

Modelos de volatilidad para el sentimiento diario de las noticias: el caso peruano*

Willy Alanya (BCRP) y Nicole Vásquez (UCS)

XLI Encuentro de Economistas - BCRP

23 de Octubre, 2023

* Las opiniones expresadas en este estudio corresponden a los autores y no deben ser atribuidas al BCRP.

Contenidos

- 1 Introducción y objetivo
- 2 Revisión de la literatura
 - Revisión de literatura: Internacional
 - Revisión de literatura: Nacional
- 3 Datos y metodología
- 4 Resultados
- 5 Conclusiones

Contenidos

- 1 **Introducción y objetivo**
- 2 Revisión de la literatura
 - Revisión de literatura: Internacional
 - Revisión de literatura: Nacional
- 3 Datos y metodología
- 4 Resultados
- 5 Conclusiones

Introducción

- En la última década, el uso de técnicas de machine learning ha experimentado un notable aumento en la gestión y análisis de datos, independientemente de su organización.
- En particular, ha habido un aumento significativo en el empleo del análisis textual en el ámbito económico.

Introducción

- Según Algaba et al. (2023), los indicadores de sentimiento derivados de encuestas mensuales o informes anuales muestran una baja precisión en la predicción del sentimiento en tiempo real basado en noticias diarias.
- En ese sentido, un índice de sentimiento de noticias económicas que proporciona datos a una frecuencia diaria ofrecería una aproximación más precisa de las expectativas de los agentes económicos.

Objetivo

- **Objetivo**

- Analizar la volatilidad de un indicador de sentimiento de noticias de alta frecuencia que permite aproximar las expectativas de la economía peruana.

Contenidos

- 1 Introducción y objetivo
- 2 Revisión de la literatura
 - Revisión de literatura: Internacional
 - Revisión de literatura: Nacional
- 3 Datos y metodología
- 4 Resultados
- 5 Conclusiones

Literatura

Internacional

- Algaba et al. (2023) demuestran que incluir mitigadores e intensificadores de palabras mejora la correlación entre el índice de sentimiento y las encuestas de expectativas económicas en la realidad belga.
- Shapiro et al. (2022) indican que uso combinado de diccionarios y valoración de palabras para mejorar la capacidad predictiva del índice en relación a las encuestas de confianza del consumidor.
- Nguyen y La Cava (2020) indican que un diccionario de palabras sencillo puede arrojar resultados similares a aquellos obtenidos mediante algoritmos de aprendizaje automático.

Literatura

Nacional

- Lahura et al. (2023) construyen un índice de legibilidad para analizar la comprensibilidad de comunicaciones emitidas por el BCRP.
- Vega y Pinelo (2022) indican que los mayores episodios de incertidumbre se encuentran asociados con los choques a lo largo de la historia económica en el último centenario del Perú.

Contenidos

- 1 Introducción y objetivo
- 2 Revisión de la literatura
 - Revisión de literatura: Internacional
 - Revisión de literatura: Nacional
- 3 Datos y metodología**
- 4 Resultados
- 5 Conclusiones

Datos

Fuente de Información

- Recopilación periódica elaborada por el Departamento de Prensa del Banco Central de Reserva.
- Período de análisis: 1 de enero de 2007 al 31 de julio de 2023.

Construcción del índice de sentimiento

1. El Departamento de Prensa del BCRP filtra una serie de artículos impresos.
2. Mediante el uso del paquete R tesseract se realiza la lectura de las noticias.
3. Las noticias en texto se traducen de español a inglés haciendo uso del traductor de Microsoft Office.
4. Se mide el sentimiento diario de las noticias usando sentometrics.

Construcción del índice de sentimiento

La medición del sentimiento de las noticias sigue la siguiente fórmula, para cada diccionario:

$$s_t = \frac{1}{W} \sum_{i=1}^W val_i \cdot palabra_i$$

Donde:

- s_t , representa el valor del indicador en el día t .
- W , representa el total de términos.
- val_i , representa los intensificadores o mitigadores.
- $palabra_i$, representa el termino i de cada diccionario.

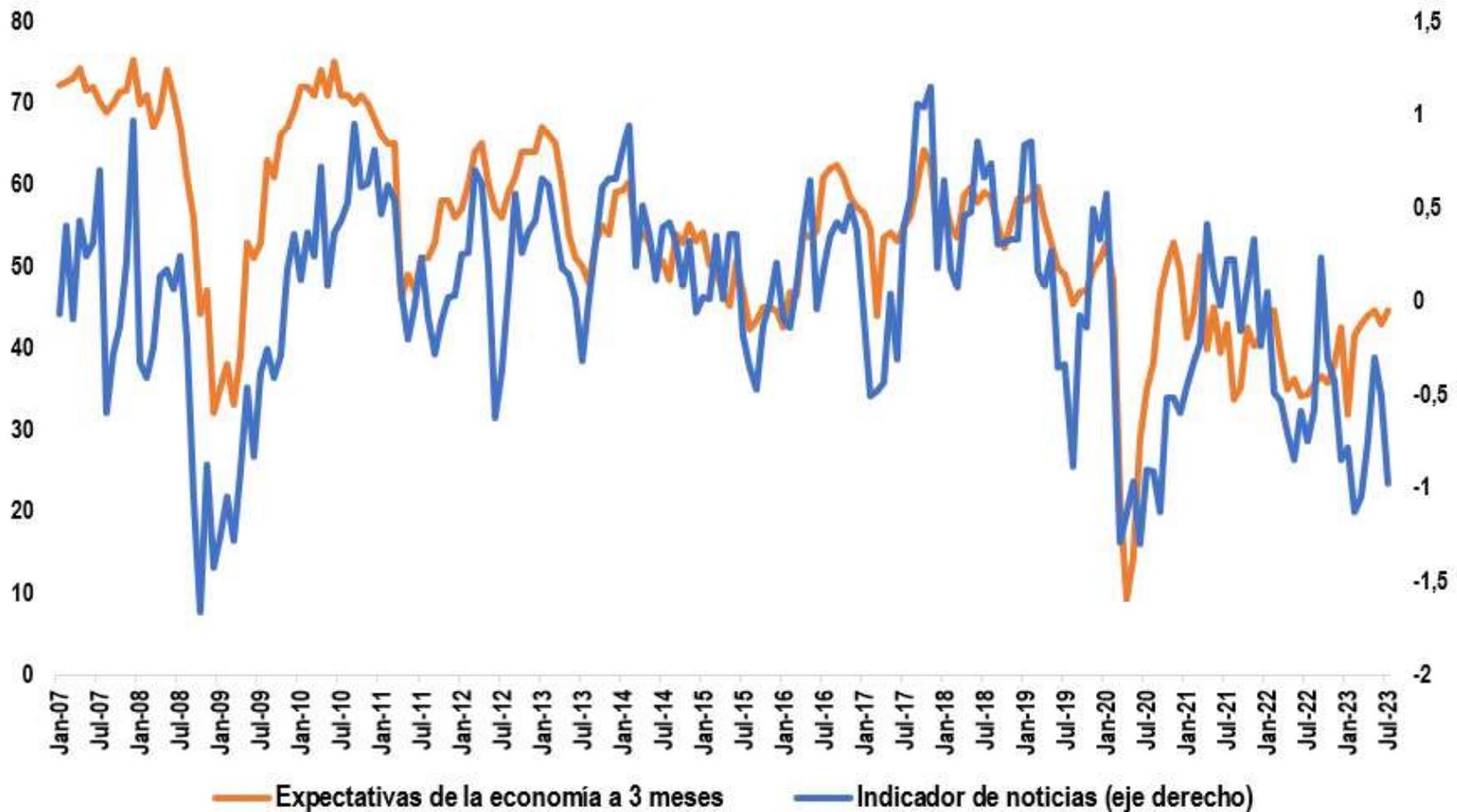
Correlaciones dinámicas con los índices de difusión

Índices de confianza BCRP	HR y LM	GI, HR y LM
Índice de venta respecto al mes anterior (t)	0,6	0,4
Índice de órdenes de compra respecto al mes anterior (t)	0,6	0,4
Índice de la situación actual del negocio (t)	0,6	0,3
Índice de expectativas de la economía a de 3 meses (t)	0,6	0,4
Índice de expectativas del sector a de 3 meses (t)	0,6	0,4
Índice de expectativas de la demanda a de 3 meses (t)	0,6	0,4
Índice de venta respecto al mes anterior (t+1)	0,6	0,4
Índice de órdenes de compra respecto al mes anterior (t+1)	0,5	0,4
Índice de expectativas de la economía a de 3 meses (t+1)	0,6	0,3
Índice de expectativas del sector a de 3 meses (t+1)	0,6	0,3
Índice de expectativas de la demanda a de 3 meses (t+1)	0,6	0,4

Nota: HR, LM y GI denotan los diccionarios de Henry (2008), Loughran y McDonald (2011) y General Inquirer, respectivamente.

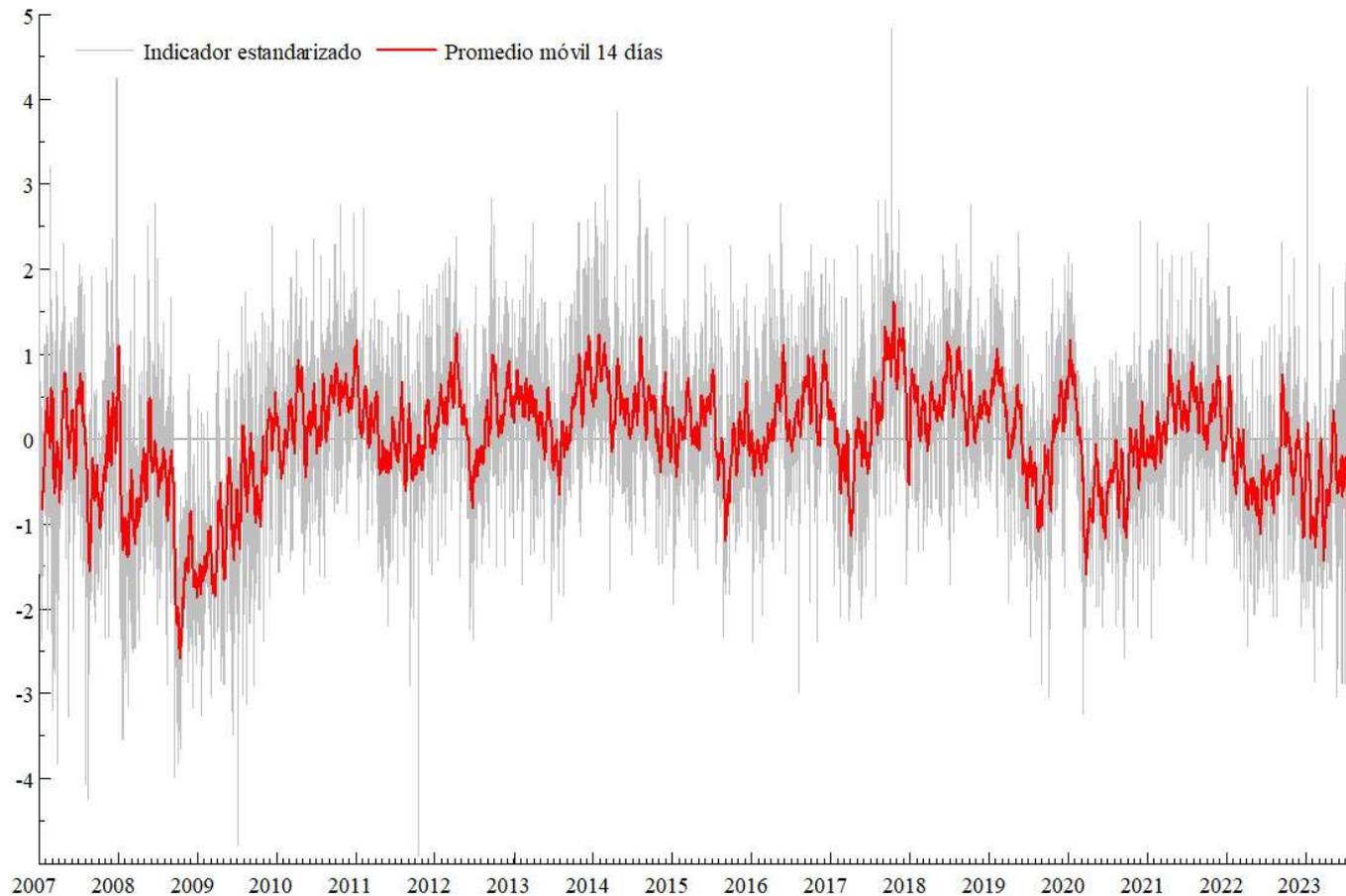
Fuente: BCRP.

Expectativas de la economía a 3 meses e indicador de noticias estandarizado



Fuente: BCRP.

Indicador de noticias estandarizado y promedio móvil



Fuente: Departamento de Prensa del BCRP.

Resultados

Estadísticos descriptivos

Estadísticos descriptivos de la proxy de volatilidad:

Estadísticas descriptivas	
Media	0.0112
Mediana	0.949
Desviación Estándar	0.0178
Mínimo	0.002
Máximo	0.331
Skewness	5.071
Kurtosis	52.033
Observaciones	6037

Elaboración Propia

Nota: La proxy se midió en base al cuadrado del indicador no estandarizado.

Midiendo la volatilidad del sentimiento

Metodología: EGARCH

La estimación para la volatilidad del sentimiento de las noticias económicas es la siguiente:

$$y_t = \sigma_t \varepsilon_t$$

$$\log \sigma_t^2 = \omega + \beta \log \sigma_{t-1}^2 + \theta_1 \left(\frac{y_{t-1}}{\sigma_{t-1}} \right) + \theta_2 \left(\left| \frac{y_{t-1}}{\sigma_{t-1}} \right| - \zeta \right)$$

$$\varepsilon_t \sim t(\eta, \nu)$$

Contenidos

- 1 Introducción y objetivo
- 2 Revisión de la literatura
 - Revisión de literatura: Internacional
 - Revisión de literatura: Nacional
- 3 Datos y metodología
- 4 Resultados**
- 5 Conclusiones

Resultados

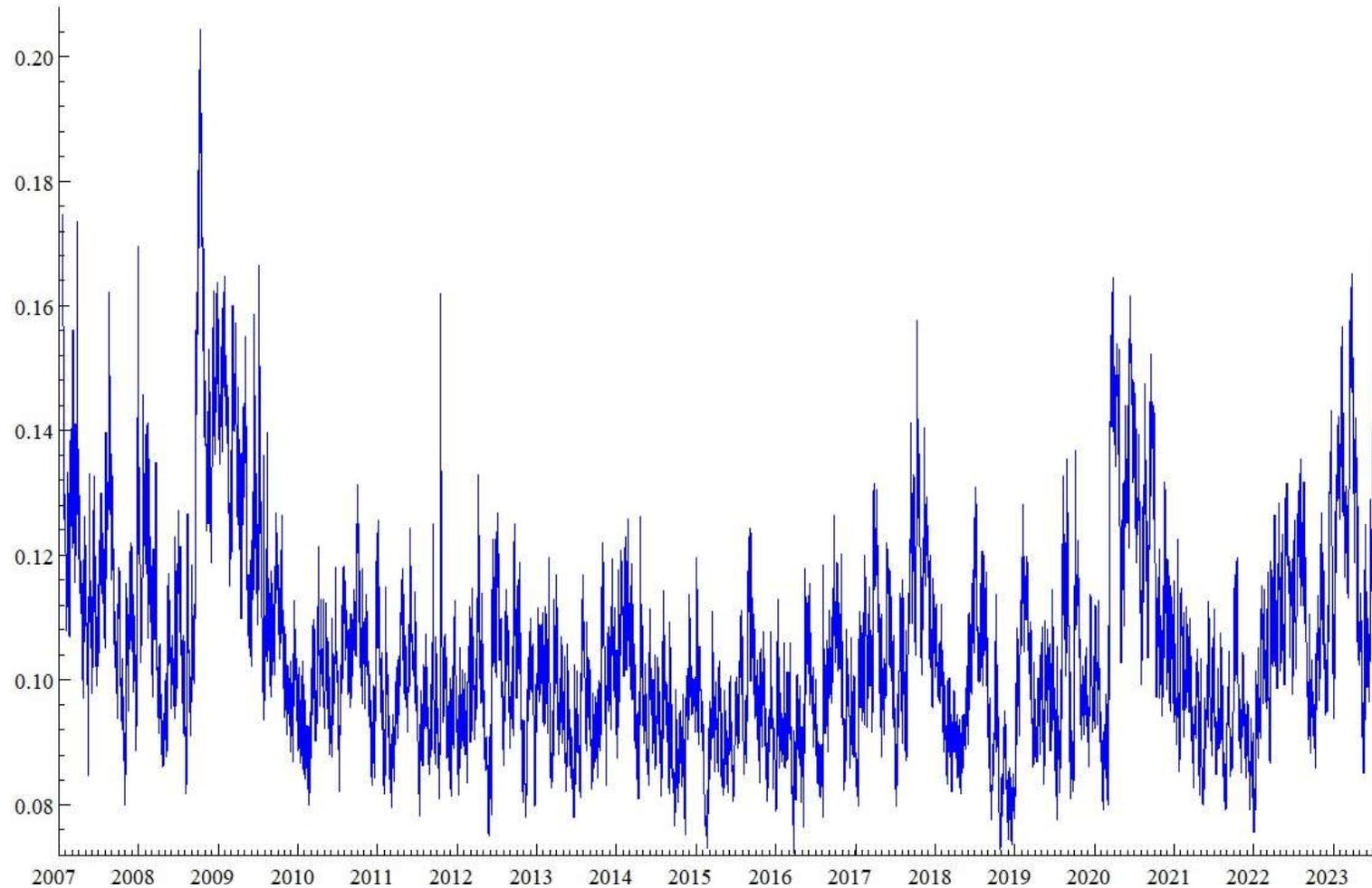
Estimados: Modelo EGARCH

Parámetro	EGARCH
ω	-0.376*** (0.140)
β	0.949*** (0.020)
θ_1	-0.021** (0.009)
θ_2	0.199*** (0.030)
η	0.051* (0.031)
ν	24.641*** (7.538)
$\log L$	5130.899
AIC	-10249.798

Elaboración Propia

Nota: *p<0.10, **p<0.05, ***p<0.01

Volatilidad EGARCH del indicador de noticias



Fuente: BCRP.

Contenidos

- 1 Introducción y objetivo
- 2 Revisión de la literatura
 - Revisión de literatura: Internacional
 - Revisión de literatura: Nacional
- 3 Datos y metodología
- 4 Resultados
- 5 Conclusiones**

Conclusiones

- Los resultados encontrados son consistentes con la literatura internacional, confirmando que utilizar información de alta frecuencia, mejora la aproximación a las expectativas económicas (Algaba et al., 2023 y Shapiro et al., 2022).
- La literatura internacional sugiere que el modelo *score driven* resulta más flexible y con un mejor ajuste a los datos dado que supera las limitaciones de la existencia de momentos incondicionales (Harvey, 2013).

Conclusiones

- El indicador de sentimiento de noticias económicas muestra una alta correlación contemporánea y adelantada con diversos índices de expectativas del BCRP.
- La mayor volatilidad suele estar estrechamente relacionada con choques negativos, principalmente eventos de carácter político y económico.

Modelos de volatilidad para el sentimiento diario de las noticias: el caso peruano*

Willy Alanya (BCRP) y Nicole Vásquez (UCS)

XLI Encuentro de Economistas - BCRP

23 de Octubre, 2023

* Las opiniones expresadas en este estudio corresponden a los autores y no deben ser atribuidas al BCRP.