

From Daily Prices to Inflation Trends

Comparing Stylized Facts in Supermarket Data with Official Inflation
Statistics

Gonzalo Bueno & Marco Vega

Octubre 2024

Índice

1 Introducción

2 Antecedentes

3 Datos y Metodología

- Recolección de Datos
- Construcción del Índice
- Filtrado de Descuentos

4 Resultados

- Índices de Precios Diarios
- Índice Suavizado con Media Móvil
- Comparación con Datos Oficiales
- Ratios y Análisis de Categorías
- Ratio Precios de Alimentos a Bienes
- Impulsos respuesta: En Línea vs. Oficiales

5 Conclusiones

Introducción

- **Motivación:** El monitoreo de la inflación se basa tradicionalmente en datos oficiales mensuales, que suavizan las fluctuaciones a corto plazo.
- **Objetivo:** Utilizar datos de supermercados obtenidos por web scraping a alta frecuencia para construir índices de precios diarios y compararlos con las medidas oficiales de inflación.
- **Contribución:** Proporciona ideas sobre cómo los datos en tiempo real pueden mejorar el monitoreo de la inflación, ofreciendo señales tempranas de movimientos de precios.

Antecedentes

- Los avances en la recolección de datos, particularmente el web scraping, permiten un seguimiento granular de los precios de miles de productos.
- Estudios previos (e.g., Cavallo, 2013) demostraron la viabilidad del uso de datos en línea para estimaciones de inflación.
- Los desafíos incluyen la volatilidad de la muestra, los efectos de descuentos y el filtrado de ruido, aspectos que aborda este estudio.

Recolección de Datos

- **Fuente:** Datos obtenidos mediante web scraping de supermercados peruanos, recolectados diariamente.
- **Período:** Abril 2020 a Septiembre 2024.
- **Tamaño de Muestra:** Más de 100,000 productos.
- **Filtrado:**
 - ▶ Se excluyeron productos con menos del 50% de disponibilidad de datos durante su periodo de vigencia en la muestra.
 - ▶ Se mantuvieron los productos con una correlación alta con su rubro correspondiente en data oficial (**supuesto:** la información oficial de inflación es más informativa que una parte de productos en la economía).

Construcción del Índice

- Se siguió la metodología de Cavallo (2013) para construir índices de precios diarios.
- **Pasos:**
 - 1 Calcular las variaciones diarias de precios como porcentajes para cada producto.
 - 2 Agregar estas variaciones utilizando promedios geométricos ponderados por igual (Índice de Jevons):

$$R_{i,t} = \left(\prod_j \frac{p_{j,t}}{p_{j,t-1}} \right)^{\frac{1}{n_{i,t}}} \quad (1)$$

donde $R_{i,t}$ es la variación para la categoría i en el día t . **Supuesto:** los pesos en el consumo de cada producto son iguales.

- 3 Acumular las variaciones diarias para formar el índice:

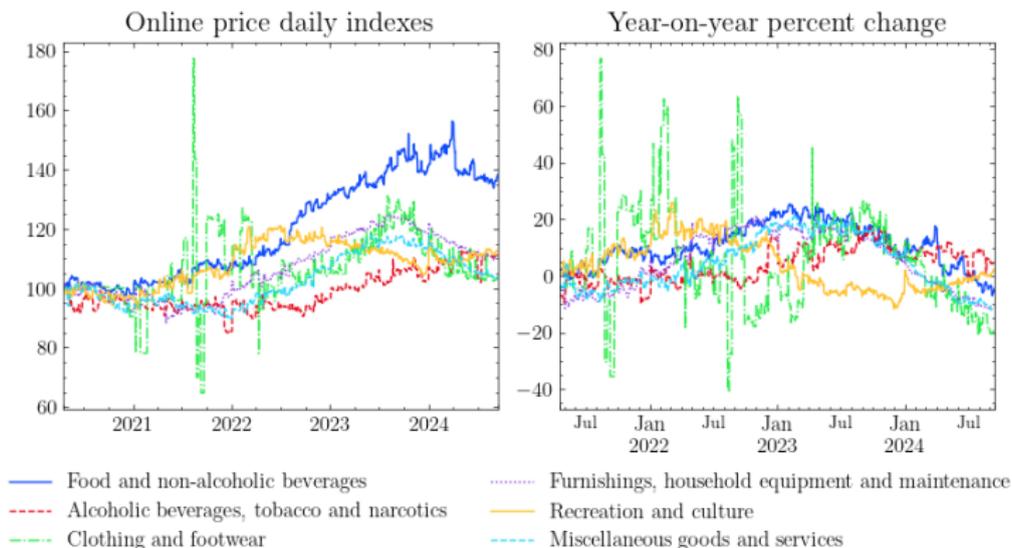
$$I_{i,t} = \left(\prod_{\tau=1}^t R_{i,t-\tau+1} \right) \times 100 \quad (2)$$

Filtrado de Descuentos

- Los datos de supermercados incluyen precios regulares y con descuento.
- **Enfoque:** Se utilizó el mínimo de los dos para reflejar el costo efectivo para el consumidor.
- Los descuentos, la sustitución de productos o la disponibilidad de stock pueden provocar *chain drift*: una situación en la que los precios que componen un índice fluctúan y luego regresan a su valor inicial, pero el índice no logra volver a su valor base. Esto puede ocurrir por la ponderación fluctuante: $\frac{1}{n_{i,t}}$.
- Se aplicó un filtro para excluir caídas de precios a corto plazo (menos de 7 días) como ruido.
- Basado en Nakamura y Steinsson (2008), distinguiendo descuentos persistentes de los temporales.

Índices de Precios Diarios

- Alta volatilidad observada en los índices diarios debido a cambios frecuentes a corto plazo.
- Categorías como ropa y calzado mostraron mayor ruido, particularmente debido a datos limitados hasta mediados de 2022.
- Los valores del índice diario proporcionaron ideas valiosas en tiempo real, pero requerían suavizado para el análisis de tendencias.



Índice Suavizado con Media Móvil

- **Enfoque de Suavizado:** Se aplicó una media móvil de 30 días.
- El suavizado ayuda a proporcionar una tendencia mensual, comparable con las estadísticas oficiales de inflación.
- Equilibra la necesidad de oportunidad con la reducción de las fluctuaciones a corto plazo.

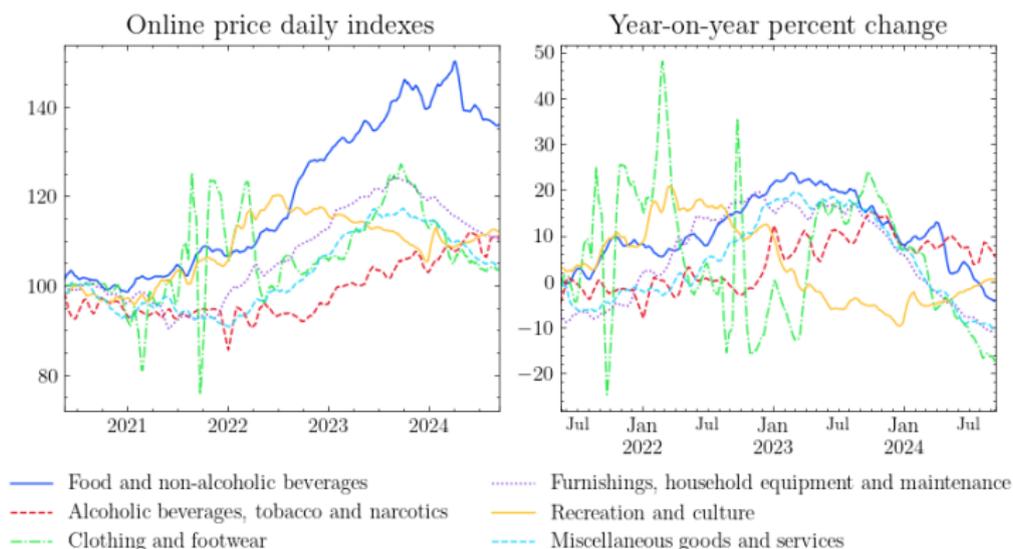


Figure: Media Móvil de 30 Días del Índice Diario del IPC

Comparación con Datos Oficiales

- Las tendencias fueron generalmente alineadas, pero los datos en línea mostraron picos y valles más pronunciados.
- Los índices en línea capturaron fluctuaciones a corto plazo que los datos oficiales mensuales pasaron por alto.

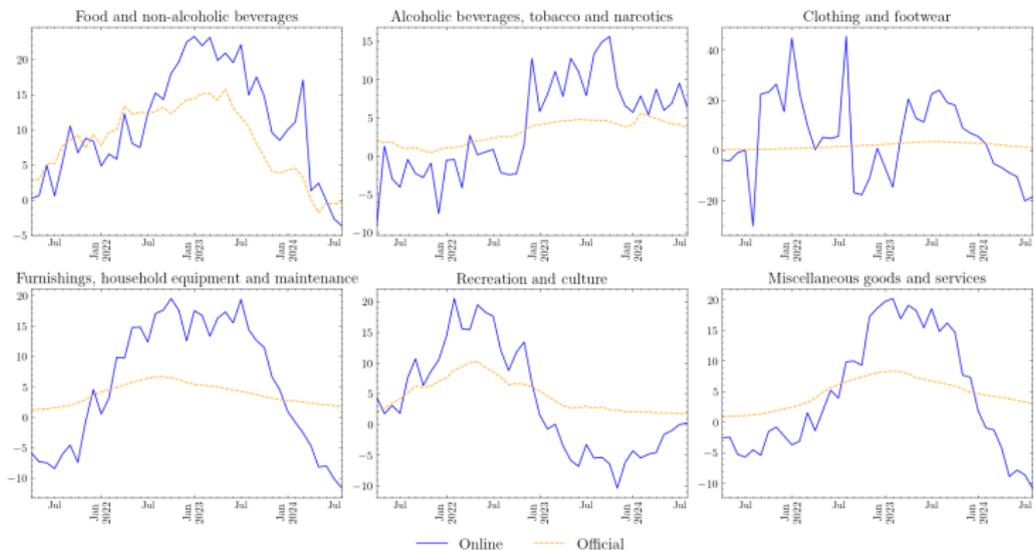


Figure: Índices en Línea vs. Índices Oficiales

Ratios Índices en Línea a Oficiales

- Se calcularon ratios para analizar discrepancias entre índices en línea y oficiales.
- **Observaciones:**
 - ▶ Los ratios tendieron a estabilizarse, sugiriendo convergencia entre datos en línea y oficiales.
 - ▶ Los alimentos y bebidas no alcohólicas mostraron ratios más altos, indicando aumentos más pronunciados en precios en línea.

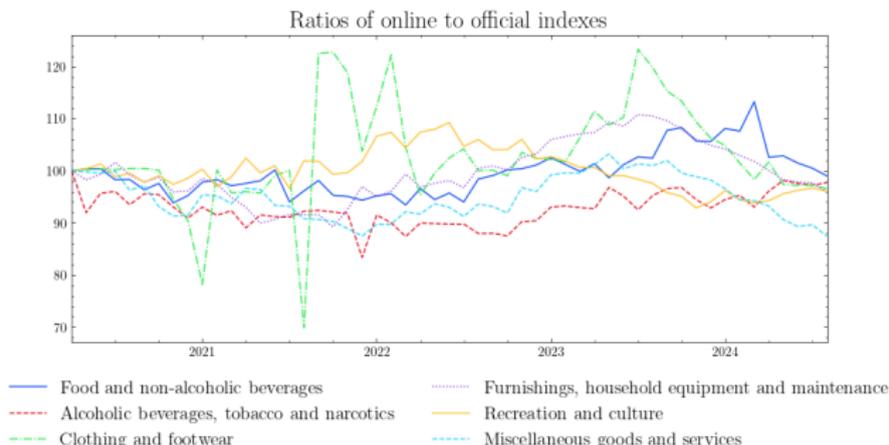


Figure: Ratios: Índices en Línea vs. Oficiales

Ratio Precios de Alimentos a Bienes

- **Objetivo:** Examinar cómo se comportan los precios de alimentos en relación con los bienes no alimentarios.
- **Resultados:**
 - ▶ Tanto los datos en línea como los oficiales muestran que los precios de los alimentos aumentaron más rápido que los de bienes, especialmente durante períodos de aumentos de inflación.
 - ▶ Los datos en línea exhibieron más volatilidad, con picos y caídas más pronunciados.
 - ▶ El ratio proporciona información sobre la divergencia persistente entre los precios de alimentos y bienes, reflejando choques de oferta externos y cambios en la demanda.



Impulsos respuesta: En Línea vs. Oficiales

- Se utilizaron IRFs para comparar la respuesta del precio de alimentos a choques económicos.
- **Conclusión Clave:** Los índices en línea mostraron reacciones más rápidas pero menos persistentes a los choques en comparación con los índices oficiales.

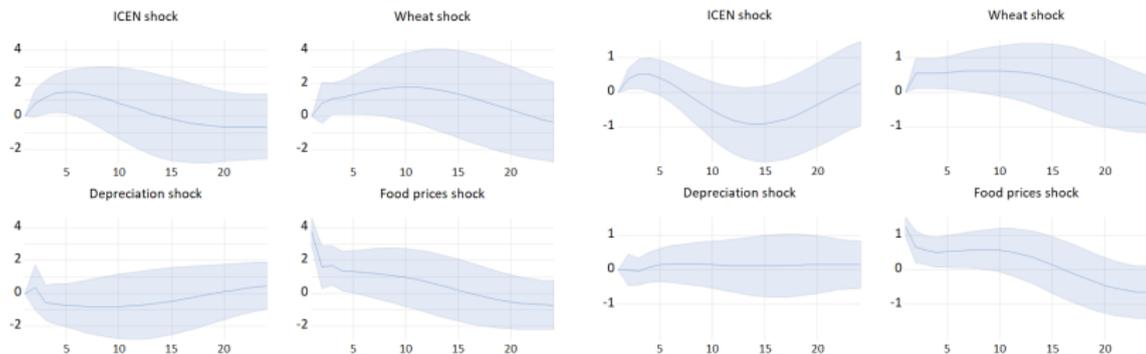


Figure: Funciones de Respuesta al Impulso: (a) Índice en Línea (izquierda) vs. (b) Índice Oficial (derecha)

Conclusiones

- **Resultados Principales:** Los datos de alta frecuencia de supermercados pueden proporcionar señales tempranas de tendencias inflacionarias, complementando las fuentes de datos tradicionales.
- **Desafíos:** Mayor ruido en los datos en línea requiere técnicas efectivas de suavizado y filtrado.
- **Futuras Direcciones:**
 - ▶ Ampliar la muestra para periodos anteriores a 2020.
 - ▶ Mejorar los algoritmos de filtrado para reducir el ruido y mejorar la precisión.
 - ▶ Explorar modelos econométricos adicionales para analizar las dinámicas de la inflación en tiempo real.

From Daily Prices to Inflation Trends

Comparing Stylized Facts in Supermarket Data with Official Inflation
Statistics

Gonzalo Bueno & Marco Vega

Octubre 2024