



FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES  
GRADO EN ADMINISTRACIÓN Y DIRECCIÓN DE EMPRESAS

# LEYENDO MÁS ALLÁ DE LAS PALABRAS

Uso del análisis de textos para medir el sentimiento  
de mercado en España

Autora:  
Ángela Pérez Albertos

Tutora:  
M<sup>a</sup> Teresa Corzo Santamaría

Madrid, Junio 2018

## RESUMEN

El Análisis de Sentimientos permite predecir el sentimiento positivo o negativo de un segmento de la población a través de análisis de textos. Una de sus aplicaciones más exitosas ha tenido lugar en el campo de las Finanzas, donde autores como Paul Tetlock, Tim Loughran o Bill McDonald lo han utilizado para medir el sentimiento del mercado estadounidense. Este trabajo ofrece un recorrido por la evolución del pensamiento financiero hasta llegar a las actuales finanzas del comportamiento, y presenta por primera vez en español, una medida del sentimiento de mercado en España obtenida a través del Análisis de Sentimientos. Para ello, se opta por un enfoque no supervisado de sentimientos y se analiza, a través de la versión en español del diccionario LIWC, la columna de “Crónica de Bolsa” del diario Expansión. Se construye así un índice que guarda una fuerte relación con dos medidas tradicionales (el ICII y el IAR), por lo que se valida el Análisis de Sentimientos como herramienta para medir el sentimiento del inversor. Estos resultados están limitados por dificultades inherentes a la técnica (como el tratamiento de la ironía o de la negación), por el uso de un diccionario no específico para las finanzas y por el análisis de una columna de alto contenido descriptivo. No obstante, se espera de este estudio que presente un punto de partida para nuevas investigaciones sobre la materia en el idioma español.

**Palabras clave:** mercados financieros, finanzas del comportamiento, sentimiento de mercado, sentimiento del inversor, análisis de textos, análisis de sentimientos.

## ABSTRACT

Sentiment Analysis is a technique used to predict if a text has a positive or a negative sentiment. Authors such as Paul Tetlock, Tim Loughran or Bill McDonald have successfully used it to measure market sentiment in the US. This paper presents an overview of the evolution of financial theories up to the currently popular behavioral finance theories and introduces for the first time in the Spanish language, a measure of market sentiment obtained through Sentiment Analysis. To that end, the Spanish version of the LIWC dictionary is used to analyze the Spanish financial column: “Crónica de Bolsa”. The resulting indicator is found to be strongly associated to two more traditional measures (the ICII and the IAR), hence validating Sentiment Analysis as a method for

the measurement of investor sentiment. These findings are limited by some challenges such as the treatment of the irony or of the negation, by the use of a general dictionary and by the analysis of a somewhat descriptive column. All in all, this paper is expected to contribute to the existing literature and to provide a basis for future studies on this matter in the Spanish language.

**Key words:** financial markets, behavioral finance, market sentiment, text analysis, text mining, sentiment analysis.

# INDICE

<b>INDICE DE SIGLAS</b> .....	<b>V</b>
<b>LISTA DE FIGURAS</b> .....	<b>VI</b>
<b>1 INTRODUCCIÓN</b> .....	<b>1</b>
1.1. Objetivos generales y específicos .....	2
1.2. Limitaciones.....	3
1.3. Partes del trabajo.....	4
<b>2 REVISIÓN BIBLIOGRÁFICA</b> .....	<b>5</b>
2.1. El sentimiento de mercado.....	5
2.1.1. <i>Hipótesis de los Mercados Eficientes vs. finanzas del comportamiento</i> ....	5
2.1.2. <i>Sentimiento de mercado: definición y sentido financiero</i> .....	7
2.1.3. <i>Herramientas para la medición del sentimiento de mercado</i> .....	10
2.2. El Análisis de Sentimientos .....	14
2.2.1. <i>Definición</i> .....	14
2.2.2. <i>Enfoque basado en aprendizaje computacional</i> .....	16
2.2.3. <i>Enfoque basado en léxicos</i> .....	17
2.3. El sentimiento de mercado y el Análisis de Sentimientos .....	22
2.3.1. <i>Estudios destacados</i> .....	22
2.3.2. <i>Un léxico específico para las Finanzas</i> .....	24
<b>3 METODOLOGÍA</b> .....	<b>26</b>
3.1. Fuentes de datos.....	26
3.2. Método .....	27
<b>4 ANÁLISIS Y DISCUSIÓN</b> .....	<b>29</b>
4.1. ¿Reflejan las noticias financieras el sentimiento de mercado?.....	30
4.2. Construcción del Índice de Sentimiento de Mercado .....	32
4.3. Comparación de medidas del sentimiento de mercado.....	37
4.3.1. <i>ISM vs. Índice de Confianza del Inversor Institucional</i> .....	37
4.3.2. <i>ISM vs. Índice de Confianza del Consumidor</i> .....	39
4.3.3. <i>ISM vs. Indicador de Sentimiento Económico</i> .....	41
4.3.4. <i>ISM vs. Índice de Aversión al Riesgo</i> .....	43
4.4. ¿Es el ISM una medida válida del sentimiento de mercado? .....	45
<b>5 CONCLUSIONES Y PROPUESTAS</b> .....	<b>47</b>
<b>6 BIBLIOGRAFÍA</b> .....	<b>50</b>
<b>7 ANEXO</b> .....	<b>58</b>

## INDICE DE SIGLAS

AS	Análisis de Sentimientos
CIS	Centro de Investigaciones Sociológicas
DAL	Dictionary of Affect in Language
EMH	Efficient Markets Hypothesis
ESI	Economic Sentiment Indicator
GI	General Inquirer
HME	Hipótesis de los Mercados Eficientes
IAR	Índice de Aversión al Riesgo
ICC	Índice de Confianza del Consumidor
ICII	Índice de Confianza del Inversor Institucional
ISE	Indicador de Sentimiento Económico
ISM	Índice de Sentimiento de Mercado
LIWC	Linguistic Inquiry
OM	Opinion Mining
PLN	Procesamiento Natural del Lenguaje
SA	Sentiment Analysis
SEC	Security Exchange Comission
TAE	Trastorno Afectivo Estacional
WSJ	Wall Street Journal

# LISTA DE FIGURAS

Figura 1. Interés de búsqueda del término “Sentiment Analysis” en Google (2008-18)

Figura 2. Ejemplo sobre el funcionamiento de un enfoque basado en aprendizaje computacional

Figura 3. Ejemplo sobre el funcionamiento de un enfoque basado en léxicos

Figura 4. Diccionarios de sentimientos existentes en español

Figura 5. Índice de Sentimientos Positivos y Negativos (2012-18)

Figura 6. Índice de Sentimientos Positivos y Negativos (2012-18) – media móvil ponderada

Figura 7. Índice de Sentimiento de Mercado (2012-18)

Figura 8. Índice de Sentimiento de Mercado e IBEX 35 (2012-18)

Figura 9. ISM, ICII e IBEX 35 (2015-18)

Figura 10. ISM mensualizado vs. ICII (2015-18)

Figura 11. Gráfico de dispersión ICII vs. ISM

Figura 12. ISM, ICC e IBEX 35 (2012-18)

Figura 13. ISM mensualizado vs. ICC (2012-18)

Figura 14. Gráfico de dispersión ICC vs. ISM

Figura 15. ESI, ISM e IBEX 35 (2012-18)

Figura 16. ISM mensualizado vs. ESI (2012-18)

Figura 17. Gráfico de dispersión ESI vs. ISM

Figura 18. IAR, ISM e IBEX 35 (2012-18)

Figura 19. IAR, ISM Negativo e IBEX 35 (2012-18)

Figura 20. ISM Negativo mensualizado vs. IAR (2012-18)

Figura 21. Gráfico de dispersión IAR vs. ISM Negativo

# 1 INTRODUCCIÓN

Las Finanzas se han nutrido de elementos propios de otras ramas del conocimiento para realizar grandes avances durante las últimas décadas. Así por ejemplo, el modelo de Markowitz surgió de la aplicación de conceptos matemáticos a las Finanzas. También el Premio Nobel de 2017 se entregó a Richard Thaler con motivo de su aportación a las finanzas del comportamiento, que toman de la psicología cognitiva sus principios básicos. Dada la gran popularidad que el “Big Data” ha alcanzado recientemente, parece razonable esperar que la explotación de grandes cantidades de datos estructurados y desestructurados produzcan el próximo gran avance en el campo de las Finanzas. Pero, ¿cómo se han aprovechado estas hasta ahora de este fenómeno? Una de las aplicaciones que más fuerza ha cobrado ha sido el uso del Análisis de Sentimientos (AS) o *Sentiment Analysis (SA)* para medir el sentimiento del inversor. Esta técnica consiste en la extracción de la subjetividad u opinión de un texto, ya sea una noticia, un informe anual o una crítica sobre un producto (Pang & Lee, 2008).

En Estados Unidos, los máximos exponentes en este campo han sido autores como Paul Tetlock, Tim Loughran o Bill McDonald. Los estudios más reconocidos tratan de extraer el sentimiento de noticias financieras de periódicos como el *Wall Street Journal* o el *New York Times*, de informes anuales de empresas o de redes sociales como Twitter, para después observar su relación con los movimientos en los precios de la Bolsa americana. Además, los esfuerzos más recientes han ido dirigidos a mejorar la precisión de este método imperfecto, dado que el lenguaje es maleable y la gente expresa sentimientos de forma compleja. Por ello, se puede afirmar que el Análisis de Sentimientos es todavía una técnica incipiente.

La mayoría de los estudios analizan textos escritos en inglés. Análisis comparables en otros idiomas son todavía escasos, siendo inexistentes en el caso del mercado español. Este trabajo presenta la primera medición del sentimiento de mercado en España a través de técnicas de AS en español. Para ello, se comienza revisando la evolución del pensamiento financiero desde la teoría expuesta por Eugene Fama hasta las actuales teorías de las finanzas del comportamiento, destacando porqué el sentimiento de mercado tiene un papel fundamental en el entendimiento del funcionamiento de los

mercados financieros. Después, se analiza la columna “Crónica de Bolsa” del periódico *Expansión*, mediante la versión en español del diccionario Linguistic Inquiry Word Count (LIWC). Este es un diccionario ampliamente utilizado por la comunidad académica tanto en inglés como en español (ver por ejemplo Newman et al. (2003), Slatcher et al. (2007) o Carrera-Fernández, Guardía-Olmos y Perú-Cebollero (2013)).

Finalmente, y con el fin de determinar la validez de la medida estimada, se compara la obtenida con las siguientes medidas tradicionales del sentimiento del inversor: el Índice de Confianza del Inversor Institucional (ICII), realizado por el grupo de *Behavioral Finance* de Icade; el Índice de Confianza del Consumidor (ICC), calculado por el Centro de Investigaciones Sociológicas (CIS); el Indicador de Sentimiento Económico (ISE), obtenido de la Comisión Europea; y un índice construido a partir de búsquedas de la herramienta Google Trends, siguiendo a Gómez Martínez (2013).

## 1.1. Objetivos generales y específicos

El propósito del presente trabajo de investigación es humilde comparado con el gran potencial de la idea propuesta. Imagínesse que viviéramos en un mundo en el que los analistas financieros pudieran tener información actualizada constantemente sobre lo que piensa y siente el mercado, que pudieran ver qué se dice en los blogs de todo el mundo, en los periódicos de todas partes y en las informes anuales de todas las empresas. Podrían aprovechar al fin, toda la información que deriva de los datos, no sólo numéricos, sino también textuales; y la ventaja la tendría aquel analista capaz de explotar esta información de manera casi inmediata. El potencial real de la idea requiere inevitablemente que la técnica de Análisis de Sentimientos sea aplicable a textos escritos en una gran variedad de idiomas. Dada la inexistencia de investigación previa en español, este trabajo espera, de manera general, presentar un punto de partida para futuras investigaciones sobre la materia en el idioma español. Además, se aspira a estudiar y recoger de manera estructurada bajo un mismo cuerpo la literatura existente sobre sentimiento de mercado y Análisis de Sentimientos, tanto en inglés como en español. Buscará, con dicho estudio, entender porqué el sentimiento del inversor es relevante y descubrir qué aplicaciones ha tenido el Análisis de Sentimientos y cómo puede aportar y enriquecer, en el campo de las finanzas, las técnicas actuales de medición del sentimiento de mercado.

Estos objetivos generales nos llevan a plantear la principal pregunta de investigación del trabajo: ¿es el Análisis de Sentimientos de noticias financieras una técnica fiable para medir el sentimiento del inversor en el mercado español? Con este fin, se establecen los siguientes dos objetivos específicos:

- Medir el sentimiento del mercado español a partir del Análisis de Sentimientos de la columna “Crónica de Bolsa” del diario Expansión mediante la versión en español de la herramienta LIWC.
- Comparar esta medida con otras medidas tradicionales del sentimiento del inversor, con el objetivo de cerciorarnos de la validez de nuestro análisis. Las medidas tradicionales escogidas son el Índice de Confianza del Inversor Institucional, el Índice de Confianza del Consumidor, el Índice de Sentimiento Económico y un índice construido a partir de búsquedas de la herramienta Google Trends.

## 1.2. Limitaciones

El análisis propuesto estudia por primera vez un área del conocimiento antes inexplorado en español. Por tanto, era de esperar que una investigación tan incipiente se encontrara con una serie de limitaciones y desafíos al llevar a cabo la tarea.

Muchos de estas limitaciones son inherentes al Análisis de Sentimientos, que es en sí misma, una técnica imperfecta. Las personas expresan sentimientos de manera compleja y las emociones son subjetivas por definición, por lo que ni los humanos coincidirían el 100% de las veces en la categorización de textos subjetivos en “positivos” o “negativos”. El uso de elementos retóricos o de la ironía dificulta aún más la tarea. Además, el AS es una técnica basada en el lenguaje, y por tanto es incapaz de interpretar el mensaje contenido en los textos. Por ejemplo, el AS nunca podrá determinar si una frase como “los tipos de interés han aumentado en un 0,25%” es positiva o es negativa, puesto que para ello es necesario interpretarla en base a conocimientos económico-financieros y no de lenguaje.

Otros desafíos en cambio, son específicos al idioma castellano. Puesto que el origen de la técnica radica en Estados Unidos, los primeros y más reconocidos recursos

para el AS existen en inglés. Aunque algunos de ellos cuentan ya con su versión en español, muchos parten de una traducción del inglés adaptada al español, lo que puede comprometer su validez. Además, todavía existen recursos en inglés inexistentes en español. Este es el caso por ejemplo del diccionario de Loughran y McDonald (2011), un diccionario de sentimientos adaptado al ámbito financiero y del que no existe ningún comparable en español. Es por ello, que este trabajo se ve obligado a utilizar un diccionario general (la versión en español del LIWC), lo que se reconoce que puede reducir la precisión de los resultados presentados. Por último, un problema añadido también por el idioma español es la complejidad del análisis de las derivaciones (femenino, masculino, verbos, adverbios...). En este trabajo, se utiliza el software de LIWC, que permite realizar el AS en español, sin necesidad de preprocesar manualmente el texto.

### 1.3. Partes del trabajo

La presente tesis se articula de la siguiente manera. A continuación, se presenta el estado de la cuestión en finanzas, respecto a relevancia, causa, impacto y medición del sentimiento del mercado; en informática, respecto a los principales estudios, aplicaciones y recursos del Análisis de Sentimientos; y en los avances que convergen ambos campos de conocimiento. Seguidamente, se indican las bases de datos y se explica el método utilizado para el análisis. En la sección 4, se presentan los resultados obtenidos en el cálculo del Índice de Sentimiento de Mercado y se compara este indicador con el resto de medidas tradicionales del sentimiento del inversor. Finalmente, la sección 5 incluye las conclusiones y cierra el trabajo con una serie de propuestas para investigaciones futuras.

## 2 REVISIÓN BIBLIOGRÁFICA

Dado que este trabajo trata de conciliar los avances realizados en dos ramas diferentes del conocimiento (las finanzas y la informática), se tratarán primero cada uno de los temas por separado (sentimiento de mercado y Análisis de Sentimientos), para después ofrecer una visión de la complementariedad de ambos.

### 2.1. El sentimiento de mercado

#### 2.1.1. Hipótesis de los Mercados Eficientes vs. finanzas del comportamiento

En 1965, Eugene Fama, Premio Nobel de Economía, comienza a desarrollar en su tesis doctoral la teoría que se convertiría en la base del entendimiento de los mercados financieros en décadas posteriores: la Hipótesis de los Mercados Eficientes (HME, o EMH por sus siglas en inglés). En su teoría, Fama (1970) establece un nexo fundamental entre información y precio. Define los mercados “informacionalmente eficientes” como aquellos en los que los precios de los activos de los mercados reflejan toda la información disponible sobre sus valores futuros. Imagine por ejemplo que una acción de la empresa A se intercambia en el mercado por 10€. Imagine ahora que usted sabe que mañana ocurrirá algo que hará la acción subir un 20% hasta los 12€. Por tanto, intenta hoy comprar la acción de A a un precio menor de 12€, con el objetivo de venderla mañana y obtener un beneficio. Al intentarlo, se encuentra con que el resto de inversores en el mercado también conocen la ocurrencia de este evento futuro y por tanto, nadie está dispuesto a venderle la acción por menos de 12€ (que es el precio de venta que podría obtener mañana). Así, el único valor al que se podría intercambiar este activo hoy es de 12€, reflejando por tanto el precio de la acción de A toda la información existente en el mercado sobre la misma.

Fama (1970) define tres modalidades de eficiencia dependiendo de la información que reflejan los precios de los activos: eficiencia débil, semi-fuerte y fuerte. La primera supone que los precios reflejan la información de transacciones previas (precios, ofertas, apuestas y volúmenes pasados) y por tanto no es posible obtener un rendimiento superior al mercado mediante el análisis técnico de este tipo de información. La segunda supone que no es posible obtenerlo mediante el análisis de cualquier información pública (incluye las anteriores además de cualquier información hecha pública por las empresas como sus

cuentas anuales). La tercera supone que en ningún caso se puede obtener un rendimiento superior al mercado pues los valores de los activos siempre reflejan cualquier información disponible sobre los mismos, ya sea pública o privada.

Una de las grandes críticas a la Hipótesis de los Mercados Eficientes ha sido aportada por las finanzas del comportamiento o *behavioral finance*. Estas observan comportamientos en el mercado, como las burbujas financieras, que no pueden ser explicados por la teoría de Fama y proponen en su lugar una teoría financiera alternativa. Los nuevos modelos propuestos por esta teoría se basan en dos premisas fundamentales (De Long et al., 1990):

- I. Los inversores pueden ser influenciados por “sentimientos”.
- II. Existen límites al arbitraje: apostar en contra de los sentimientos de los inversores es caro y conlleva un alto riesgo<sup>1</sup>.

Es importante destacar en este sentido la comparación de ambas premisas con la teorías clásica. En primer lugar, la HME sí contempla la existencia de inversores irracionales pero considera que son minoritarios en el mercado (Thaler, 1999). Defiende que las expectativas de los inversores predominantes en el mercado son racionales y que ante la aparición de nueva información, estos actualizan sus expectativas de forma inmediata. Puesto que los inversores racionales dominan el mercado, el mercado se comporta de forma racional. Las finanzas del comportamiento en cambio, toman de la psicología la defensa de que la racionalidad humana es un supuesto que se aleja de la realidad y que la toma de decisiones de la mayoría de agentes del mercado se ve fuertemente influenciada por emociones, sentimientos y sesgos cognitivos (Barberis & Thaler, 2002). Sólo hemos de mirar a nuestro alrededor y pensar en las personas que conocemos que tienen inversiones en Bolsa: ¿podemos considerarlos inversores sofisticados? ¿Están al día de toda la información disponible sobre sus inversiones? ¿La analizan de forma adecuada? O incluso, ¿tienen la cultura financiera suficiente para analizarla de forma adecuada? La gran novedad de la afirmación consiste no obstante en que puede aplicar, no sólo a inversores individuales, sino también a profesionales.

---

<sup>1</sup> Para leer más sobre los límites al arbitraje, véase el trabajo de Shleifer y Vishny (1997)

En segundo lugar, la HME considera que no hay límites al arbitraje, por lo que la actuación de los inversores irracionales no provoca ineficiencias en el mercado. Dado que si ocurre un desajuste temporal de los precios de un activo, resultado de las expectativas irracionales de un número de inversores, todos los inversores racionales conocen esta información, entonces todos intentan aprovecharse de ella. Cuando intentan arbitrar el mercado, el precio del activo infravalorado sube, al aumentar su demanda; y el precio del activo sobrevalorado baja, al aumentar su oferta. Así, la oportunidad de obtener retornos extraordinarios desaparece: el arbitraje de los inversores racionales a los inversores irracionales hace que los precios vuelvan a su valor fundamental (Shleifer & Vishny, 1997). Por contra, las finanzas del comportamiento observan que este modelo teórico se aleja de la realidad. Entre otros límites al arbitraje de los inversores mencionan, por ejemplo, las restricciones de capital o los horizontes temporales de inversión corto placistas.

Es así como con la llegada de las finanzas del comportamiento y la defensa de la existencia de los límites al arbitraje, comienza a cobrar relevancia el concepto de sentimiento del inversor. En el siguiente apartado se analiza en mayor detalle su origen, concepto e impacto en los mercados.

### 2.1.2. Sentimiento de mercado: definición y sentido financiero

El primer académico en abrir el camino a la discusión sobre el sentimiento de mercado fue Fischer Black (1986) en su artículo “Noise”. Black habla del “ruido” o la información aleatoria en la que incorrectamente basamos nuestras decisiones financieras con la esperanza de obtener un retorno. Argumenta que este ruido “hace posibles los intercambios en Bolsa” pues añade liquidez, “pero también hace el mercado ineficiente” (Black, 1986). Distingue entre los inversores que operan según el ruido o *noise traders* y los inversores que operan según información válida o *information traders*. Esta distinción o distinciones similares serán ampliamente utilizadas en la literatura posterior sobre finanzas del comportamiento (De Long et al., 1990; Nofsinger, 2005; Shleifer & Vishny, 1997). Así, los *noise traders* son aquellos inversores susceptibles de cometer errores cognitivos mientras que los *information traders* son aquellos capaces de tomar las decisiones correctas con la información adecuada, pudiendo así aprovecharse y obtener mayores retornos a expensas de los errores de los *noise traders*.

Los estudios de finales de los 80 y principios de los 90 de Black (1986), De Long et al. (1990), Campbell y Kyle (1993) o Shleifer y Vishny (1997) entre otros, pusieron en voga la importancia de los límites al arbitraje y del ruido en la explicación de las ineficiencias de los mercados. El concepto de ruido se asimilará posteriormente al concepto de sentimiento del inversor, definido ampliamente como “las creencias sobre flujos de caja y riesgos de inversión que no están justificadas por la información disponible” (Baker & Wurgler, 2007). En otras palabras, el sentimiento del inversor es una medida de las expectativas irracionales de los inversores sobre el precio de los activos. Agregado, y cuando existen fuertes correlaciones entre el sentimiento de unos y otros inversores, conforma lo que se conoce como sentimiento de mercado. Las incógnitas que surgen alrededor de este concepto se podrían resumir en las tres grandes preguntas contestadas a continuación.

### *¿Qué factores influyen en el sentimiento de mercado?*

Deténgase el lector un momento a pensar en las acciones que tienen él o sus allegados en Bolsa. ¿Qué le llevó a decidirse por esa(s) empresa(s) en concreto? ¿Y a comprar ese volumen de acciones? ¿Y en ese momento determinado? Quizás tenga acciones en el Santander porque tiene usted una cuenta en ese banco. O tiene acciones en la líder en tendencias ecológicas Patagonia, porque comulga con sus valores. O tiene acciones en Amazon, porque leyó en un boletín financiero que las acciones del gigante tecnológico seguirán subiendo en Bolsa. Y quizás las compró justo después de haber recibido una buena noticia, como un ascenso.

En general, las finanzas del comportamiento argumentan que los sesgos cognitivos, las emociones y los estados de ánimo que influyen en nuestra toma de decisiones en el día a día, influyen también en nuestra toma de decisiones en el ámbito financiero. También el sentimiento general de la sociedad (es decir, el de todos los agentes económicos tanto inversores como no inversores) tiene la capacidad de afectar nuestras decisiones de inversión, puesto que no vivimos aislados sino en constante interacción con nuestro entorno. Todos estos factores son los que determinan el sentimiento del inversor (Nofsinger, 2005) y afectan a su vez a nuestra percepción de la rentabilidad y de la incertidumbre. Así, un sentimiento alto (bajo) está relacionado con mayores niveles de

optimismo (pesimismo), sobreestimación (subestimación) del éxito y subestimación (sobreestimación) del riesgo, y con una tendencia al alza (a la baja) en los precios. En las frecuentemente utilizadas palabras inglesas, hablaríamos de un sentimiento *bullish* o *bearish* respectivamente.

*¿Qué impacto tiene el sentimiento de mercado en la Bolsa?*

No es ya materia de debate el hecho de que nuestras emociones y estado de ánimo afectan nuestras decisiones financieras<sup>2</sup>. No obstante, sí lo es el impacto que estas decisiones irracionales tienen en los mercados. Este debate se ha desarrollado en torno a cuatro puntos fundamentales. Primero, el impacto del sentimiento en los diferentes tipos de inversores. En general, existe un amplio consenso en torno al hecho de que los inversores no profesionales son más susceptibles de verse influenciados por el sentimiento. No obstante, Shiller y Pound (1989) defienden que el sentimiento del inversor afecta también a los inversores profesionales e institucionales. Demuestran en su estudio que las conversaciones que los inversores tanto individuales como institucionales intercambian con sus conocidos (i.e. jefes, vecinos, familia, compañeros de trabajo) afectan en gran medida a sus decisiones de inversión.

Segundo, el impacto del sentimiento en los diferentes tipos de activos. Intuitivamente, podemos deducir que el sentimiento del inversor tendrá un mayor impacto en los precios de los activos de Bolsa que son intercambiados en mayor medida por los *noise traders* o inversores irracionales. Siguiendo una lógica similar, parece también razonable concluir que los activos más influenciados por el sentimiento del inversor sean aquellos cuyo precio esté rodeado de mayor incertidumbre ya que para calcular su valor habrá que hacer un mayor número de asunciones relevantes (e.g. las hipótesis sobre el tipo de interés o el ratio de crecimiento esperado son susceptibles de verse influenciadas por emociones, estado de ánimo o sesgos cognitivos). Esto es lo que demuestran Baker y Wurgler (2007) sobre acciones de empresas jóvenes, con gran volatilidad, no rentables y que no pagan dividendos.

Tercero, el impacto del sentimiento en los precios de los activos. Brown y Cliff, (2005) demuestran que el sentimiento del inversor afecta a los precios de los activos en

---

<sup>2</sup> En este respecto, véase también el estudio experimental de Harding y He (2011)

el mercado. En concreto, demuestran que los inversores más optimistas (pesimistas) tienden a mover los precios de los activos por encima (debajo) de su valor fundamental, y que este error en la valoración sólo se revierte en el largo plazo. Directamente relacionado con el precio de los activos está la volatilidad del mercado. En este ámbito, destaca el estudio de Shiller (1981), quien presenta evidencia convincente de que la excesiva volatilidad de los mercados no puede en ningún caso resultar de la acción de agentes racionales.

Cuarto y último, el impacto del sentimiento en la rentabilidad obtenida. Saunders, (1993) y Hirshleifer y Shumway (2003) demuestran que los días soleados están asociados con mayores retornos en Bolsa, lo que se atribuye al mayor nivel de optimismo de los inversores fruto del buen tiempo. Esto es porque el humor es otro reconocido factor de influencia del sentimiento: la gente que tiende a estar de mejor humor y a ser más optimista tiende también a sobreestimar las probabilidades de éxito de sus inversiones y a subestimar el riesgo asumido, así como a realizar más intercambios en Bolsa.

### *¿Cómo identificar y medir el Sentimiento de Mercado?*

Dado que esta pregunta es el elemento central de esta tesis (y realmente la clave del debate en torno al sentimiento de mercado), se le dedicará un apartado propio en la siguiente sección.

#### 2.1.3. Herramientas para la medición del sentimiento de mercado

El sentimiento del mercado no es fácil de medir, dado que por naturaleza, es una medida subjetiva, determinada por una gran variedad de factores y que agrega los diferentes estados de una multitud de agentes. Aunque no existe amplio consenso sobre la validez de una medida en concreto, son muchos los estudios que han abordado la tarea de medición del sentimiento de mercado. Baker y Wurgler (2007) recogen en su trabajo estos intentos, agrupándolos en dos grandes bloques: los que utilizan medidas directas y los que utilizan medidas indirectas.

### *Medidas directas*

Las medidas directas tratan de identificar los sesgos de inversores individuales (el sentimiento del inversor) y posteriormente agregarlos para medir su impacto en los mercados. Entre estas medidas destacan:

- I. Las encuestas a inversores, en los que estos responden directamente a preguntas sobre sus creencias. La hipótesis fundamental en este caso, es que sus respuestas se corresponderán con sus acciones en el mercado, las cuales se traducirán en patrones observables de movimiento de los valores financieros. No obstante, existe controversia en torno a esta hipótesis pues existe gran evidencia de que lo que la gente dice que cree, muchas veces dista de en lo que realmente cree; y por tanto lo que dice no es reflejo fiel de lo que hará. Como ejemplos de este tipo de medición encontramos las encuestas sobre la actitud de los inversores de Robert Shiller realizadas desde 1989<sup>3</sup>, el Sentiment Index publicado en Barron y recogido por la American Association of Individual Investors desde 1987<sup>4</sup> o en el caso del mercado español, el ICII elaborado por el grupo de investigación en *Behavioral Finance* de la Universidad Pontificia Comillas desde octubre de 2015<sup>5</sup>.
- II. Los estudios sobre predicciones de los inversores, en los que se analizan las predicciones de boletines o *newsletters* financieras para después clasificarlas en “*bullish*” o en “*bearish*”. Esta es la medida de sentimiento que utilizarán en sus estudios Clarke y Statman (1998) o Brown y Cliff (2005).

### *Medidas indirectas*

Las medidas indirectas miden el sentimiento agregado del mercado a través de diferentes *proxys*. Entre estas medidas destacan:

---

<sup>3</sup> Los índices de confianza derivados de la información extraída de las encuestas están disponibles en <https://som.yale.edu/faculty-research/our-centers-initiatives/international-center-finance/data/stock-market-confidence>

<sup>4</sup> La encuesta está disponible en <http://www.aaii.com/o/sentimentsurvey>

<sup>5</sup> La explicación sobre la elaboración y evolución del ICII puede encontrarse en De Benito Lobo, de los Ríos Sastre y Lazcano Benito (2017)

- I. Los *proxys* del estado anímico del inversor. Por ejemplo, Kamstra, Kramer y Levi (2003) investigan el efecto de las estaciones en el mercado en lo que llaman el efecto del trastorno afectivo estacional (TAE)<sup>6</sup>. En su estudio, demuestran que el retorno en los mercados está fuertemente correlacionado con la cantidad de horas de luz del día en otoño e invierno. En otro curioso estudio, Edmans, García y Norli (2007) miden el sentimiento del inversor en base a los resultados de partidos de fútbol. Así, encuentran una fuerte asociación entre las pérdidas del equipo de un país en partidos importantes y malos retornos en acciones de empresas pequeñas de ese mismo país, en el día siguiente al partido. Gómez Martínez y Prado Román (2014) encuentran resultados similares en los países cuya selección de fútbol ha ganado alguna vez el Mundial, entre los que estudian también a España. Otros autores utilizan el Índice de Confianza del Consumidor como *proxy* del estado anímico del inversor, bajo el supuesto de que este estará fuertemente correlacionado con el sentimiento del inversor (Schmeling, 2009).
- II. Las búsquedas en Internet. Por ejemplo, Da, Engelberg y Gao (2015) demuestran en su estudio la capacidad predictiva de información extraída de Google Trends sobre la volatilidad del mercado. Gómez Martínez (2013) también utiliza Google Insights (posteriormente Google Trends) para elaborar un Índice de Aversión al Riesgo (IAR) en base a las búsquedas de términos económico-financieros que correlacionan negativamente con la evolución de los mercados, para posteriormente observar su capacidad predictiva sobre la evolución del Eurostoxx 50.
- III. Los descuentos o *premiums* de los *closed-end funds* o fondos de inversión con capital fijo<sup>7</sup>. Este descuento o *premium* es la diferencia entre el valor neto de los activos del fondo y el valor del fondo en el mercado. Puesto que los *closed-end funds* son inversiones populares entre inversores individuales (ofrecen atractivas distribuciones de capital), Lee, Shleifer y Thaler (1991) y Neal y Wheatley (1998) han defendido que la media de sus *premiums* es una medida fiable del sentimiento del inversor.

---

<sup>6</sup> Toman este término del trastorno médico TAE, una condición médica por la que la persona sufre episodios de depresión durante una época del año, generalmente en invierno, cuando los días tienen menos horas de luz. Esta condición también está asociada con una mayor aversión al riesgo

<sup>7</sup> Los *closed-end funds* son un tipo de sociedad que emite un número fijo de acciones que posteriormente son intercambiadas en los mercados de capital

- IV. El volumen de activos intercambiados o la liquidez del mercado. Por ejemplo, Baker y Stein (2003) utilizan esta medida como un proxy para el sentimiento del inversor en circunstancias en las que hay restricciones a la venta en descubierto. En este contexto, los inversores irracionales tienden a comprar más acciones cuando son optimistas, siendo los agentes predominantes en el mercado y añadiendo mayor liquidez.
- V. Finalmente, otros proxys utilizados han sido por ejemplo los flujos de caja de los fondos de inversión (Frazzini & Lamont, 2015), el comportamiento de inversores individuales (Greenwood & Nagel, 2008), el volumen de OPAs y de sus retornos durante el primer día (Baker, Stein, & Wurgler, 2003) o la volatilidad de las opciones (Whaley, 2000). Baker y Wurgler (2007) construyen en su estudio un índice de sentimiento del inversor mediante la combinación de varias de las anteriores variables y posteriormente Corredor, Ferrer y Santamaría (2013) replican el análisis para el mercado español.

En suma, podemos observar que los intentos de medición del sentimiento del inversor han sido múltiples y variados, siendo en su gran mayoría aplicados al mercado estadounidense. La autora sólo tiene conocimiento de la existencia de tres intentos de medición del sentimiento de mercado en el contexto español: el ICII por el grupo de *Behavioral Finance* de Icade, los resultados de la selección española de fútbol por Gómez Martínez y Prado Román (2014) y un índice compuesto de cinco variables por Corredor et al. (2013).

## 2.2. El Análisis de Sentimientos

### 2.2.1. Definición

En la Era de la Información, cantidades ingentes de información son producidas día a día por los agentes económicos. El Big Data hace referencia a este conjunto de datos cuyo procesamiento requiere, por su elevado tamaño, velocidad y variedad, de nuevas y avanzadas herramientas de análisis (McKinsey Global Institute, 2011). Una de las herramientas que surge para combatir esta tarea es el análisis de textos, *text analysis* o *text mining*, definido como el procesamiento de textos no estructurados para la extracción de información relevante y para su transformación en información útil para la toma de decisiones. El Análisis de Sentimientos es un tipo de análisis de textos que se enmarca dentro del área del Procesamiento Natural del Lenguaje (PLN).

El Análisis de Sentimientos (AS), *Sentiment Analysis* (SA) u *Opinion Mining* (OM)<sup>8</sup> consiste en la extracción de la polaridad, valencia u opinión de un texto. En este trabajo, nos referimos a la polaridad, valencia u opinión para indicar la inclinación positiva, negativa o neutra de una palabra o texto (Molina González et al., 2015). Normalmente, el Análisis de Sentimientos captura un factor evaluativo (i.e. ¿es el texto positivo/negativo/neutro?) y/o un factor de intensidad (e.g. del 0 al 5, ¿cómo de positivo/negativo/neutro es el texto?) y/o una clasificación dentro de una lista predeterminada de categorías (e.g. ¿puede el texto clasificarse dentro de las categorías felicidad/tristeza/enfado/sorpresa?).

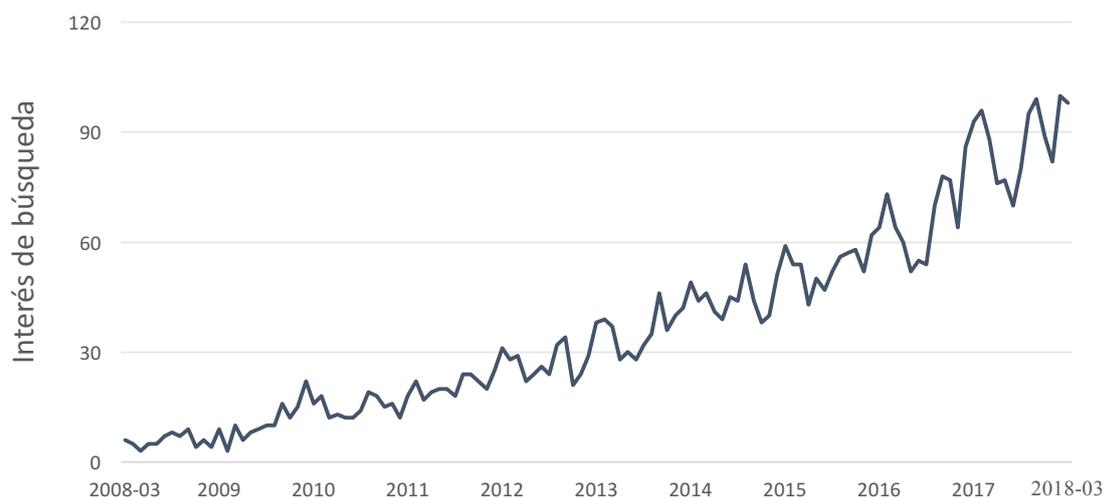
El AS es útil porque permite entender la percepción, opiniones y sentimientos de una parte o de la totalidad de la población. Su importancia reside en la explotación de forma automatizada (y potencialmente en tiempo real) de grandes cantidades de valiosa información hasta ahora no analizadas. Puede utilizarse como análisis complementario para la toma de decisiones o para un mejor entendimiento del consumidor o de la población. Así por ejemplo, puede ser utilizado por una productora de cine para conocer la opinión de los espectadores sobre su última película, analizando sus comentarios en foros de internet; o por el gobierno de un país para inferir las grandes preocupaciones de

---

<sup>8</sup> Aunque algunos autores diferencian entre el SA y el OM, para el propósito de esta tesis se seguirá a Pang y Lee (2008) y se consideran sinónimos.

sus ciudadanos a partir del análisis de sus tweets; o por el Ministerio de Sanidad para identificar posibles brotes de enfermedades entre la población mediante el análisis de sus búsquedas en google. Todo indica a que su potencial no ha sido todavía completamente explotado, aunque ha sido una técnica ya utilizada en ámbitos como la política (por ejemplo en Tasmajan et al. (2010), en Choy et al. (2011) o en Martínez-Cámara et al. (2014)), la inteligencia de mercado (por ejemplo en Turney (2002)) o la psicología (por ejemplo en Pennebaker, Francis y Booth (2001) o en Emerging Technology (2018)). Como medida de su creciente popularidad, podemos fijarnos en la figura 1, en la que se muestra el interés creciente durante la última década en las palabras “Sentiment Analysis”.

Figura 1. Interés de búsqueda del término “Sentiment Analysis” en Google (2008-18)



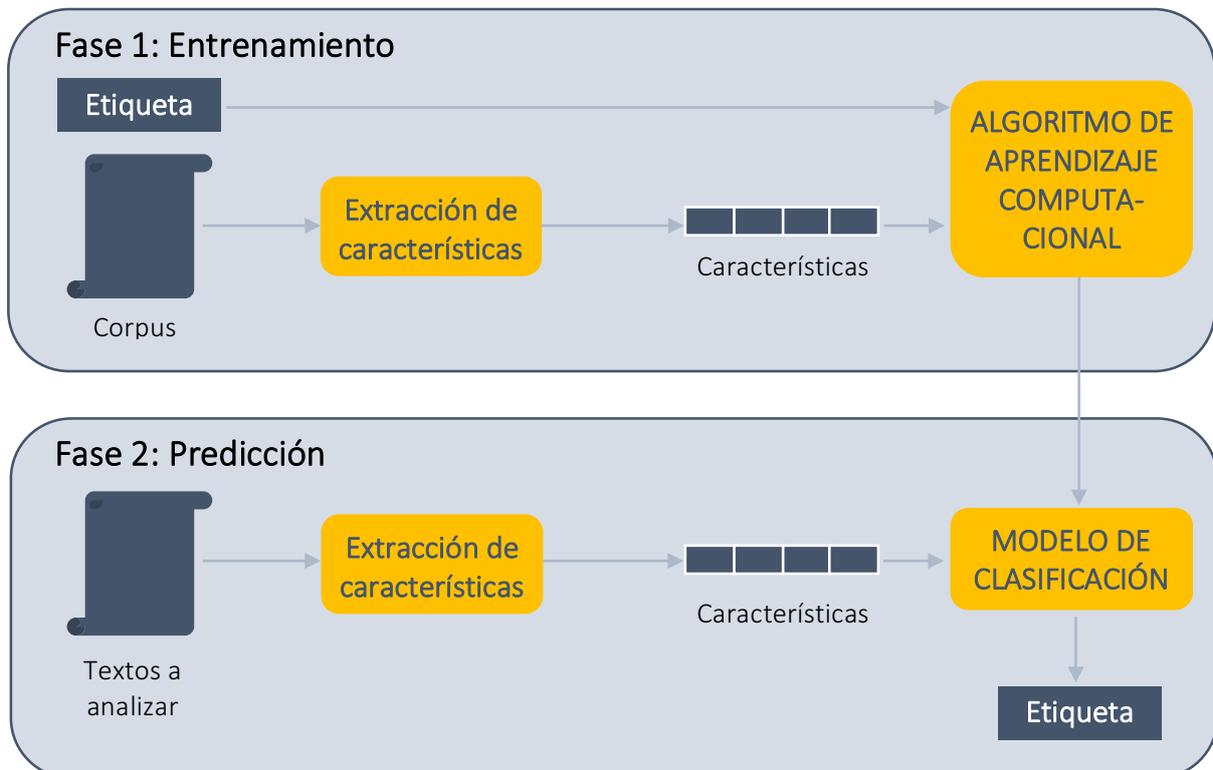
Fuente: basado en información extraída de Google Trends el 31 de marzo de 2018.  
El interés de búsqueda alcanza un valor de 100 cuando la popularidad de un término es máxima.

La tarea de extraer sentimientos de textos se ha abordado principalmente desde dos enfoques diferentes, el aprendizaje computacional y los léxicos; y en ocasiones desde la combinación de ambos. En este trabajo, hemos decidido optar en nuestro análisis por la aproximación semántica, cómo se argumentará en la sección 3, por lo que el grueso de esta sección será la descripción de esta segunda técnica.

### 2.2.2. Enfoque basado en aprendizaje computacional

Los enfoques basados en aprendizaje computacional, también conocidos como enfoques de *machine learning* o análisis supervisado de sentimientos, parten de un conjunto determinado de textos (corpus), los cuales llevan ya asociados o etiquetados unas características semánticas sobre su polaridad (Molina González et al., 2015). Cada texto de este corpus se representa en un vector de palabras, *bag of words* o n-gramas y se utiliza para entrenar a un algoritmo a identificar la polaridad o sentimiento de otros textos. Posteriormente, este algoritmo se utiliza para identificar la polaridad de los textos que son objeto de nuestro análisis y que no pertenecen a nuestro corpus inicial. Esto significa que se necesitan dos conjuntos de textos: el corpus de entrenamiento y los textos a clasificar. Algunas de las técnicas más utilizadas dentro de este enfoque son Support Vector Machines, Naïve Bayes o en las investigaciones más avanzadas, Deep Learning. En la figura 2, se incluye un esquema para ilustrar el funcionamiento de esta técnica.

Figura 2. Ejemplo sobre el funcionamiento de un enfoque basado en aprendizaje computacional



Fuente: adaptado de Bird, Klein y Loper (2009)

Pang, Lee y Vaithyanathan (2002) fueron unos de los primeros autores en utilizar este enfoque en su análisis de la polaridad de reseñas de películas, clasificando las mismas

en positivas o negativas. También ha obtenido un gran reconocimiento la investigación de Turney (2002), quien desarrolló un algoritmo capaz de predecir con una precisión del 74% si las críticas en Internet recomiendan o no cuatro productos diferentes: coches, bancos, películas y destinos de viaje. En política, ha sido una técnica muy utilizada, pues puede entenderse la gran valía que conocer la opinión pública tiene para las estrategias de los partidos políticos. Así, fue utilizada de manera pionera por Tuzmajan et al. (2010) para medir el sentimiento político durante las elecciones federales alemanas de 2009 y por Choy et al. (2011) para predecir los resultados de las elecciones en Singapur en 2011.

En España, su uso ha sido más bien escaso. Por ejemplo, Dubiau y Ale (2013) clasifican críticas gastronómicas y Martínez-Cámara et al. (2014) comparan la eficiencia de diferentes meta-clasificadores en la identificación de la polaridad de críticas sobre películas en Mucho Cine. En el ámbito político, esta técnica ha sido aplicada por Arcila-Calderón et al. (2017), para clasificar en tiempo real las opiniones políticas de tweets en español. Además, cabe destacar que cada año se celebra el Taller de Análisis de Sentimientos de la SEPLN (TASS) en español, que pretende ser un foro de debate sobre los avances del Análisis de Sentimientos en nuestra lengua (se puede ver por ejemplo a Hurtado, Pla y Buscaldi (2015), a García Cumbreiras y Villena Román (2016) o a García Cumbreiras et al. (2016)).

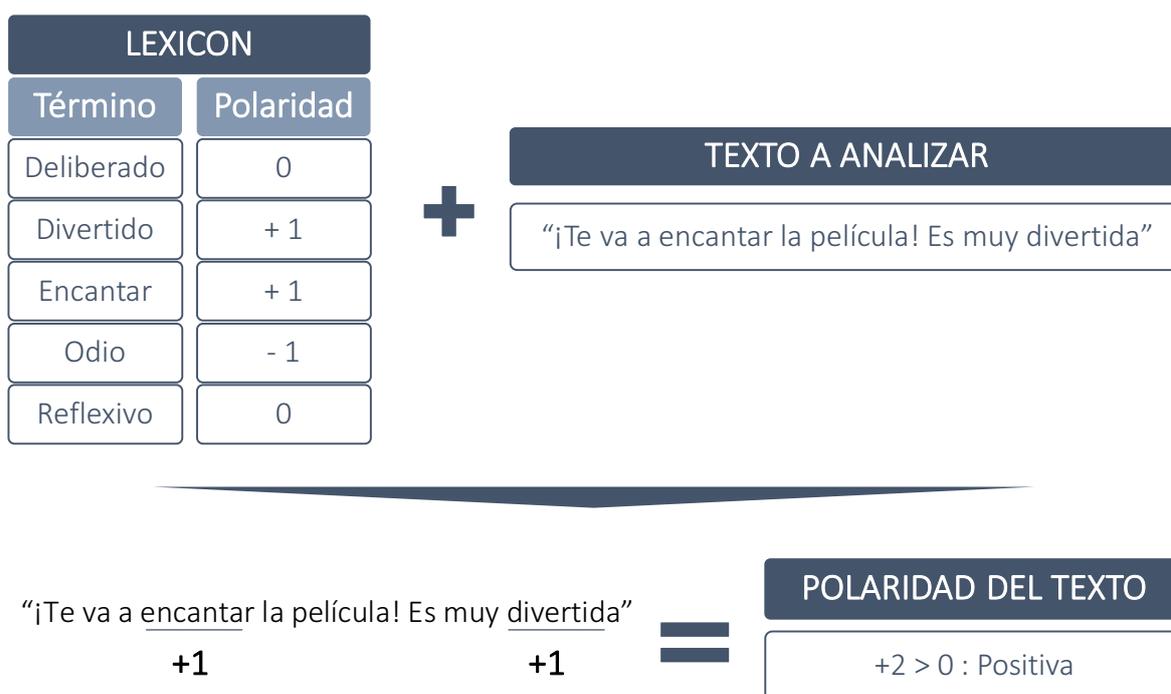
Por lo general, su rendimiento ha demostrado ser superior al del enfoque basado en léxicos (presentado a continuación), pues es capaz de tener en cuenta el contexto y las sutilezas del lenguaje (Miranda & Guzmán, 2017). Esta técnica tiene el inconveniente de que necesita un corpus previo al análisis con una gran cantidad de textos para garantizar la validez de los resultados, además de requerir de conocimientos avanzados en análisis de datos (Baviera Puig, 2017).

### 2.2.3. Enfoque basado en léxicos

Los enfoques basados en léxicos, también conocidos como enfoques semánticos o análisis no supervisados de sentimientos, parten del uso de un diccionario de términos o *lexicon*, en los que cada término lleva asociada una orientación semántica de polaridad u opinión. A diferencia del aprendizaje computacional, este enfoque no se basa en algoritmos sino en diccionarios semánticos que contienen palabras cargadas de valencia,

como los adjetivos o los adverbios, ya catalogadas según su sentimiento (Taboada et al., 2011). Para llevar a cabo un Análisis de Sentimientos con este enfoque, los textos que son objeto de nuestro análisis son pre-procesados y divididos en palabras. Posteriormente, se comprueba la aparición de los términos de nuestro *lexicon* en nuestros textos, se ignoran las palabras que no pertenecen al *lexicon*, y se suman los valores de la polaridad de los términos que sí aparecen<sup>9</sup>. De esta manera, se identifica la polaridad de nuestros textos. Podemos ilustrar esta técnica mediante el ejemplo representado en la figura 3, en la que se analiza la polaridad de una frase en base a la clasificación de las palabras de un diccionario en negativo (-1), neutro (0), positivo (+1). En este ejemplo, se considerará la frase positiva si la suma de la polaridad de sus términos es mayor que 0, y negativa si es menor.

Figura 3. Ejemplo sobre el funcionamiento de un enfoque basado en léxicos



Fuente: elaboración propia

En este sencillo ejemplo podemos ver de manera simplificada como funciona el enfoque basado en léxicos. No obstante, los análisis más avanzados incorporan una mayor

<sup>9</sup> Existen diferentes técnicas o procedimientos para calcular la valencia total de un texto y la suma simple de la polaridad de sus términos no es sino la más sencilla. Ver en este respecto a Jiménez Zafra et al. (2015)

complejidad, por lo que requieren de un pre-procesamiento de los textos con técnicas que abordan el tratamiento de las derivaciones, el género y el número (i.e. “encantar” vs. “encantó”) y el impacto de los términos modificadores (i.e. “muy”, “poco”) o de la negación (i.e. “no”, “tampoco”).

Cabe ahora preguntarse, ¿qué diccionarios de sentimientos existen? Los dos diccionarios cuyo uso en investigación es más extendido son el Harvard General Inquirer y el LIWC, aunque también se incluyen otros diccionarios alternativos.

### *El Harvard General Inquirer*

El Harvard General Inquirer (GI) es el diccionario de Análisis de Sentimientos por excelencia<sup>10</sup>. Fue desarrollado en 1963 y está disponible sólo en inglés. Este diccionario clasifica casi 12.000 palabras en “positivas”, “negativas” o “neutras”, además de en más de 100 categorías como “placer”, “viaje”, “política” o “tiempo”. Las categorías no son ni mutuamente exclusivas ni completamente exhaustivas, por lo que cada palabra puede pertenecer a varias categorías o a ninguna. Por ejemplo, la palabra *acceptable*, aceptable en español, está clasificada como “positiva”, “afectiva” y “evaluativa”; mientras que la palabra *criminal*, o criminal en español, está clasificada como “negativa”, “hostil”, “legal” y “rol”. El GI ha sido utilizado en numerosos estudios como en Turney (2002) o en el trabajo seminal de Tetlock (2007).

### *El Linguistic Inquiry and Word Count*

El LIWC es otro de los diccionarios más usados en investigación<sup>11</sup>. Fue diseñado en 2001 por Pennebaker et al. (2001) y está disponible en varios idiomas, contando con una versión beta en español (Ramírez-Esparza et al., 2007). La versión en español clasifica más de 7.000 palabras en 72 categorías entre las que se incluye “positivo”, “negativo”, “preposiciones”, “pronombres”, “tiempo” o “enojo”. Aunque su origen se encuentra en el campo de la psicología, en un estudio sobre la mejora de la salud de personas tras escribir sobre sus experiencias traumáticas, su uso se ha extendido con éxito

---

<sup>10</sup> El GI puede accederse de manera gratuita en <http://www.wjh.harvard.edu/~inquirer/>

<sup>11</sup> El LIWC puede accederse previo pago de una licencia en <http://liwc.wpengine.com>

a otros campos como el análisis de literatura, de narraciones personales o de conversaciones cotidianas. El LIWC es utilizado por ejemplo por Newman et al. (2003) para distinguir entre historias reales e historias falsas, descubriendo que las personas que mienten, evitan hablar con pronombres en primera persona o por Slatcher et al. (2007) para comparar el lenguaje utilizado por aspirantes a la Presidencia y a la Vice-presidencia de Estados Unidos en las elecciones de 2004. Su versión en español es utilizada por Carrera-Fernández, Guardía-Olmos y Perú-Cebollero, (2013) para analizar las intervenciones en debates televisados de los candidatos del PP y del PSOE a las elecciones generales de 2008 y 2011; y de manera similar por Fernández-Cabana, Rúas-Araújo y Alves-Pérez (2014) con los tweets y discursos electorales de los siete candidatos a las elecciones del Parlamento de Galicia de 2012. También lo utilizan Salas-Zaráte et al. (2014) para evaluar el desempeño del diccionario en la clasificación de críticas de películas y de productos tecnológicos.

#### *Otros diccionarios semánticos*

Además del GI y el LIWC, también encontramos en la web otros diccionarios generales como el Dictionary of Affect in Language (DAL), que asigna una valencia del 1 (negativo) al 3 (positivo), el AFINN lexicon, que asigna una valencia de -5 (muy negativo) a +5 (muy positivo) o el diccionario de Bing Liu, que contiene en torno a 6.800 palabras clasificadas en positivas o negativas. Estos están disponibles sólo en inglés.

En español, aunque los recursos son todavía escasos, cada vez empiezan a proliferar más herramientas para el Análisis de Sentimientos (ver figura 4). Además de la ya mencionada versión del LIWC, encontramos una versión en español del diccionario SentiStrength, diseñado específicamente para analizar textos encontrados en redes sociales en la web. Por su parte, Molina-González, Martín-Valdivia y Perea-Ortega (2013) diseñan una lista de 2.509 palabras positivas y 5.626 negativas (iSOL) a partir del lexicón de Bing Liu. También Díaz Rangel, Suárez Guerra y Sidorov (2014) presentan su propio diccionario en español, con más de 2.000 palabras clasificadas en seis emociones básicas y asociadas a una ponderación o porcentaje de probabilidad de ser utilizadas con un contenido emocional. Todos estos diccionarios están diseñados de forma independiente al dominio de aplicación.

Figura 4. Dicionarios de sentimientos existentes en español

Diccionario	Nº aproximado de términos	Clasificación
Versión en español del LIWC	2.300	“Positivas”, “negativas” y más de 70 categorías
Versión en español del DAL	2.500	“Placer”, “activación”, “metáfora”
NRC	14.000	8 emociones básicas
SentiStrength*	26.000	Valor de -5 (muy negativo) a 5 (muy positivo)
iSOL	8.000	2.509 palabras positivas y 5.626 palabras negativas
Díaz Rangel, Suárez Guerra y Sidorov (2014)**	2.000	6 emociones básicas Asignación a una ponderación o porcentaje de probabilidad de ser utilizadas con un contenido emocional

Fuente: elaboración propia.

\*SentiStrength es un diccionario que contiene vocabulario y expresiones específicas de redes sociales

\*\* Díaz Rangel, Suárez Guerra y Sidorov son autores mexicanos

## 2.3. El sentimiento de mercado y el Análisis de Sentimientos

La conexión entre la información y los precios en Bolsa ha sido un elemento central de las teorías financieras desde que Fama formulara la ya expuesta Hipótesis de los Mercados Eficientes. En la Era de la Información, el problema no es ya encontrar información relevante. Para los analistas financieros, la ventaja la tendrá aquel que sea capaz de procesar grandes cantidades de información en el menor tiempo posible. Los inversores acuden a tres fuentes de información para realizar sus predicciones sobre el mercado: predicciones de analistas, variables cuantitativas sobre los estados financieros de las empresas y descripciones cualitativas de las actividades económicas de las compañías (Tetlock, 2007). Tanto la primera como la tercera fuente contienen datos de tipo cualitativos. Es por esto que resulta fácil entender porqué el análisis de textos y el Análisis de Sentimientos han alcanzado gran popularidad en el mundo de las finanzas (Ahern & Sosyura, 2014).

Tras haber introducido la literatura relativa al análisis del sentimiento de mercado en finanzas, y al Análisis de Sentimientos en informática, esta sección pretende destacar aquellos estudios que han hecho converger ambos avances. Esto es, en esta sección se presentará la literatura que ha tratado de medir el sentimiento de mercado mediante el Análisis de Sentimientos.

### 2.3.1. Estudios destacados

Todos los estudios que se presentan a continuación analizan el mercado de valores americano. A conocimiento de la autora, ningún artículo hasta la fecha ha tratado de medir el sentimiento del mercado español a partir del Análisis de Sentimientos.

Mención especial en este respecto merece la labor de Paul Tetlock, actualmente profesor en la Escuela de Negocios de Columbia. Tetlock publica en 2007 en el *Journal of Finance* su famoso análisis de la columna del *Wall Street Journal* (WSJ): “Abreast of the Market”. En este estudio, Tetlock es el primero en demostrar que el contenido de los medios puede ser utilizado como un indicador del sentimiento del inversor y que este tiene capacidad predictiva sobre los movimientos en Bolsa. Para ello, extrae el contenido

de las publicaciones de la citada columna del WSJ, escrita entre 1984 y 1999, y gracias al uso del General Inquirer elabora un índice de negatividad. A grandes rasgos, su artículo demuestra tres hipótesis: altos niveles de negatividad predicen bajadas en los precios de mercado, seguidas de una reversión a los precios fundamentales; niveles inusualmente altos o bajos de negatividad predicen altos volúmenes de intercambios en Bolsa; y retornos bajos llevan a niveles altos de negatividad. En un estudio posterior, Tetlock, Saar-Tsechansky y Macskassy (2008) replican el mismo análisis con todas las noticias publicadas en el *Wall Street Journal* y el *Dow Jones News Service* entre 1980 y 2004 sobre las 500 empresas pertenecientes al S&P 500. En este nuevo artículo, tratan de demostrar cómo el análisis del lenguaje aporta información adicional a los datos cuantitativos sobre los beneficios y retornos de mercado de las compañías.

Tetlock no es el único en relacionar el contenido de las noticias financieras con eventos de los mercados. Boudoukh et al. (2013) defienden en su análisis que no todas las noticias son relevantes, pues algunas no contienen información fundamental sobre los valores de mercado. Demuestran mediante una herramienta de Análisis de Sentimientos con enfoque semántico conocida como *The Stock Sonar*, que las noticias que sí son relevantes tienen en efecto un impacto relevante en los mercados financieros. García (2013) estudia el contenido de dos columnas del *New York Times* publicadas entre 1905 y 1958 y demuestra que tienen capacidad de predecir los retornos en Bolsa pero sólo durante los periodos de recesión. Argumenta que esto posiblemente se deba a que los agentes económicos incorporan la información contenida en las noticias financieras a su toma de decisiones de manera diferente si es una época de recesión o si es una época de expansión. Las noticias publicadas por las propias empresas también son estudiadas por Lee et al. (2013), quienes analizan los informes 8-K<sup>12</sup> de compañías públicas del S&P 500 y verifican su relación con su capitalización bursátil o por Demers y Vega (2010), quienes construyen una medida de optimismo a raíz de la información textual de los informes financieros publicados trimestralmente por las empresas públicas.

En lugar de las noticias financieras, otros autores optan por medir el sentimiento de mercado mediante el análisis del contenido de redes sociales o de blogs en la web. Por ejemplo, Bollen, Mao y Zeng (2011) y Zhang, Fuehres y Gloor (2011) lo hacen en base

---

<sup>12</sup> Un informe 8-K es un informe utilizado por una empresa para notificar a sus inversores sobre eventos extraordinarios como el cambio de Director Ejecutivo o una declaración de bancarrota

a tweets; y Das y Chen (2007) lo hacen en base al contenido de foros de discusión de Yahoo sobre una serie de reconocidas empresas, para lo que utilizan técnicas de aprendizaje computacional.

### 2.3.2. Un léxico específico para las Finanzas

Gran parte de los citados estudios realizan el Análisis de Sentimientos a partir de diccionarios de polaridad generales. Esto es, de léxicos que han sido concebidos para su uso en una multitud de ámbitos diferentes, desde la psicología hasta el marketing y que son independientes del dominio en el que son aplicados. Los diccionarios semánticos GI y LIWC presentados en la sección 2.2.3. forman parte de este tipo de diccionarios. No obstante, muchas palabras tienen una polaridad u otra dependiendo de su contexto o dominio. Véase la palabra “larga” en el siguiente ejemplo:

Crítica de un restaurante: “Esperamos una hora porque la cola era muy *larga*”

Crítica de una cámara de fotos: “La batería es muy *larga* y durará semanas”

Mientras que en la primera crítica, la palabra “larga” probablemente se asocie con una opinión negativa del restaurante, en la segunda crítica, probablemente lo haga con una opinión positiva de la cámara de fotos. Pues bien, en un contexto financiero, este problema es extremadamente común. Loughran y McDonald (2011) son los primeros en observar en su artículo “When is a Liability not a Liability?” que palabras como “impuestos”, “costes”, “depreciar” o “pasivo” son clasificadas por el Harvard-IV-4 como negativas, pero en un contexto financiero son términos puramente descriptivos. Así, señalan que el uso de diccionarios generales puede llevar a conclusiones equivocadas. Por ello, desarrollan un diccionario adaptado al ámbito financiero a partir de las palabras más frecuentemente utilizadas en los informes 10-K<sup>13</sup> de empresas americanas entre 1994 y 2007. Este nuevo diccionario clasifica más de 1.200 palabras en cinco categorías: positivo, negativo, litigio, modales fuertes y modales débiles. Al comparar su diccionario con el Harvard-IV-4 en su capacidad para predecir anuncios de retornos, concluyen que su diccionario realiza mejor la tarea y que por tanto, para medir el sentimiento de mercado

---

<sup>13</sup> Un informe 10-K es un informe anual que muestra el desempeño financiero de una compañía. La Security Exchange Commission (SEC) de los Estados Unidos requiere su publicación y su contenido es más extenso que el de los informes anuales que toda empresa está obligada a enviar a sus inversores.

a través de Análisis de Sentimientos, conviene utilizar un léxico adaptado al ámbito financiero.

Su descubrimiento tuvo así importantes repercusiones sobre los posteriores avances en la materia. García (2013) usa el diccionario de Loughran y McDonald en su estudio y Jegadeesh y Wu (2013) actualizan su lista de palabras, asignando a cada término una polaridad cuantificada. Argumentan que es también muy importante asignar a cada palabra la “intensidad” adecuada, pues no todas las palabras negativas son igual de negativas.

## 3 METODOLOGÍA

### 3.1. Fuentes de datos

Este trabajo analiza un corpus compuesto por más de 1.400 noticias de la columna “Crónica de Bolsa” del diario español Expansión. Las noticias contenidas en el corpus fueron publicadas de manera diaria, con excepción de fines de semana y festivos, entre febrero de 2012 y marzo de 2018. “Crónica de Bolsa” es una columna de tipo financiero, publicada en la sección de “mercados” del Expansión, que informa al final de cada día (se publica normalmente en torno a las 19:30h) sobre los acontecimientos ocurridos en materia económico-financiera y sobre las últimas novedades del IBEX 35, el índice bursátil del mercado español. El título de cada noticia suele hacer referencia directa al movimiento del IBEX 35 durante ese día, mientras que el cuerpo de la misma suele seguir una estructura relativamente fija en la cual se abordan movimientos en la Bolsa española, movimientos en otros índices europeos, movimientos en los precios de las acciones de las empresas más grandes del IBEX 35, situación de los precios del petróleo y el mercado de renta fija y de divisas. Cada noticia contiene una media de 735 palabras. En el anexo (A.1. y A.2.) se incluyen dos de ellas a modo de ejemplo. Por último, la columna pretende ser el vehículo informativo principal del diario para informar sobre los eventos del mercado, por lo que no debe de confundirse con artículos o blogs de opinión escritos desde la perspectiva personal de un experto. No obstante, la vocación informativa de la columna no impide que se utilice en ella un vocabulario subjetivo a partir de cuál es posible extraer sentimientos.

A partir de este corpus, se construye un Índice de Sentimiento de Mercado (ISM). La idoneidad del uso de esta columna para la tarea aquí abordada resulta de tres factores. El primero, la importancia del diario Expansión como el periódico nacional líder en materia económica y financiera, lo que nos permite suponer que se encuentra entre los que mayor alcance tienen (llega a un mayor número de personas) y entre los que mayor influencia puede llegar a tener entre inversores (sus análisis gozan de reconocimiento por ser el periódico líder en la materia). El segundo, el contenido de la columna, principalmente enfocado a analizar los movimientos en la Bolsa española. El tercero y último, la frecuencia con que es publicada, permitiendo así obtener un índice diario para captar el sentimiento de mercado.

Por otro lado, para la comparación del Índice de Sentimiento de Mercado obtenido con otras medidas, se han utilizado datos sobre el Índice de Confianza del Inversor Institucional, proporcionados directamente por el grupo de investigación en *Behavioral Finance* de Icade; sobre el Índice de Confianza del Consumidor, obtenidos directamente del CIS; sobre el Indicador de Sentimiento Económico, obtenido de la base de datos de la Comisión Europea<sup>14</sup>; y finalmente sobre el Índice de Aversión al Riesgo, obtenido siguiendo a Gómez-Martínez (2013) a través de la herramienta Google Trends<sup>15</sup>.

## 3.2. Método

Para el análisis de los datos propuestos, se ha procedido de la siguiente manera. Primero, se descargaron las noticias de “Crónica de Bolsa” del Expansión, mediante técnicas de *web scrapping* gracias a la herramienta de análisis Python. Esta tarea pudo realizarse gracias a la ayuda del profesor de la Universidad y miembro del grupo de *Behavioral Finance*, Mahmoud Aymo. A continuación, se analizaron todos los textos utilizando LIWC. Esto es, se optó por hacer un análisis no supervisado de sentimientos o por un enfoque basado en léxicos, además de por un diccionario de sentimientos general y no adaptado al dominio de las finanzas. Esta elección se debe a que este enfoque y el LIWC han sido los más empleados en técnicas de Análisis de Sentimientos en español y a que la inexistencia de un diccionario en español específico para las finanzas obliga a la utilización de uno general. No obstante, se ha estudiado la clasificación realizada por LIWC de las palabras más comunes en materia económico-financiera y el único potencial problema encontrado es la clasificación como positivas de algunas palabras neutras (e.g. “valor” o “incrementar”) y la no clasificación de otras palabras cargadas de sentimiento (e.g. “bajista” o “alcista”) (ver anexo A.3. para un mayor detalle). Para la utilización del LIWC, es necesario obtener una licencia que da acceso tanto a un software que permite realizar de manera automática el análisis de textos como a los diccionarios en sus diferentes versiones. Para este trabajo, se utilizó la versión del diccionario en español. De esta manera, se calculó a través de LIWC el porcentaje de palabras positivas y negativas

---

<sup>14</sup> Disponible en <https://data.europa.eu/euodp/es/data/dataset/c04BuUz6WXIQGJkHPwLug/resource/7c7a33ee-6460-44c8-9d7a-c363f8d391db>

<sup>15</sup> Herramienta accesible de manera gratuita en <https://trends.google.es/trends/?geo=ES>

que contienen los textos analizados y se obtuvo así, una medida del sentimiento del mercado a partir de la suma de ambos porcentajes en cada texto analizado.

Además, la elección del LIWC como software de análisis radica también en la conveniencia de su uso, pues permite realizar el análisis de textos sin necesidad de preprocesar el texto. En técnicas de preprocesamiento en inglés, habitualmente se opta por “lemmatizar” las palabras, esto es, eliminar la derivación para mantener simplemente la raíz, e interpretar el sentimiento contenido en la misma. Por ejemplo, véase el verbo “perder”, o “to lose” en inglés. El inglés no cuenta con derivaciones verbales, por lo que ya se utilice la primera persona, la tercera, el plural o el singular, se empleará la misma palabra. En caso de que se utilice la palabra “perdedor” o “loser”, se reduciría la misma a “lose”. Ambas se interpretarían como negativas. El idioma español no sólo añade complejidad porque es necesario tener en cuenta todas las derivaciones posibles (“perdimos”, “pierd-o”, “perd-eréis”), si no también porque al reducir la palabra a su raíz, podemos perder información importante sobre el sentimiento que expresa. Esto puede verse por ejemplo, en los verbos futuros y condicionales, que en inglés se expresan mediante la adición de una nueva palabra (“will” y “would” respectivamente), pero en español se expresan mediante nuevas derivaciones (“yo perderé” o “yo perdería”). La clave en este respecto es responder a dos preguntas sobre las consecuencias de obviar esta complejidad añadida y replicar el preprocesamiento del texto que se haría en inglés: ¿se perdería tanta información sobre el sentimiento como para no poder determinarlo con precisión? O, ¿se obtendría información errónea que llevaría a determinar el sentimiento del texto de modo incorrecto? En este trabajo, se asume que la respuesta a ambas preguntas es negativa, y se procede a realizar el análisis con LIWC.

Finalmente, para la comparación de los índices seleccionados, se han realizado principalmente análisis cuantitativos como la correlación o los gráficos de dispersión, empleando para ello la herramienta Microsoft Excel.

## 4 ANÁLISIS Y DISCUSIÓN

Este trabajo busca medir el sentimiento del mercado español a partir del análisis no supervisado de sentimientos de noticias financieras. Para ello, utiliza un cuerpo de noticias de la columna “Crónica de Bolsa” del periódico Expansión y un software informático que incorpora un diccionario de polaridad en español conocido como LIWC. Esta es la primera vez que un estudio de este tipo se lleva a cabo para el mercado financiero español, por lo que sus resultados serán de importancia para la comunidad académica, y se espera también, serán los primeros de una multitud de investigaciones posteriores sobre la materia.

Una de las claves del análisis es encontrar un método para validar la fiabilidad de nuestra medida como *proxy* del sentimiento de mercado. La dificultad de la tarea en realidad es enorme, pues el sentimiento de mercado es por definición subjetivo y no existe ningún indicador objetivo con el que comparar nuestra medida. ¿Cómo han resuelto otros estudios este problema? Por un lado, en los estudios que realizan trabajos comparables para el mercado americano, lo común es buscar relaciones entre la medida encontrada y movimientos en Bolsa (i.e. retornos del mercado o valores de los activos). No obstante, esta relación simplemente verifica el link entre las noticias financieras y los mercados, sin verificar realmente si el sentimiento encontrado en las noticias es una medida fiable del sentimiento del inversor. Por otro lado, en los estudios sobre análisis de sentimientos desarrollados en otros contextos, es común tener una medida “fiable” y “objetiva” de la opinión de los usuarios. Así por ejemplo, muchos autores determinan la precisión de sus análisis comparando sus clasificaciones a partir del análisis de los textos, con la evaluación hecha por los propios usuarios (i.e. puntuación dada o número de estrellas del 1 al 5 marcadas), mediante el análisis de precisión, métodos de recuperación o medidas F. Otra alternativa sería la clasificación manual de los textos analizados, es decir, clasificar uno a uno los artículos obtenidos en positivos, negativos o neutros; para después comparar la clasificación realizada por un ser humano, con la realizada por el algoritmo y verificar así su precisión. No obstante, la cantidad de recursos que tendrían que dedicarse a esta tarea hicieron que se descartase como opción.

En este estudio se ha optado por una cuarta vía. Con el objetivo de centrar nuestro esfuerzo en la medición del sentimiento del inversor, se ha comparado la medida de sentimiento obtenida a partir del AS con otras medidas tradicionales del sentimiento del inversor. Estas son el ICII, el ICC, el ESI y el IAR, un índice construido a partir de búsqueda de una serie de palabras en Google, siguiendo el análisis realizado por Gómez Martínez (2013). La razón por la que se eligen estas y no otras medidas es por disponibilidad de los datos. No obstante, se reconoce que habría sido también interesante ver su relación con otras medidas tradicionales del inversor como los descuentos o premiums de los *closed-end funds*, los flujos de caja de los fondos de inversión, el volumen de activos intercambiados, el volumen de OPAs y de sus retornos durante el primer día o la volatilidad de las opciones. El objetivo del análisis aquí planteado es identificar si los diferentes indicadores están midiendo lo mismo, es decir, si pueden considerarse medidas fiables del sentimiento del inversor y si, cómo planteábamos, el análisis de sentimientos de noticias financieras puede aportar nuevos avances en este campo.

#### 4.1. ¿Reflejan las noticias financieras el sentimiento de mercado?

Antes de comenzar el análisis propuesto, resulta necesario responder a una pregunta que es la base de la principal línea de investigación de este trabajo: ¿reflejan las noticias financieras el sentimiento de mercado? Las noticias se definen como aquella información que está disponible para el gran público. En el contexto financiero, hablamos de información sobre una compañía, su rentabilidad, su funcionamiento operativo o cualquier noticia que aporte información sobre el mercado (Yu, 2014). ¿Cuáles son los mecanismos que subyacen en la relación entre las noticias y el mercado de valores? Esta relación puede ser de tres naturalezas distintas, por lo que a continuación se presentan tres supuestos bajo los cuales las noticias financieras y el sentimiento de mercado guardan una relación.

- I. Las noticias contienen información ya existente en el mercado. Por tanto, el contenido de las noticias refleja información ya asimilada por los inversores y es reflejo de su percepción sobre la misma. En este caso, el AS daría como resultado directamente el sentimiento del mercado: un sentimiento positivo

(negativo) se encontraría después de una subida (bajada) en los precios y estaría seguido de una reversión en el corto/medio plazo.

- II. Las noticias contienen información novedosa sobre el mercado. Por tanto, los inversores reaccionarán de una manera u otra ante esta información novedosa. En este caso, el AS de las noticias serviría para predecir el sentimiento de mercado: un sentimiento positivo (negativo) estaría seguido de una subida (bajada) en los precios y de una reversión en el corto plazo.
- III. Las noticias, contengan información novedosa o no, influyen en el sentimiento de los inversores. Esto es porque transmiten irremediamente de manera subjetiva la ocurrencia de los eventos. El poder de las noticias financieras reside en su capacidad para focalizar la atención de los lectores en una serie de noticias pre-seleccionadas (llamando a la aparición de los sesgos de disponibilidad o de atención), para transmitir información sobre diferentes tipos de eventos e influenciar nuestras creencias mediante la interpretación de dichos eventos (Tetlock, 2016). En este caso, el AS de las noticias nos serviría también para predecir el sentimiento de mercado ante eventos específicos, bajo el supuesto de que sabemos cómo reaccionan los inversores ante determinado tipo de información (i.e. noticias que contienen un sentimiento negativo provocan un sentimiento negativo en el mercado).

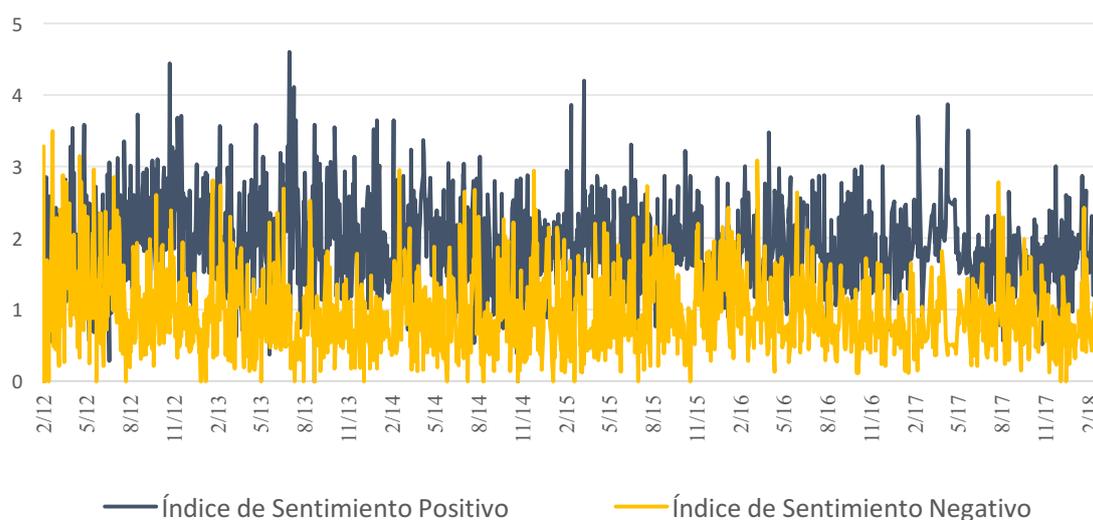
En el primer caso, las noticias reflejan el sentimiento, mientras que en el segundo lo causan y en el tercero lo refuerzan. ¿Cómo saber qué supuesto es el correcto? Depende de la noticia. Como defendían Boudoukh et al. (2013) no todas las noticias contienen información fundamental sobre los valores de mercado y como argumenta Tetlock (2011), los precios reaccionan en menor medida ante noticias antiguas, pero aún así siguen reaccionan en exceso a la información contenida en las mismas (según la HME, dado que no es nueva información, la reacción debería ser nula). En resumen, podemos extraer las siguientes dos conclusiones: (1) existe un link bidireccional entre las noticias y los mercados; y (2) los precios de los mercados incorporan de las noticias tanto la información fundamental como el ruido o el sentimiento (Tetlock, 2016).

## 4.2. Construcción del Índice de Sentimiento de Mercado

Procedemos ahora en esta sección a la medición del sentimiento del mercado financiero español a partir del AS de las más de 1.400 noticias de la columna “Crónica de Bolsa”. Tras obtener la licencia para usar el LIWC, este trae incorporado además del propio diccionario de sentimientos, un software informático para analizar cada uno de los textos en las diferentes categorías incluidas en el diccionario. Las dos categorías que nos atañen para la tarea son “emopos” y “emoneg”. En la primera categoría, LIWC indica el porcentaje de palabras de cada texto que son positivas; mientras que en la segunda, señala el porcentaje de palabras de cada texto que son negativas.

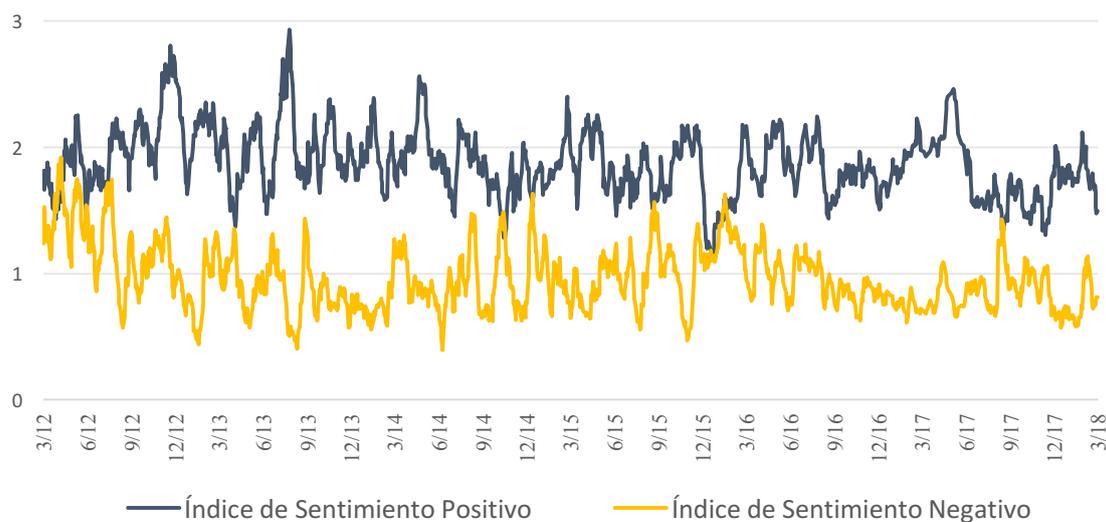
A continuación, se muestran los resultados obtenidos de cada una de las categorías. Por la manera en la que se ha construido el índice, la serie temporal obtenida sufre grandes fluctuaciones y es mucho más volátil que cualquier índice construido en base a encuestas. Por ello, se presentan dos figuras: la 5, en la que se presentan los resultados originales, entendiéndose que la granularidad del sentimiento diario es también una de las características más atractivas del ISM; y la 6, en la que se presenta una media móvil ponderada, que permite reducir las fluctuaciones a corto plazo para resaltar en su lugar las tendencias a largo plazo.

Figura 5. Índice de Sentimientos Positivos y Negativos (2012-18)



Fuente: elaboración propia basada en el análisis de noticias del Expansión mediante la herramienta LIWC

Figura 6. Índice de Sentimientos Positivos y Negativos (2012-18) – media móvil ponderada



Fuente: elaboración propia basada en el análisis de noticias del Expansión mediante la herramienta LIWC

Pueden realizarse las siguientes apreciaciones sobre los gráficos presentados. Primero, puede observarse que el porcentaje de palabras contenidas en las noticias y que LIWC cataloga como positivas o negativas es relativamente pequeño. Así, el mínimo se encuentra en 0% mientras que el máximo se encuentra en 4,60% (porcentaje de palabras positivas en la columna del 15 de julio de 2013). Este es un porcentaje pequeño sobre todo considerando que LIWC categoriza de media, un 72% de la totalidad de palabras contenidas en cada artículo<sup>16</sup>. Esto podría resultar o bien del hecho de que la columna sea más descriptiva (y por lo tanto tuviera menor carga de sentimiento) de lo que en un inicio se planteó; o bien del hecho de que el diccionario LIWC no sea apropiado por no ser específico para el mundo financiero. Se descarta la segunda opción, puesto que, como se introdujo en la sección 2.3.2, Loughran y McDonald (2011) reivindican la necesidad de utilizar un diccionario financiero principalmente porque diccionarios tradicionales clasificaban como negativos, términos que son neutrales en el ámbito financiero (i.e. “deuda”). Se comprobó que esto no ocurría con LIWC (ver anexo A.3.) y en todo caso, se concluye que este efecto resultaría en un porcentaje mayor de palabras clasificadas como negativas, y no en los resultados encontrados. Sí ha de aclararse no obstante, que LIWC no clasifica palabras claves cargadas de sentimiento en un ámbito financiero, como “alcista” o “bajista”. A pesar por tanto, de que el bajo porcentaje de palabras cargadas de

<sup>16</sup> Las palabras contenidas en el diccionario y no clasificadas ni como positivas ni como negativas, están asignadas a otras categorías gramaticales como “Artículo”, “Verbos” o de contenido como “Social” o “Movimiento”

sentimientos muestre indicios de que se ha realizado el análisis de noticias con relativamente escaso contenido emocional, se continuará el análisis pues se considera que el porcentaje y la diferencia entre el índice positivo y negativo, son lo suficientemente relevantes como para extraer conclusiones válidas sobre el análisis perseguido.

Segundo, podemos observar un claro sesgo hacia lo positivo. Esto era de esperar dada la forma de expresión de la negación en español. Así, tómesese como ejemplo las siguientes frases extraídas de “Crónica de Bolsa” en diferentes fechas:

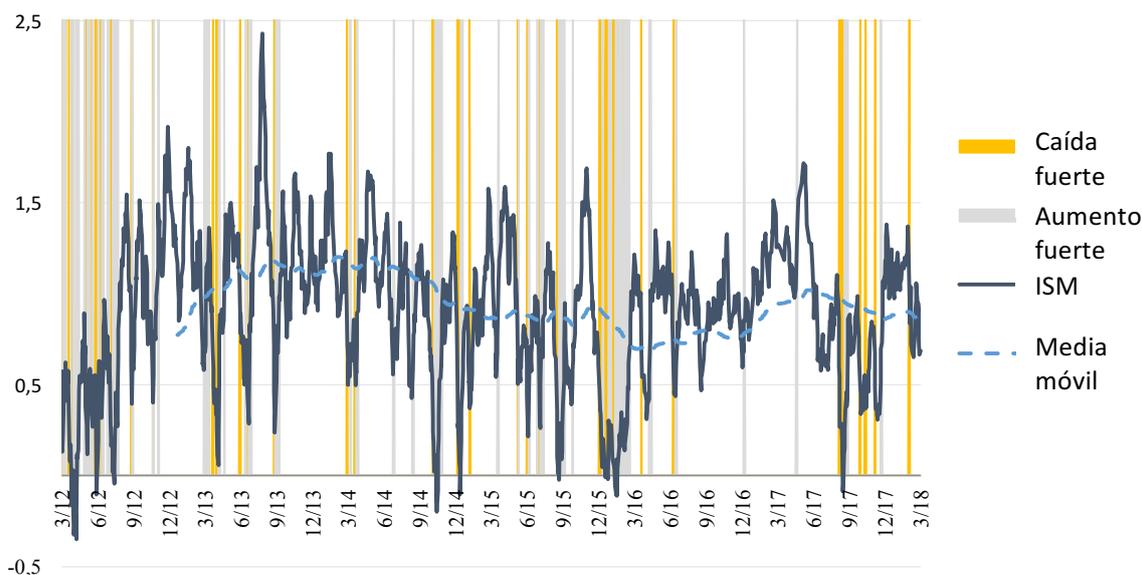
15 enero 2013: “La mejora de la deuda pública no ha tenido reflejo en el Ibex.”

6 febrero 2018: “Los mercados de deuda y de divisas mantienen la presión, y los resultados empresariales no ayudaron demasiado a recuperar el optimismo”.

Ambas frases contienen palabras clasificadas como positivas por LIWC (“mejora” en la primera, y “optimismo” en la segunda), pero el mensaje es no obstante, negativo. Este problema, consecuencia de la negación de términos positivos es frecuente en el Análisis de Sentimientos. En cambio, el problema contrario (negar una palabra negativa, lo que resulta en un sentimiento positivo), es un recurso lingüístico mucho menos común (Loughran & McDonald, 2011). Este sesgo natural hacia lo positivo puede además verse agravado por la clasificación de LIWC como positivas de palabras neutras como “incremento” o “valor”. Existen estudios, como el de Jiménez Zafra et al. (2015), que abordan exclusivamente como tratar la negación en estos análisis. Otros artículos, como los propios Loughran y McDonald (2011) o como Gómez-Martínez (2013) optan por analizar exclusivamente el impacto en los mercados del uso de términos cargados de valencia negativa. Por ello, se ha considerado importante mostrar en la figura 6 el ISM, desagregado en positivo y negativo.

En la siguiente figura 7, se presenta el ISM, obtenido tras substrair el índice negativo al índice positivo. Esta vez, se presenta de nuevo la media móvil ponderada de los resultados.

Figura 7. Índice de Sentimiento de Mercado (2012-18)



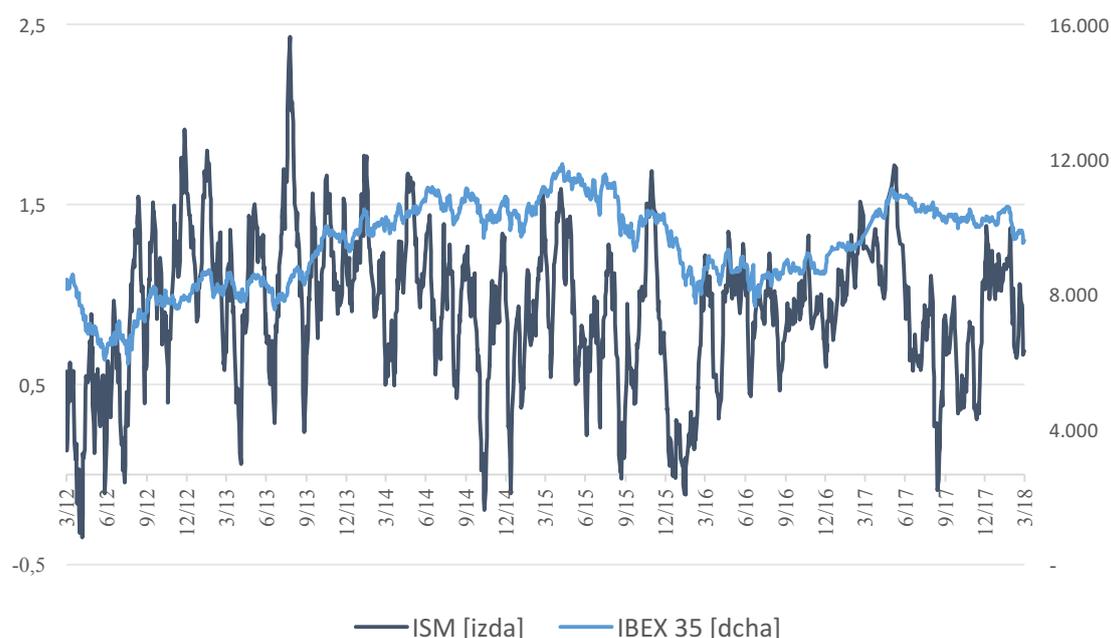
Fuente: elaboración propia basada en el análisis de noticias del Expansión mediante la herramienta LIWC

¿Cómo se deben interpretar entonces los resultados obtenidos? Un ISM positivo (negativo) resulta de la utilización de una mayor (menor) proporción de palabras positivas frente a negativas. Gráficamente, es muy fácil observar el sesgo hacia lo positivo comentado en los anteriores párrafos. Por ello, no tendría sentido simplemente interpretar un índice positivo como el reflejo de un sentimiento positivo de mercado. En su lugar, se presenta una media móvil anual y se propone la siguiente interpretación: índices por encima de la media móvil anual reflejan un sentimiento positivo de mercado, que a su vez se traduce en una actitud optimista entre los inversores (ya sea como reflejo, causa o reforzamiento), en una sobreestimación de la rentabilidad y en una menor aversión al riesgo. Esto, según la teoría expuesta, estaría relacionado con alzas en los precios. La interpretación de índices por debajo de la media móvil sería exactamente la contraria: estos reflejarían un sentimiento negativo de mercado, que se traduce en una actitud pesimista, en una subestimación de la rentabilidad y en una mayor aversión al riesgo, con la consecuente bajada en los precios de los activos. Además, se incluye en el gráfico los días en los que el sentimiento cambia de manera abrupta respecto al día anterior, lo que podría ser de utilidad para analizar movimientos diarios en el Índice con respecto a movimientos diarios en el IBEX 35.

Así, para verificar estas interpretaciones, se incluye en la figura 8, una comparativa de la evolución del ISM y de la Bolsa española. Analicemos por ejemplo dos

de los extremos pronunciados del índice: el extremo positivo que toma en julio de 2013, y el negativo que toma en torno a enero de 2016. El primero (sentimiento muy positivo) coincide con un movimiento al alza en el IBEX 35, que no volvería a caer a niveles similares hasta comienzos del 2016. En cambio, el segundo (sentimiento muy negativo) coincide con la bajada que sufrió el IBEX de los 10.453 puntos en noviembre de 2015 a los 7.920 puntos en febrero de 2016.

Figura 8. Índice de Sentimiento de Mercado e IBEX 35 (2012-18)



Fuente: ISM - elaboración propia basada en el análisis de noticias del Expansión mediante la herramienta LIWC  
IBEX 35 – extraídos de Yahoo Finance

Puesto que, como se ha establecido ya, el propósito de este trabajo no es relacionar el índice construido con la Bolsa (el desarrollo de esta análisis sería motivo de otro artículo), no se ahondará más en esta cuestión. Por contra, para continuar con la validación de los resultados obtenidos, se compara en el siguiente apartado, el ISM con otras medidas tradicionales de sentimiento de mercado.

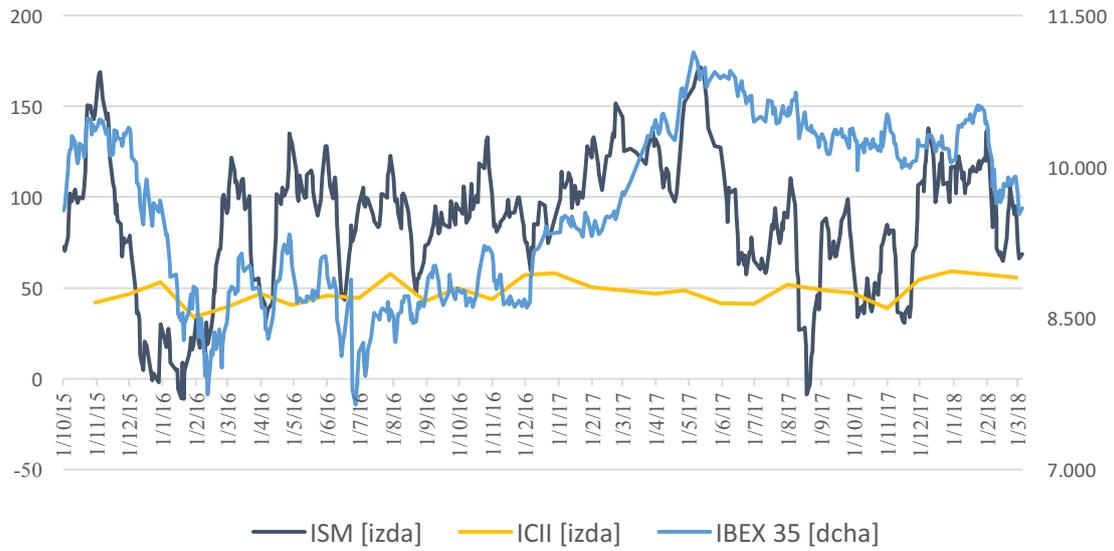
## 4.3. Comparación de medidas del sentimiento de mercado

### 4.3.1. ISM vs. Índice de Confianza del Inversor Institucional

El Índice de Confianza del Inversor Institucional (ICII) es un índice elaborado por el grupo de investigación en *Behavioral Finance* de la Universidad Pontificia Comillas, en colaboración con la revista *Funds People*. El índice está disponible desde octubre de 2015 y se publica desde entonces de forma mensual. El ICII puede tomar valores entre el 0 y el 100, y se calcula en base a las respuestas obtenidas de una encuesta realizada principalmente a varones de entre 40 y 50 años, con experiencia en el sector bancario, y más en concreto en el mundo de la inversión. Además, el ICII se compone a partir de cinco índices: un índice de perspectivas, un índice de valoración, un índice de capacidad de recuperación a corto plazo, un índice de capacidad de recuperación a medio plazo y un índice de riesgo de crash (De Benito Lobo et al., 2017).

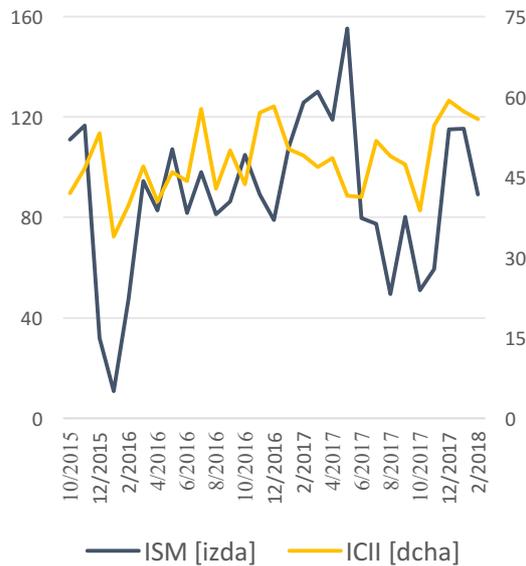
La diferencia en las frecuencias de medida entre el Índice de Sentimiento de Mercado, obtenido a través del Análisis de Sentimientos y el ICII, dificulta en un principio la comparación de ambos índices (figura 9). No obstante, suavizando el índice del ISM y transformándolo en un índice mensual utilizando la media de los resultados diarios, se puede comenzar a observar que ambas medidas siguen unos patrones similares (figura 10): cuando el ISM aumenta (mayor optimismo en el mercado), también lo hace el ICII (mayor confianza en el mercado); y viceversa. En ocasiones, como en la subida de octubre a noviembre de 2016, y en la bajada de abril a mayo de 2017, el ICII parece adelantarse al sentimiento obtenido de las noticias financieras. Esto podría deberse al efecto de las expectativas de evolución del mercado incorporado en el ICII. Esta hipótesis se ha testado mediante dos estrategias diferentes, por las que intuimos que este efecto puede de hecho tener una influencia en el fenómeno observado. Por un lado, se comprobó la correlación entre ambos índices en el mismo mes y entre el ICII y el ISM retardado un mes, y no se encontró ninguna mejora en el movimiento conjunto entre ambas variables (la correlación pasó de 0,217 a 0,148). Por otro lado, no obstante, se eliminó el Índice de perspectivas para el cómputo del ICII y se comprobó una mejora en la correlación (de 0,217 pasó a 0,274) y en el  $R^2$  del gráfico de dispersión (de 0,0471 pasó a 0,0749).

Figura 9. ISM, ICII e IBEX 35 (2015-18)



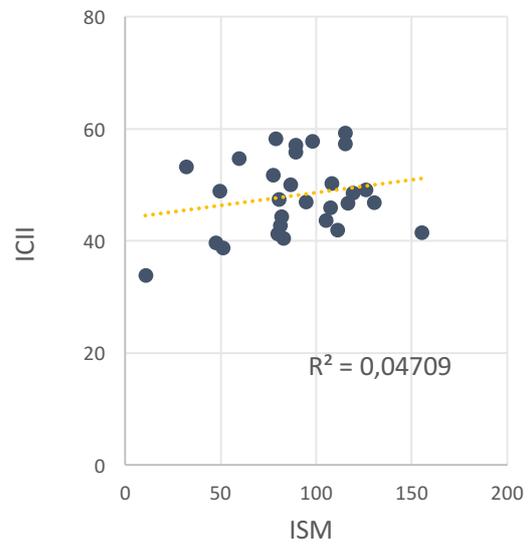
Fuente: ISM - elaboración propia basada en el análisis de noticias del Expansión mediante la herramienta LIWC. Escala adaptada a los resultados del ICII. ICII – datos del grupo *Behavioral Finance* de la Universidad Pontificia Comillas. IBEX 35 – datos extraídos de *Yahoo Finance*

Figura 10. ISM mensualizado vs. ICII (2015-18)



Fuente: ISM - elaboración propia basada en el análisis de noticias del Expansión mediante la herramienta LIWC. Resultados diarios mensualizados y escala adaptada a los resultados del ICII. ICII – datos del grupo *Behavioral Finance* de la Universidad Pontificia Comillas

Figura 11. Gráfico de dispersión ICII vs. ISM



Fuente: elaboración propia basada en el análisis de noticias del Expansión mediante la herramienta LIWC y en los datos del ICII del grupo *Behavioral Finance* de la Universidad Pontificia Comillas

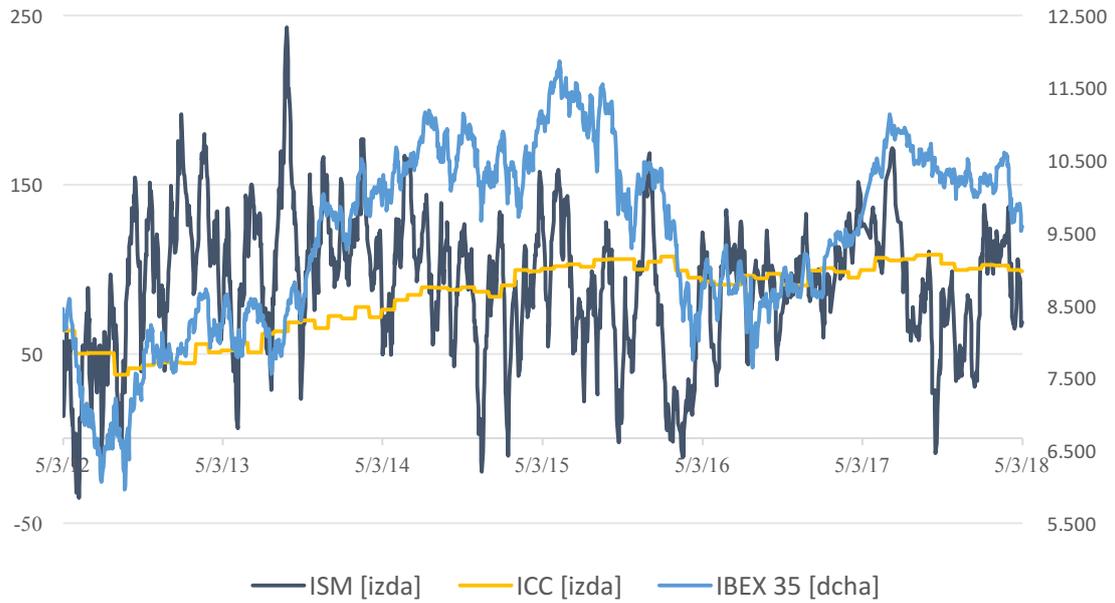
Aunque gráficamente ambos índices parecen representar la misma realidad (figura 10), y por tanto, podría concluirse que ambos reflejan el sentimiento de mercado, estadísticamente la relación entre ambos es todavía débil. Se cree que esto puede deberse a que entre enero y septiembre de 2017, ambos índices presentan movimientos completamente independientes. Este es un periodo alcista en el mercado español, y si recordamos los resultados que García (2013) presentaba en su estudio, podríamos concluir que los inversores se ven influenciados por las noticias financieras de manera diferentes dependiendo de si es una época de recesión o de expansión. Si esto fuera así, en épocas de recesión o en épocas bajistas, el ICII estaría más fuertemente correlacionado con el ISM, pues los inversores prestarían más atención a la información contenida en las noticias, mientras que en épocas alcistas, ambos índices se moverían de forma independiente. La comprobación de esta hipótesis se ve limitada por el pequeño número de observaciones con las que todavía se cuentan del ICII, que empezó a publicarse hace apenas dos años y medio.

#### 4.3.2. ISM vs. Índice de Confianza del Consumidor

El Índice de Confianza del Consumidor (ICC) es un indicador elaborado por el CIS. El ICC está disponible desde septiembre de 2004 y se publica desde entonces de forma mensual. Este índice puede tomar valores entre el 0 y el 200 y se calcula en base a las respuestas obtenidas de una encuesta realizada a 1.200 individuos mayores de 16 años. Su finalidad es anticipar la evolución del consumo privado y se compone a partir de preguntas sobre la valoración de los consumidores sobre la evolución económica actual y reciente y sobre sus expectativas respecto a la potencial evolución en el futuro inmediato (CIS, 2018).

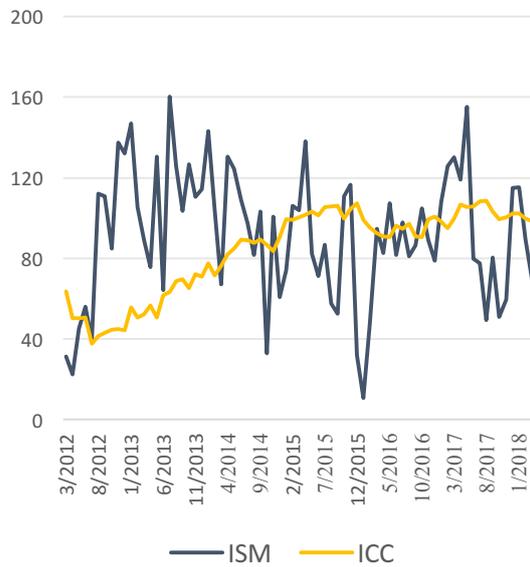
El ICC fue propuesto por Schelling (2009) como proxy del estado anímico del inversor, afirmando que el sentimiento del consumidor debía estar fuertemente correlacionado con el sentimiento del inversor. No obstante, en las figuras 12 y 13 podemos observar que la similitud entre la evolución del ISM y del ICC es prácticamente inexistente y por tanto parecen estar reflejando sentimientos completamente independientes. De hecho, la correlación entre ambas medidas es cercana a nula (-0,1056).

Figura 12. ISM, ICC e IBEX 35 (2012-18)



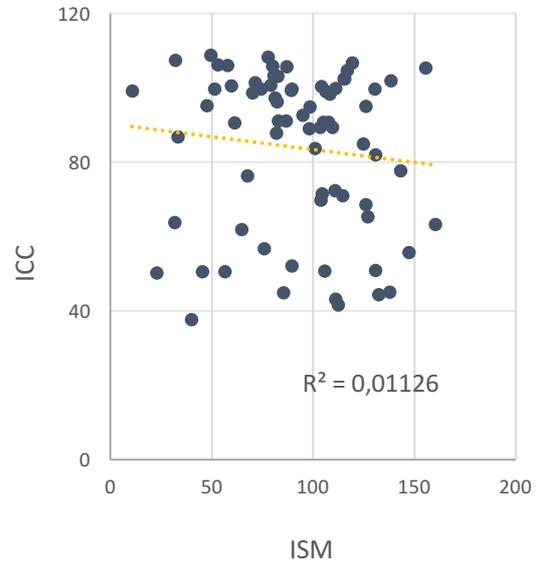
Fuente: ISM - elaboración propia basada en el análisis de noticias del Expansión mediante la herramienta LIWC. Escala adaptada a los resultados del ICC. ICC – datos obtenidos del CIS. IBEX 35 – datos extraídos de Yahoo Finance.

Figura 13. ISM mensualizado vs. ICC (2012-18)



Fuente: ISM - elaboración propia basada en el análisis de noticias del Expansión mediante la herramienta LIWC. Resultados diarios mensualizados y escala adaptada a los resultados del ICC. ICC – datos obtenidos del CIS

Figura 14. Gráfico de dispersión ICC vs. ISM



Fuente: ISM - elaboración propia basada en el análisis de noticias del Expansión mediante la herramienta LIWC y en los datos del ICC obtenidos del CIS

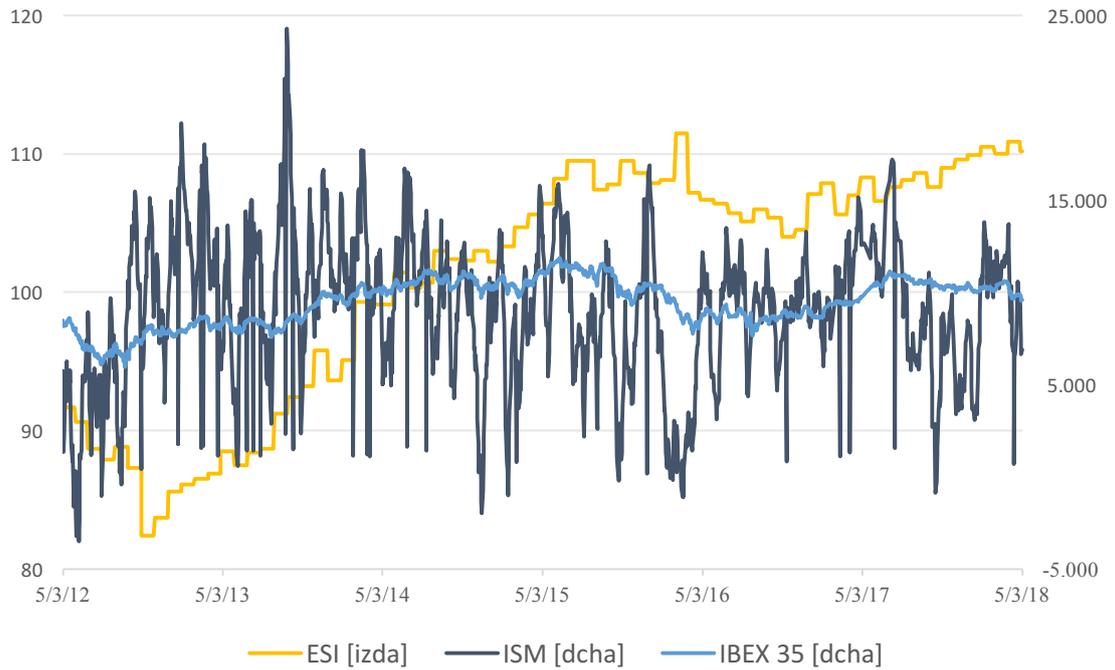
Puesto que la frecuencia de publicación del ICC vuelve a ser mensual, se repite el mismo análisis realizado para comparar el ICII, mensualizando los resultados del ISM para observar patrones a largo plazo con mayor claridad. La figura 13 muestra los resultados de dicho análisis, revelando que el ICC parece ser una medida mucho más estable en el tiempo que el ISM, en el que se observan grandes fluctuaciones de manera casi mensual. Así mientras la desviación típica del ISM es de 40,7, la del ICC es de tan sólo 21,4. Una posible explicación a este resultado es que los consumidores no actualizan sus expectativas ni sus perspectivas ante la economía de manera ni tan regular ni tan rápida como harían los inversores ante cambios en el mercado. De esta manera, el ICC recogería su sentimiento general respecto al rumbo de la economía (en un marco más general y más a largo plazo), que el sentimiento recogido por el ISM, mucho más influenciado por eventos con consecuencias en el futuro inmediato.

#### 4.3.3. ISM vs. Indicador de Sentimiento Económico

El Indicador de Sentimiento Económico (ISE, o ESI por sus siglas en inglés) es un indicador elaborado por la Dirección General de Asuntos Económicos y Financieros de la Comisión Europea para todos los países miembro de la Comisión. El ESI está disponible desde 1985, aunque para España no comenzará a calcularse hasta abril de 1987. Desde entonces, se publica de forma mensual. El ESI se calcula como un índice con base 100, por lo que valores de 100 indican la media a largo plazo de la confianza económica en cada mercado, mientras que los valores por debajo de 100 indican una desconfianza en la economía frente al nivel histórico. Este indicador se obtiene a partir de entrevistas sobre la situación actual y futura de la economía y se compone de cinco indicadores parciales: de confianza del sector industrial (40%), servicios (30%), consumidores (20%), construcción (5%) y comercio (5%).

Las figuras 15 y 16 revelan una situación muy similar a la encontrada en la comparación entre el ICC y el ISM. De nuevo, el ESI y el ISM parece ser indicadores completamente independientes, y guardando una relación prácticamente nula (correlación de -0,1418 en este caso). El ESI parece ser también una medida mucho más estable que el ISM de la percepción de la población ante la situación económica (de hecho tanto el ESI como el ICC presentan su sentimiento más negativo a finales de 2012, momento a partir del cual la recuperación económica se traduce en una mejora en el sentimiento de ambos indicadores; y la correlación entre ambos es de 0,9722).

Figura 15. ESI, ISM e IBEX 35 (2012-18)



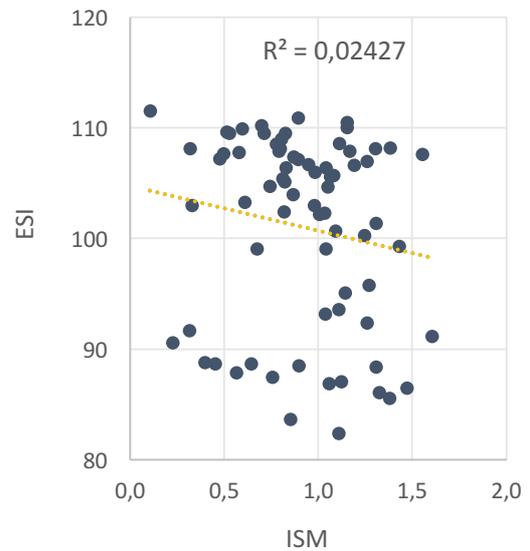
Fuente: ESI – datos obtenidos de la Comisión Europea. ISM - elaboración propia basada en el análisis de noticias del Expansión mediante la herramienta LIWC. Escala adaptada a los resultados del IBEX 35. IBEX 35 – datos extraídos de Yahoo Finance.

Figura 16. ISM mensualizado vs. ESI (2012-18)



Fuente: ISM - elaboración propia basada en el análisis de noticias del Expansión mediante la herramienta LIWC. Resultados diarios mensualizados. ESI – datos obtenidos de la Comisión Europea.

Figura 17. Gráfico de dispersión ESI vs. ISM



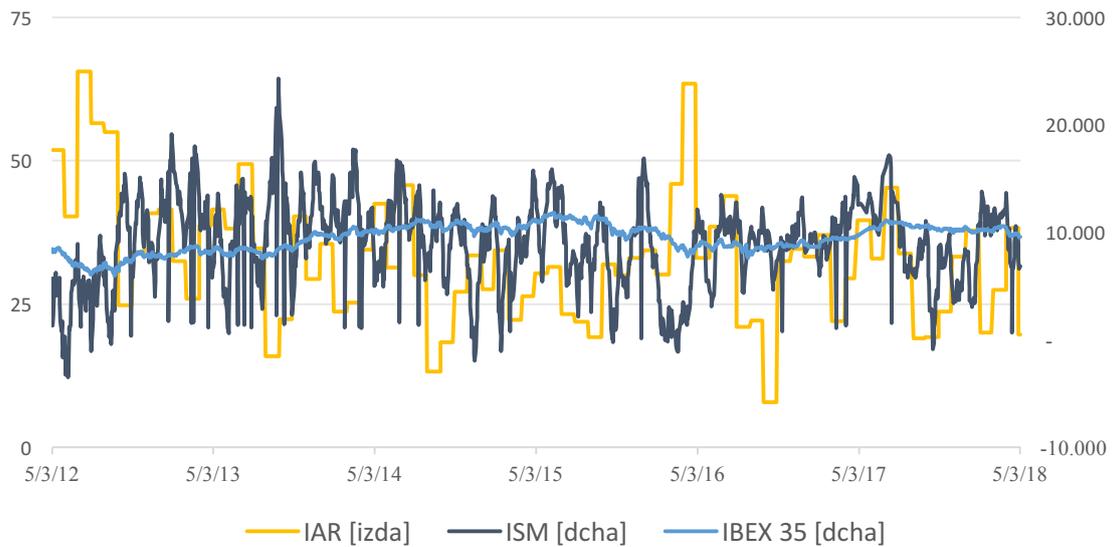
Fuente: elaboración propia basada en el análisis de noticias del Expansión mediante la herramienta LIWC y en los datos del ESI obtenidos de la Comisión Europea

#### 4.3.4. ISM vs. IAR

Las búsquedas en Google Insights (actual Google Trends) fueron utilizadas por Gómez Martínez (2013) para construir un Índice de Aversión al Riesgo (IAR) a nivel europeo. El IAR fue construido en 2013 a partir de las búsquedas de los términos “bear market”, “mercado bajista”, “economic recession”, “recesión económica”, “crisis económica”, “economic crisis”, “crisis financiera”, “financial crisis”, “short selling” y “venta en corto”. El autor elige estas búsquedas pues son las que determina en su artículo que tienen capacidad predictiva negativa sobre las cotizaciones del Eurostoxx. Para el propósito de este trabajo, se ha replicado el IAR con las búsquedas en google de los mismos términos, excepto “mercado bajista” y “venta en corto”, por insuficiencia en el número de búsquedas. Además, se ha limitado el ámbito geográfico a España (en lugar de Europa, como se hacía en la construcción del IAR original). La herramienta de Google proporciona así de forma gratuita un indicador mensual de la popularidad de las búsquedas escogidas. Como resultado de hacer una media entre la popularidad de las distintas búsquedas, se obtiene un indicador que toma valores entre el 0 y el 100.

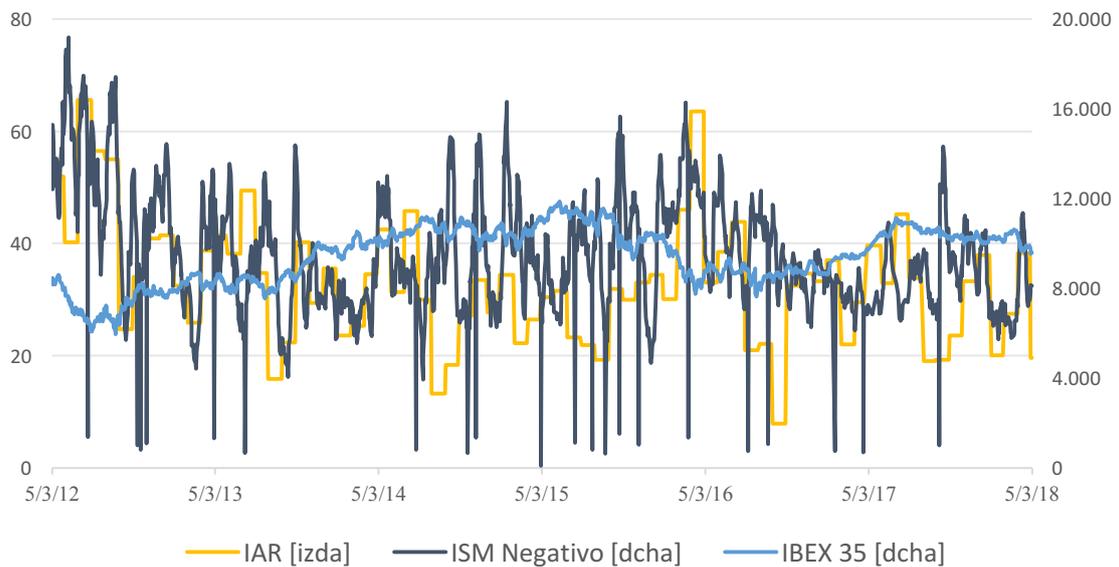
La figura 18 presenta la evolución del ISM, del nuevo IAR construido y del IBEX 35, entre marzo de 2012 y marzo de 2018, en la que no se ve claro un movimiento conjunto entre ambas medidas de sentimiento (su correlación es negativa, de -0,2670). Dado que el IAR es una medida negativa de sentimiento (por razón de su construcción), se decide en su lugar, comparar el ISM negativo con el IAR (figura 19), y la relación encontrada en este caso es mucho más fuerte. La correlación entre estos dos indicadores es de 0,4956, y se puede apreciar de forma gráfica en la figura 19, como subidas y caídas en el ISM y en el IAR coinciden en su mayoría a lo largo del periodo estudiado.

Figura 18. IAR, ISM e IBEX 35 (2012-18)



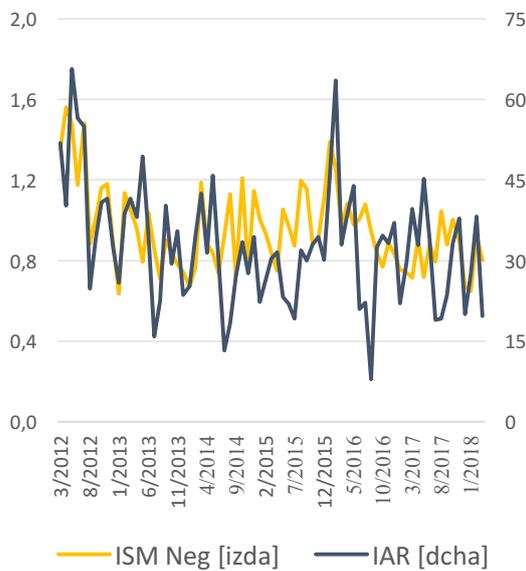
Fuente: IAR – elaboración propia basada en el análisis de búsquedas de Google Trends siguiendo a Gómez-Martínez (2013). ISM - elaboración propia basada en el análisis de noticias del Expansión mediante la herramienta LIWC. Escala adaptada a los resultados del IBEX 35. IBEX 35 – datos extraídos de *Yahoo Finance*

Figura 19. IAR, ISM Negativo e IBEX 35 (2012-18)



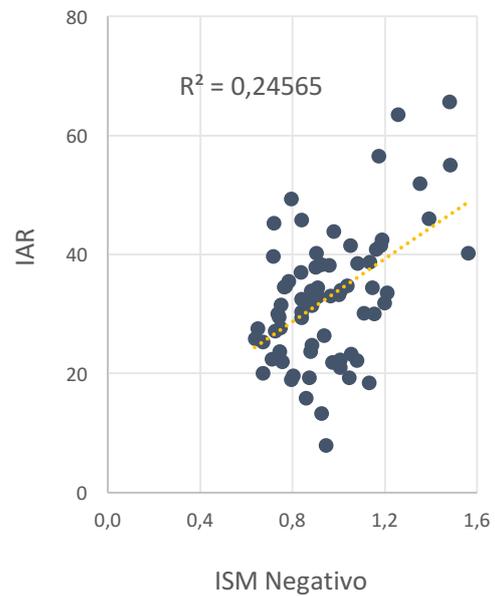
Fuente: IAR – elaboración propia basada en el análisis de búsquedas de Google Trends siguiendo a Gómez-Martínez (2013). ISM Negativo - elaboración propia basada en el análisis de noticias del Expansión mediante la herramienta LIWC. Escala adaptada a los resultados del IBEX 35. IBEX 35 – datos extraídos de *Yahoo Finance*

Figura 20. ISM Negativo mensualizado vs. IAR (2012-18)



Fuente: ISM Negativo - elaboración propia basada en el análisis de noticias del Expansión mediante la herramienta LIWC. Resultados mensualizados. IAR – elaboración propia basada en el análisis de búsquedas de Google Trends.

Figura 21. Gráfico de dispersión IAR vs. ISM Negativo



Fuente: elaboración propia basada en el análisis de búsquedas de Google Trends y en el análisis de noticias del Expansión mediante la herramienta LIWC

#### 4.4. ¿Es el ISM una medida válida del sentimiento de mercado?

Como recapitulación de los resultados obtenidos podemos concluir que el ISM es más volátil que las medidas tradicionales del sentimiento de mercado, dado que es más sensible a la ocurrencia de eventos con consecuencias negativas/positivas en el corto y medio plazo. No obstante, no debe olvidarse que el ISM es la única medida hasta ahora conocida que puede calcularse de manera diaria (frente a la frecuencia mensual del resto de medidas estudiadas), y que esto podría presentarse como una gran ventaja a la hora de analizar los cambios en el mercado frente a cambios en el sentimiento del inversor, de manera granular. Además, se ha observado que el ISM no guarda relación alguna ni con el ICC ni con el ESI. Como se argumentó en secciones anteriores, esto puede deberse a que los estudios que proponen estos índices como *proxys* del sentimiento del inversor, demuestran su validez al verificar que existe una relación entre la medida encontrada y los movimientos en Bolsa (i.e. retornos del mercado o valores de los activos). No

obstante, esta relación simplemente verifica el link entre el índice y los mercados, sin verificar realmente si el sentimiento encontrado refleja de hecho el sentimiento del inversor. No obstante, sí se ha observado que el ISM varía de forma conjunta con el ICII (a excepción del periodo de expansión en 2017) y con el IAR. Dado que estas son dos medidas basadas directamente, de una manera u otra, en las acciones (ya sean respuestas a preguntas o búsquedas en internet) de inversores individuales o profesionales, este descubrimiento se considera lo suficientemente relevante como para validar el ISM como una medida fiable del sentimiento del inversor, y el AS como una técnica que enriquece las actuales herramientas de medición del mismo.

## 5 CONCLUSIONES Y PROPUESTAS

El Análisis de Sentimientos es una técnica de análisis de textos cuyo objetivo consiste en determinar si el sentimiento de un texto, ya sea una crítica de un producto, una noticia de un periódico o un comentario en una red social, tiene una inclinación positiva, negativa o neutra. Autores como Paul Tetlock, Tim Loughran o Bill McDonald fueron pioneros en aplicar esta técnica al ámbito financiero para medir el sentimiento de mercado. Sus hallazgos han aportado importantes avances al mundo de las Finanzas, dado que el concepto de sentimiento de mercado ha cobrado actualmente y con el auge de las finanzas del comportamiento, más importancia que nunca para entender el comportamiento de los mercados. Asimismo, el debate en torno a este concepto gira en torno a tres preguntas: qué factores causan el sentimiento de mercado, qué impacto tiene en el mismo en la Bolsa y cómo puede identificarse y medirse. La última de estas preguntas es sin duda la que mayor controversia ha despertado entre los académicos. Puede entenderse entonces que el AS, como herramienta de medición, tenga el potencial de enriquecer en gran medida los análisis actuales de los mercados financieros.

Este trabajo es el primero en presentar la medición del sentimiento de mercado en España a través de técnicas de Análisis de Sentimientos. Para ello, se ha hecho una revisión de la literatura existente en tres ámbitos. Primero, aquella relativa a los avances desde la teoría clásica de Fama hasta las teorías contemporáneas de finanzas del comportamiento y que han llevado a considerar actualmente el sentimiento del inversor, como una de las claves a tener en cuenta para comprender tanto las decisiones individuales de inversores como el comportamiento de los mercados financieros en general. Segundo, aquella relativa a los avances en el campo de la informática, destacando aquellos enfoques y aplicaciones más relevantes de la técnica presentada. Tercero, aquella relativa a los estudios que han hecho converger ambos avances, es decir, que han tratado de medir el sentimiento de mercado a través de Análisis de Sentimientos. Todos los estudios revisados estudian el mercado estadounidense, y no se ha tenido conocimiento de ninguno que replique el análisis para la Bolsa española.

Para la tarea propuesta, se ha elegido analizar la columna “Crónica de Bolsa” del diario Expansión, el principal vehículo a través del cual el reconocido periódico informa a sus lectores sobre los acontecimientos acaecidos cada día en los mercados financieros.

Se ha optado además por un enfoque basado en léxicos y por el uso de la versión en español del diccionario LIWC, dado que este ha sido uno de los más utilizados por los académicos en España. Como resultado, se ha obtenido un Índice de Sentimiento de Mercado que captura de manera fiable el sentimiento del inversor en España. La validación del índice construido se ha realizado a través de su comparación con otras medidas tradicionales del sentimiento de mercado. Mientras que el ISM no parece guardar relación ni con el ICC, ni con el ESI (medidas más estables y generales del sentimiento de la población con respecto a la economía); sí parece hacerlo con el ICII y con el IAR (medidas más volátiles y específicas del sentimiento de los inversores). Esta conclusión está limitada en el caso del ICII por la corta trayectoria de este indicador (publicado desde hace dos años y medio). Cabe además destacar una de las grandes ventajas del ISM frente al resto de medidas tradicionales: su frecuencia. Mientras que otros indicadores se calculan de manera mensual, el ISM se puede calcular diariamente, lo que resulta en un índice más volátil, pero que a la vez permite realizar análisis granulares sobre la relación entre el sentimiento de mercado y los movimientos en Bolsa, con el potencial de trazar así las estrategias de inversión apropiadas.

El análisis realizado y los resultados encontrados están sujetos a una serie de limitaciones. Algunas de ellas son inherentes al Análisis de Sentimientos, una técnica en sí imperfecta. Estas son por ejemplo, la dificultad de extraer sentimientos de textos que usan la ironía u otros recursos retóricos o de interpretar sentimientos derivados de los contenidos encontrados en los textos. También lo son el tratamiento de la negación, los intensificadores u otros aspectos que suelen abordarse en el preprocesamiento de los textos. Otras limitaciones, en cambio, resultan de realizar el análisis en el idioma español. Puesto que este trabajo es pionero en la tarea, los recursos más adecuados y necesarios para el análisis (esto es, un diccionario de sentimientos aplicado al dominio financiero), son inexistentes en español. Se reconoce por tanto que el uso del LIWC, un diccionario general, a pesar de los mecanismos de control implementados, puede haber comprometido la validez/precisión de los resultados. Además, se ha observado que el contenido emocional de la columna elegida es relativamente escaso, lo que de nuevo, ha podido dificultar la extracción con precisión del sentimiento de las noticias.

Se espera de este estudio que presente un punto de partida para nuevas investigaciones sobre la materia en el idioma español. Las posibilidades de las futuras líneas de trabajo son por tanto infinitas. Por un lado, se anima a futuros estudios a trabajar

en la mejora del análisis aquí propuesto, ya sea mediante el uso de un diccionario específico a la aplicación al ámbito financiero o mediante el uso de un corpus diferente (otra columna, un blog de opinión...). Por otro lado, se consideran también una multitud de posibles análisis que enriquecerían la validez de la nueva medida de sentimiento de mercado aquí construida: estudiar en mayor profundidad su relación con el IBEX 35, analizar en detalle si la reacción de los mercados ante las noticias es diferente en el corto y en el largo plazo, observar si los inversores reaccionan de manera distinta ante noticias de contenido negativo o positivo, identificar si existen sentimientos específicos (miedo, incertidumbre, asombro) capaces de predecir movimientos en Bolsa mejor que simplemente el sentimiento positivo o negativo...

En conclusión, el Análisis de Sentimientos ha demostrado ser una innovadora técnica capaz de reflejar de manera fiable el sentimiento de los mercados, también en el caso español. Este es un descubrimiento con un gran potencial para responder a preguntas sobre la relación entre las noticias financieras y las decisiones de los inversores y sobre el funcionamiento de los mercados financieros en general.

## 6 BIBLIOGRAFÍA

Ahern, K., & Sosyura, D. (2014). Who writes the news? Corporate press Releases During Merger Negotiations. *Journal of Finance* , 69 (1), 241-291.

Arcila-Calderón, C., Ortega-Mohedano, F., Jiménez-Amores, J., & Trullenque, S. (2017). Análisis supervisado de sentimientos políticos en español: clasificación en tiempo real de tweets basada en aprendizaje automático. *El profesional de la información* , 26 (5), 973-982.

Black, F. (1986). Noise. *The Journal of Finance* , 41 (3), 528-543.

Baker, M., & Stein, J. (2003). Market Liquidity as a sentiment indicator. *Journal of Financial Markets* , 7 (2004), 271-299.

Baker, M., & Wurgler, J. (2007, June). Investor Sentiment in the Stock Market. *National Bureau of Economic Research* , 1-30.

Baker, M., Stein, J., & Wurgler, J. (2003). When Does the Market Matter? Stock Prices and the Investment of Equity-Dependent Firms. *The Quarterly Journal of Economics* , 118 (3), 969–1005.

Barberis, N., & Thaler, R. (2002). *A survey of behavioral finance*. Massachusetts: National Bureau of Economic Research.

Baviera Puig (2017). Técnicas para el Análisis de Sentimiento en Twitter : Aprendizaje Automático Supervisado y SentiStrength. *Revista de Comunicación Digital* , 1 (3), 33-50.

Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing with Python*. Sebastopol, CA: O'Reilly Media.

Bollen, J., Mao, H., & Zeng, X. J. (2011). Twitter mood predicts the stock market. *Journal of Computational Science* , 2 (1), 1-8.

Boudoukh, J., Feldman, R., Kogan, S., & Richardson, M. (2013). Which news moves stock prices? A textual analysis. *NBER Working Paper No. w18725*.

Bradshaw, T. (2018, febrero 7). *Snap beats Wall Street's forecasts for first time*. Recuperado de: <https://www.ft.com/content/eaf997f8-0b86-11e8-8eb7-42f857ea9f09>

Brown, G., & Cliff, M. (2005). Investor Sentiment and Asset Valuation. *The Journal of Business* , 78 (2), 405-440.

Campbell, J., & Kyle, A. (1993). Smart money, noise trading and stock price behavior. *Review of Economic Studies* , 60 (1), 1-34.

Carrera-Fernández, M., Guardía-Olmos, J., & Però-Cebollero, M. (2013). Psicología y lenguaje en política: los candidatos a la Presidencia del Gobierno y su estilo lingüístico. *Anuario de Psicología* , 43 (1), 39-52.

Choy, M., Cheong, M., Laik, M.-N., & Shung, K. P. (2011). *A sentiment analysis of Singapore presidential election 2011 using Twitter data with census correction*.

CIS. (2018). Nota de Investigación: La construcción del Indicador de Confianza del Consumidor. Recuperado de: [http://www.cis.es/cis/export/sites/default/-Archivos/NotasdeInvestigacion/NI006\\_ICC\\_Informe.pdf](http://www.cis.es/cis/export/sites/default/-Archivos/NotasdeInvestigacion/NI006_ICC_Informe.pdf)

Clarke, R., & Statman, M. (1998, May/June). Bullish or Bearish? *Financial Analysts Journal* , 63-72.

Corredor, P., Ferrer, E., & Santamaría, R. (2013). El sentimiento del inversor y las rentabilidades de las acciones. El caso español. *Revista española de Financiación y Contabilidad* , 42 (158), 211-237.

Da, Z., Engelberg, J., & Gao, P. (2015). The Sum of All FEARS Investor Sentiment and Asset Prices. *The Review of Financial Studies* , 28 (1), 1-32.

Das, S., & Chen, M. (2007). Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web. *Management Science* , 53 (9), 1375-1388.

De Long, B., Shleifer, A., Summers, L., & Waldmann, R. (1990). Noise Trader Risk in Financial Markets. *The Journal of Political Economy* , 98 (4), 703-738.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise Trader Risk in Financial Markets. *The Journal of Political Economy* , 98 (4), 703-738.

De Benito Lobo, L., de los Ríos Sastre, S., & Lazcano Benito, L. (2017, octubre 16). *Una aproximación al sentimiento del inversor en el mercado español: el Índice de Confianza del Inversor Institucional (ICII)*. Recuperado de: <https:// analisisfinanciero.ieaf.es/publicaciones-la-revista-analisis-financiero-4/gestion-patrimonial/85-una-aproximacion-al-sentimiento-del-inversor-en-el-mercado-espanol-elindice-de-confianza-del-inversor-institucional-icii>

Demers, E., & Vega, C. (2010). *Soft Information in Earnings Announcements: News or Noise?* INSEAD.

Díaz Rangel, I., Suárez Guerra, S., & Sidorov, G. (2014). Creación y evaluación de un diccionario marcado con emociones y ponderado para el español. *Revista semestral de lingüística, filología y traducción* , 29, 31-46.

Dubiau, L., & Ale, J. M. (2013). 14th Argentine Symposium on Artificial Intelligence. (pp. 36-47). Buenos Aires: ASAI.

Edmans, A., García, D., & Norli, Ø. (2007). Sports Sentiment and Stock Returns. *The Journal of Finance* , 62 (4), 1967-1998.

Emerging Technology. (2018, febrero 5). *100,000 happy moments. What makes people happy?* Recuperado de: <https://www-technologyreview-com.cdn.ampproject.org/c/s/www.technologyreview.com/s/610159/100000-happy-moments/amp/>

Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The journal of finance* , 25 (2), 383-417.

Fernández-Cabana, M., Rúas-Araújo, J., & Alves-Pérez, T. (2014). Psicología, lenguaje y comunicación: análisis con la herramienta LIWC de los discursos y tweets de los

candidatos a las elecciones gallegas de 2012. *The UB Journal of Psychology* , 44 (2), 169-184.

Frazzini, A., & Lamont, O. (2015). *Dumb money: mutual fund flows and the cross-section of stock returns*. Massachusetts: National Bureau of Economic Research.

García Cumbreiras, M. Á., Martínez-Cámara, E., Villena Román, J., & García Morera, J. (2016). TASS 2015 – The Evolution of the Spanish Opinion Mining Systems. *Procesamiento de Lenguaje Natural* , 56, 33-40.

García Cumbreiras, M. A., & Villena Román, J. (2016). TASS 2015 - The Evolution of the Spanish Opinion Systems. *Procesamiento de Lenguaje Natural* , 56, 33-40.

García, D. (2013). Sentiment during recessions. *The Journal of Finance* , 68 (3), 1267-1300.

Greenwood, R., & Nagel, S. (2008). *Inexperienced Investors and Bubbles*. Massachusetts: National Bureau of Economic Research.

Gómez Martínez, R. (2013). Señales de inversión basadas en un índice de aversión al riesgo. *Investigaciones Europeas de Dirección y Economía de la Empresa* , 19, 147-157.

Gómez Martínez, R., & Prado Román, C. (2014). Sentimiento del inversor, selecciones nacionales de fútbol y su influencia sobre sus índices nacionales. *Revista Europea de Dirección y Economía de la Empresa* , 23, 99-114.

Henríquez Miranda, C., & Guzmán, J. (2016). Una Revisión Sobre el Análisis de Sentimientos en Español. *Tecciencia* , 12 (22), 35-48.

Hirshleifer, D., & Shumway, T. (2003). Good Day Sunshine: Stock Returns and the Weather. *The Journal of Finance* , 58 (3), 1009-1032.

Hurtado, L. F., Pla, F., & Buscaldi, D. (2015). ELiRF-UPV en TASS 2015: Análisis de sentimientos en Twitter. *Proceedings of TASS 2015: Workshop on sentiment analysis at Sepln*, (pp. 75-79).

- Jegadeesh, N., & Wu, D. (2013). Word power: A new approach for content analysis. *Journal of Financial Economics* , 110 (3), 712-729.
- Jiménez Zafra, S., Martínez Cámara, E., Martín Valdivia, T., & Molina González, D. (2015). Tratamiento de la Negación en el Análisis de Opiniones en Español. *Procesamiento del Lenguaje Natural* , 54, 37-44.
- Johnson, E., & Tversky, A. (1983). Affect, Generalization and the Perception of Risk. *Journal of Personality and Social Psychology* , 45 (1), 20-31.
- Kamstra, M., Kramer, L., & Levi, M. (2003). Winter Blues: A Sad Stock Market Cycle. *American Economic Review* , 93 (1), 324-343.
- Lee, C., Shleifer, A., & Thaler, R. (1991). Investor Sentiment and the Closed-End Fund Puzzle. *The Journal of Finance* , 46 (1), 75-109.
- Lee, H., Surdeanu, M., MacCartney, B., & Jurafsky, D. (2013). On the Importance of Text Analysis for Stock Prediction.
- Loughran, T., & McDonald, B. (2011). When is a Liability not a Liability? *The Journal of Finance* , 66 (1), 35-65.
- McKinsey Global Institute. (2011). *Big Data: The next frontier for innovation, competition, and productivity*. Recuperado de: [http://www.mckinsey.com/insights/business\\_technology/big\\_data\\_the\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation).
- Martínez-Cámara, E., Martín-Valdivia, T., Molina-González, D., & Perea-Ortega, J. (2014). Integrating Spanish Lexical Resources by Meta-classifiers for polarity classification. *Journal of Information Science* , 1-19.
- Molina González, D., Martínez Cámara, E., Martín Valdivia, T., & Jiménez Zafra, S. (2015). eSOLHotel: Generación de un lexicón de opinión en español adaptado al dominio turístico. *Procesamiento del Lenguaje Natural* , 54, 21-28.

- Molina-González, M. C., Martín-Valdivia, M. T., & Perea-Ortega, J. M. (2013). Semantic orientation for polarity classification in Spanish reviews. *Expert Systems with Applications* , 40 (18), 7250-7257.
- Neal, R., & Wheatley, S. (1998). Do Measures of Investor Sentiment Predict Returns? *The Journal of Financial and Quantitative Analysis* , 33 (4), 523-547.
- Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin* , 29, 665-675.
- Nofsinger, J. (2005). Social Mood and Financial Economics. *Journal of Behavioral Economics* , 6 (3), 144-160.
- Ogneva, M. (2010, abril 19). *How companies can use sentiment analysis to improve their business*. Recuperado de: [https://mashable.com/2010/04/19/sentiment-analysis/#rfW3F2\\_ZH5qB](https://mashable.com/2010/04/19/sentiment-analysis/#rfW3F2_ZH5qB)
- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment Classification using Machine Learning Techniques. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 79-86). Philadelphia: Association for Computational Linguistics.
- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval* , 2 (1), 1-135.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). Linguistic inquiry and word count (LIWC).
- Ramírez-Esparza, N., Pennebaker, J. W., García, F. A., & Suriá, R. (2007). La psicología del uso de las palabras: un programa de computadora que analiza textos en español. *Revista Mexicana de Psicología* , 24 (1), 85-99.
- Slatcher, R., Chung, C., Pennebaker, J., & Stone, L. (2007). Winning words: Individual differences in linguistic style among U.S. presidential and vice presidential candidates. *Journal of Research in Personality* , 41 (1), 63-75.

- Schiller, R. (1984). Stock Prices and Social Dynamics. *Brookings Papers on Economic Activity* , 2, 457-498.
- Schmeling, M. (2009). Investor sentiment and stock returns: some international evidence. *Journal of Empirical Finance* , 16, 394-408.
- Salas-Zaráte, M. d., López-López, E., Valencia-García, R., Aussenac-Giller, N., Almela, Á., & Alor-Hernández, G. (2014). A study on LIWC categories for opinion mining in Spanish reviews. *Journal of Information Science* , 40 (6), 749-760.
- Saunders, E. M. (1993). Stock Prices and Wall Street Weather. *The American Economic Review* , 83 (5), 1337-1345.
- Shleifer, A., & Vishny, R. (1997). The Limits of Arbitrage. *The Journal of Finance* , 52 (1), 35-55.
- Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review* , 71 (3), 421-36.
- Shiller, R. J., & Pound, J. (1989). Survey Evidence on Diffusion of Investment Among Institutional Investors. *Journal of Economic Behavior and Organization* , 12, 47-66.
- Snap Inc. (2018). *Snap Inc. Reports Fourth Quarter and Full Year 2017 Results*. Annual Report, Venice, California.
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics* , 37 (2), 267-307.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology* , 29 (1), 24-54.
- Tetlock, P. (2011). All the News That's Fit to Reprint: Do Investors React to Stale Information? *The Review of Financial Studies* , 24 (5), 1481-1512.
- Tetlock, P. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance* , 62 (3), 1139-1168.

Tetlock, P. (2016). The Role of Media in the Stock Market. *News and Finance Conference*. Columbia Business School.

Tetlock, P., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying Language to Measure Firms' Fundamentals. *The Journal of Finance* , 63 (3), 1437-1466.

Thaler, R. (1999). The End of Behavioral Finance. *Financial Analysts Journal* , 55 (6), 12-17.

Turney, P. (2002). Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 417-424.

Tusmajan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting elections with Twitter. What 140 characters reveal about political sentiment. *Pros of the 4th Intl AAAI conf on weblogs and social media*, 10, pp. 178-185.

Whaley, R. (2000). The Investor Fear Gauge. *The Journal of Portfolio Management Spring 2000* , 26 (3), 12-17.

Yu, X. (2014). *Analysis of News Sentiment and its Application to Finance*. PhD Thesis, Brunel University, School of Information Systems, Computing and Mathematics.

Zhang, X., Fuehres, H., & Gloor, P. (2011). Predicting Stock Market Indicators Through Twitter “I hope it is not as bad as I fear”. *Procedia - Social and Behavioral Sciences* , 26, 55-62.

## 7 ANEXO

### *A.1. Ejemplo 1 de noticia publicada en la Crónica de Bolsa*

“La volatilidad se ha apoderado de las bolsas europeas, que han sufrido fuertes oscilaciones en las últimas sesiones. El Ibex ha cedido hoy un 1,34% hasta los 8.639,20 puntos y solo conoce los números rojos desde la elección de Donald Trump. El selectivo español se ha dejado un 1,7% en su segunda semana de caídas. Las ventas también se prolongan en la deuda, con la rentabilidad del bono español al filo del 1,5%.

Los mercados no consiguen aplacar las turbulencias desatadas desde el triunfo de Donald Trump en las elecciones estadounidenses. Con la actividad en Wall Street algo reducida por la celebración en EEUU del 'Día de los Veteranos' las oscilaciones han vuelto a ser la tónica en las bolsas europeas. A pesar de la festividad, el Dow Jones ha cerrado en un nuevo récord por encima de los 18.800 puntos, aprovechando el efecto del triunfo de Donald Trump y ha firmado su mejor semana en cinco años al anotarse un 5,4%. En cambio, el S&P 500 ha cedido un 0,1%, aunque en la semana se ha acabado dejando un 3,8%.

Por su parte, la Bolsa española ha contado las sesiones por caídas. Desde que se conoció al nuevo inquilino de la Casa Blanca, el Ibex ha perdido un 3,3%. Hoy el Ibex se ha dejado un 1,34% hasta los 8.639,20 puntos y en su segunda semana de recortes ha cedido un 1,7%.

Los grandes bancos se han convertido hoy en los principales lastres del Ibex. Santander se ha anotado un descenso del 4,04% y BBVA se ha dejado un 2,28%, con lo que pierde los 6 euros por acción. Popular se ha depreciado un 0,21% y amplía así su brecha por debajo del euro por acción. Otros blue chips del Ibex han acusado también las bajadas que persisten en divisas latinoamericanas como el peso mexicano. Telefónica ha recortado un 2,70% y Repsol, un 2,68%, castigado además por el fuerte descenso en el precio del crudo. De los grandes, solo Inditex ha escapado de los números rojos, pero por la mínima (+0,02%). La corrección ha alcanzado también a Arcelor (-3,04%) y a Técnicas Reunidas (-4,87%), mientras que Gamesa ha salido del bache con los resultados como revulsivo. Después de perder un 11,7% en las tres sesiones anteriores, la mejora de previsiones ha provocado una remontada del 6%. Dos bancos medianos han esquivado las pérdidas. Bankia ha repuntado un 3,66% y Sabadell, un 1,08%.

Las Bolsas europeas también comenzaron la jornada al alza. Pero un nuevo cambio de rumbo, habitual en las últimas sesiones, borró el optimismo inicial. Solo el Dax ha logrado cerrar en positivo (+0,36%) y el Mib italiano, prácticamente plano (+0,04%). La que peor parada ha salido ha sido la Bolsa de Londres, que se ha depreciado casi un 1,5% en paralelo a las subidas de la libra. El goteo de resultados empresariales ha continuado, con protagonistas hoy como Allianz. La aseguradora ha repuntado un 1,4% tras superar las previsiones con sus cuentas del tercer trimestre, gracias en parte a la mejora en su gestora Pimco. Otra compañía alemana, Air Berlin, ha recibido con frialdad (-0,2%) la publicación de sus resultados.

La jornada comenzó con avances moderados en la renta variable asiática. El Nikkei, desinflado un 5,3% el miércoles y disparado un 6,1% ayer, recuperó la calma con un leve repunte del 0,18%. En la Bolsa de Shanghai las ganancias alcanzaron el 0,8%.

La oleada de ventas ha continuado en el mercado de deuda pública. El interés exigido al bund alemán amplía su escalada hasta el 0,3%. Esta subida evita un mayor deterioro en la prima de riesgo de España, que se mantiene por debajo de los 120 puntos básicos, a pesar de que el interés del bono a diez años ha llegado a tocar el 1,5%. El triunfo de Donald Trump ha disparado las previsiones de inflación y de subidas de los tipos de interés en EEUU.

La última jornada de la semana consolida las subidas del dólar. El euro vuelve a bajar de los 1,09 dólares, más cerca de los niveles con los que comenzó el año. La libra sale reforzada de la semana. Alcanza los 1,26 dólares, y en su cruce con el euro la divisa comunitaria se repliega hasta los 86 peniques.

El mercado del petróleo sufre un nuevo varapalo. La OPEP augura un mayor superávit de oferta en 2017 debido al aumento de producción. El barril de Brent baja hasta los 44 dólares, mientras que el barril tipo West Texas, de referencia en EEUU, se sitúa en 43 dólares. La corrección continúa en el precio del oro. El metal precioso se desinfla hasta los 1.230 dólares la onza. Nada más conocerse el triunfo de Donald Trump en las elecciones estadounidenses, el oro se disparó por encima de los 1.320 dólares.”

Fuente: El Expansión, 11 noviembre 2016

## *A.2. Ejemplo 2 de noticia publicada en la Crónica de Bolsa*

“El último aluvión de resultados ha desinflado a Grifols y disparado a Meliá en un Ibex que ha cerrado la última sesión de febrero con una caída del 0,61% hasta los 9.840,30 puntos. El selectivo español salda el mes con una caída del 5,8%, en su peor mes desde junio de 2016, cuando tuvo lugar el referéndum sobre el Brexit. En el año cede un 2%. El dólar roza máximos de mes y medio frente al euro.

La Bolsa española cierra el mes más convulso de los últimos tiempos, el del resurgir de la volatilidad y el que supuso la entrada oficial de Wall Street en fase de corrección. El Ibex finalizó enero con subidas del 4%, en 10.451 puntos. La escalada que se remontaba a 2017 se frenó en seco con el cambio de mes. La corrección con la que arrancó febrero desinfló al índice selectivo español hasta un 8%, al registrar mínimos anuales en los 9.581 puntos. El Ibex ha cerrado la última sesión del mes con una caída del 0,61% hasta los 9.840,30 puntos. El selectivo salda febrero con una caída del 5,8% en su peor mes desde junio de 2016 tras la celebración del referéndum sobre el Brexit.

El cierre del mes coincide con el fin de la temporada de presentación de resultados en la Bolsa española. La consecuencia ha sido el mayor aluvión de cuentas de la temporada. Empresas como Endesa, Repsol, Merlin, Amadeus, Acerinox y Grifols, entre otras, han publicado sus cifras correspondientes al ejercicio 2017. A este listado se unen otras compañías del Ibex como Aena, Meliá e Indra que han recibido hoy el veredicto del mercado a los resultados publicados al cierre de ayer. Los analistas de JPMorgan han reaccionado a las cifras de Indra con una mejora de valoración, hasta los 12 euros por acción y sus títulos han sumado un 0,53%.

Este aluvión de resultados ha propiciado los mayores movimientos en el Ibex. En el ranking de las mayores caídas del día sobresale Grifols, que ha recortado un 4,93% y ha cerrado por debajo de los 23 euros por acción como respuesta a unos ingresos inferiores a lo esperado. En el extremo opuesto destacan las subidas, del 7,82%, con las que Meliá ha recibido sus cuentas y sus previsiones optimistas para 2018. También con fuertes compras se han recibido las de Repsol (+3,48%) y Merlin (+3,61%). Al margen de los resultados, dos pesos pesados como Inditex (-1,15%) y Telefónica (-0,34%) han perdido los niveles de 25 y 8 euros por acción, respectivamente. Del resto de pesos pesados, Santander ha cedido un 1,10% y BBVA, un 0,86%.

Las Bolsas europeas han cerrado con caídas similares a los del parqué español. La apertura alcista de Wall Street permitió a la mayoría de las plazas del Viejo Continente girar al alza. Sin embargo, la posterior vuelta a los números rojos al otro lado del Atlántico impuso finalmente las ventas en Europa. El Dow Jones ha cedido un 1,5%, hasta 25.028 puntos, y ha perdido un 4,3% en febrero, su peor mes desde enero de 2016. El S&P 500 ha recortado un 1,11%, hasta 2.713 puntos, y se ha dejado un 3,9% mensual en su peor mes en dos años. En cuanto al Nasdaq Composite, ha caído un 0,78%, hasta 7.273 puntos, con lo que ha aumentado su recorte mensual al 1,87%, su mayor caída desde octubre de 2016.

La agenda del día ha deparado otra oleada de resultados empresariales en Europa. Entre las compañías que han publicado sus cuentas correspondientes al ejercicio

2017 sobresalen Bayer, AstraZeneca, Admiral, ITV y Ahold. Las caídas han rozado el 2% en el caso del gigante químico alemán Bayer. Los números rojos se han impuesto igualmente en la cotización de la británica ITV, desinflada un 7,6% un día después de subir al calor de la oferta de Comcast sobre Sky. A cambio, Dialog Semiconductor se ha revalorizado un 6,6% en la Bolsa de Londres como premio a la publicación de sus resultados.

Las presiones bajistas se reactivaron ayer en Wall Street, con descensos superiores al 1%, y han continuado esta mañana en Asia, con pérdidas del 1,44% en el Nikkei de Tokio. La primera comparecencia en el Congreso de EEUU del nuevo presidente de la Reserva Federal, Jerome Powell, puso en guardia de nuevo a los inversores. Sus mensajes adquirieron un claro tono optimista. La confianza en la fortaleza del crecimiento es extensible a sus perspectivas sobre la inflación. Con este panorama, Powell advirtió de la conveniencia de evitar un sobrecalentamiento de la economía estadounidense, lo que se traducirá en nuevas subidas de los tipos de interés. El mercado, tras su mensaje, valora de nuevo la posibilidad de que sean cuatro, en lugar de tres, las subidas que aprobará la Fed en 2018.

El interés exigido al bono de Estados Unidos a diez años cae al 2,87% tras superar ayer el 2,90%, en línea con sus máximos de cuatro años. El PIB de EEUU ha confirmado hoy un crecimiento del 2,5% en el cuarto trimestre de 2017. En Europa, en cambio, el interés del bund alemán se afianza por debajo del 0,7%, mientras que en España la rentabilidad exigida al bono a diez años resiste cerca del 1,55%. La prima de riesgo repite por debajo de los 90 puntos básicos.

Los mensajes de la Reserva Federal refuerzan el rebote del dólar. La divisa estadounidense anticipa nuevas subidas de los tipos de interés por parte de la Fed, mientras que la moneda comunitaria ha recibido sin sorpresas el dato del IPC de la eurozona correspondiente a febrero. El repunte del 1,2%, lejos aún del objetivo del BCE próximo al 2%, se ha ajustado a lo esperado. El euro ha llegado a caer por debajo del nivel de los 1,22 dólares, sus cotas más bajas desde mediados de enero. La libra también cede terreno, y baja de los 1,38 dólares.

La remontada del dólar contribuye al freno del precio del petróleo. El barril de Brent baja de los 66 dólares y el barril de tipo West Texas, de referencia en EEUU, cotiza por debajo de los 63 dólares.”

Fuente: El Expansión, 28 de febrero de 2018

A.3. Clasificación de términos económico-financieros según LIWC

<b>Término</b>	<b>Clasificación</b>
<b>Activo</b>	"MecCog", "Causa", "Relativ", "Movim"
<b>Aumentar</b>	"EmoPos", "Relativ", "Movim"
<b>Baja</b>	"Muerte"
<b>Bolsa</b>	"Dinero"
<b>Caer</b>	"Relativ", "Movim", "Espacio"
<b>Cierre</b>	"Verbos", "Present", "Subjuntiv", "MecCog", "Insight"
<b>Coste</b>	"Dinero"
<b>Cuenta</b>	"Verbos", "Present", "VerbosEl", "Trabajo", "Dinero"
<b>Dividendo</b>	"Dinero"
<b>Descenso</b>	"Relativ", "Espacio"
<b>Deuda</b>	"Dinero"
<b>Impuesto</b>	"Dinero"
<b>Incrementar</b>	"EmoPos"
<b>Interés</b>	"MecCog", "Dinero"
<b>Inversión</b>	"Dinero"
<b>Mercado</b>	"Trabajo"
<b>Renta</b>	"Placer", "Hogar", "Dinero"
<b>Resultado</b>	"MecCog", "Insight", "Causa"
<b>Subida</b>	"Relativ", "Movim", "Trabajo", "Logro"
<b>Valor</b>	"Afect", "Emopos", "Dinero"
<b>Venta</b>	"Dinero"

Fuente: elaboración propia basada en la versión en español del diccionario LIWC. Otros términos también comunes en el ámbito económico-financiero como "alcista", "bajista", "depreciar", "pasivo", "recortes", "riesgo" o "volatilidad" no están incluidos en el diccionario