Borradores de ECONOMÍA

Heterogeneidad en la fijación de precios en Colombia:Análisis de sus determinantes a partir de modelos de conteo

Por:
Juan Carlos Parra A.
Martha Misas A
Enrique López E.

Núm. 628
2010

BANCO DE LA REPÚBLICA
COLOMBIA

tá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Col

Heterogeneidad en la fijación de precios en Colombia: Análisis de sus determinantes a partir de modelos de conteo*

Juan Carlos Parra A.
School of Economics and Managment
CREATES, Aarhus University
jparra@creates.au.dk

Martha Misas A.
Facultad de Ciencias Económicas y Administrativas
Pontificia Universidad Javeriana
mmisas@javeriana.edu.co

Enrique López E.
Unidad de Investigaciones
Banco de la República
elopezen@banrep.gov.co

Noviembre de 2010

Resumen

En este documento se estudian los determinantes de la heterogeneidad observada en la flexibilidad de precios, empleando los resultados encontrados en una encuesta directa por Misas et al. (2009). Para esto se utilizan los modelos de conteo y se diseñan e implementan un conjunto de pruebas de especificación y de selección de modelos que garantizan una correcta inferencia estadística. En términos generales, los determinantes más significativos a la hora de explicar el grado de flexibilidad de los precios, son las características del producto, los acuerdos contractuales y el sector económico al que pertenece la firma. Por su parte, técnicas de mercadeo, como la posibilidad de discriminar precios entre compradores, son menos significativas. También se encuentra que la existencia de leyes o decretos administrativos, la existencia de líderes en precios y los movimientos de costos no explican la heterogeneidad en el proceso de fijación de los precios.

Palabras Clave: Formación de precios, flexibilidad de precios, modelos de conteo, bootstrapping, Colombia.
Clasificación JEL: C12, C21, C25, C51, E31, L10

---

*La serie Borradores de Economía es una publicación de la Subgerencia de Estudios Económicos del Banco de la República. Los trabajos son de carácter provisional, las opiniones y posibles errores son responsabilidad exclusiva de los autores y sus contenidos no comprometen al Banco de la República ni a su Junta Directiva. Los autores agradecen a Alejandro Gaviria y a Juan Barco por su valiosa colaboración a lo largo de esta investigación.

1. Introducción

De tiempo atrás la teoría económica se ha interrogado acerca de la posible existencia de precios rígidos; también sobre la magnitud del fenómeno y las implicaciones de su existencia[1]. El fenómeno de precios rígidos ha sido estudiado tanto en la macroeconomía como en la microeconomía[2], donde la discusión se ha centrado alrededor de la existencia o no de eficiencia en las asignaciones de recursos que genera su presencia. Independientemente de esta preocupación, la literatura especializada ha concluido que el comportamiento de los mercados es diferente dependiendo de si los precios son rígidos o flexibles.

Hoy en día la evidencia empírica no deja dudas acerca de la existencia de las rigideces de precios. La escuela clásica ha aceptado ese rasgo keynesiano para conformar la llamada Nueva Síntesis Neoclásica (o Neokeynesiana). Las características más importantes de los Modelos Neokeynesianos son la optimización intertemporal, las expectativas racionales, la competencia imperfecta y las rigideces de corto plazo en los precios y salarios que causan distorsiones en los mercados. Las distorsiones llevan a pérdidas de bienestar, las cuales se pueden contener por el accionar de la autoridad monetaria ya sea por una modificación de la tasa de interés o del volumen de dinero que irriga a la economía. Sin embargo, una política monetaria exitosa requiere de una profunda comprensión de la magnitud y el carácter de esas rigideces y de las implicaciones que conlleva incorporarlas en la modelación macroeconómica.

Los esquemas de rigidez de precios más estudiados, en términos de sus implicaciones de política y de bienestar, han sido los de Taylor (1980) y Calvo (1983). Mientras que el modelo de Taylor está basado en duraciones fijas de los precios, el modelo de Calvo incluye algún grado de incertidumbre, y define una probabilidad para determinar el grado de rigidez de aquellos. Cada fijador de precios o salarios está dotado con una probabilidad predefinida para ajustar esos valores en cada período. Aquellos con una baja probabilidad pueden tener restringida su capacidad de reaccionar a un choque exógeno, aún cuando lo quisieran hacer. Otro grupo de determinadores de precios y salarios podrían, por el contrario, ajustarse más rápido. Con la existencia de mercados en competencia monopolística que le permitan a las firmas tener algún grado de poder de mercado, ambas especificaciones permiten dar fundamentos microeconómicos a la curva de Phillips Neokeynesiana. Ésta es ampliamente utilizada hoy en día por muchos

[1]"*That the price of linen and woollen cloth is liable neither to such frequent nor to such great variations as the price of corn, every man's experience will inform him*". Esta cita de Adam Smith hecha por Gordon (1981) ilustra la larga historia que en la teoría económica tiene el problema del ajuste en los precios.

[2]Las primeras investigaciones relacionadas datan del período de la post crisis de los años 30 con las ideas de Means (1936), quien desde el pensamiento microeconómico sugirió la presencia de precios "administrados" o inflexibles, que de acuerdo con sus hipótesis tenían efectos perversos sobre el funcionamiento del sistema económico, y eran en últimas los responsables de que el *laissez faire* no funcionara perfectamente. Su análisis despertó el interés de otros economistas por el estudio del ajuste de los precios. Ruggles (1955) realiza un recuento de las distintas definiciones de flexibilidad e inflexibilidad de los precios existentes hasta la primera mitad del siglo pasado.

bancos centrales para realizar análisis de política monetaria y pronosticar el comportamiento de los principales agregados macroeconómicos. Dichos esquemas de formación de precios hacen parte de un gran conjunto de teorías que se han desarrollado durante las últimas décadas para explicar el lento ajuste de los precios.

La validez empírica de estas consideraciones teóricas ha sido investigada extensivamente. Bils y Klenow (2004) realizaron uno de los primeros estudios comprehensivos sobre los cambios de los precios y presentaron información relacionada sobre los precios al consumidor en Estados Unidos. Los autores encuentran que los modelos tradicionales con rigideces de precios *á la* Calvo o *á la* Taylor predicen por exceso la persistencia y por defecto una menor volatilidad de la inflación de bienes de consumo que registran una menor frecuencia de ajuste de precios. Su principal hallazgo, en efecto, se refiere a que la rigidez de precios es heterogénea entre las firmas. Por ejemplo, los precios de la energía o de los alimentos son muy volátiles en el tiempo, mientras que los precios de los bienes sofisticados de consumo durable exhiben un relativo alto grado de rigidez.

A pesar de que heterogeneidades en las rigideces de precios están ampliamente documentadas, muchos de los modelos de equilibrio general dinámicos y estocásticos (DSGE por sus siglas en inglés) utilizados por los bancos centrales incorporan solo un sector dotado con algún grado de rigidez. Sin embargo, la presencia de heterogeneidad puede generar dinámicas distintas a aquellas obtenidas en los modelos de agente representativo. En efecto, Carvalho (2006) muestra como en un modelo de precios rígidos con heterogeneidad sectorial en la frecuencia de cambio de precios, los efectos de choques de política monetaria tienen efectos reales más persistentes y de mayor duración que en un modelo de agente representativo con firmas idénticas y con rigideces nominales y reales similares. De otro lado, la existencia de diferencias sectoriales en el proceso de fijación de precios puede tener implicaciones en la formulación de la política monetaria. Al respecto, Aoki (2001), muestra que el banco central debería elegir como objetivo de política monetaria un índice de precios que asigne un mayor peso relativo a aquellos sectores (o regiones en el caso de Benigno (2004)) cuyos precios sean menos flexibles, con el objeto de maximizar el bienestar de la sociedad. Morales y Jaramillo (1995) empleando datos para Colombia llegan a conclusiones similares y sugieren que las decisiones de política monetaria del banco central se basen en una desagregación del IPC que controle por heterogeneidad en el grado de flexibilidad de los precios.

Dada la relevancia de la heterogeneidad en el proceso de fijación de precios para la banca central, este trabajo pretende agregar elementos de análisis para la elaboración y conducción de la política monetaria. En particular, el documento explora el papel de un conjunto de factores a la hora de explicar los diferentes grados de flexibilidad de precios registrados por las industrias colombianas. El aspecto más novedoso del trabajo es la utilización de modelos de conteo para el estudio de dicha flexibilidad. Para la implementación de esta técnica se emplean los resultados de la encuesta realizada entre 2007 y 2008 por Misas et al. (2009) a un conjunto de empresas colombianas. Los datos fueron recolectados con entrevistas personales a los responsables de la determinación de precios dentro de las 786 firmas que conforman la muestra seleccionada. La base de datos comprende a un amplio rango de empresas con características de mercado específicas. También se documenta, en dicha base de datos, el reconocimiento por parte de los empresarios de teorías tradicionales de precios rígidos que serían relevantes para su actividad. Con la información extraída de la encuesta, los autores concluyen que las industrias del sector primario, agricultura, caza y pesca, son más flexibles que aquellas empresas del sector industrial, lo cual sugiere cierto grado de heterogeneidad en el proceso de fijación de precios. El análisis que acá se realiza parte del estudio de la distribución de frecuencias de ajustes de precios reportadas por los encuestados[3].

Julio (2010) y Zárate (2010) estudian éste fenómeno previamente para Colombia a partir de estimaciones de funciones Hazard con los micro datos provenientes del IPC y del IPP, respectivamente. Los autores encuentran como determinantes de la heterogeneidad en la fijación de precios, el tipo de minorista que reporta, variables de estado como la inflación sectorial y el PIB, la existencia de riesgos competitivos debido a la decisión de aumentar o disminuir los precios, la durabilidad del bien, entre otros.

El documento se desarrolla en cinco partes, de las cuales la primera es esta introducción. El resto del artículo se organiza de la siguiente manera. La sección 2 introduce los modelos econométricos de conteo que permiten estudiar el grado de flexibilidad de los precios y el detalle de sus determinantes. La sección 3 describe un conjunto de pruebas estadísticas que permiten verificar los supuestos sobre los cuales se construyen los modelos econométricos estudiados. La sección 4 describe los principales resultados y finalmente la sección 5 concluye.

2. Modelos de conteo

En esta sección se presentan los modelos de conteo utilizados para examinar los factores determinantes de la heterogeneidad en las rigideces de los precios en Colombia. Tal vez con la excepción de Álvarez y Hernando (2006) la técnica no ha sido muy aplicada para el examen de datos de precios; sin embargo, se constituye en una herramienta muy apropiada en la medida en que permite relacionar la frecuencia del número de cambios con un grupo de variables explicativas.

[3]Es importante anotar que en los resultados de la encuesta de precios se detectaron valores atípicos. En consecuencia y como se explica en el apéndice A, para realizar algunos de los ejercicios fue necesario remover dichas observaciones.

2.1. Enfoque econométrico para la modelación del ajuste de precios

El estudio de los determinantes de la heterogeneidad en la frecuencia de los cambios de precios que se presenta en la sección 4 emplea un amplio conjunto de información proveniente de los resultados de la encuesta realizada a 786 firmas colombianas elegidas a través de un procedimiento de muestreo aleatorio estratificado por Misas et al. (2009). En términos generales, se cuenta con un conjunto de observaciones de corte transversal $\{(Y_i, \mathbf{X}_i), i = 1, \ldots, N\}$ las cuales se suponen independientes e idénticamente distribuidas. Y corresponde a un vector $(N \times 1)$ de variables aleatorias discretas que toman únicamente valores enteros no negativos y que denotan el número de cambios de precios realizados por las compañías entrevistadas durante los doce meses previos a la encuesta. Dada la naturaleza del vector aleatorio, se dice que las variables que conforman a Y son variables de conteo. Por su parte, $\mathbf{X}$ corresponde a una matriz $(N \times K)$ de variables explicativas que se supone no estocástica, de tal manera que la i-ésima fila está conformada por la información de las K variables correspondientes al i-ésimo individuo, $\mathbf{x}_i' = (x_{i1}, x_{i2}, \ldots, x_{ik})$. Las K variables que conforman dicha matriz se describen en la sección 2.2.

El Cuadro 1 presenta la distribución de frecuencias del número de cambios de precios registrado por las firmas encuestadas. El 17.43 % de las firmas no cambiaron sus precios en los doce meses previos a la realización de la encuesta, mientras que el 57.51 % realizaron como máximo dos cambios. Si bien el número de firmas con cambios superiores a dos es reducido, se deben tener en cuenta en la modelación. En particular, las colas de la distribución son relevantes en el estudio de la heterogeneidad de precios en sectores como la agricultura, sector en el cual se encuentra un mayor número de cambios de precios por año. Adicionalmente, el hecho de contar con observaciones en las colas de la distribución puede sugerir el uso de una especificación econométrica frente a otra, como se detalla más adelante.

En función de la variabilidad de Y, existen varias alternativas para estudiar su relación con las variables contenidas en la matriz de diseño $\mathbf{X}$. El modelo de regresión de Poisson es el punto de partida para modelar datos de conteo. Sea y_i, con $y_i = 0, 1, 2, \ldots$, una realización de una variable aleatoria con función de distribución de probabilidad de Poisson. Así, $Y_i \sim \mathcal{P}(m_i)$ y la probabilidad de ocurrencia de un evento particular, y_i, viene dada por:

$$\Pr[Y_i = y_i \mid m_i] = \frac{\exp[-m_i]\, m_i^{y_i}}{m_i!}, \qquad i = 1, \ldots, N \; ; \; y_i = 0, 1, 2, \ldots \tag{1}$$

donde m_i corresponde a la tasa de ocurrencia media del evento en estudio. Bajo el supuesto de que la variable aleatoria se distribuye efectivamente Poisson, entonces se verifica la propiedad de equidispersión según la cual, los dos primeros momentos de la distribución son iguales. Estos se definen como:

Cuadro 1: Frecuencia de ajuste de precios

No. de cambios	Conteo	Frecuencia (%)	Frec. Acum. (%)
0	137	17.43	17.43
1	299	38.04	55.47
2	153	19.47	74.94
3	55	7.00	81.93
4	29	3.69	85.62
5	19	2.42	88.04
6	16	2.04	90.59
7	4	0.51	91.60
8	8	1.02	91.60
9	0	0.00	91.60
10	9	1.15	92.75
11	0	0.00	92.75
12	20	2.54	95.29
13-24	14	1.78	97.07
25-50	14	1.78	98.85
51-99	4	0.51	99.36
100+	5	0.64	100
Total	786	100	

Fuente: Elaboración propia

$$\mathbb{E}\left(Y_i\right) = \operatorname{Var}\left(Y_i\right) = m_i \tag{2}$$

Con el objeto de introducir covariables en la probabilidad de ocurrencia de los eventos, así como garantizar una media no negativa, es usual parametrizar el valor esperado condicional del modelo de regresión de Poisson de la siguiente manera:

$$m\left(\mathbf{x}_i\right) = \exp\left(\mathbf{x}_i'\beta\right) \tag{3}$$

El modelo de regresión de Poisson permite modelar con esta especificación cierto grado de heteroscedasticidad observable. Con una correcta especificación de la media, los parámetros del modelo pueden ser estimados consistentemente por medio del método de máxima verosimilitud (MV). Sin embargo, para obtener estimadores eficientes, se debe verificar el supuesto de que los datos están efectivamente distribuidos Poisson. Como se presenta en la sección 3.1, es posible validar la correcta especificación de la media y por lo tanto la consistencia de los estimadores. Sin embargo, probar la correcta especificación de toda la función de densidad es más complejo.

En caso de que los datos no presenten una distribución de Poisson, aún es posible obtener estimadores consistentes y eficientes por medio del método de pseudo-máxima verosimilitud (PMV). Bajo este enfoque, una incorrecta especificación del proceso generador de datos resulta irrelevante y la eficiencia se puede alcanzar realizando una estimación robusta de la matriz de varianza-covarianza. Los estimadores de las matrices de varianza-covarianza bajo los dos enfoques de estimación vienen dados por:

$$\hat{\Sigma}(\beta)_{MV} = \sum_{i=1}^{N} \hat{m}(\mathbf{x}_i)\, \mathbf{x}_i \mathbf{x}_i^{'} \tag{4}$$

$$\hat{\Sigma}(\beta)_{PMV} = \left(\sum_{i=1}^{N} \hat{m}(\mathbf{x}_i)\, \mathbf{x}_i \mathbf{x}_i'\right)^{-1} \left(\sum_{i=1}^{N} \hat{\omega}_i \mathbf{x}_i \mathbf{x}_i'\right) \left(\sum_{i=1}^{N} \hat{m}(\mathbf{x}_i)\, \mathbf{x}_i \mathbf{x}_i'\right)^{-1} \tag{5}$$

donde $\hat{\omega}_i = \hat{\omega}(\mathbf{x}_i)$ es una estimación de la varianza condicional de Y_i, $\omega(\mathbf{x}_i) = \mathrm{Var}(Y_i \mid \mathbf{x}_i)$, $\hat{m}(\mathbf{x}_i) = \exp\left(\mathbf{x}_i'\hat{\beta}\right)$ y $\hat{\beta}$ es la solución al sistema no lineal, $\sum_{i=1}^{N}(y_i - \exp(\mathbf{x}_i'\beta))\,\mathbf{x}_i = \mathbf{0}$, resultante de las K condiciones de primer orden del problema de log-máxima o log-pseudo-máxima verosimilitud de acuerdo con el supuesto de independencia entre las observaciones:

$$\ln\mathcal{L}(\beta) = \sum_{i=1}^{N} \left\{ y_i \mathbf{x}_i'\beta - \exp\left(\mathbf{x}_i'\beta\right) - \ln y_i! \right\} \tag{6}$$

Con el supuesto de una media condicional exponencial correctamente especificada es posible mostrar que:

$$\hat{\beta} \overset{a}{\sim} \mathcal{N}\left(\beta, \hat{\Sigma}(\beta)_j\right),\ j = MV, PMV \tag{7}$$

Es importante anotar que: 1) los dos estimadores, MV y PMV, son idénticos pero con matrices de varianza-covarianza diferentes y 2) si el proceso generador de datos de la variable de conteo es efectivamente Poisson, entonces se tiene que $\hat{\Sigma}(\beta)_{MV} = \hat{\Sigma}(\beta)_{PMV}$ dado que $\omega(\mathbf{x}_i) = m(\mathbf{x}_i)$.

Una vez estimado el modelo de regresión de Poisson, este puede ser utilizado para estudiar el efecto marginal (EM) de cambios en alguna de sus covariables sobre el valor esperado del fenómeno estudiado. Dada la especificación exponencial de la media condicional, el efecto marginal de un cambio en una covariable continua, x_j, viene dado por:

$$\hat{\mathrm{EM}}_j = \frac{\partial \mathbb{E}\,(\hat{Y} \mid \mathbf{x})}{\partial x_j} = \hat{\beta}_j \exp\left(\mathbf{x}'\hat{\beta}\right) \tag{8}$$

donde x_j denota el j-ésimo regresor contenido en $\mathbf{x}$. El efecto parcial es función de los parámetros estimados y de todos los regresores. Por lo tanto la magnitud del efecto parcial depende de $\mathbf{x}$ y de $\hat{\beta}$. En este sentido, el parámetro $\hat{\beta}_j$ puede ser interpretado como una semielasticidad, ya que por una unidad de cambio en x_j, la media condicional aumenta en un múltiplo de $\hat{\beta}_j$. Es de señalar que si x_j está expresado en logaritmos, $\hat{\beta}_j$ puede ser interpretado como una elasticidad.

Si por el contrario, el interés radica en calcular el efecto marginal de una variable discreta, d, sobre la media condicional, entonces:

$$\hat{\mathrm{EM}}_j = \mathbb{E}\left(Y \mid \hat{\mathbf{z}}, d = 1\right) - \mathbb{E}\left(Y \mid \hat{\mathbf{z}}, d = 0\right) \tag{9}$$

donde $\mathbf{z}$ denota todos los regresores diferentes al j-ésimo. Dado que la construcción de los EM dependen del punto donde se evalúen, existen diferentes maneras de calcularlos. Es usual construir EM como el promedio de los efectos marginales sobre cada individuo (PEM), como el efecto en el individuo promedio de los regresores (EMM), o como el efecto sobre un agente representativo (EMR).

Los resultados también pueden ser empleados para calcular el cambio porcentual en el conteo esperado ante cambios de ζ unidades en x_j, manteniendo todas las demás variables constantes:

$$\Delta_j\,\% = 100 \times \frac{\mathbb{E}\left(Y \mid \mathbf{x}, x_j + \zeta\right) - \mathbb{E}\left(Y \mid \mathbf{x}, x_j\right)}{\mathbb{E}\left(Y \mid \mathbf{x}, x_j\right)} = 100 \times \left[\exp\left(\hat{\beta}_j \times \zeta\right) - 1\right] \tag{10}$$

Usualmente, el modelo de regresión de Poisson resulta muy restrictivo dado el supuesto de equidispersión. El problema fundamental es que la distribución está parametrizada en términos de un único parámetro escalar, m, de tal forma que todos los momentos de son funciones de este. En general, los datos de conteo utilizados en aplicaciones empíricas rechazan este supuesto. En particular, es usual encontrar casos en donde la varianza excede la media, situación que se conoce como sobredispersión[4].

La sobredispersión es una característica con consecuencias similares a la falla del supuesto de homoscedasticidad en el modelo clásico de regresión lineal. Dada una especificación correcta para la media condicional, los estimadores MV y de PMV de Poisson siguen siendo consistentes. Sin embargo, es importante controlar por posible sobredispersión ya que en exceso, ésta puede llevar a errores estándar más pequeños y por lo tanto a estadísticos t inapropiados. Una posible forma de corregir la presencia de sobredispersión es emplear el estimador robusto de la matriz de varianza-covarianza dado en (5) bajo algún tipo de parametrización de ω_i consistente con el exceso de varianza.

Una alternativa para modelar la presencia de sobredispersión en los datos parte de la causa misma que la genera. Al respecto, Cameron y Trivedi (1998) muestran como la existencia de heterogeneidad adicional a la explicada por los regresores, es decir, aquella no observable, puede causar este fenómeno[5]. El modelo estándar para capturar la sobredispersión es el modelo Binomial Negativo, el cual parte de la siguiente función de distribución de densidad para Y_i:

[4]También es posible encontrar situaciones en que la media condicional excede la varianza condicional, situación que se conoce como subdispersión.

[5]Sin embargo, pueden existir otras causas que generen la sobredispersión en los datos y que conlleven a estimadores no solo no eficientes sino también inconsistentes.

$$\Pr\left(Y_i = y_i \mid m\left(\mathbf{x}_i\right), \delta_0\right) = \frac{\Gamma\left(y_i + \delta_0^{-1}\right)}{\Gamma\left(\delta_0^{-1}\right)\Gamma\left(y_i+1\right)}\left(\frac{\delta_0^{-1}}{m\left(\mathbf{x}_i\right)+\delta_0^{-1}}\right)^{\delta_0^{-1}}\left(\frac{m\left(\mathbf{x}_i\right)}{m\left(\mathbf{x}_i\right)+\delta_0^{,1}}\right)^{y_i} \tag{11}$$

con $\delta_0 \geq 0$ y $y_i = 0, 1, 2, \ldots$. En caso que $\delta_0 = 0$, entonces la función de distribución (11) colapsa a la de Poisson. Con esta representación de la función de densidad, los conteos son obtenidos como un proceso de Poisson (en el cual los eventos son serialmente independientes), con parámetro de media aleatorio. Este se define como $\tilde{m}\left(\mathbf{x}_i\right) = m\left(\mathbf{x}_i\right) \cdot u_i$, donde $m\left(\mathbf{x}_i\right) = \exp\left(\mathbf{x}_i'\beta\right)$ y $u_i = \exp\left(\upsilon\right)$ y υ representa la heterogeneidad no observada la cual se considera en si misma una variable aleatoria con función de densidad definida[6].

Los dos primeros momentos de la distribución Binomial Negativa que se obtienen al considerar la presencia de heterogeneidad no observada son:

$$\begin{aligned} \mathbb{E}\left[Y_i \mid \mathbf{x}_i, \delta_0\right] &= m\left(\mathbf{x}_i\right) && (12)\\ \operatorname{Var}\left[Y_i \mid \mathbf{x}_i, \delta_0\right] &= \omega_i\left(\mathbf{x}_i\right) = m\left(\mathbf{x}_i\right)\left(1+\delta_0 m\left(\mathbf{x}_i\right)\right) && (13)\end{aligned}$$

donde $\mathbb{E}\left(Y_i \mid \mathbf{x}_i, \delta_0\right) \leq \operatorname{Var}\left(Y_i \mid \mathbf{x}_i, \delta_0\right)$. El parámetro δ_0 mide la dispersión de los datos respecto a la media. Valores más grandes de δ_0 indican una mayor dispersión. Los parámetros del modelo Binomial Negativo, β y δ_0, pueden ser estimados de manera consistente a través del método de máxima verosimilitud, el cual permite obtener al mismo tiempo un estimador eficiente de la matriz de varianza-covarianza. Los estimadores $\hat{\beta}$ y $\hat{\delta}_0$ corresponden a la solución del sistema de ecuaciones no lineal de dimensión $K \times 1$:

$$\begin{aligned} \sum_{i=1}^{N} \frac{y_i - \exp\left(\mathbf{x}_i'\beta\right)}{1+\exp\left(\mathbf{x}_i'\beta\right)\delta_0}\mathbf{x}_i &= 0 && (14)\\ \sum_{i=1}^{N}\left[\frac{1}{\delta_0^2}\left\{\ln\left(1+\delta_0\exp\left(\mathbf{x}_i'\beta\right)\right) - \sum_{j=0}^{y_i}\frac{1}{\left(\delta_0^{-1}+j-1\right)}\right\} + \frac{y_i-\exp\left(\mathbf{x}_i'\beta\right)}{\delta_0\left(1+\delta_0\exp\left(\mathbf{x}_i'\beta\right)\right)}\right] &= 0 && (15)\end{aligned}$$

que resultan de la función de log-verosimilitud:

$$\ln\mathcal{L}\left(\beta, \delta_0\right) = \sum_{i=1}^{N}\left\{\sum_{j=1}^{y_i}\ln\left(\delta_0^{-1}+j-1\right) - \ln y! - \left(y_i+\delta_0^{-1}\right)\ln\left[1+\exp\left(\mathbf{x}_i'\beta\right)\delta_0\right] + y_i\ln\left(\delta_0\right) + y_i\mathbf{x}_i'\beta\right\} \tag{16}$$

La interpretación de los parámetros estimados es similar a aquella que se realiza sobre el modelo de regresión de Poisson.

[6]La distribución de probabilidad dada por (11) se obtiene partiendo del hecho de que $Y_i \sim \mathcal{P}\left(\tilde{m}\left(\mathbf{x}_i\right)\right)$ y suponiendo que $u_i \sim \Gamma\left(\delta_0\right)$

Una vez estimados los parámetros es posible construir diferentes medidas de evaluación de ajuste, similares a aquellas que se realizan para el modelo clásico de regresión lineal. Si bien los modelos utilizados son no lineales, es posible elaborar un conjunto de residuales que midan que tan distante están los valores predichos de la variable dependiente de sus valores efectivamente observados. Dicho análisis sirve para detectar valores atípicos, ajuste deficiente en alguna parte de la distribución, o problemas en la especificación del modelo.

Sea:

$$r_i = y_i - \hat{m}(\mathbf{x}_i) \tag{17}$$

el residuo básico el cual, de acuerdo con el supuesto de Poisson, es asimétrico y heteroscedástico. Dadas estas características, resulta conveniente definir otras medidas residuales para el análisis de los resultados. Así, el residual de Pearson resulta luego de corregir por heteroscedasticidad el residuo básico:

$$r_{Pi} = \frac{y_i - \hat{m}(\mathbf{x}_i)}{\sqrt{\hat{\omega}(\mathbf{x}_i)}} \tag{18}$$

siendo $\hat{\omega}(\mathbf{x}_i)$ una estimación de la varianza condicional, $\omega(\mathbf{x}_i)$, de Y_i. En el caso del modelo de Poisson se tiene que $\omega(\mathbf{x}_i) = m(\mathbf{x}_i)$ y en el caso del modelo Binomial Negativo se verifica (13). En muestras grandes, el residual de Pearson presenta media cero y varianza constante e igual a uno, pero continúa siendo asimétrico[7].

Finalmente, el residuo más utilizado en la literatura especializada, debido a la cantidad de modelos sobre los cuales puede ser aplicado, es el residuo Deviance, el cual se define como:

$$r_{Di} = \text{signo}\left(y_i - \hat{m}(\mathbf{x}_i)\right)\sqrt{2\left[\mathcal{L}(y_i) - \mathcal{L}(\hat{m}(\mathbf{x}_i))\right]} \tag{19}$$

donde $\mathcal{L}(\hat{m}(\mathbf{x}_i))$ es la log-densidad de Y_i evaluada en $m(\mathbf{x}_i) = \hat{m}(\mathbf{x}_i)$ y $\mathcal{L}(y_i)$ es la log-densidad evaluada en $m(\mathbf{x}_i) = y_i$.

Las diferentes medidas residuales también pueden ser utilizadas en la construcción de medidas de bondad de ajuste global de los modelos estimados. Estas medidas corresponden a sumas ponderadas de residuos, similares a las empleadas en el modelo clásico de regresión lineal. Una medida estándar para evaluar el ajuste de los modelos corresponde al estadístico de Pearson, el cual se define como:

$$P = \sum_{i=1}^{N} \frac{\left(y_i - \hat{m}(\mathbf{x}_i)\right)^2}{\hat{\omega}(\mathbf{x}_i)} \tag{20}$$

[7] Si el interés radica en encontrar residuales que se aproximen a la normalidad, es posible utilizar el residuo Anscombe. Ver Cameron y Trivedi (1998)

Siguiendo a Cameron y Trivedi (1998), con una correcta especificación de la media, $P > N-K$ puede ser interpretado como evidencia a favor de sobredispersión en el modelo básico de Poisson[8]. Visto de otra manera, $P \neq N-K$ indica una incorrecta especificación de la media condicional.

Una medida alternativa para evaluar la bondad del ajuste parte del residuo (19). La suma de residuos al cuadrado, conocido como el estadístico Deviance, se define como:

$$D\left(y, \hat{m}\left(\mathbf{x}\right)\right) = \sum_{i=1}^{N} r_{Di}^2 \tag{21}$$

el cual también es empleado en la detección de valores atípicos tal y como se presenta en el apéndice A. A partir de este estadístico es posible definir una medida de pseudo-R^2:

$$R_{DEV}^2 = 1 - \frac{D\left(y, \hat{m}\left(\mathbf{x}\right)\right)}{D\left(y, \bar{y}\right)} \tag{22}$$

donde $D\left(y, \hat{m}\left(\mathbf{x}\right)\right)$ el el estadístico Deviance del modelo ajustado y $D\left(y, \bar{y}\right)$ es el estadístico Deviance del modelo únicamente con intercepto. El estadístico R_{DEV}^2 permite medir la reducción en el Deviance debido a la inclusión adicional de regresores en el modelo estimado. Una medida similar, obtenida a partir de los residuos de Pearson es el R^2 de Pearson. Para modelos con función de varianza $\omega\left(m\left(\mathbf{x}\right), \delta_0\right)$, dicha medida se define como:

$$R_P^2 = 1 - \frac{\sum_{i=1}^{N} \frac{\left(y_i - \hat{m}\left(\mathbf{x}_i\right)\right)}{\omega\left(\hat{m}\left(\mathbf{x}_i\right), \delta_0\right)}}{\sum_{i=1}^{N} \frac{\left(y_i - \hat{m}_0\right)}{\omega\left(\hat{m}_0, \delta_0\right)}} \tag{23}$$

donde $\hat{m}_0$ es la media estimada en el modelo que contiene únicamente intercepto.

Finalmente, otra medida de ajuste ampliamente utilizada consiste en el pseudo-R^2 construido a partir de la ganancia potencial que se genera en la función de log-verosimilitud al incluir regresores adicionales. Esta medida se define como:

$$R^2 = 1 - \frac{\mathcal{L}_{ajustado}}{\mathcal{L}_0} \tag{24}$$

donde $\mathcal{L}_{ajustado}$ es el valor de la log-verosimilitud del modelo ajustado y $\mathcal{L}_0$ es el valor de la log-verosimilitud del modelo únicamente con constante. Para el caso de variables discretas finitas, como en el caso de variables de conteo, esta medida tiene la propiedad de que $0 \leq R^2 < 1$.

2.2. Determinantes del grado de flexibilidad de precios

En la teoría de la organización industrial, el análisis de la heterogeneidad en el proceso de fijación de precios se ha relacionado con el grado de concentración industrial. Para Carlton

[8] Por su parte, $P < N-K$ puede interpretarse como posible presencia de subdispersión.

(1989), esta relación tiene su origen en el estudio de las diferentes estructuras de mercado cómo mecanismo para analizar la fijación de precios por parte de las firmas. Los modelos básicos de asignación de recursos, competencia perfecta, monopolio y oligopolio, permiten construir un primer conjunto de determinantes del precio. En términos generales, los cambios de precios en que incurren los productores dependen de las condiciones de demanda y oferta, las cuales se recogen a través de sus elasticidades y por lo tanto, de los distintos grados de concentración, y de la estructura de costos de las firmas así como, de los choques de costos a los que están sujetas.

Aún con las grandes implicaciones que tienen los modelos básicos de mercado mencionados anteriormente, algunos autores han encontrado evidencia a favor de mecanismos que vacían los mercados distintos al sistema de precios, lo cual podría explicar el hecho de que los precios no varíen en la cantidad esperada ante cambios en los determinantes inicialmente sugeridos. De acuerdo con Carlton (1989): "...la relación entre los cambios de precios y los cambios de costos varía con la forma de la curva de demanda y por lo tanto, no es posible hacer afirmaciones generales acerca de la variabilidad de los precios con relación a la variabilidad en los costos, basados únicamente en el hecho de que el mercado sea competitivo o monopólico. Más aún, ya que sabemos que los oligopolios se encuentran en medio del espectro entre competencia perfecta y monopolios puros, las simples teorías no permiten hacer predicciones diferenciadas de la flexibilidad de precios (para grandes cambios en costos) que dependan exclusivamente del grado de competencia del mercado" (traducción propia).

Al respecto Backman (1940) realiza una descripción minuciosa sobre los posibles factores que pueden afectar la flexibilidad de los precios. Estos se pueden agrupar en siete categorías las cuales incluyen a su vez otro conjunto de variables. Los grandes determinantes de la heterogeneidad en el proceso de fijación de precios son, de acuerdo con el autor: i) las características del producto, ii) la existencia de leyes o decretos administrativos, iii) la concentración del control, iv) las técnicas de mercadeo, v) los hábitos y las costumbres, vi) los arreglos contractuales y vii) la estructura del mercado.

El conjunto de variables explicativas empleado en este ejercicio corresponde a un subconjunto de las respuestas obtenidas por Misas et al. (2009). Si bien dicho conjunto no contiene información sobre todos los determinantes delineados por Backman (1940), permite realizar una aproximación a las principales variables. Este documento pretende evaluar la importancia de cada uno de estos factores en la explicación de la heterogeneidad en la formación de precios en Colombia. Una descripción detallada de las variables incluidas en los modelos, así como algunas estadísticas descriptivas se encuentran en el apéndice B y C, respectivamente.

En primer lugar, se incluyen características del producto que afectan tanto la demanda como la oferta de los bienes. En términos de los factores que afectan la demanda se incluye el tipo de bien producido ($final$), y la importancia de la demanda a la hora de cambiar el precio

(c_dda). Del lado de la oferta, se incluye la importancia de los cambios en los costos laborales (c_clab), en los costos de materias primas (c_cmp), en los costos financieros (c_cfin), en los costos de la energía y los combustibles (c_cce), cambio en los impuestos (c_imp) y cambios en la tasa de cambio (c_tc). Este grupo de variables, excepto el tipo de bien, se denominan disparadores de precios tanto por el lado de la demanda como por el de la oferta.

De otro lado, se incluye una variable que recoge la posible existencia de regulación o limitación en la fijación de precios por parte del productor, ($restr\_set$), con el propósito de aproximar el segundo grupo de determinantes. Por su parte, la concentración del control hace referencia al poder monopólico por parte de un conjunto de empresarios, lo que Means (1936) llamó precios administrados. Como proxy de este grupo se emplea la existencia o no de líderes de precios en la industria (*líder*) y el número de competidores que perciben los empresarios en su industria ($comp2$). Respecto a las técnicas de mercadeo, se incluye una variable que diferencia entre productores que discriminan precios entre compradores (*discr*) y otra que mide la importancia atribuida por los encuestados a los cambios en calidad de los bienes antes que a los cambios en precios, cuando se enfrentan situaciones que presionan cambios en los segundos (*calidad*).

Para capturar arreglos contractuales, bien sean explícitos o implícitos, se incluye el porcentaje de ventas que realizan los productores con clientes con quienes esperan realizar futuras transacciones (v_lp), y dos variables categóricas que indican la importancia atribuida por los empresarios a los contratos formales, (*explícito*), o informales (*implícito*) para no modificar sus precios. Dentro de la categoría de hábitos, se incluye la regla de revisión de precios que siguen los encuestados ($state1$). Finalmente, la importancia de la demanda interna (v_dom), el tamaño de la firma, tanto por empleados (*empl*) como por activos (*grande*), el tipo de industria (*sector*) y el mecanismo de fijación de precios, bien sea basado en el precio de los competidores ($estr\_com$) o en los costos más un margen ($estr\_mon$) hacen alusión a la estructura de mercado a la cual pertenecen las empresas.

Siguiendo a de de Munnik y Xu (2007) se realiza un análisis preliminar entre la frecuencia de ajuste de precios reportada por los encuestados y cada uno de los determinantes recién descritos, empleando para ello la base de datos completa, es decir, sin remover los valores atípicos. Se llevan a cabo dos ejercicios. El primero consiste en la prueba no paramétrica de igualdad de poblaciones de Kruskall y Wallis aplicada sobre las categorías de cada una de los posibles factores explicativos. De otro lado, se analiza, por medio del coeficiente de correlación de Spearman, la existencia de asociación entre cada uno de estos factores y la frecuencia de cambio de precios. Los resultados para todas las variables, excepto para los disparadores de precios, junto a algunas estadísticas adicionales se presentan en el Cuadro 2.

Cuadro 2: Factores que influencian la frecuencia de ajuste de precios

Determinantes del ajuste de precios	Número de observaciones	Número medio de cambios	≤1 cambio de precio al año (% de firmas)	> 52 cambios de precios al año (% de firmas)
Total Muestra	786	4	431	6
Caracteristicas del producto				
Tipo de bien*[a,d]				
Final	539	5	431	6
Intermedio / Capital	247	3	427	0
Leyes o decretos administrativos				
Regulación / limitaciones para fijar precio*				
Si	442	4	435	2
No	344	5	436	4
Concentración de control				
Existencia de lideres				
Si	171	5	87	2
Si, y se reconoce como lider	144	2	86	0
No	471	5	263	4
Percepción de competencia[a,d]				
Poca (Menos de 5 competidores)	206	2	131	0
Mucha (Más de 5 competidores)	580	5	305	6
Técnicas de mercadeo				
Discriminación de precios entre compradores[a,d]				
Si	553	8	292	5
No	233	4	144	1
Cambio en calidad del producto sobre cambio en precios[b,e]				
Poco y no importante	355	7	183	6
Importante y muy importante	431	3	253	0
Arreglos contractuales				
Porcentaje de ventas totales con clientes de largo plazo				
0 - 25 %	95	12	49	4
26 - 50 %	51	2	27	0
51 - 75 %	97	3	55	0
76 - 100 %	543	4	305	2
Contratos explícitos				
Poco y no importante	288	5	158	2
Importante y muy importante	498	4	278	4
Contratos implícitos[a,d]				
Poco y no importante	226	7	108	5
Importante y muy importante	560	3	328	1

(*Continuación*)

Hábitos y costumbres				
Regla de revisión de precios^{a,d}				
Tiempo dependiente	562	2	340	0
Estado dependiente	224	10	96	6
Estructura de mercado				
Importancia demanda interna (% de ventas)^{e}				
0 - 25 %	29	2	17	0
26 - 50 %	46	2	23	0
51 - 75 %	89	4	54	0
76 - 100 %	622	5	342	6
Número de empleados^{c,f}				
0 - 100	524	5	302	5
101 - 500	214	5	114	1
501 - 1000	31	3	16	0
1001 - 5000	15	7	4	0
Más de 5000	2	2	2	0
Tamaño de la firma (valor de los activos)^{a,d}				
No grande	510	4	300	4
Grande	276	5	136	2
Sector económico^{a,d}				
Agricultura	85	23	15	6
Pesca	8	9	3	0
Industria	693	2	413	0
Estrategia de fijación de precios*				
Precio basado en competidores^{a,d}	733	4	419	5
Costo más un margen ^{a,d}	576	5	302	5

*El número de cambios de precio no suma el total puesto que la respuesta a la pregunta no era excluyente. Ver apéndice A. ^{a} indica el rechazo de la hipótesis nula de igualdad de poblaciones al 99 % de confianza. ^{b} indica el rechazo de la hipótesis nual de igualdad de poblaciones al 95 %. ^{c} indica rechazo de la hipótesis nula de igualdad de poblaciones al 90 %. Por su parte, ^{d,e,f} indican rechazo de la hipótesis nula de correlación nula al 1 %, 5 % y 10 %, respectivamente.

Las pruebas indican una fuerte relación entre la frecuencia de cambio de precios y el tipo de bien producido, el grado de competencia percibido por los empresarios, la discriminación de precios entre diferentes compradores, la existencia de acuerdos implícitos entre vendedores y compradores, el tipo de regla de revisión de precios, el tamaño de la firma, el sector económico y la estrategia de fijación de precios. También se encuentra relación, aunque en menor medida, con el número de empleados de la firma, la importancia de la demanda interna y la importancia de cambios en la calidad del producto sobre cambios en precio. Finalmente, las pruebas de correlación e igualdad poblacional no sugieren asociación de la frecuencia de ajuste con la existencia de regulación en la fijación de los precios, la existencia de líderes de precios, el porcentaje de ventas que las firmas realizan con clientes que consideran de largo plazo y la existencia de contratos explícitos. Si bien estos resultados dan luces sobre los factores que determinan las decisiones de precios por parte de las empresas colombianas, estos pueden llevar a conclusiones erróneas al no considerar todos los factores de manera simultánea. Por lo tanto, dichos resultados se complementan con aquellos que se pueden obtener de los modelos de regresión de conteo.

3. Pruebas de especificación

Para poder alcanzar una estimación consistente de los coeficientes de regresión y poder realizar una correcta inferencia estadística sobre los mismos, es necesario garantizar a su vez una correcta especificación de la media condicional. De otro lado, si bien la estimación de las probabilidades condicionales puede ser obtenida por métodos no paramétricos, los modelos de regresión presentados en la sección anterior requieren de un supuesto acerca de la forma funcional de la distribución de probabilidad condicional. Surge entonces la necesidad de aplicar pruebas que permitan validar la especificación de la media condicional y de la distribución de probabilidad condicional. En esta sección se presenta un conjunto de pruebas estadísticas desarrolladas por Álvarez y Delgado (2002), las cuales se basan en los resultados obtenidos por Stute (1997) y Andrews (1997).

Las pruebas parten del conjunto de información $\{(Y_i, \mathbf{X}_i), i = 1, \ldots, N\}$, el cual se supone es aleatorio, independiente e idénticamente distribuido. Tal como se expuso en la sección anterior, la probabilidad condicional de ocurrencia de un evento se expresa como:

$$P_y(\mathbf{x}_i) = \Pr(Y_i = y_i \mid \mathbf{X}_i = \mathbf{x}_i) \tag{25}$$

y puede ser modelada a partir de dos formas parámetricas alternativas: Poisson y Binominal Negativa, las cuales, en su orden, pueden ser representadas de la siguiente manera:

$$P_y(\mathbf{x}_i) = \frac{\exp(-m(\mathbf{x}_i))\, m(\mathbf{x}_i)^{y_i}}{y_i!} \tag{26}$$

$$P_y(\mathbf{x}_i) = \frac{\Gamma\left(y_i + \delta_0^{-1}\right)}{\Gamma\left(\delta_0^{-1}\right)\Gamma\left(y_i+1\right)}\left(\frac{\delta_0^{-1}}{m(\mathbf{x}_i)+\delta_0^{-1}}\right)^{\delta_0^{-1}}\left(\frac{m(\mathbf{x}_i)}{m(\mathbf{x}_i)+\delta_0^{-1}}\right)^{y_i} \tag{27}$$

siendo $m(\mathbf{x}_i) = \mathbb{E}(Y_i \mid \mathbf{x}_i)$ y $\delta_0 > 0$.

3.1. Prueba sobre la media condicional

Un punto central en la regresión de datos de conteo es establecer la correcta especificación de la media condicional. Así, la prueba propuesta por Álvarez y Delgado (2002) que se presenta a continuación verifica si existe o no evidencia para rechazar el planteamiento de la siguiente hipótesis nula, cuyas alternas representan esquemas no paramétricos:

$$H_0 : \mathbb{E}(Y \mid \mathbf{X}) = m(\mathbf{X}; \beta_0), \qquad \beta_0 \in \mathbf{B} \tag{28}$$

donde $\mathbf{B}$ corresponde al espacio de parámetros tal que $\mathbf{B} \subset \mathbb{R}^K$. Dicha hipótesis nula puede ser reescrita como:

$$H_0 : T(\mathbf{X}, \beta_0) = 0 \qquad \beta_0 \in \mathbf{B}, \tag{29}$$

siendo

$$\begin{aligned} T(\mathbf{x}, \beta_0) &= \int_{-\infty}^{\mathbf{x}} \left(\mathbb{E}(Y \mid \mathbf{X} = \mathbf{u}) - m(\mathbf{u}; \beta_0)\right) dF_x(\mathbf{u}) \\ &= \mathbb{E}\left\{\left[Y - m(\mathbf{x}; \beta_0)\right] \mathbf{1}(\mathbf{X} \leq \mathbf{x})\right\} \end{aligned} \tag{30}$$

donde $F_x(\cdot)$ corresponde a la función de distribución y $\mathbf{1}(A)$ a la función indicadora del evento A. Dado que $\mathbf{X}_i$ contiene múltiples regresores, la función indicadora se representa de la siguiente manera: $\mathbf{1}(\mathbf{X}_i \leq \mathbf{x}_i) = \prod_{j=1}^{K} \mathbf{1}(\mathbf{X}_{ij} \leq \mathbf{x}_j)$. Así, $T(x; \beta_0)$ corresponde a la diferencia entre la curva de regresión parámetrica bajo la hipótesis nula y las alternativas no parámetricas.

Un estimador consistente de $T(\mathbf{x}; \beta_0)$ bajo la nula, H_0, es:

$$\hat{T}_N(x) = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - m\left(\mathbf{X}_i; \hat{\beta}_N\right)\right)\mathbf{1}(\mathbf{X}_i \leq \mathbf{x}_i) \tag{31}$$

La prueba se basa en la estadística de Cramér-Von Mises:

$$\hat{C}_N = \frac{1}{N}\sum_{i=1}^{N}\left[N^{\frac{1}{2}}\hat{T}_N(\mathbf{X}_i)\right]^2 \tag{32}$$

De acuerdo con Stute (1997), la estimación de la distribución límite de $\hat{C}_N$ se lleva a cabo a través de la aplicación de un "*Wild-Bootstrapping*" sobre los residuales. El procedimiento se detalla a continuación.

Procedimiento "Wild-Bootstrapping"

1. A partir del conjunto de información original $\mathcal{T}_N = \{(Y_i, \mathbf{X}_i), i = 1, \ldots, N\}$ se obtiene un estimador consistente de β_0, $\hat{\beta}_N$ y se construye el estadístico $\hat{C}_N$.

2. Se genera de forma aleatoria una variable *iid* acotada, $V_i,\ i = 1, \ldots, N$, tal que $\mathbb{E}[V_i] = 0$ y $\mathbb{E}\left[V_i^2\right] = 1$, la cual debe conllevar la propiedad de asimetría de las perturbaciones subyacentes de la regresión. En la práctica se selecciona para V_i una distribución discreta con valores $-\frac{(\sqrt{5}+1)}{2}$ y $\frac{(\sqrt{5}+1)}{2}$ y con puntos de masa en $\frac{(\sqrt{5}+1)}{2\sqrt{5}}$y $\frac{(\sqrt{5}-1)}{2\sqrt{5}}$, respectivamente, similar a aquella presentada en el apéndice A.

3. Se construye un nuevo conjunto de información $\mathcal{T}_N^* = \{(Y_i^*, \mathbf{X}_i), i = 1, \ldots, N\}$ a partir del conjunto de información original donde $y_i^* = m\left(\mathbf{x}_i; \hat{\beta_N}\right) + \epsilon_i^*$, y $\epsilon_i^* = \hat{\epsilon_i} V_i$ con $\hat{\epsilon_i} = y_i - m\left(\mathbf{x}_i; \hat{\beta_N}\right)$.

4. A partir de la nueva muestra $\mathcal{T}_N^*$ se reestima el modelo y se encuentra $\hat{\beta}_N^*$ y se calcula la versión de remuestreo de $\hat{C}_N$ de la siguiente forma:

$$\hat{C}_N^* = \frac{1}{N} \sum_{i=1}^{N} \left[N^{\frac{1}{2}} \hat{T}_N^* (\mathbf{X}_i)\right]^2 \tag{33}$$

   donde

$$\hat{T}_N^* (x) = \frac{1}{N} \sum_{i=1}^{N} \left(Y_i^* - m\left(\mathbf{X}_i; \hat{\beta}_N^*\right)\right) \mathbf{1} (\mathbf{X}_i \leq \mathbf{x}_i) \tag{34}$$

5. Los pasos 2 a 4 se repiten B veces, generando $\mathcal{T}_N^{*(b)}, b = 1, \ldots, B$ muestras o nuevos conjuntos de información y para cada uno de ellos la versión "*bootstrap*" de $\hat{C}_N$, $\hat{C}_N^{*(b)}, b = 1, \ldots, B$. Los valores críticos de la prueba son estimados a través de los cuantiles condicionales de $\hat{C}_N^{*(b)}, b = 1, \ldots, B$. Así, a un nivel de significancia α, el valor crítico "*bootstrap*" puede ser aproximado por $\frac{1}{B} \sum_{b=1}^{B} \mathbf{1} \left(\hat{C}_N^{*(b)} > \hat{C}_{\alpha N}^{*(b)}\right) = \alpha$ y el "*bootstrap p-value*" se define como $p_B^* - value = \frac{1}{B} \sum_{b=1}^{B} \mathbf{1} \left(\hat{C}_N^{*(b)} > \hat{C}_{\alpha N}^{*(b)}\right)$. Así, existe evidencia para rechazar H_0 cuando $\hat{C}_N > \hat{C}_{\alpha N}^{*(b)}$o $p_B^* - value < \alpha$.

3.2. Prueba de ajuste de probabilidades

Con el objeto de evaluar la capacidad predictiva del modelo de conteo estimado se procede a probar la bondad de ajuste bajo el supuesto de una función de distribución de probabilidad particular. La hipótesis nula de esta prueba se puede plantear como:

$$H_0 : P_y\left(\mathbf{X}\right) = P_y\left(\mathbf{X};\theta_0\right), \qquad \theta_0 \in \Theta \text{ y } \forall y = 0, 1, 2, \ldots \tag{35}$$

donde P_y es una función de distribución paramétrica y θ_0 es un vector desconocido de parámetros perteneciente al espacio de parámetros Θ. La anterior hipótesis nula se puede reescribir como:

$$H_0 : M\left(y, \mathbf{X}, \theta_0\right) = 0, \qquad \theta_0 \in \Theta \text{ y } \forall y = 0, 1, 2, \ldots \tag{36}$$

siendo

$$\begin{aligned} M\left(y, \mathbf{x}, \theta_0\right) &= \int_{-\infty}^{\mathbf{x}} \left\{\mathbb{E}\left[\mathbf{1}\left(Y = y \mid \mathbf{X} = \mathbf{u}\right)\right] - P_y\left(\mathbf{u};\theta_0\right)\right\} dF_x\left(\mathbf{u}\right) \\ &= \mathbb{E}\left\{\left[\mathbf{1}\left(Y = y\right) - P_y\left(\mathbf{X};\theta_0\right)\right]\mathbf{1}\left(\mathbf{X} \leq \mathbf{x}\right)\right\} \end{aligned} \tag{37}$$

La hipótesis alternativa plantea un proceso generador de datos diferente a un proceso de conteo. La función $M\left(y, \mathbf{x}, \theta_0\right)$ puede ser estimada de manera consistente a través de:

$$\hat{M}_N\left(y, x\right) = \frac{1}{N}\sum_{i=1}^{N}\left\{\mathbf{1}\left(Y_i = y_i\right) - P_y\left(\mathbf{x}_i;\hat{\theta}_N\right)\right\}\mathbf{1}\left(\mathbf{X}_i \leq \mathbf{x}_i\right) \tag{38}$$

Siendo $\hat{\theta}_N$ el estimador de máxima verosimilitud de θ_0. Con dicha medida se construye el siguiente estadístico de Cramér-von Mises:

$$\bar{C}_N = \frac{1}{N}\sum_{i=1}^{N}\left[N^{\frac{1}{2}}\hat{M}_N\left(y_i, \mathbf{x}_i\right)\right]^2 \tag{39}$$

Andrews (1997) demuestra la consistencia de las pruebas paramétricas bajo "*bootstrap*". Dicho procedimiento se lleva a cabo cumpliendo los siguientes pasos:

1. Con base en la muestra original $\mathcal{T}_N = \left\{\left(Y_i, \mathbf{X}_i\right), i = 1, \ldots, N\right\}$ se estima bajo H_0 y mediante el método de máxima verosimilitud a $\hat{\theta}_N$ a partir del cual se calcula la estadística $\bar{C}_N$.

2. Se construye una nueva muestra $\mathcal{T}_N^* = \left\{\left(Y_i^*, \mathbf{X}_i\right), i = 1, \ldots, N\right\}$ con base en la muestra original. Las realizaciones de la variable aleatoria discreta Y_i^* son generadas a partir de

la distribución $P_y\left(\mathbf{X}_i;\hat{\theta}_N\right)$ para cada $\mathbf{X}_i$.

3. Con la nueva muestra se estiman nuevamente por el método de máxima verosimilitud a $\hat{\theta}_N^*$ y posteriormente a $\bar{C}_N^*$ como sigue:

$$\bar{C}_N^* = \frac{1}{N}\sum_{i=1}^{N}\left[N^{\frac{1}{2}}\hat{M}_{Nn}^*\left(y_i,\mathbf{x}_i\right)\right]^2 \tag{40}$$

siendo

$$\hat{M}_N^*\left(y,x\right) = \frac{1}{N}\sum_{i=1}^{N}\left\{\mathbf{1}\left(Y_i^* = y_i\right) - P_y\left(\mathbf{x}_i;\hat{\theta}_N^*\right)\right\}\mathbf{1}\left(\mathbf{X}_i \leq \mathbf{x}_i\right) \tag{41}$$

4. Los pasos 2 y 3 se llevan a cabo B veces, construyendo nuevas muestras $\mathcal{T}_N^{*(b)}, b = 1,\ldots,B$ y calculando en cada una de ellas $\bar{C}_N^{*(b)}, b = 1,\ldots,B$. Tanto los valores críticos, como los "*p-values*" se generan de la misma manera que en la prueba anterior.

3.3. Prueba de ajuste para puntos particulares

Dada su parametrización, los modelos tradicionales de conteo, como el modelo de Poisson y el Binomial Negativo, pueden generar ajustes deficientes en algunos valores particulares de la variable dependiente. En particular, es bien sabido que el modelo de regresión de Poisson ajusta por defecto la ocurrencia de eventos iguales a cero. Para sobrellevar esta deficiencia, la literatura ha desarrollado los modelos de inflación de ceros. Sin embargo, el exceso puede presentarse en cualquier otro valor, distinto del cero[9].

La prueba de decisión binaria desarrollada por Álvarez y Delgado (2002) es empleada en este trabajo para dar luces sobre la necesidad de construir en un futuro modelos de mezclas que permitan corregir excesos en algunas de las frecuencias de cambios de precios en caso que los modelos tradicionales no ajusten adecuadamente. Lo anterior se logra evaluando la bondad del ajuste en aquellos conteos particulares bajo una especificación paramétrica particular. Específicamente, esta prueba se basa en la comparación de un resultado observado en algún conteo en particular y la proporción esperada de dicho conteo bajo la hipótesis nula del modelo. Para la exposición de la prueba, se supone que el exceso de conteos se presenta en cero (no afecta la generalización de la prueba). La probabilidad condicional del dato de conteo igual a cero puede expresarse como una expectativa condicional, así:

$$P_0\left(x\right) = \Pr\left(Y = 0 \mid \mathbf{X} = \mathbf{x}\right) = \mathbb{E}\left[\mathbf{1}\left(Y = 0 \mid \mathbf{X} = \mathbf{x}\right)\right] \tag{42}$$

[9]Para un ejemplo, véase Melkersson y Rooth (2000).

La hipótesis nula en este caso plantea:

$$H_0 : P_0(\mathbf{X}) = P_0(\mathbf{X}; \theta_0) \quad \theta_0 \in \Theta \tag{43}$$

la cual puede ser reescrita como:

$$H_0 : M(0, \mathbf{X}; \theta_0) = 0 \tag{44}$$

siendo:

$$M(0, \mathbf{x}; \theta_0) = \int_{-\infty}^{\mathbf{x}} \{\mathbb{E}[\mathbf{1}(Y = 0 \mid \mathbf{X} = \mathbf{u})] - P_0(\mathbf{u}; \theta_0)\} dF_x(\mathbf{u})$$

$$= \mathbb{E}\{[\mathbf{1}(Y = 0) - P_0(\mathbf{X}; \theta_0)] \mathbf{1}(\mathbf{X} \leq \mathbf{x})\} \tag{45}$$

De la misma forma que las pruebas anteriores, Stute (1997) y Andrews (1997) proponen el uso de la estadística Cramér-von Mises, $\tilde{C}_N$, definida de la siguiente manera:

$$\tilde{C}_N = \frac{1}{N} \sum_{i=1}^{N} \left[N^{\frac{1}{2}} \hat{M}_N(0, \mathbf{x}_i)\right]^2 \tag{46}$$

El término $\hat{M}_N(0, \mathbf{x}_i)$ se calcula mediante la siguiente expresión

$$\hat{M}_N(0, \mathbf{x}_i) = \frac{1}{N} \sum_{i=1}^{N} \left\{\mathbf{1}(y_i = 0) - P_0\left(\mathbf{x}_i; \hat{\theta}_0\right)\right\} \mathbf{1}(\mathbf{X}_i \leq \mathbf{x}_i) \tag{47}$$

Igualmente, en este caso se propone un test basado en la técnica "*bootstrap*", la cual se implemente de la siguiente forma:

1. A partir de la muestra original $\mathcal{T}_N = \{(Y_i, \mathbf{X}_i), i = 1, \ldots, N\}$ se estima considerando el método de máxima verosimilitud y H_0 a $\hat{\theta}_N$ y se calcula la estadística $\tilde{C}_N$.

2. Se construye, con base en la muestra original, una nueva muestra $\mathcal{T}_N^* = \{(Y_i^*, \mathbf{X}_i), i = 1, \ldots, N\}$. La variable y_i^* es generada como una variable aleatoria Bernoulli con parámetro $p = P_0\left(\mathbf{X}_i; \hat{\theta}_N\right)$ para cada $\mathbf{X}_i$.

3. Con la nueva muestra se estiman $\hat{\theta}_N^*$ y posteriormente se construye $\tilde{C}_N^*$ así:

$$\tilde{C}_N^* = \frac{1}{N} \sum_{i=1}^{N} \left[N^{\frac{1}{2}} \hat{M}_N^*(0, \mathbf{X}_i)\right]^2 \tag{48}$$

donde

$$\hat{M}_N^* = \frac{1}{N}\sum_{I=1}^{N}\left\{\mathbf{1}\left(y_i^*=0\right) - P_0\left(\mathbf{X}_i;\hat{\theta}_N^*\right)\right\}\mathbf{1}\left(\mathbf{X}_i \leq \mathbf{x}_i\right) \tag{49}$$

4. Los pasos 2 y 3 se llevan a cabo B veces, construyendo nuevas muestras $\mathcal{T}_N^{*(b)}, b = 1,\dots,B$ y calculando en cada una de ellas $\tilde{C}_N^{*(b)}, b = 1,\dots,B$. tTanto los valores críticos como los "$p-values$" se generan de la misma manera que en las primera prueba anteriormente explicada.

4. Resultados

Las estimaciones puntuales de los modelos detallados en la sección 2.1 se presentan en el Cuadro 4. La columnas (2) y (7) muestran los resultados de las estimaciones máximo verosímiles de los modelos no lineales de Poisson y Binomial Negativo que se obtuvieron al emplear la base de datos sin valores atípicos. Las columnas restantes presentan los valores de los estadísticos t y la significancia estadística de dichas estimaciones. En la mayoría de los casos, las magnitudes y direcciones de los coeficientes estimados son similares, aunque la significancia estadística varía sustancialmente de modelo a modelo. Las estimaciones están acompañadas de medidas de R^2 construidas a partir de (24), los estadísticos Deviance, los valores máximos de las funciones de log-verosimilitud y el criterio de información consistente de Akaike.

Cuadro 4: Resultados modelos de conteo

	POISSON					BIN. NEG.		
Variable	coef.	t-ratio (MV)	Sig.	t-ratio (PMV)	Sig.	coef.	t-ratio	Sig.
CONST.	1.334	6.64	***	3.09	***	1.410	4.17	***
FINAL	-0.311	-6.52	***	-2.84	***	-0.392	-4.90	***
C_DDA	0.119	8.02	***	3.88	***	0.099	4.12	***
C_CLAB	-0.052	-2.86	***	-1.30		-0.043	-1.43	
C_CFIN	-0.068	-3.55	***	-1.46		-0.053	-1.71	***
C_CMP	0.042	2.13	**	0.89		0.028	0.87	
C_CCE	0.052	2.93	***	1.16		0.031	1.00	
C_IMP	-0.018	-0.99		-0.55		-0.004	-0.13	
C_TC	-0.003	-0.25		-0.09		0.009	0.44	
RESTR_SET	0.080	1.69	*	0.79		0.088	1.12	
LIDER	-0.049	-1.54		-0.74		0.020	0.38	
COMP2	0.185	3.30	***	1.51		0.184	2.01	**
DISCR	0.144	2.67	***	1.16		0.161	1.84	*
CALIDAD	0.096	1.96	**	0.88		0.065	0.80	
V_LP	0.002	3.00	***	1.40		0.002	1.74	*
EXPLICITO	-0.014	-0.27		-0.13		-0.007	-0.08	
IMPLICITO	-0.343	-6.36	***	-2.84	***	-0.298	-3.16	***
STATE1	0.498	10.43	***	4.47	***	0.540	6.59	***
V_DOM	0.002	1.98	**	0.89		0.002	1.03	
EMPL	0.101	2.87	***	1.24		0.167	1.95	*
GRANDE	0.261	5.17	***	2.62	***	0.171	1.90	*
SECTOR	-0.476	-14.42	***	-6.41	***	-0.498	-7.42	***
ESTR_MON	-0.208	-2.65	***	-1.17		-0.182	-1.18	
ESTR_COM	-0.070	-1.27		-0.51		0.005	0.06	
$\ln \delta_0$	-	-		-		-0.547	-6.324	
R^2		0.2289					0.0894	
R^2_P		0.4834					0.2768	
P		2837.953					-	
D		1965.667					752.448	
$\ln \mathcal{L}$		-1810.9955					-1483.984	
CAIC[a]		3805.5044					3151.4814	

Fuente: Cálculos de los autores.
[a] Criterio de información consitente de Akaike: $CAIC = -2 \ln \mathcal{L} + (1 + \ln N) K$;*** indica significancia individual al 1 %, **al 5 % y * al 10 %

Por su parte, el Cuadro 5 presenta el promedio de los efectos marginales estimados para cada variable[10] y el cambio porcentual en la frecuencia de ajuste de precios ante cambios en

[10]El promedio de los efectos marginales para la variable x_j, se construye como sigue: $\text{PEM}_j =$

cada una de las variables manteniendo las demás variables constantes.

Cuadro 5: Efectos marginales promedio (PEM) y variación porcentual

	POISSON			BIN. NEG.		
Variable	EM_j	$\Delta_j\,\%$	Sig.	EM_j	$\Delta_j\,\%$	Sig.
$FINAL$	-0.839	-26.8	***	-1.066	-32.4	***
C_DDA	0.306	12.7	***	0.253	10.4	***
C_CLAB	-0.135	-5.1	***	-0.109	-4.2	***
C_CFIN	-0.176	-6.6	***	-0.135	-5.1	
C_CMP	0.107	4.3	**	0.073	2.9	***
C_CCE	0.134	5.4	***	0.078	3.1	
C_IMP	-0.046	-1.8		-0.011	-0.4	
C_TC	-0.008	-0.3		0.024	0.9	
$RESTR\_SET$	0.205	8.4	*	0.224	9.2	
$LIDER$	-0.126	-4.8		0.050	2.0	
$COMP2$	0.453	20.3	***	0.451	20.2	
$DISCR$	0.359	15.5	***	0.399	17.5	**
$CALIDAD$	0.247	10.1	**	0.167	6.7	*
V_LP	0.006	0.2	***	0.006	0.2	
$EXPLICITO$	-0.037	-1.4		-0.019	-0.7	*
$IMPLICITO$	-0.927	-29.1	***	-0.799	-25.8	
$STATE1$	1.378	64.5	***	1.501	71.6	***
V_DOM	0.006	0.2	**	0.005	0.2	***
$EMPL$	0.260	10.6	***	0.427	18.1	
$GRANDE$	0.689	29.8	***	0.447	18.7	*
$SECTOR$	-1.223	-37.9	***	-1.274	-39.2	*
$ESTR\_MON$	-0.579	-18.8	***	-0.500	-16.6	***
$ESTR\_COM$	-0.183	-6.8		0.014	0.5	

Fuente: Cálculo de los autores, *** indica significancia individual al 1 %, **al 5 % y * al 10 %

En la sección 4.1, se implementan las diferentes pruebas presentadas en la sección 3, así como otras medidas de evaluación de modelos, con el propósito de elegir el modelo más relevante a la hora de explicar la frecuencia de ajuste de precios. A partir del modelo seleccionado, en la sección 4.2 se presenta el análisis de los determinantes del grado de flexibilidad de precios.

4.1. Evaluación de la bondad de ajuste

Una vez obtenidas las estimaciones de los coeficientes de regresión y de sus errores estándar respectivos se aplican diferentes métodos de selección de modelos para elegir aquel que mejor describa la frecuencia de ajuste de precios observada. Primero se lleva a cabo la prueba de

$\frac{1}{N}\sum_{i=1}^{N}\hat{\beta}_j \exp\left(\mathbf{x}_i'\hat{\beta}\right)$.

especificación de la media condicional desarrollada en la sección 3.1 bajo la siguiente hipótesis nula:

$$H_0 : \mathbb{E}\left(Y \mid \mathbf{X}\right) = \exp\left(\mathbf{X}'\beta\right) \tag{50}$$

La prueba emplea la técnica de *bootstrapping* con 5000 replicaciones y tomando el estimador de Poisson como valor inicial de los parámetros de la media. Los resultados que se presentan en el Cuadro 6 sugieren que no se existe evidencia estadística para rechazar la hipótesis nula, a ningún nivel de significancia usual. Dado que la media condicional está correctamente especificada, es posible concluir que bajo los dos modelos de regresión, las estimaciones de los parámetros son consistentes aún si el verdadero proceso generador de datos no es Poisson o Binomial Negativo. Sin embargo, este resultado aún no permite discriminar entre las dos especificaciones dado que presentan la misma media condicional como se muestra en las ecuaciones (26) y (27).

Cuadro 6: Pruebas de especificación de media y distribución condicional

		Valores Críticos*			
	Estadístico C_n	1 %	5 %	10 %	$p_B - valor$
Media Condicional					
$H_0 : \mathbb{E}\left(Y \mid \mathbf{X}\right) = \exp\left(\mathbf{X}'\beta\right)$	0.0419	0.1044	0.0920	0.0856	0.9602
Distribución Condicional					
$H_0 : P_y\left(x\right) =$ Poisson	0.0035	0.0122	0.0107	0.0098	0.9284
$H_0 : P_y\left(x\right) =$ Binominal Negativa	0.0034	0.0143	0.0124	0.0114	0.9782

Fuente: Cálculos de los autores.

*Construidos a partir de un procedimiento de *Bootstrapping* con 5000 replicaciones. Las hipótesis alternas corresponden a la negación de la nulas.

Las medidas de bondad de ajuste R^2 basadas en los incrementos de la función de log-verosimilitud y en los residuos de Pearson sugieren que el modelo de regresión de Poisson es superior al modelo Binomial Negativo. Sin embargo, medidas tales como el Deviance y el Criterio de Información Consistente de Akaike, sugieren lo contrario. Adicionalmente, el criterio Pearson indica una mala especificación del modelo de regresión de Poisson ya que $P \neq N - K$. En particular, $P > N - K$, lo que sugiere una incorrecta especificación de la función de varianza.

La anterior conclusión se valida por medio de una prueba de sobredispersión, la cual parte del hecho de que el modelo Binomial Negativo colapsa al modelo de Poisson cuando $\delta_0 = 0$. De esta forma, es posible probar la presencia de sobredispersión por medio de la verificación de la hipótesis nula $H_0 : \delta_0 = 0$, versus la alterna que plantea sobredispersión, $H_1 : \delta_0 > 0$. La prueba se lleva a cabo por medio de un test de razón de verosimilitud con distribución

Cuadro 7: Pruebas de sobredispersión

$H_0 : \delta_0 = 0$ vs. $H_1 : \delta_0 > 0$	Estadístico	Valor-p
Razón de verosimilitud	654.05	0.000
Prueba t	4.73	0.000

Fuente: Cálculo de los autores.

asintótica corregida por posible truncamiento del coeficiente δ_0 en cero. El estadístico de prueba viene dado por $G^2 = 2\left(\ln \mathcal{L}_{BN} - \ln \mathcal{L}_P\right)$, donde $\ln \mathcal{L}_{BN}$ corresponde al valor máximo de la función de log-verosimilitud del modelo Binomial Negativo y $\ln \mathcal{L}_P$ al valor máximo de la función de log-verosimilitud del modelo de Poisson. El estadístico G^2 tiene una distribución mixta entre una chi cuadrado sin grados libertad y una chi cuadrado con un grado de libertad. Los resultados de la prueba se presentan en el Cuadro 7, de donde se concluye que el modelo presenta sobredispersión a cualquier nivel de significancia estadística. Esta conclusión favorece el modelo Binomial Negativo sobre el modelo de Poisson.

Cameron y Trivedi (1998) sugieren un método alterno para probar la existencia de sobredispersión bajo la hipótesis nula de equidisperión, $\delta_0 = 0$, a partir de la función de varianza obtenida del modelo Binomial Negativo, $\text{Var}\left(Y \mid \mathbf{X}\right) = m\left(\mathbf{x}\right) + \delta_0 m\left(\mathbf{x}\right)^2$. La prueba se realiza a través de una regresión auxiliar de la variable $\left\{\left(y_i - \hat{m}\left(\mathbf{x}_i\right)\right)^2 - y_i\right\} / \hat{m}\left(\mathbf{x}_i\right)$ sobre la media estimada, $\hat{m}\left(\mathbf{x}_i\right)$ sin incluir intercepto. A partir de la estimación se realiza una prueba t de significancia individual sobre el coeficiente del regresor. El resultado que se presenta en el Cuadro 7 es similar al obtenido por medio de la prueba de razón de verosimilitud.

Una medida adicional de selección de modelos ampliamente utilizada en datos de conteo consiste en comparar las probabilidades estimadas, bajo diferentes especificaciones, con las frecuencias de conteos efectivamente observadas. Las primeras se construyen a partir de los parámetros estimados:

$$\hat{P}\left[Y_i = y_i \mid m\left(\mathbf{x}_i\right)\right] = \frac{\exp\left(-\hat{m}\left(\mathbf{x}_i\right)\right) \hat{m}\left(\mathbf{x}_i\right)^{y_i}}{y_i!} \tag{51}$$

$$\hat{P}\left[Y_i = y_i \mid m\left(\mathbf{x}_i\right), \delta_0\right] = \frac{\Gamma\left(y_i + \hat{\delta}_0^{-1}\right)}{\Gamma\left(\hat{\delta}_0^{-1}\right)\Gamma\left(y_i + 1\right)} \left(\frac{\hat{\delta}_0^{-1}}{\hat{m}\left(\mathbf{x}_i\right) + \hat{\delta}_0^{-1}}\right)^{\hat{\delta}_0^{-1}} \left(\frac{\hat{m}\left(\mathbf{x}_i\right)}{\hat{m}\left(\mathbf{x}_i\right) + \hat{\delta}_0^{;1}}\right)^{y_i} \tag{52}$$

donde $\hat{m}\left(\mathbf{x}_i\right) = \exp\left(\mathbf{x}_i' \hat{\beta}\right)$.

Los resultados que se exponen en el Cuadro 8 son mixtos. Para frecuencias de ajustes bajas como el cero, uno y dos, el modelo de Poisson presenta un mejor ajuste desde un punto de vista numérico. Sin embargo, el modelo Binomial Negativo hace un mejor trabajo para frecuencias

Cuadro 8: Probabilidades estimadas

Conteo	Frecuencia Observada	Frecuencia Ajustada	
		Poisson	Bin. Neg.
0	0.178	0.172	0.279
1	0.388	0.249	0.230
2	0.199	0.210	0.158
3	0.071	0.139	0.103
4	0.038	0.083	0.067
5	0.025	0.048	0.044
6	0.021	0.029	0.030
7	0.005	0.018	0.021
8	0.010	0.012	0.015
10	0.012	0.007	0.008
12	0.026	0.004	0.005
15	0.005	0.002	0.002
20	0.013	0.001	0.001
30	0.006	0.000	0.000
33	0.001	0.000	0.000

Fuente: Cálculo de los autores
Nota: Únicamente se presentan los ajustes de las frecuencias efectivamente observadas.

más altas. Este ejercicio se complementa con el procedimiento de *bootstrapping* descrito en la sección 3.2 que busca estudiar la función de distribución condicional de todos los datos bajo las distintas especificación econométricas. Los resultados de este procedimiento se presentan en el Cuadro 6 para el modelo de Poisson y el modelo Binomial Negativo. Si bien existen diferencias puntuales entre los ajustes de ambos modelos, como recién se mencionó, dichas diferencias no son estadísticamente significativas toda vez que el procedimiento de *bootstrapping* indica que no existe evidencia para rechazar ambas especificaciones probabilísticas.

Finalmente, se procede a evaluar la bondad de ajuste bajo ambas especificaciones en los datos de conteo con mayor frecuencia relativa en la muestra. Para ello se emplea el procedimiento descrito en la sección 3.3 y sus resultados se presentan en el Cuadro 9. Es posible concluir que el modelo Binomial Negativo presenta un mejor ajuste de los diferentes cambios de precios respecto al modelo de Poisson toda vez que el primero es capaz de predecir apropiadamente frecuencias de ajuste unitaria, mientras que el segundo no lo hace.

Si bien, las estimaciones econométricas y las pruebas estadísticas, diferentes a la de ajuste de valores particulares, no son contundentes para la selección de un único modelo, la existencia de sobredispersión en la función de varianza y el correcto ajuste de la frecuencia particular nos lleva a elegir al modelo Binomial Negativo sobre el modelo de Poisson, aun cuando es plausible emplear el modelo de regresión de Poisson con matriz de varianza-covarianza robusta. Sin embargo, este último no es apropiado toda vez que la especificación del modelo sugiere

Cuadro 9: Bondad de ajuste conteos particulares

Conteo (c)	$H_0 : P_{y=c}(\mathbf{X}) = Poisson$ Estadístico $\tilde{C}_N$	$H_0 : P_{y=c}(\mathbf{X}) = Bin.\ Neg.$ Estadístico $\tilde{C}_N$
0	0.0078*	0.0033***
1	0.0043	0.0067*
2	0.0091*	5.3943*
3	0.0071*	3.4362*
4	0.0029*	2.3980*
5	1.6919*	1.6952*
6	0.6848*	0.6405*

*Indica significancia estadística al 1 %, ** al 5 % y *** al 10 %. Los niveles de signficancia se construyen a partir de un procedimiento de bootstrapping con 5000 replicaciones.

la posible existencia de heterogeneidad no observada. Es de señalar que la interpretación económica que se presenta a continuación se basa en los resultados obtenidos mediante el modelo Binomial Negativo.

4.2. Análisis económico de los resultados

Esta sección resume los principales resultados económicos que se derivan de la estimación de los modelos analizados en la sección anterior. Si bien, la naturaleza de la variable dependiente sugiere modelos de conteo de Poisson para estudiar sus determinantes, la batería de pruebas presentadas previamente indican que el proceso que gobierna el grado de flexibilidad de precios es mejor descrito por el modelo Binomial Negativo. Así, con base en esta especificación estadística se analizan los resultados económicos.

En términos generales, se encuentra que las características del producto, el grado de competencia que enfrenta la firma, los acuerdos contractuales y el sector económico al que pertenece la firma son los determinantes más significativos a la hora de explicar el grado de flexibilidad de los precios. Por su parte, algunas técnicas de mercadeo, como la posibilidad de discriminar precios, son significativas a un nivel de significancia menor. Los resultados también permiten concluir que la existencia de leyes o decretos administrativos, la existencia de líderes en precios y los movimientos de costos no explican la heterogeneidad en el proceso de fijación de los mismos.

Respecto a las características del producto, se encuentra que los productores de bienes finales cambian con menos frecuencia el precio de sus bienes. Es decir, los precios de los bienes finales son más rígidos que aquellos asociados a los bienes de capital e intermedios. Manteniendo todo lo demás constante, los primeros cambian el precio de su producto un 32.4 % menos que los segundos. Adicionalmente, se encuentra un efecto positivo entre la importancia atribuida a los choques de demanda y la frecuencia con que se ajustan los precios. En otras palabras, aquellos empresarios más sensibles a movimientos en la demanda por sus bienes, cambian los precios de sus productos un 10.4 % más que aquellos empresarios que no consideran tan relevante las variaciones en su demanda. Por su parte, el único componente de la estructura de costos que afecta el grado de flexibilidad de los precios de forma significativa es el financiero. Dado que los costos financieros suelen pactarse con tasas fijas de mediano o largo plazo, es decir, corresponden a costos que no cambian frecuentemente, los resultados sugieren que aquellas industrias que enfrentan unos costos más elevados modifican sus precios un 5.1 % menos.

Resulta interesante el hecho de que los cambios en los costos laborales y los cambios en los precios de las materias primas no son factores determinantes del grado de flexibilidad de precios, aún cuando la dirección del efecto marginal obtenido en las estimaciones es el correcto. En particular, es de esperar que aquellas firmas donde los costos laborales son relevantes tiendan a reportar un reducido número de cambios en precios dadas las rigideces implícitas en los salarios. Así mismo, se espera que aquellos sectores donde las materias primas juegan un papel importante en la producción, muestren un mayor grado de flexibilidad en los precios ya que los precios de los insumos cambian frecuentemente, tal y como se concluyó anteriormente. Resultados similares son encontrados por Álvarez y Hernando (2006) para un conjunto de países de la Unión Europea cuando utilizan toda las observaciones de la muestra.

El segundo gran determinante de la frecuencia de ajuste está asociado a la concentración de poder o control. Los resultados sugieren que el grado de competencia que perciben los empresarios conlleva a cambios de precios más frecuentes. *Ceteris paribus*, firmas que se enfrentan a una mayor competencia cambian su precio un 20.2 % más frecuentemente que aquellas que perciben menos competencia. Igualmente, la posibilidad de discriminar precios entre compradores implica un 17.5 % más flexibilidad en los precios.

Respecto a los arreglos contractuales se encuentra que una mayor cantidad de ventas con clientes que las firmas consideran de largo plazo implica una mayor flexibilidad en el ajuste de precios. Si bien esto puede resultar contra intuitivo, su efecto marginal es relativamente bajo y la relación exacta entre la existencia de vínculos entre productores y consumidores es capturada correctamente por la existencia de contratos implícitos. En efecto, el establecimiento de acuerdos tácitos entre firmas y compradores reduce el número de cambios de precios en un 25.8 %. Por el contrario, la existencia de contratos explícitos no es estadísticamente significativa, y su efecto marginal sobre la frecuencia de ajuste de los precios es mínimo. Esto sugiere que la posibilidad de fijar precios explícitamente a través de acuerdos contractuales por intervalos de tiempo no explica el comportamiento generalizado que se observa en las frecuencias de ajuste de precios de las empresas colombianas.

De otro lado, las firmas que revisan los precios de sus productos de acuerdo al estado de la economía son más flexibles que aquellas que emplean reglas dependientes del tiempo. Relativo al tamaño de la firma, los resultados sugieren que las firmas más grandes, tanto en términos de empleados como en términos del valor de sus activos totales, son más flexibles. Por el contrario, las pequeñas y medianas empresas son más rígidas a la hora de cambiar los precios. En efecto, las primeras cambian los precios un 18.3 % más frecuentemente que las segundas, en promedio. El sector de la agricultura resulta ser el más flexible toda vez que a medida que el grado de manufactura aumenta, las firmas son más rígidas para ajustar sus precios. Más específicamente, por cada 100 cambios de precios que se registran en el sector de la agricultura, el sector industrial solo cambia 20 veces su precio.

Finalmente, las firmas que emplean reglas de fijación de precios basadas en costos más un margen tienden a ser menos flexibles a la hora de ajustar el precio. Este resultado es consistente toda vez que las firmas que usualmente aplican estas reglas son aquellas que se ubican en mercados poco competidos. Por el contrario, las firmas que siguen reglas de fijación de precios en donde el precio se fija con base en el precio de la competencia, tienden a ser más flexibles.

5. Conclusiones

Recientemente se ha publicado en Colombia una amplia evidencia sobre la formación de los precios a nivel microeconómico, motivada por la falta de conocimiento sobre su grado de rigidez y sus posibles efectos en la conducción de la política monetaria. Uno de los principales hallazgos de esa agenda de investigación se relaciona con que la rigidez de precios no es constante entre productos, sino que por el contrario es variable. El principal aporte de este trabajo es profundizar en el examen de este fenómeno conocido como heterogeneidad en las rigideces de precios, utilizando para ello los modelos de conteo aplicados a los resultados de una encuesta cualitativa realizada durante 2007 y 2008, con la cual se interrogó a una muestra representativa de firmas colombianas acerca de la forma como fijaban sus precios. Este es un enfoque relativamente novedoso para abordar esta problemática y tiene la ventaja de permitir el estudio del grado de flexibilidad de los precios, aproximado por la frecuencia de cambio de los mismos, y de establecer que factores de mercado y caracteristicas propias de la empresa afectan las frecuencias de ajuste de precios.

Se encontraron tres resultados principales. El primero se refiere a que el proceso que gobierna el grado de flexibilidad de los precios en Colombia es mejor descrito por un modelo Binomial Negativo. Lo anterior se debe a la existencia de una gran variabilidad no observable en el número de cambios de precios por año, que excede significativamente la media, lo cual invalida la inferencia estadística en el modelo básico de Poisson. De otro lado, se halló que la frecuencia de cambio de los precios está afectada por varios factores. Los más importantes dentro de estos son las características del producto, el grado de competencia que enfrenta la firma, los acuerdos contractuales y el sector de la economía en el cual opera la firma. De otro lado, se encontró que la existencia de leyes o decretos administrativos, la existencia de líderes en precios y los movimientos de costos no explican la heterogeneidad en el proceso de fijación de los precios.

Una enseñanza final de este trabajo es que sus resultados deberían tenerse en cuenta para la modelación macroeconómica y la evaluación de la política monetaria. En efecto, los resultados sugieren que las firmas ajustan sus precios de forma tal que buscan obtener el mayor beneficio posible basadas en características propias de la firma y de mercado, i.e., basadas en sus microfundamentos. Estos últimos hacen que las firmas sean diferentes, lo cual se traduce en heterogeneidad a la hora de fijar sus precios. La extensión de la heterogeneidad sugiere que la formación de los precios puede modelarse entonces como una mezcla de comportamientos *à la* Calvo y *à la* Taylor. En la política monetaria debe tenerse en cuenta que los agentes estado-dependientes se caracterizan por tener una mayor flexibilidad a la hora de revisar sus precios, ellos pueden reaccionar con mayor presteza, mientras que aquellos que siguen reglas tiempo dependientes reaccionaran lentamente, con la consecuencia de inducir algún rezago en el impacto de la política monetaria.

Referencias

Álvarez, B. y Delgado, M. A. (2002). Goodness-of-fit techniques for count data models: an application to the demand for dental care in spain. *Empirical Economics*, 27(3):543–567.

Álvarez, L. y Hernando, I. (2006). Competition and price adjustment in the euro area. Banco de España Working Papers 0629, Banco de España.

Andrews, D. W. K. (1997). A conditional kolmogorov test. *Econometrica*, 65(5):1097–1128.

Aoki, K. (2001). Optimal monetary policy responses to relative-price changes. *Journal of Monetary Economics*, 48(1):55–80.

Backman, J. (1940). The causes of price inflexibility. *The Quarterly Journal of Economics*, 54(3):474–489.

Benigno, P. (2004). Optimal monetary policy in a currency area. *Journal of International Economics*, 63(2):293–320.

Bils, M. y Klenow, P. J. (2004). Some evidence on the importance of sticky prices. *Journal of Political Economy*, 112(5):947–985.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, 12(3):383–398.

Cameron, A. C. y Trivedi, P. K. (1998). *Regresion Analysis of Count Data*. Number 30 in Econometric Society Monographs. Cambridge University Press, 6th edición.

Carlton, D. W. (1989). The theory and the facts of how markets clear: Is industrial organization valuable for understanding macroeconomics? En Schmalensee, R. y Willig, R., editores, *Handbook of Industrial Organization*, volume 1 of *Handbook of Industrial Organization*, chapter 15, pages 909–946. Elsevier.

Carvalho, C. (2006). Heterogeneity in price stickiness and the real effects of monetary shocks. *The B.E. Journal of Macroeconomics*, 0(1).

de Munnik, D. y Xu, K. (2007). Micro foundations of price-setting behaviour: Evidence from canadian firms. Working Papers 07-31, Bank of Canada.

Dixon, W. J. (1950). Analysis of extreme values. *Annals of Mathematical Statistics*, 21(4):488–506.

Gordon, R. J. (1981). Output fluctuations and gradual price adjustment. *Journal of Economic Literature*, 19(2):493–530.

Hawkins, D. (1980). *Identification of Outliers*. Springer.

Julio, J. M. (2010). Heterogeneidad observada y no observada en la formación de precios del ipc colombiano. Borradores de Economía 597, Banco de la República.

Means, G. C. (1936). Notes on inflexible prices. *The American Economic Review*, 26(1):23–35. Supplement, Papers and Proceedings of the Forty-eighth Annual Meeting of the American Economic Association.

Melkersson, M. y Rooth, D.-O. (2000). Modeling female fertility using inflated count data models. *Journal of Population Economics*, (13):189–203.

Misas, M., López, E., y Parra, J. C. (2009). La formación de precios en las empresas colombianas: evidencia a partir de una encuesta directa. Borradores de Economía 569, Banco de la República de Colombia.

Morales, P. y Jaramillo, C. F. (1995). Estructura del Índice de precios al consumidor: Algunas implicaciones para el análisis de la inflación. Borradores de Economía 39, Banco de la República.

Ruggles, R. (1955). The nature of price flexibility and the determinants of relative price changes in the economy. En *Business Concentration and Price Policy*, NBER Chapters, pages 439–504. National Bureau of Economic Research, Inc.

Stute, W. (1997). Nonparametric model checks for regression. *The Annals of Statistics*, 25(2):613–641.

Taylor, J. B. (1980). Aggregate dynamics and staggered contracts. *Journal of Political Economy*, 88(1):1–23.

Williams, D. A. (1987). Generalized linear model diagnostics using the deviance and single case deletions. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 36(2):181–191.

Zárate, H. M. (2010). Reglas de fijación de precios de los productores colombianos: evidencia a partir de los modelos de duración con micro datos del ipp. Borradores de Economía 600, Banco de la República.

A. Detección y remoción de valores atípicos

Los resultados obtenidos en la encuesta de precios sugieren la posible existencia de valores atípicos en la frecuencia de ajuste de precios. Tal y como se observa en el Cuadro 1, un 2.93 % de los encuestados aseguró cambiar el precio de su principal producto más de 25 veces en el último año previo a la realización de la encuesta. Intuitivamente, es posible que exista una gran cantidad de cambios de precios en sectores como la agricultura, donde se observan estructuras de mercado más competitivas. Sin embargo, tras analizar caso por caso se encuentra que los cambios de precios de dichas empresas o bien no pertenecen a este sector o en caso de pertenecer, su comportamiento es diferente al de otras empresas que producen productos similares en mercados comparables. Por lo tanto, en esta sección se desarrolla una prueba estadística para detectar la presencia de valores extremos.

Hawkins (1980) define los valores atípicos o *outliers* como observaciones que se desvían significativamente de otras observaciones hasta el punto de hacer pensar que estas pudieron haber sido generadas por otro proceso generador de datos. Por su parte, Dixon (1950) los define como "observaciones que generan dudas en el investigador". El interés en detectar y remover estas observaciones atípicas se fundamenta en los efectos perjudiciales que estas pueden tener en la estimación puntual de los parámetros y en la inferencia estadística. En efecto, la existencia de *outliers* puede generar correlaciones espúreas entre variables, así como incrementos en los errores tipo I y tipo II.

Para evaluar el hecho de que la i-ésima observación de la variable dependiente sea un valor atípico se emplea una variación del método no paramétrico desarrollado por Williams (1987). De acuerdo con el autor, basta con analizar la reducción, G_i, que se obtiene en el estadístico Deviance, $D\left(y, \hat{m}\left(\mathbf{x}\right)\right)$, cuando la i-ésima observación es eliminada. Se dice entonces que la observación i es un *outlier,* si el estadístico G_i asociado, es el máximo de todos los estadísticos calculados, el cual se denota por $\max G_i$. La construcción del estadístico es computacionalmente costosa cuando el conjunto de observaciones es elevado. Sin embargo, Williams (1987) muestra como la contribución a G_i de la i-ésima observación viene dada por d_i^2, y la contribución de las $N-1$ observaciones restantes es aproximada por $h_i r_{Pi}^2$. De esta manera se tiene entonces que:

$$G_i \approx r_{Gi} = \text{signo}\left(y_i - \hat{m}\left(\mathbf{x}_i\right)\right) \cdot \sqrt{\left(1-h_i\right) r_{Di}^2 + h_i r_{Pi}^2} \tag{53}$$

donde h_i es el i-ésimo elemento de la matriz $\mathbf{H}$ definida como[11] $\mathbf{H} = \mathbf{W}^{1/2}\mathbf{X}\left(\mathbf{X}'\mathbf{W}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{W}^{1/2}$, siendo $\mathbf{W} = \text{diag}\left[\omega_i\right]$. Es de señalar que, en el caso del modelo de regresión de Poisson, $\omega_i = m_i$.

La prueba basada en el estadístico $\max r_{Gi}$ se realiza por medio de una aproximación de su distribución de probabilidad a través de un ejercicio de *bootstrapping* debido a que

[11] En el modelo clásico de regresión lineal, $\mathbf{H}$ se conoce como la matriz "gorro" dado que $\hat{y} = \mathbf{H}y$.

su distribución no está definida para variables aleatorias que se alejan de la normalidad. El procedimiento se detalla a continuación:

1. A partir de la muestra original $\mathcal{T}_N = \{(Y_i, \mathbf{X}_i), i = 1, \ldots, N\}$ se lleva a cabo la estimación del modelo de regresión con datos de conteo considerando como proceso generador de datos la especificación de Poisson o Binomial Negativa.

2. Una vez estimado el modelo, se construyen los residuales básicos $r_i = y_i - \hat{m}(\mathbf{x}_i)$ y los residuales r_{G_i} dados por (53), los cuales son utilizados para construir la estadística que permite identificar la presencia de valores atípicos.

3. Se genera una variable aleatoria Bernoulli *iid*, υ_i, con media igual a $\left(\sqrt{5}+1\right)/2\sqrt{5}$, la cual se transforma de la siguiente manera: si $\upsilon_i = 1$ entonces $\upsilon_i = -\left(\sqrt{5}-1\right)/2$. En caso contrario, $\upsilon_i = \left(\sqrt{5}+1\right)/2$. Es importante anotar que se hace uso de esta variable aleatoria con el objeto de mantener las propiedades de asimetría de la variable de conteo.

4. Se construye un nuevo conjunto de información $\mathcal{T}_N^* = \{(Y_i^*, \mathbf{X}_i), i = 1, \ldots, N\}$ a partir del conjunto de información original $\mathcal{T}_N$ con realizaciones de la variable dependiente $y_i^* = m\left(\mathbf{x}_i; \hat{\beta_N}\right) + \epsilon_i^*$, donde $\epsilon_i^* = \hat{r}_i \upsilon_i$.

5. A partir de la nueva muestra, $\mathcal{T}_N^*$, se estima el modelo y se calcula nuevamente el residual $r_{G_i}^*$.

6. Los pasos 3 a 5 se llevan a cabo B veces, construyendo nuevas muestras $\mathcal{T}_N^{*(b)}, b = 1, \ldots, B$ y calculando para cada una de ellas $r_{G_i}^{*(b)}, b = 1, \ldots, B$. Los valores críticos de la prueba son estimados a través de los cuantíles condicionales de $r_{G_i}^{*(b)}$. Así, a un nivel de significancia α, el valor crítico del *bootstrapping* puede ser aproximado por $r_{\alpha G_i}^{*(b)}$, de tal forma que $\frac{1}{B}\sum_{b=1}^{B} \mathbf{1}\left(r_{G_i}^{*(b)} > r_{\alpha G_i}^{*(b)}\right) = \alpha$.

Con 5000 replicaciones del procedimiento es posible concluir con un nivel de confianza del 99 %, que todas aquellas observaciones asociadas a cambios de precios superiores de 36 veces en los últimos doce meses previos a la realización de la encuesta, son valores atípicos. Este resultado implica eliminar de la muestra 16 de las 786 observaciones. De esta manera, el análisis que se presentan en la sección 4 se lleva a cabo con una muestra corregida por la presencia de dichas realizaciones de la variable dependiente.

B. Definición y descripción de variables

Cuadro 10: Descripción de todas las variables explicativas

Grupo	Nombre	Descripción	Categórica
Características del producto	*final*	Tipo de bien producido: 1=Final, 0=Intermedio o de capital	Si
	c_dda	Importancia de cambios en la demanda a la hora de modificar el precio: 2=No o poco importante,... 8=Muy importante	Si
	c_clab	Importancia de cambios en los costos laborales a la hora de modificar el precio: 2=No o poco importante,... 8=Muy importante	Si
	c_cfin	Importancia de cambios en los costos financieros a la hora de modificar el precio: 2=No o poco importante,... 8=Muy importante	Si
	c_cmp	Importancia de cambios en los costos de materias primas a la hora de modificar el precio: 2=No o poco importante,... 8=Muy importante	Si
	c_cce	Importancia de cambios en los costos de la energía y los combustible a la hora de modificar el precio: 2=No o poco importante,... 8=Muy importante	Si
	c_imp	Importancia de cambios en los impuestos y otras cargas tributarias a la hora de modificar el precio: 2=No o poco importante,... 8=Muy importante	Si
	c_tc	Importancia de cambios en la tasa de cambio a la hora de modificar el precio: 2=No o poco importante,... 8=Muy importante	Si
Leyes o decretos administrativos	*restr_set*	Enfrenta algún tipo de límitación diferente a la competencia u otras fuerzas del mercado a la hora de fijar el precio: 1=Si, 0=No	Si
Concentración de control	*lider*	Existencia empresas lideres en la fijación de precios: 0=No existen, 1=Si existen, 2=Soy el lider	Si
	*comp*2	Percepción de competencia: 0=Poca competencia, 1=Mucha competencia	Si
Técnicas de mercadeo	*discr*	Existen diferentes precios para diferentes compradores: 1=Si, 0=No	Si
	calidad	Cambios en la calidad del producto antes que modificar el precio: 1=Importante y muy importante, 0=No o poco importante	Si
Arreglos contractuales	*v_lp*	Porcentaje de ventas totales que realiza con clientes que considera de largo plazo	No
	explicito	Importancia de los contratos escritos como restricción al cambio de precios: 1=Importante y muy importante, 0=No o poco importante	Si
	implicito	Importancia de los contratos tácitos como restricción al cambio de precios: 1=Importante y muy importante, 0=No o poco importante	Si

(*Continuación*)

Hábitos y costumbres	*state*1	Regla de revisión de precios: 1=Estado dependiente, 0=Tiempo dependiente	Si
Estructura del mercado	*v_dom*	Porcentaje de ventas que se realiza en el mercado interno	No
	empl	Número de empleados (en miles)	No
	grande	Tamaño de la firma vía valor de los activos totales: 1=Grande, 0=No grande (pequeña y mediana)	Si
	sector	Sector económico: 1=Agricultura, 2=Pesca, 3=Industria	Si
	estr_mon	Precios basados en un margen sobre el costo: 1=Importante y muy importante, 0=No o poco importante	Si
	estr_com	Precios basados en los precios de los competidores: 1=Importante y muy importante, 0=No o poco importante	Si

C. Estadísticas descriptivas

Cuadro 11: Estadísticas descriptivas

	Media	Desviación	Mín	Máx
Frecuencia del ajuste	2.57	4.28	0	36
Caracteristicas del producto				
Bien final (=1)	0.68	0.47	0	1
Leyes o decretos administrativos				
Regulación / limitaciones para fijar precio (=1)	0.57	0.50	0	1
Concentración de control				
Existencia de lideres	0.59	0.78	0	2
Elevada competencia (=1)	0.73	0.44	0	1
Técnicas de mercadeo				
Discriminación de precios entre compradores (=1)	0.7	0.46	0	1
Cambio en calidad del producto muy importante (=1)	0.56	0.50	0	1
Arreglos contractuales				
% de ventas con clientes de largo plazo	75.32	30.44	0	100
Contratos explícitos muy importantes (=1)	0.64	0.48	0	1
Contratos implícitos muy importantes (=1)	0.72	0.45	0	1
Hábitos y costumbres				
Estado dependiente (=1)	0.28	0.45	0	1
Estructura de mercado				
% de ventas domésticas	86.79	22.22	0	100
Número de empleados	168.97	496.88	1	9000
Grande (=1)	0.35	0.48	0	1
Sector económico	2.81	0.58	1	3
Precio basado en competidores	0.94	0.24	0	1
Costo más un margen	0.73	0.44	0	1

Fuente: Cálculo de los autores