

P

atronos de precios de ALIMENTOS PROCESADOS

GONZALO BUENO* Y LAURA OLIVERA**

Este artículo realiza una primera exploración a la base de datos de precios de supermercados y clasifica el comportamiento de los alimentos procesados a partir de criterios sobre su volatilidad y frecuencia de actualización.



* **Especialista, Departamento de Estadísticas de Precios del BCRP**
gonzalo.bueno@bcrp.gob.pe



** **Supervisor líder, Subgerencia de Gestión de Riesgos del BCRP**
laura.olivera@bcrp.gob.pe

INTRODUCCIÓN

El análisis de la inflación se puede nutrir de estudiar y comprender el proceso de formación de precios individuales, especialmente cuando se cuenta con información de alta frecuencia. Es por ello que, para complementar la información que el Banco Central de Reserva del Perú (BCRP) utiliza para el análisis de los índices de precios, desde el año 2020 se realiza la descarga de precios en línea de productos vendidos en supermercados.

El presente artículo se enfoca en el análisis de un grupo particular de productos: los alimentos procesados. A pesar de que estos productos forman parte de la categoría de precios de alimentos y energía (AE)¹, que en conjunto refleja el componente más volátil de la inflación, el subgrupo de alimentos procesados presenta características que los relacionan al componente más estable de la inflación, sobre el cual actúa la política monetaria. Así, en periodos de baja inflación, estos precios no varían con mucha frecuencia²; no obstante, durante un periodo de inflación alta y persistente, como el que se ha experimentado desde mediados de 2021, estos productos pueden comenzar a presentar variaciones de precios al alza de manera más frecuente.

Desde 2021 la variación interanual de los precios de los alimentos procesados se ha incrementado y ha llegado a su pico en agosto de 2022 (13,3 por ciento), según los índices de precios publicados por el Instituto Nacional de Estadística e Informática (INEI). En la misma ventana de tiempo, de acuerdo con la información descargada de precios de supermercados, la mediana de

las variaciones a 365 días de los productos que corresponden a tal categoría se incrementó y llegó a su valor máximo en setiembre de 2022 (12 por ciento). De este modo, se observa una tendencia similar en ambos conjuntos de datos (correlación de 71,4 por ciento).

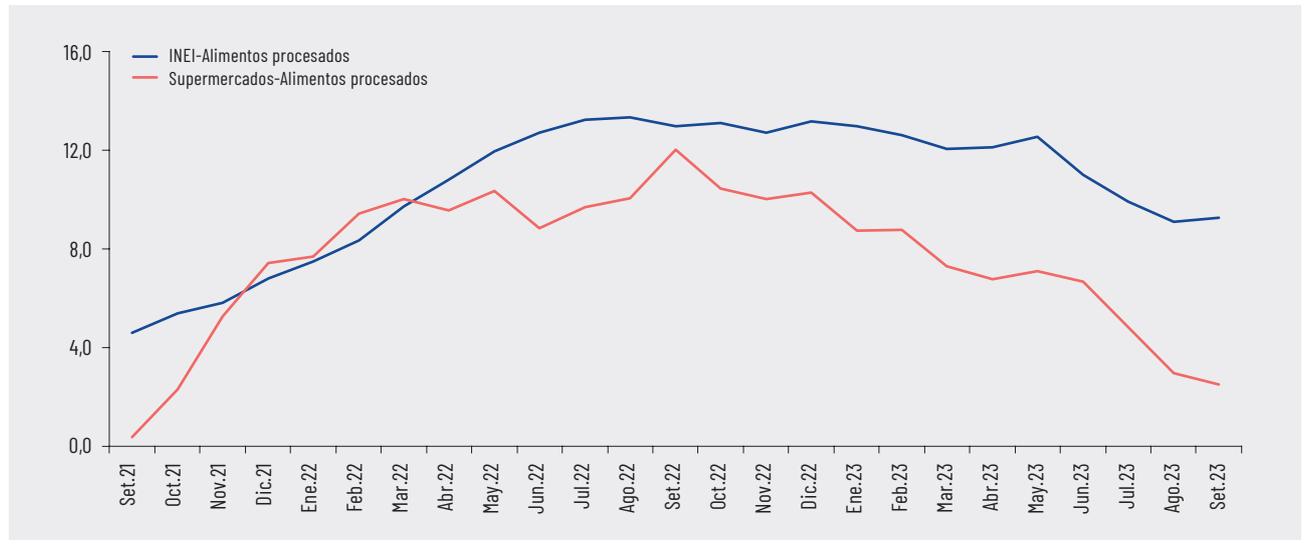
Este artículo tiene como objetivo brindar una primera exploración a los datos de precios de supermercados. De esta forma, se profundiza en una aplicación de la información disponible para el análisis coyuntural de la inflación y también se abre las puertas a futuros análisis para sintetizar resultados relevantes para el trabajo de la institución.

DATOS

Mediante el uso de técnicas de *web scraping* se obtuvo información y precios, desde abril de 2020, de todos los productos vendidos a través de la página web de supermercados. La descarga de datos se realizó mediante la librería "Selenium" de Python, con la que se obtiene de manera automática la información básica de los registros de cada producto en la web: (i) nombre, (ii) precio, (iii) marca, (iv) la categoría y (v) el código SKU³. Con respecto a los precios, los productos ofertados en los sitios web de los supermercados pueden tener tres precios: (i) el precio regular⁴, (ii) el precio "online", mejor precio o precio con descuento (según la estructura HTML suelen recibir estas etiquetas) y (iii) el precio con descuento especial usando una tarjeta particular.

Al 17 de octubre de 2023, se han obtenido alrededor de 554 mil registros únicos de variedades de pro-

GRÁFICO 1 ■ Mediana de variaciones a 365 días de la media móvil de precios de supermercados vs. la variación a 12 meses del indicador de alimentos procesados del INEI



FUENTE: INEI Y SUPERMERCADOS. ELABORACIÓN PROPIA.

1 El BCRP clasifica los índices de precios en 2 grandes categorías: precios de alimentos y energía (AE) e inflación sin alimentos y energía (SAE). Estas dos categorías reflejan los precios más y menos volátiles de la economía, respectivamente.
 2 Los alimentos procesados son parte de la inflación subyacente, según algunas metodologías de cálculo, por lo que su evolución no estaría afectada por situaciones coyunturales.
 3 Es la categoría asignada dentro de la estructura del sitio web del supermercado.
 4 El precio regular solo aparece cuando en el sitio web se indica que el precio *online* implica un descuento. No se ha realizado una comprobación física de que el precio regular corresponda al precio con el que el producto se vende en las tiendas.

GRÁFICO 2 ■ Esquema de la clasificación del consumo individual por finalidades

Nivel 1	División	Alimentos y bebidas no alcohólicas					
Nivel 2	Grupo	Alimentos					
Nivel 3	Clase	Pan y cereales					
Nivel 4	Subclase	Arroz de todas las clases					
Nivel 5	Rubro	Arroz					
Nivel 6	Producto	Arroz envasado					
Nivel 7	Artículo	Arroz envasado extra					
Nivel 8	Variedad específica	Arroz paisana bolsa naranja 1 kg					

FUENTE: INEI.
ELABORACIÓN PROPIA.

ductos que alguna vez han sido vendidos, sin importar si aún siguen disponibles⁵. Actualmente el flujo de información nueva en la base de datos incluye alrededor de 170 mil registros diarios. Del total del flujo diario, 14 mil productos corresponden a alimentos procesados.

Para poder facilitar el análisis de precios se agrupó los productos de acuerdo con la Clasificación del Consumo Individual por Finalidades (CCIF o COICOP, por sus siglas en inglés), que es la clasificación de referencia internacional del gasto de los hogares⁶. El INEI utiliza esta clasificación desde enero 2022 para estructurar la canasta de consumo con la que se construye el índice de precios al consumidor (IPC). Cada registro (producto listado en los sitios web de los supermercados) de la base de datos corresponde a la unidad de medida más pequeña de la CCIF, que es la variedad. Luego ello se va agregando en diferentes categorías, de las cuales el nivel de producto y rubro son las más comunes para realizar el análisis de precios.

La clasificación de las variedades en productos y rubros de la canasta de consumo del IPC se realizó utilizando técnicas de *machine learning*, con un modelo de aprendizaje supervisado de *text classification*⁷ (Armand et al., 2017). Para el caso de los alimentos, de un total de 22 mil registros correspondientes a un supermercado, se clasificó manualmente el 50 por ciento para construir una muestra de entrenamiento, que luego se utilizó para que el modelo clasifique el 50 por ciento

restante, como también los datos de otros supermercados.

SELECCIÓN DE SERIES

Primero, se seleccionó alrededor de 5 mil variedades de productos correspondientes a alimentos procesados que contaban con las series más largas para el análisis⁸. Posteriormente, se limitó la muestra de análisis a mil variedades de alimentos procesados que corresponden a ocho tipos de productos para simplificar el análisis.

PATRONES

Para clasificar las series seleccionadas en patrones de formación de precios se han tomado dos criterios: la volatilidad diaria de las series y la frecuencia de actualizaciones de precios.

La volatilidad de las series se relaciona con la presencia de descuentos, que parece ser un elemento importante dentro del proceso de formación de precios en los supermercados. Sin embargo, estas fluctuaciones constantes de precios dificultan la identificación de cambios persistentes de niveles. Por ende, para obtener el nivel de precio estable de cada producto se aplicó un algoritmo de reconocimiento de descuentos basado en la metodología de Nakamura y Steinsson (2008). Esta plantea distinguir periodos de descuentos para detectar el nivel de precio “estable”, detectando patrones de precios en formas V simétricas y asimétri-

5 Las principales causas para que un registro deje de actualizarse son que su venta haya sido interrumpida o que su presentación haya cambiado, lo cual genera que se convierta en un registro nuevo. Generalmente, en periodos de alta inflación, las empresas pueden optar por incrementar precios o reducir el tamaño de sus productos. Estas nuevas presentaciones se registran como un producto nuevo en la base de datos.

6 La CCIF fue adoptada por la Comisión de Estadísticas de las Naciones Unidas en 1999 y ha sido revisada por este mismo organismo en 2018 como el estándar internacional aceptado para la construcción de índices de precios.

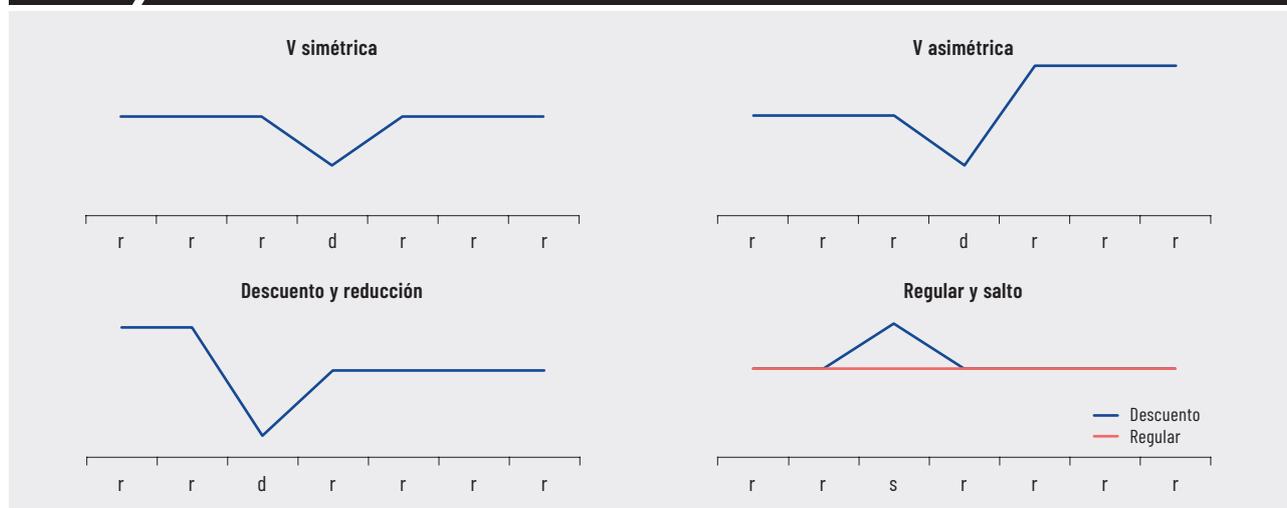
7 La librería disponible en Python se denomina *fasttext*.

8 Se descartaron series que contaban con menos del 70 por ciento de observaciones en toda la muestra. A pesar de que contar con series más largas nos permite identificar regularidades en los datos, este método simple para filtrar las series significa la pérdida de datos de (i) productos que cambian de presentación o (ii) de productos nuevos que ingresan al mercado y cuentan con pocas observaciones. Está pendiente afrontar esta limitación estableciendo un modo automatizado de empalme de series de precios de productos con cambios de presentación.

CUADRO 1 ■ Supuestos y parámetros elegidos para la parametrización

Supuesto	Explicación	Elección
Duración máxima de descuentos	Define el tiempo máximo que dura un descuento para distinguirlo de reducciones en los precios.	100 días
Duración máxima de descuentos que coinciden con reducciones	En caso de descuentos simultáneos a una reducción de precios regulares, se establece un límite de tiempo para diferenciarlo de reducciones de precios.	30 días
Duración máxima de un salto	El filtro detecta saltos de precios hacia arriba y, en caso de que duren menos de una ventana establecida, no se considera como incremento en el precio regular.	10 días

GRÁFICO 3 ■ Esquema de descuentos en forma de V simétrica y asimétrica



NOTA: LA SIGLA "R" DENOMINA A LOS PRECIOS REGULARES; "D", A LOS PRECIOS CON DESCUENTO; Y "S", AL SALTO DE PRECIO.

cas. Los patrones son simétricos cuando luego del descuento se regresa al mismo valor del precio original, mientras que, cuando son asimétricos, el rebote puede ser a valores superiores. Esta metodología se potencia cuando se usan datos de alta frecuencia como los trabajados en este artículo. En el Cuadro 1 se explica la calibración del filtro.

Para determinar la volatilidad de las series, se calculó un indicador basado en el coeficiente de variabilidad:

$$Volatilidad_i = \sqrt{\frac{\sum_{t=1}^T (x_{it} - \bar{x}_{it})^2}{\sum_{t=1}^T \frac{x_{it}}{T}}} * 100$$

La desviación estándar se calcula usando la diferencia del precio en niveles y tomando su valor ancla estable. Este indicador de dispersión se divide entre el promedio de toda la segunda serie. Esta medida permite capturar la volatilidad generada por saltos y su magnitud, eliminando el ruido por la presencia de tendencias crecientes.

Para calcular la frecuencia de actualización se promedió el número de veces que un precio "estable" o "ancla" cambia por trimestre (hacia arriba o abajo). De tal forma, este indicador mide el grado de ajuste limpio de descuentos.

Se identificaron cinco patrones en la formación de precios de alimentos procesados. Primero se distinguieron los productos que no presentan descuentos. Luego, mediante un algoritmo de *k-means clustering*⁹ se distinguieron tres grupos. Finalmente, se separaron los productos que no han sufrido actualizaciones. La variable que define los grupos es la volatilidad. Se nota en el Gráfico 4 cómo la mayoría de alimentos procesados se actualiza en promedio trimestralmente entre 0 y 1 veces.

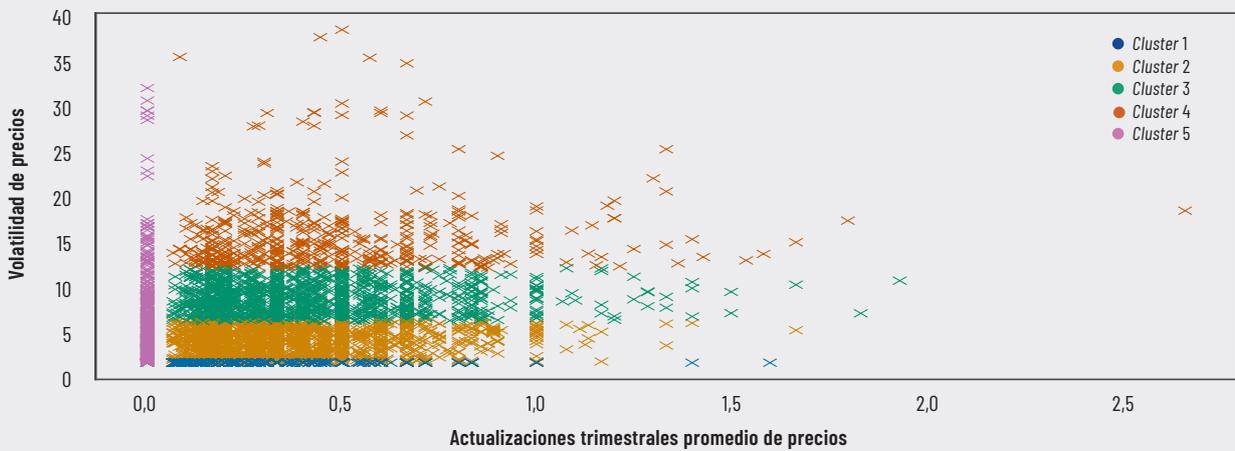
Estos resultados no pueden ser extrapolados como una guía para el comportamiento general de todos los precios, pues, además de supermercados, hay otras fuentes de información que no son objeto de análisis de este artículo que sí influyen en el gasto familiar.

Para la presentación de las series de productos se ha censurado el nombre para preservar la confidencialidad de los datos. Así, la denominación pasa a ser el rubro al cual fue clasificada la variedad acompañada de un número que diferencia marcas e información de tamaños o sabores. Por ejemplo, un producto de nombre "Galletas marca ABC sabor chocolate de 60 gramos" pasa a "Galletas rellenas 1 sabor 1 pequeña".

En los gráficos de la siguiente sección se presenta las series en niveles (azul), para observar la volatilidad diaria, las medias móviles de 30 días (negro) y el pre-

9 Este es un algoritmo de *machine learning* que clasifica observaciones en *clusters* de acuerdo a su comportamiento. En este caso, respecto a la volatilidad y a las actualizaciones trimestrales.

GRÁFICO 4 ■ Relación entre volatilidad y actualizaciones de precios



FUENTE: SUPERMERCADOS.
ELABORACIÓN PROPIA.

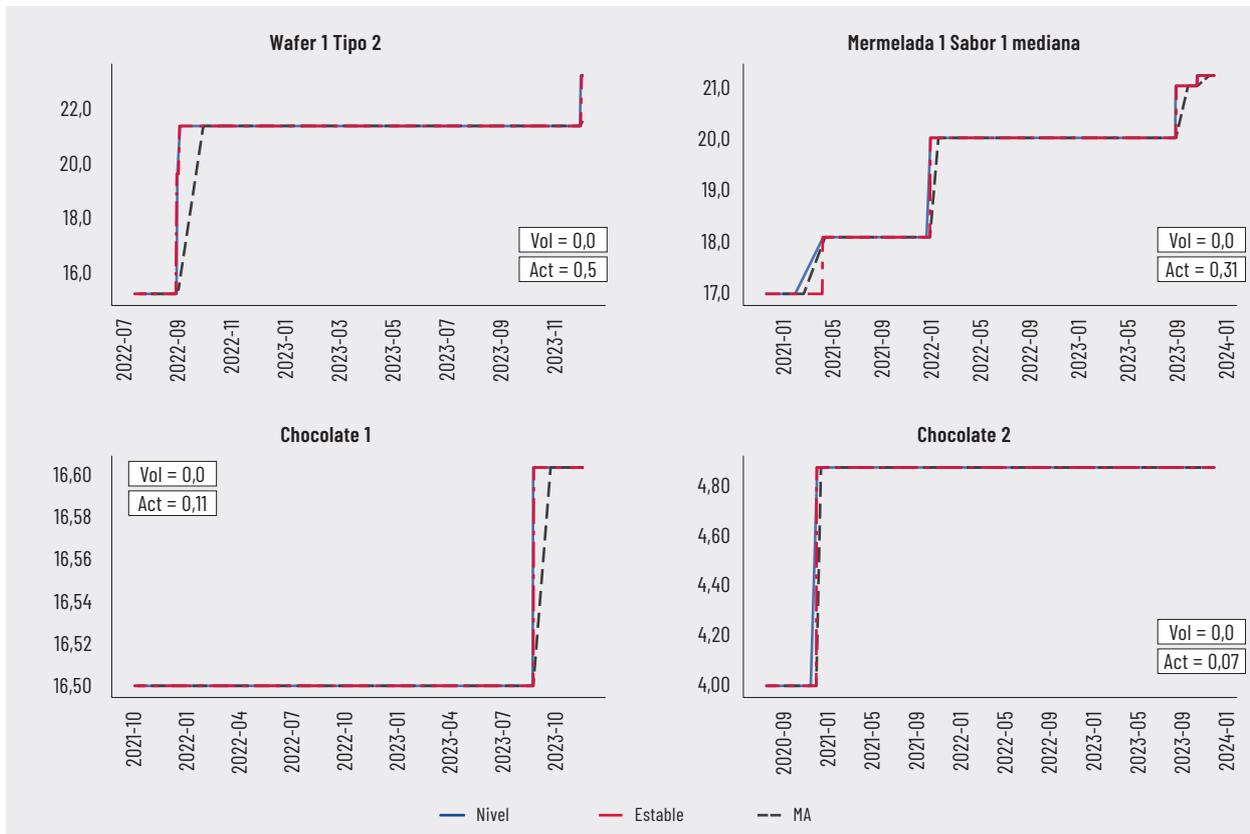
cio estable o ancla al que reierten las series (rojo) para identificar puntos de inflexión de carácter más permanente y facilitar el análisis de tendencias.

i. Cluster 1: series de precios sin volatilidad

Estos productos se caracterizan por no presentar descuentos que generen volatilidad en las series, pero sí ajustes en los niveles estables de precios. Así, se aseme-

jan a una forma de escalera ascendente. Algunos productos demoran mucho tiempo en presentar un cambio. En el caso del "chocolate 1" pasaron casi dos años antes de subir 10 céntimos. Estos productos parecerían seguir un comportamiento de *always on discount*. La media móvil demora un periodo de 30 días en actualizarse al nuevo valor, pero no está sujeto a variación por descuentos.

GRÁFICO 5 ■ Ejemplos del cluster 1



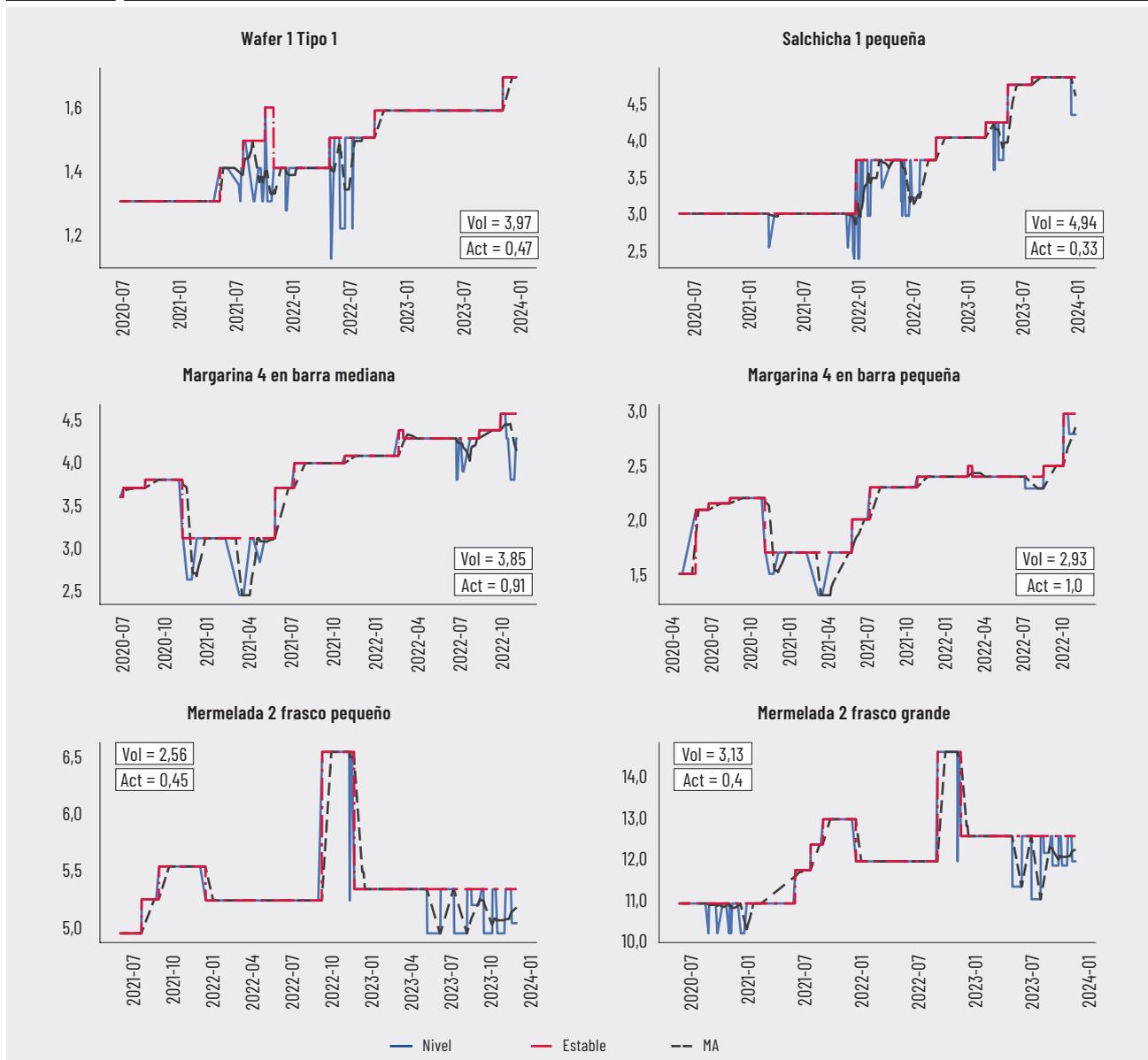
NOTA: VOL Y ACT SON LOS INDICADORES DE VOLATILIDAD Y ACTUALIZACIÓN, RESPECTIVAMENTE. NIVEL, MA Y ESTABLE SON LOS PRECIOS EN NIVELES, EN MEDIA MÓVIL 30 DÍAS Y REGULARES SEGÚN EL ALGORITMO DE DESCUENTOS, RESPECTIVAMENTE. ELABORACIÓN PROPIA.

ii. Cluster 2: Series de precios con volatilidad baja

Estos productos se caracterizan por tener una volatilidad interdiaria baja debido a la poca presencia de descuentos. Esto da la apariencia de una evolución tendencial en la serie. El comportamiento por tamaño dentro de un mismo tipo de producto puede variar. Por ejemplo, en el caso de la "mermelada 2", a pesar de tener dos presentaciones (frasco grande y pequeño), tiene un comportamiento de volatilidad y actualización similar. Por otro lado, las presenta-

ciones pequeñas de "margarina 4" varían su precio regular con mayor frecuencia y a un nuevo nivel proporcionalmente mayor que el tamaño mediano. Esto se evidencia en el valor del indicador de frecuencia de actualización que difiere entre presentaciones, lo cual implica mayor cantidad de saltos y mayor intensidad en estos respecto al valor promedio. Una explicación es la no linealidad en el precio por unidad. Por ejemplo, la presentación barra pequeña llegó a valer 2,99 soles (0,03 soles por gramo) que difiere de la presentación mediana de 4,50 (0,02 soles por

GRÁFICO 6 | Ejemplos del cluster 2



NOTA: VOL Y ACT SON LOS INDICADORES DE VOLATILIDAD Y ACTUALIZACIÓN, RESPECTIVAMENTE. NIVEL, MA Y ESTABLE SON LOS PRECIOS EN NIVELES, EN MEDIA MÓVIL 30 DÍAS Y REGULARES SEGÚN EL ALGORITMO DE DESCUENTOS, RESPECTIVAMENTE. ELABORACIÓN PROPIA.

gramo). Ello podría tener implicancias sobre el gasto en alimentos de la población de menores ingresos, ya que las presentaciones pequeñas son consumidas por aquellas familias que viven de ingresos diarios y no pueden planificar para comprar en cantidades mayores.

iii. Cluster 3: Series de precios de volatilidad media

En el caso de este *cluster*, los precios tienen una mayor cantidad de descuentos. La "salchicha 2 pequeña" presenta una frecuencia de actualización mayor que la presentación grande, pero una menor volatilidad. Esta particularidad podría reflejar compensación de márgenes entre diferentes variedades de un mismo producto, elevando los precios en aquellas presentaciones cuya elasticidad de demanda sea menor. Sin embargo, si los aumentos de precios son persistentes a nivel de toda la economía, se

podrían ver obligados a ajustar incluso en las variedades que se han mantenido estables o rezagadas.

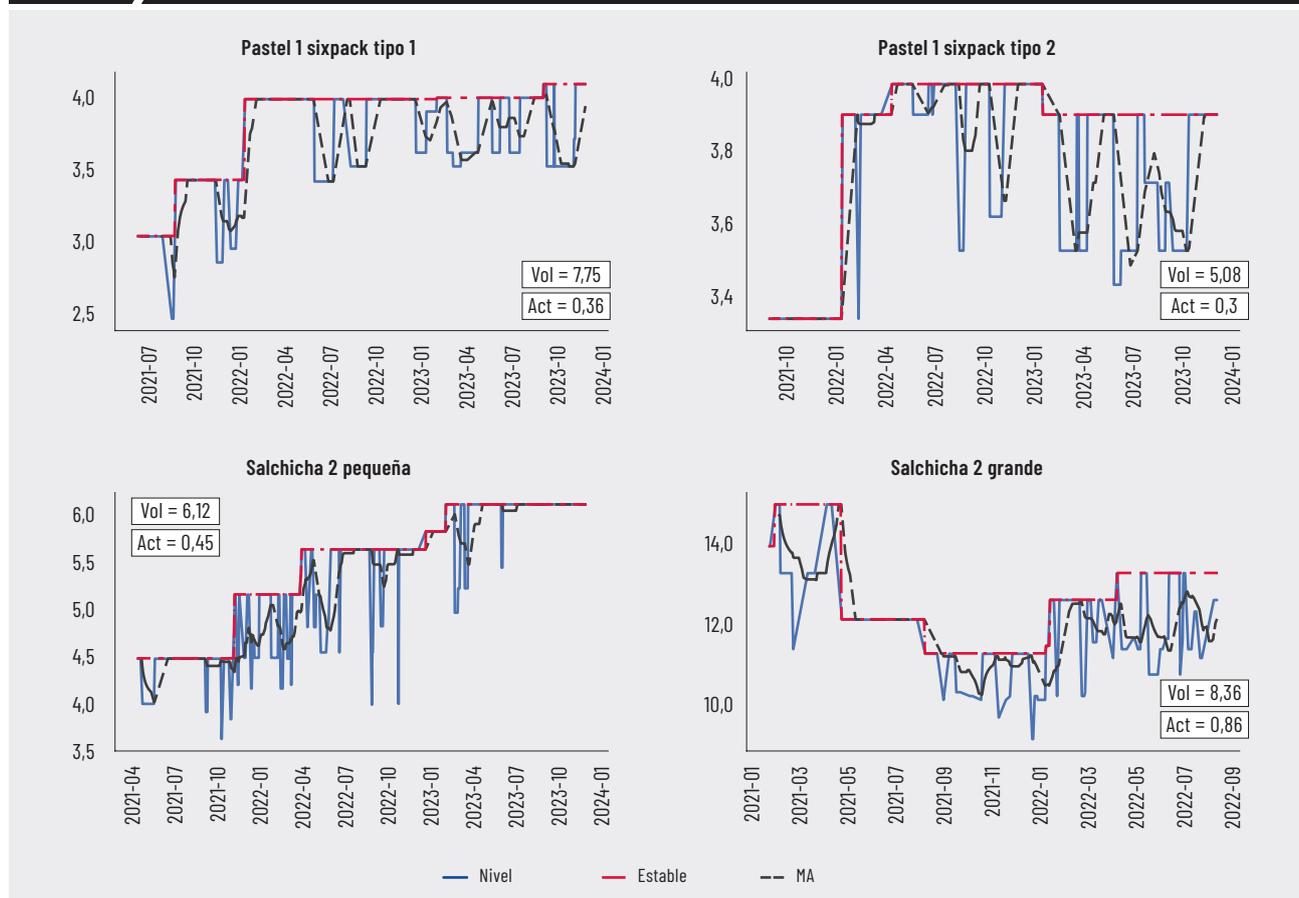
iv. Cluster 4: Series de precios con volatilidad alta

En estos productos, la volatilidad es la mayor de toda la muestra. Así, entran constantemente en periodos de descuentos. Es por ello que la media móvil se vuelve más volátil al capturar los precios bajos a pesar de no observarse un cambio en el precio estable. Algunos productos de este *cluster* presentan valores *outliers* que parecen ser errores desde el origen de datos.

v. Cluster 5: Series sin cambio en el nivel estable de precios

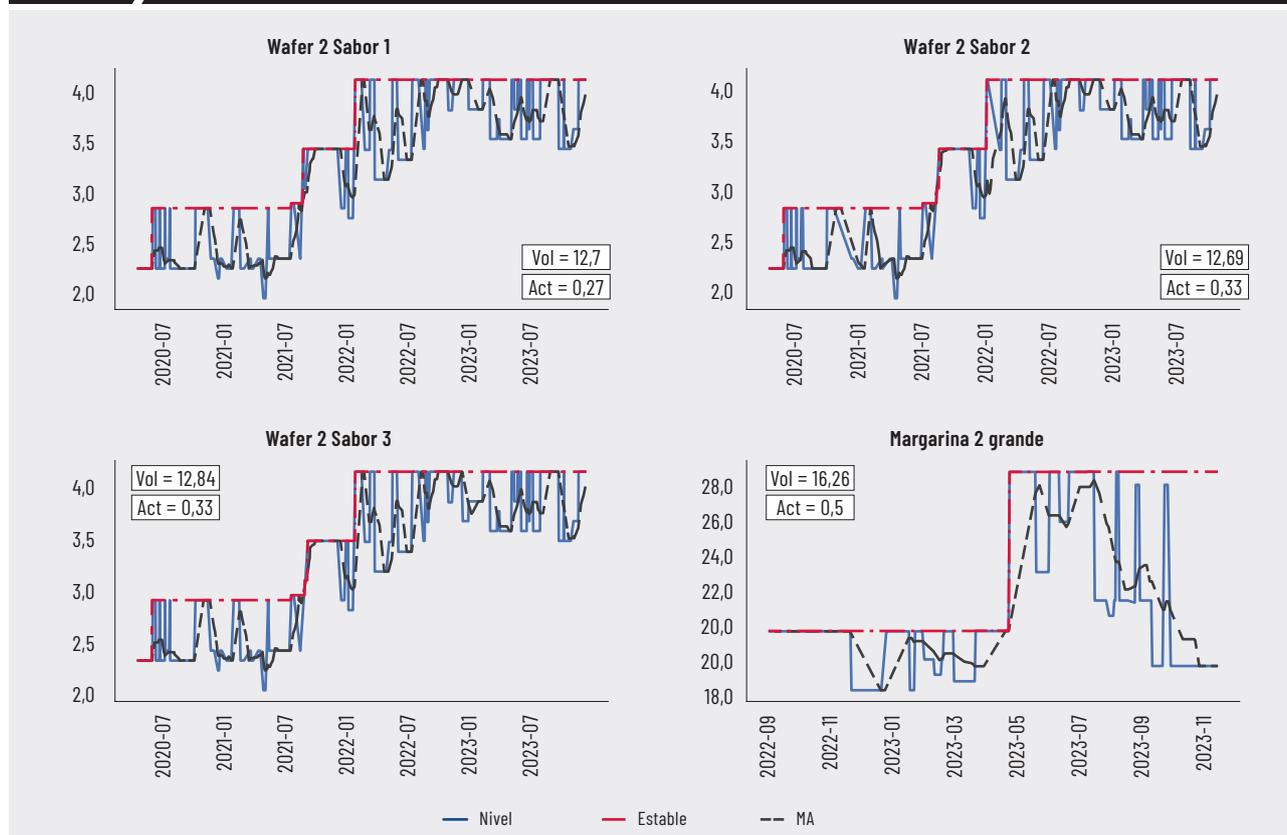
Este grupo se caracteriza por no tener cambios en el nivel estable de precios. En algunos casos puede no haber descuentos (como en el caso de las ha-

GRÁFICO 7 ■ Ejemplos del cluster 3



NOTA: VOL Y ACT SON LOS INDICADORES DE VOLATILIDAD Y ACTUALIZACIÓN, RESPECTIVAMENTE. NIVEL, MA Y ESTABLE SON LOS PRECIOS EN NIVELES, EN MEDIA MÓVIL 30 DÍAS Y REGULARES SEGÚN EL ALGORITMO DE DESCUENTOS, RESPECTIVAMENTE. ELABORACIÓN PROPIA.

GRÁFICO 8 | Ejemplos del cluster 4



NOTA: VOL Y ACT SON LOS INDICADORES DE VOLATILIDAD Y ACTUALIZACIÓN, RESPECTIVAMENTE. NIVEL, MA Y ESTABLE SON LOS PRECIOS EN NIVELES, EN MEDIA MÓVIL 30 DÍAS Y REGULARES SEGÚN EL ALGORITMO DE DESCUENTOS, RESPECTIVAMENTE. ELABORACIÓN PROPIA.

rinas 1 “grande” y “pequeña”). Sin embargo, ello no es característico del grupo. La “galleta 1” y la “mermelada 3 pequeña” han tenido variaciones que duran menos de 100 días.

Un producto con un precio estático puede ser reflejo de 3 circunstancias: (i) que no hay presiones de costos en los precios de los insumos, (ii) que la empresa productora está reduciendo sus márgenes para evitar perder mercado ante una eventual alza de precios o (iii) una ganancia de eficiencia en la producción o un deterioro en la calidad del producto.

CONCLUSIÓN

Los datos de alta frecuencia de precios de supermercados ofrecen una mirada al proceso de formación de precios. En este artículo se analizaron cinco patrones dentro del grupo de alimentos procesados. Los grupos y productos varían respecto a su volatilidad (número de descuentos) y la actualización (frecuencia de ajuste de precios).

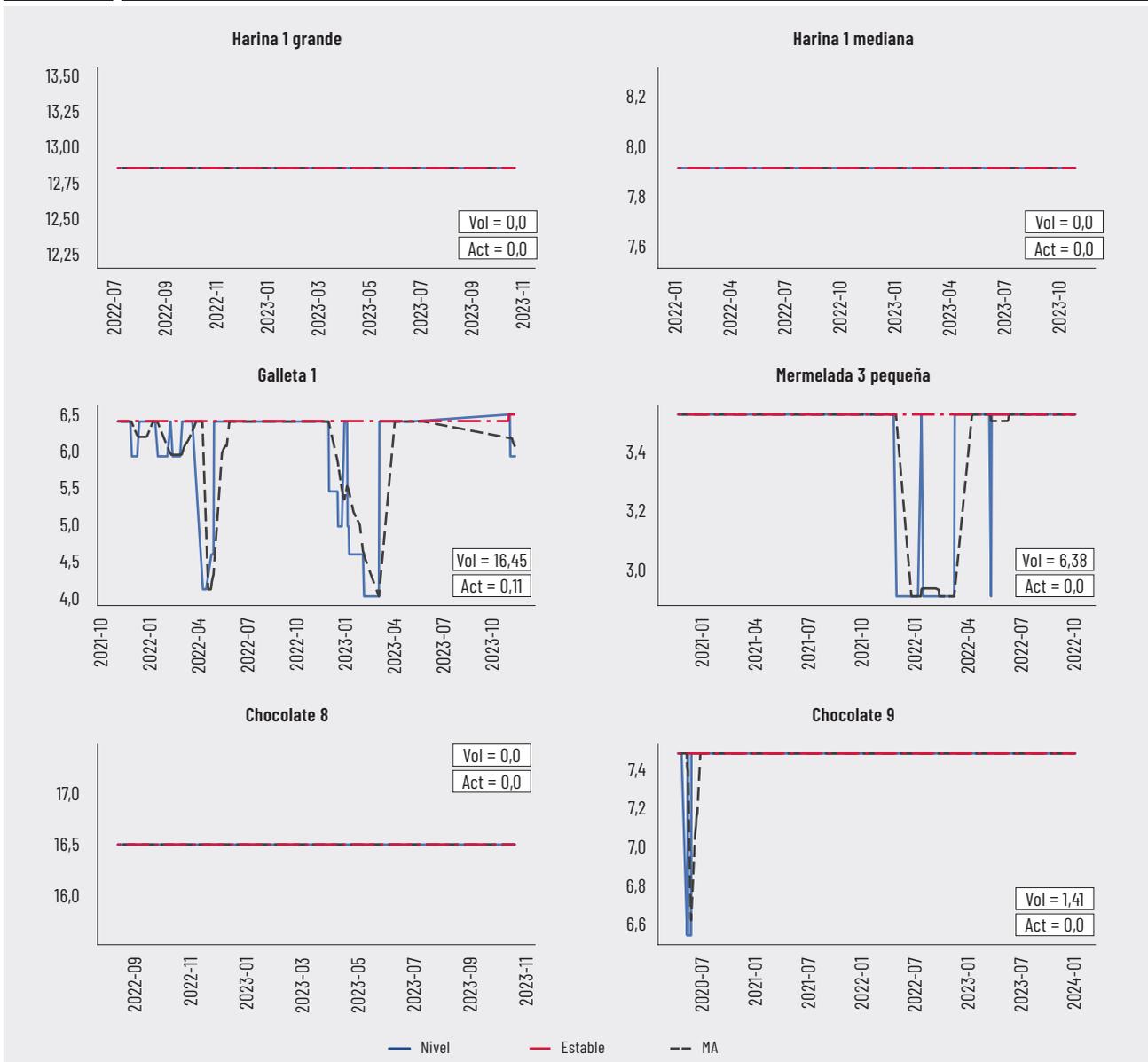
Hay tres resultados importantes de este estudio. Primero, los cambios de precios pueden variar en comportamiento dependiendo del tamaño de las variedades de productos. Ello a su vez podría

estar asociado a las elasticidades de gasto en alimentos de la población de menores ingresos, que consume presentaciones de menor tamaño. Esta particularidad podría reflejar una compensación de márgenes entre diferentes variedades de un mismo producto, elevando los precios en aquellas presentaciones cuya elasticidad de demanda sea menor. En caso de inflación generalizada, las variedades estables o rezagadas tendrían una actualización de precios.

En segundo lugar, hay productos cuyos precios se mantienen estáticos. Ello puede deberse a (i) ausencia de presiones de costos en los precios de los insumos, (ii) reducción de márgenes para evitar perder mercado ante una eventual alza de precios o (iii) una ganancia de eficiencia en la producción o un deterioro en la calidad del producto.

Finalmente, en productos de mayor volatilidad, la media móvil captura los movimientos de descuentos en los precios y no necesariamente los cambios en precios estables. Ello no necesariamente implica que el indicador tiene un problema, sino que su lectura debe estar asociada al precio efectivo que paga un consumidor y no siempre a los niveles tendenciales.

GRÁFICO 9 ■ Ejemplos del cluster 5



NOTA: VOL Y ACT SON LOS INDICADORES DE VOLATILIDAD Y ACTUALIZACIÓN, RESPECTIVAMENTE. NIVEL, MA Y ESTABLE SON LOS PRECIOS EN NIVELES, EN MEDIA MÓVIL 30 DÍAS Y REGULARES SEGÚN EL ALGORITMO DE DESCUENTOS, RESPECTIVAMENTE. ELABORACIÓN PROPIA.

Para el trabajo realizado, ha sido fundamental la diferenciación de periodos de descuentos y de precios regulares. El algoritmo implementado ha permitido generar series que capturan mejor la tendencia de los precios. Así, resalta la importancia de diferenciar movimientos en los precios que correspondan a volatilidad y no a saltos de la tendencia. Otro pendiente de la investigación es estudiar la relación de precios con las unidades de medida. No se puede asumir necesariamente linealidad, pues diferentes presentaciones pueden estar sujetas a descuento por mayores tamaños como también a mayor ruido en el caso de las más pequeñas.

REFERENCIAS

- Armand, J., Edouard, G., Piotr, B., & Tomas, M. (2017). *Bag of Tricks for Efficient Text Classification*. Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers.
- Instituto Nacional de Estadística e Informática (2023). *Índices de precios al consumidor en Lima Metropolitana*.
- Nakamura, E., & Steinsson, J. (2008). Five Facts about Prices: A Re-evaluation of Menu Cost Models. *The Quarterly Journal of Economics*, 1415-1464.