toma el valor de 1 si el valor del cilindraje es menor al quintil 1, y así sucesivamente. Se creó la variable GNB, que toma el valor de 1 si el color del auto es gris, negro o blanco (47 % tiene alguno de estos colores en la muestra). Finalmente se creó la variable Nuevo, la cual toma como 1 si el carro fue importado como nuevo y 0 como usado.

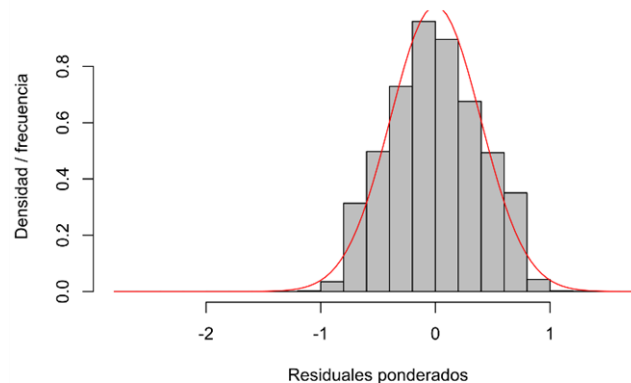
Ante la presencia de heteroscedasticidad en el modelo por MCO, usando la base del MOP, se procedió con el ajuste de un modelo por MCP. Las ponderaciones de las observaciones fueron estimadas por medio de un ajuste auxiliar de regresión robusta, usando el *método de Hubber*. A continuación, la ecuación teórica en la que se base el modelo econométrico ajustado: ***Precio = β_0 + Marca + Tamaño + Antigüedad + Color + Carro Nuevo.***

¹ Los coeficientes asociados a estas variables no fueron significativos en la estimación del modelo robusto inicial para el cálculo de las ponderaciones.

La especificación final de modelo econométrico sigue una forma funcional Log – Nivel. Esto implica que cambios en una unidad de las variables predictoras se interpretan como cambio en puntos porcentuales de la variable dependiente (precio), igual al producto del valor del coeficiente multiplicado por 100 ($\beta \cdot 100$).

Al examinar gráficamente la distribución de los residuales ponderados (**Figura 1**), se tiene que estos siguen una distribución muy similar a una campana de Gauss. De igual manera, al aplicar la prueba de Jarque-Bera sobre la distribución de los residuales ponderados, se rechaza la hipótesis nula (X -cuadrado = 8218.3, $gl = 2$, $p < 2.2e-16$), con lo que se concluye que no hay evidencia que indique que los residuales ponderados no sigan una distribución normal.

Figura 1. Histograma de residuales ponderados y campana de Gauss



Fuente: elaboración propia con base en conjunto de datos del MOP

Por otra parte, el conjunto de datos construidos ($n = 869$), por medio de web scraping fue usado para ajustar una serie de modelos, con base en la información provista y variables generadas para el ajuste. Esto con el objeto de examinar si la lógica transversal identificada en el análisis de la base del MOP, se observaba de igual forma en el modelo ajustado usando el conjunto de datos base de construcción propia.

Análisis y discusión de los resultados

El modelo por MCP usando los datos del parque vehicular del MOP es estadísticamente significativo (Estadístico-F = $5.453e+04$, con 21 y 470,052 gl ; valor-p: $< 2.2e-16$) y presenta un R-cuadrado múltiple de 0.709. Esto da cuenta de

la validez del modelo. A continuación, se presentan los coeficientes para cada una de las variables incluidas en la estimación. Todas son significativas con un alfa de 0.001.

Cuadro 1. Coeficientes del modelo econométrico ajustado

Predictores	Coeficientes	Error estándar	t
Intercepto	8.778	0.002	4124.201
<i>Nuevo</i>	0.800	0.002	393.725
BMW	0.365	0.006	57.139
<i>DistCil</i>	0.179	0.001	318.896
<i>Otros</i>	0.055	0.003	15.989
GEO	-0.062	0.006	-9.674
PEUGEOT	-0.074	0.008	-9.159
VOLKSWAGEN	-0.089	0.006	-16.198
<i>Tiempo</i>	-0.106	0.000	-473.238
GBN	-0.140	0.001	-116.297
HONDA	-0.146	0.002	-63.502
NISSAN	-0.301	0.002	-145.814
SUZUKI	-0.338	0.004	-76.268
JEEP	-0.346	0.005	-66.819
SCION	-0.349	0.005	-64.120
MAZDA	-0.365	0.003	-117.663
MITSUBISHI	-0.399	0.003	-146.169
CHEVROLET	-0.521	0.003	-175.887
HYUNDAI	-0.524	0.002	-241.700
FORD	-0.531	0.003	-166.813
KIA	-0.553	0.002	-253.311
DODGE	-0.735	0.007	-103.283

Fuente: elaboración propia con base en conjunto de datos del MOP

Se estima que, en promedio, un auto nuevo importado sería 80 % más caro que uno usado, sin considerar ninguna variable adicional y considerando la marca base como Toyota. A pesar de la gran magnitud de este estimador, esto no resulta inesperado si se considera que se asume que un auto nuevo no ha enfrentado ningún tipo de depreciación, por lo cual resguarda todo su valor y cuenta con el Premium de primer uso. También, al momento de transar un carro usado, no se consideran sobrepuestos, tales como todos los impuestos que si deben computarse al vender y matricular un carro por primera vez.

Finalmente, debe destacarse que la condición para matricular un automotor es simplemente que su motor funcione, por lo que es posible contar con la importación de autos (y registrarlos al parque vehicular), aun cuando requieran fuertes inversiones para su puesta a punto, lo cual implicaría reducciones considerables en su valor.

Solo dos de las categorías de marcas consideradas (BMW y Otros) se valoran de mejor forma que la marca base (Toyota). Esto parece lógico, ya que la marca BMW suele considerarse de gama alta, y en la categoría Otros, queda recogido el efecto de otras marcas, plausiblemente estimadas como de "lujo", tales como Porsche y Mercedes-Benz. Por el contrario, 87.5 % de las marcas incluidas (15 de 17), en promedio, son menos apreciadas que la marca de base, Toyota. En promedio este descuento es de 35 %.

Que el automóvil sea de color gris, blanco o negro (GBN) disminuye el precio del auto en 14 %. Al ser los colores más comunes, este resultado podría ser contra intuitivo, ya que estos atributos se refieren a las preferencias más recurrentes entre los consumidores.

El comercio de autos tiene lugar en el mercado de los bienes durables, por lo que la compra de un activo similar no es tan recurrente como en los bienes de consumo. Lo anterior aumenta la probabilidad que el consumidor sea menos sensible a los costos de búsqueda para identificar la compra que maximiza su utilidad.

Por lo anterior, el efecto negativo de la variable GBN puede interpretarse como una mayor oferta de carros con estos colores en el mercado, lo cual permite, a un consumidor con preferencia específica a uno de ellos, escoger entre un parque mayor de posibles autos a adquirir, lo cual reduciría el poder de negociación de los vendedores y, por tanto, reducir el precio.

El tamaño del motor del auto está relacionado de forma positiva con su valor. Por cada aumento de una unidad en la variable *DistCil* el precio del carro aumenta en 17.9 %. Esto es lógico con la idea que autos más grandes (por ejemplo, camionetas) son consideradas como un bien superior, en comparación con los sedan (Vanherle

& Vergeer, 2016). Esta variable presenta limitaciones para comprender otros posibles elementos a considerar. Por ejemplo, existen vehículos cuyas carrocerías atienden a las de camionetas, pero cuyo cilindraje es el mismo al de un sedán promedio.

Es plausible pensar que, de cara al consumidor, el indicativo sobre la superioridad como bien, entre un activo y otro, radique en su tamaño físico y no en su cilindraje. Si bien esta variable recoge parte de la variación producida por el hecho que un sea camioneta, es posible que con este modelo se subestime el valor ajustado de automotores con carrocería de un auto grande, pero con cilindrada promedio.

Se encuentra evidencia que indica que con cada año transcurrido entre el año de fabricación y la importación del automóvil (*Tiempo*), el carro pierde 10.6 % de su valor. Esto indicaría que al cabo de 10 años un activo de esta clase tendría un valor nulo, lo cual es implausible.

Respecto a lo anterior, es posible pensar que la función de pérdida de valor de un auto no corresponda a una relación completamente lineal, y que en su lugar siga la forma de una U invertida: en los primeros años se registrarían grandes pérdidas de valor (al perder la condición de nuevo), luego esta se estabilizaría para dar paso a períodos en las que el auto vuelva a ganar valor (autos clásicos). Esta acotación sugiere explorar la validez de estas premisas, por medio de extensiones al presente modelo, las cuales incluyan variables categóricas para controlar por diferentes períodos de longevidad, permitiendo el ajuste de modelos de pendientes múltiples.

Conclusiones y recomendaciones

El modelo presentado brinda información para comprender las variables relacionadas con la variación de los precios de los carros en El Salvador. Para todas las marcas no consideradas de lujo, Toyota es a la que se le asigna mayor valor, puesto que el resto de marcas cuentan con coeficientes negativos, respecto a esa marca.

Se establece que autos con más antigüedad pierden más valor en el tiempo, pero que es necesario explorar posibles relaciones no lineales entre tiempo y precio. Que

un carro sea de color gris, negro o blanco se asocia con menor valor, presumiblemente por su abundancia en el mercado.

La manipulación de los datos permitió advertir vacíos sustanciales en el conjunto de datos que pudiesen ayudar a ajustar mejor el modelo, tales como el tipo de transmisión y combustible; el número de asientos y puertas del automotor, sobre las cuales existen datos, pero de manera incompleta, en la fuente del MOP.

Se ha comprobado que las variables sugeridas por la literatura si aplican al mercado salvadoreño, luego de ajustar un modelo de MCP con la base del MOP. No obstante, al estimar un modelo hedónico con las variables del conjunto de datos por medio de web scraping, se obtuvieron modelos válidos (significativos de forma conjunta y a nivel de coeficientes), pero con R cuadrados menores a 0.1.

Dado lo anterior, es posible pensar que los precios listados en la página de clasificados consultada no son función de los atributos o valor del carro, sino más bien que este precio ofrecido está fuertemente influenciado por el vendedor, quien asigna una perturbación al precio (hacia arriba o hacia abajo), en función de su disposición y expectativa de negociación por la venta de un carro. Estos diferenciales son posibles y surgen debido a las asimetrías de información planteadas por (Akerlof, 1970). Con base en las conclusiones anteriores se postulan las siguientes recomendaciones:

1. Exigir que los datos requeridos en el formulario del Viceministerio de Transporte, sobre la matrícula de autos por primera vez sea siempre llenada y registrada de manera completa. Además, se sugiere estandarizar y ampliar el conjunto de opciones de la casilla "CLASE" de dicho formulario, de tal suerte que se puedan tipificar los automóviles en función del tamaño y formato de su carrocería (sedan, camioneta, etc.).
2. Que instituciones con competencia en Defensa del Consumidor diseñen y pongan en práctica un sistema de certificación para sitios de precios en línea de carros usados con los cuales se estandarice el tipo y la calidad de la información publicada. Esto sería de beneficio para consumidores, ya que señalaría a los

sitios web con mayor confiabilidad y los portales de precios certificados podrían atraer más tráfico, tanto de consumidores como de vendedores, con lo cual podrán incrementar su potencial de monetización de la información con al que cuentan.

3. Promover el análisis cuantitativo, por parte de otros investigadores e instituciones relevantes, del mercado de autos en El Salvador, y así promover la estimación de modelos predictivos de los precios, los cuales puedan ser alojados en sitios web y consultados por el público general. Esto para contribuir a reducir la brecha de información entre consumidores y proveedores, tal como lo proponen (Pal, Sundararaman, Arora, Kohli, & Palakurthy, 2018).

Bibliografía

- Akerlof, G. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, 488-500.
- Cumhur, E., & Senturk, I. (2009). Hedonic Analysis of Used Car Prices in Turkey. *International Journal of Economic Perspectives*, 141-149.
- Ministerio de Obras Públicas de El Salvador. (11 de 11 de 2018). *Datos MOP*. Obtenido de Parque vehicular: <http://datos.mop.gob.sv/?q=dataset/parque-vehicular>
- Molina, K. (31 de Julio de 2017). La industria automotriz detrás de los carros usados. *El Diario De Hoy*.
- Pal, N., Sundararaman, D., Arora, P., Kohli, P., & Palakurthy, S. (2018). How much is my car worth? A methodology for predicting used cars prices using Random Forest. *Future of Information and Communications Conference (FICC)*, 1-6.
- Ramírez, G., Uriarte, J., Delbianco, F., & Larrosa, J. (2017). Un modelo hedónico de precios en línea de automóviles usados en Argentina. *Revista de métodos cuantitativos para la economía y la empresa*, 25-53.
- Sampieri, R. H. (2014). *Metodología de la investigación*. México DF: McGrawHill Education.
- Vanherle, K., & Vergeer, R. (2016). *Data gathering and analysis to improve the understanding of 2nd hand car and LDV markets and implications for the cost effectiveness and social equity of LDV CO2 regulations*. Bruselas: DG Climate Action.
- Wooldridge, J. (2010). *Introducción a la econometría. Un enfoque moderno*. Ciudad de México: Cengage Learning Editores.