

---

# Modelos de estimación de la probabilidad de negociación informada: una comparación metodológica en el mercado español\*

**David Abad\*\***

Dpto. Economía Financiera  
Universidad de Alicante

**Antonio Rubia\*\*\***

Dpto. Economía Financiera  
Universidad de Alicante

**Resumen:** Caracterizar el grado de asimetría informativa ocupa un papel predominante en la literatura de microestructura moderna. En este trabajo, se revisa y analiza la idoneidad de los métodos existentes para estimar la probabilidad de negociación informada [Easley *et al.*, 1996; Nyholm 2002, 2003]. El análisis empírico toma como referencia el mercado español. La evidencia obtenida señala que el modelo de régimen cambiante de Nyholm (2002, 2003) no ofrece medidas consistentes con los efectos de asimetría informativa. El análisis sobre el mercado español revela además una mayor probabilidad de negociación informada en los activos menos líquidos como consecuencia de una reducción drástica en el número de agentes que negocia por liquidez. Esta evidencia sugiere un fuerte comportamiento de aversión al riesgo de selección adversa.

**Palabras clave:** Información asimétrica, selección adversa, probabilidad de negociación informada, PIN.

**Clasificación JEL:** C21, C52, D82.

**Abstract:** Determining the degree of informational asymmetry is a major topic in the literature of modern microstructure. In this paper, we review and analyze the suitability of the models for estimating the probability of informed trading [Easley *et al.*, 1996; Nyholm, 2002, 2003]. The empirical analysis is carried out on the Spanish market. We find evidence suggesting that the regime-switching model by Nyholm (2002, 2003) does not provide esti-

---

\* Los autores agradecen a Marina Balboa, Gonzalo Rubio, Roberto Pascual, un evaluador anónimo y a los participantes de un seminario de investigación en la Universidad de las Islas Baleares sus comentarios y sugerencias. Los autores agradecen la financiación del proyecto BEC2002-03797. Este artículo se completó durante la estancia de investigación de Antonio Rubia en the *Anderson School at UCLA*, cuya hospitalidad es gratamente reconocida.

\*\* Facultad de Económicas, Universidad de Alicante, CP 03080, España. Tel/Fax. (34) 965903621. Correo electrónico: [goliat@ua.es](mailto:goliat@ua.es).

\*\*\* Facultad de Económicas, Universidad de Alicante, CP 03080, España. Tel/Fax. (34) 965903621. Correo electrónico: [antonio.rubia@ua.es](mailto:antonio.rubia@ua.es).

mates consistent with the effects of asymmetry. The specific analysis on the Spanish market reveals a higher likelihood of the informed trading for the less-frequently traded assets as a consequence of the dramatic fall in the number of liquidity traders. This issue suggests a strong degree of aversion to the risk of adverse selection.

**Key words:** Asymmetric information, adverse selection, probability of informed trading, PIN.

**JEL Classification:** C21, C52, D82.

## 1. INTRODUCCIÓN

Una de las principales áreas de interés en microestructura se centra en la problemática ligada a la existencia de asimetrías informativas en el conjunto de inversores. Este fenómeno tiene profundas implicaciones para la liquidez de los activos y, en términos generales, resulta esencial para la correcta comprensión del mecanismo de ajuste de precios. La literatura financiera ha reconocido dos tipos de inversores atendiendo a su nivel de información. Por un lado, los agentes informados disponen de información privada que les permite obtener provecho económico mientras el precio no refleja el valor fundamental del activo. Por otro lado, los agentes desinformados negocian principalmente por motivos de liquidez, atendiendo únicamente a información públicamente disponible y a sus creencias personales, sin por ello afectar permanentemente al valor fundamental. Dado que ambos tipos de agentes operan simultáneamente en el sistema de negociación, no es posible identificar en términos de certeza su naturaleza informativa. Este hecho genera un riesgo de selección adversa para los oferentes de liquidez, que se arriesgan a negociar frente un agente con un conocimiento más preciso sobre el valor del activo, y por tanto, a pérdidas potenciales.

La existencia de asimetrías informativas conlleva claras repercusiones económicas para los participantes del mercado, con independencia de la estructura organizativa de éste. Por ello, ha surgido un creciente interés por caracterizar el grado de asimetría en los activos negociados y sus consecuencias para los inversores, principalmente desde dos grandes perspectivas. Por un lado, se ha podido demostrar teóricamente que la existencia de asimetrías informativas es una de las causas que subyacen en la formación de la horquilla de precios y que ésta incluye al menos un componente que pretende retribuir al oferente de liquidez por hacer frente al riesgo de selección adversa (Copeland y Galai, 1983; Glosten y Milgrom, 1985; Handa, Schwartz y Tiwari, 2003). La literatura centrada en la identificación de los componentes de la horquilla es muy extensa, y una revisión completa está más allá del objetivo de este trabajo. Destacan los modelos iniciales basados en la covarianza de los precios de transacción (ver Roll, 1984; Choi, Salandro y Shastri, 1988; George, Kaul y Nimalendran, 1991; Stoll, 1989; Lin, 1992; Huang y Stoll, 1994; Lin, Sanger y Booth, 1995) y los modelos más recientes basados en la variable indicadora de la negociación (entre otros, Glosten y Harris, 1988; Madhavan y Smidt, 1993; Huang y Stoll, 1997; Madhavan, Richardson y Roomans, 1997). Sin embargo, existen dudas sobre la capacidad de la mayoría de estos modelos para medir adecuadamente el componente de selección adversa.<sup>1</sup>

<sup>1</sup> Hasbrouck (1991a, 1991b) propone una metodología dinámica basada en el vector autorregresivo (VAR). Según este autor, los modelos que no tienen en cuenta este tipo de relaciones (como los modelos de variable indicadora) podrían fallar al tratar de captar cambios permanentes en precios. Por su parte, el análisis empírico de Van Ness, Van Ness y Warr (2001) analiza el comportamiento de cinco modelos de descomposición de horquilla, encontrando que la mayoría de ellos generan estimaciones del componente de selección adversa que no guardan la relación esperada con varias variables proxy de información.

---

Por otro lado, la literatura financiera ha planteado procedimientos para caracterizar los efectos derivados de asimetrías informativas a través de la probabilidad de negociar frente a un agente informado [PIN]. La ventaja de esta metodología es que proporciona una medida directa del grado de asimetría, aunque a costa de un planteamiento técnico más complejo. La dificultad de este problema no es trivial, y la literatura existente es relativamente escasa. La contribución original se debe a Easley, Kiefer, O'Hara y Paperman (1996) [EKOP], quienes desarrollaron una medida PIN centrada en el desequilibrio entre compras y ventas en un determinado periodo de tiempo. Recientemente, Nyholm (2002, 2003) plantea una metodología alternativa basada en una generalización del modelo de variable indicadora ampliamente utilizado en la literatura de descomposición de horquilla. El autor asume que la revisión en el valor fundamental del activo se condiciona al estado (informado o no informado) del inversor, modelizado a través de una variable latente. El principal atractivo de este último planteamiento reside en que intenta proporcionar una medida PIN para cada transacción, lo que en principio sugiere un potencial empírico mucho mayor que el modelo EKOP.

La principal motivación del presente trabajo es ofrecer una revisión de los procedimientos de estimación de la probabilidad de negociación informada, estableciendo un análisis comparativo sobre un mercado dirigido por órdenes, como es el mercado español. Con ello, el análisis ofrece varias perspectivas empíricas. En primer lugar, el análisis comparativo aporta evidencia clara sobre la idoneidad de los procedimientos analizados. Si bien la metodología EKOP goza de general aceptación y ha sido utilizada en distintas aplicaciones, mucho menos es sabido acerca del comportamiento empírico del procedimiento de régimen cambiante descrito por Nyholm y su relación con respecto al modelo EKOP. No existe evidencia acerca de su correcto funcionamiento, más allá de la aportada por el propio autor. De manera similar a los estudios de Van Ness *et al.* (2001), centrados en varios modelos de descomposición de horquilla, y Chung y Li (2003) para la relación entre la medida del modelo de EKOP y la estimación del componente de selección adversa, este trabajo pretende valorar y aportar mayor evidencia sobre la validez empírica de las alternativas existentes, hasta el momento, para aproximar la probabilidad de negociación informada.

En segundo lugar, la literatura centrada en aspectos microestructurales relevantes del mercado español es relativamente escasa. La literatura existente se ha centrado en la caracterización del componente de selección adversa con distintas finalidades (ver Acosta, Osorno y Rodríguez, 2000; Gómez-Sala y Yagüe, 2002; Farinós, García e Ibáñez, 1999; Mínguez y Martín, 2003), pero no existe en nuestro conocimiento ningún trabajo previo centrado en la medida PIN. Mediante este análisis, se ofrece una caracterización de la frecuencia con la que la información es incorporada en los precios, así como de la composición característica de los inversores interesados en el activo. Este conocimiento es valioso tanto para los organizadores de mercado, como para los propios inversores. Por ejemplo, y en la medida que el riesgo de selección adversa está condicionado por la información revelada por las compañías, la correcta caracterización del grado de asimetría informativa es valiosa a la hora de establecer y evaluar políticas de revelación adecuadas.

El planteamiento de Nyholm (2002) resulta aplicable a mercados dirigidos por órdenes, como se remarca en el trabajo original, si bien, dada la novedad del modelo, la evidencia previa se limita a la aportada por el autor para el NYSE, un mercado híbrido basado en la figura del especialista, intermediarios, y libro de órdenes.<sup>2</sup> Similarmente, el planteamiento del modelo

---

2 Nótese que el modelo es una extensión directa de la regresión de variable indicadora, que fundamenta buena parte de los modelos de descomposición de horquilla, utilizados tanto en mercados con creador de mercado como en mercados dirigidos por órdenes. Ver, entre otros, Ahn *et al.*, (2002) y las referencias allí citadas, para una aplicación de estos modelos sobre mercados dirigidos por órdenes.

EKOP resulta aplicable a cualquier mercado en el que exista creación de mercado (oferta de liquidez) y asimetría informativa. Así, la literatura financiera previa presenta evidencia empírica tanto para el NYSE (entre otros, Easley *et al.*, 1996), como para mercados dirigidos por órdenes (entre otros, Brockman y Chung, 2000a, 2000b, 2002). La razón fundamental es que existen grandes similitudes entre ambos tipos de mercado en el proceso de negociación, especialmente cuando el trasfondo es el de asimetría informativa. En primer lugar, los oferentes de liquidez de los mercados dirigidos por órdenes actúan como creadores de mercado de facto, de forma similar a los *scalpers* de los mercados de futuros. Además, el comportamiento del oferente de liquidez responde a pautas similares ante el problema de selección adversa, dado que los individuos racionales intentarán eludir las pérdidas esperadas por negociar frente a inversores con mejor información.<sup>3</sup> Por último, los modelos estructurales considerados asumen una enorme abstracción en los supuestos en los que se basan. Por ejemplo, el modelo EKOP está diseñado para capturar los rasgos más sobresalientes del proceso de negociación bajo información asimétrica, y no contempla ningún otro elemento más allá de los relacionados con el problema de selección adversa (como existencia de inventarios) y la hipótesis usual en estos modelos de aprendizaje bayesiano. Ambas hipótesis resultan plausibles tanto en mercados dirigidos por precios como en mercados dirigidos por órdenes.<sup>4</sup>

La selección muestral utilizada en este trabajo es similar a la propuesta en Easley *et al.* (1996) para datos procedentes del mercado NYSE. Tras agrupar los activos del mercado español atendiendo al nivel de actividad, la medida PIN se caracteriza para cada uno de ellos atendiendo a las dos metodologías de estimación. La evidencia obtenida es consistente en mostrar que los activos más atractivos en términos de actividad están sometidos a un menor grado de asimetría y a una menor probabilidad de negociación informada. Además, se perfilan ciertas características en la negociación típica del mercado español, como un fuerte comportamiento de aversión al riesgo de selección adversa, que indudablemente repercuten en el proceso de negociación y en la liquidez de los activos. Finalmente, se analiza la capacidad de las medidas resultantes para aproximar el grado de asimetría informativa. El análisis comparativo revela que el procedimiento de Nyholm, potencialmente más atractivo, no genera estimaciones con buenas propiedades, lo que arroja severas dudas sobre la idoneidad de este modelo, al menos, en su formulación original.

El trabajo se organiza como sigue. La sección segunda presenta y discute los principales rasgos de los modelos de estimación que se emplean. La sección tercera revisa las principales características microestructurales del mercado bursátil español, presenta los criterios de selección muestral empleados, y la información descriptiva usual. Los resultados

3 El agregado del libro de órdenes desempeña una función similar a la de los creadores de mercado. Además, con independencia de la arquitectura de mercado, se observan patrones similares en el comportamiento de la negociación. Por ejemplo, los oferentes de liquidez del mercado dirigido por órdenes gestionan activamente las dimensiones de la liquidez del mercado, al igual que los creadores de mercado, ajustando de forma continua horquillas y profundidad (ver, entre otros, Biais, Hillion y Spatt, 1995). Se ha demostrado de forma teórica cómo una parte de la horquilla implícita de los mercados dirigidos por órdenes es atribuible al problema de selección adversa (Handa *et al.*, 2003), de manera que se han justificado de la misma forma efectos intradiarios comunes a todo tipo de estructura (como la mayor horquilla observada cuando el mercado abre).

4 Madhavan (1992) demuestra que la secuencia de precios es igualmente informativa en ambas arquitecturas si se garantiza la libre entrada a la oferta de liquidez. Para la finalidad de los modelos analizados, resulta creíble que la información contenida en la estructura de negociación del mercado dirigido por órdenes lleva a similares conclusiones que la observable en mercados por precios.

---

de la estimación de los dos modelos y sus implicaciones se discuten en la sección cuarta. El análisis del comportamiento de los modelos se presenta en la sección quinta. Finalmente, la sección sexta resume y concluye.

## **2. MODELOS DE PROBABILIDAD DE NEGOCIACIÓN INFORMADA**

El planteamiento común de las dos metodologías asume que una parte de los inversores cuenta con información privada relevante para el valor del título, bien porque son los primeros en recibir una señal informativa, bien por que son capaces de hacer una mejor interpretación de una misma información pública. En cualquier caso, un agente informado intentará obtener un beneficio de su ventaja comprando (vendiendo) después de detectar una señal positiva (negativa) y mientras que las cotizaciones permanezcan sin ajustarse a la nueva información. Son los agentes informados los que con su actuación fuerzan al precio de los títulos a converger hacia su valor fundamental.

Junto con los agentes informados, los agentes desinformados negocian por razones de liquidez o de cartera, siguiendo una pauta de comportamiento no predecible. En agregado muestran un comportamiento ruidoso generando desviaciones aleatorias del precio con respecto del valor fundamental. Los modelos asumen que los oferentes de liquidez son conscientes de la existencia de asimetría informativa, pero no pueden distinguir si una transacción se origina por un agente informado o no. Como consecuencia, sufrirán pérdidas cuando negocien frente a un agente mejor informado, que intentarán minimizar actualizando los precios a la nueva información tan rápido como sea posible.<sup>5</sup> De esta manera, la horquilla reflejará, al menos, un componente derivado de esta fricción y su tamaño estará relacionado positivamente con el grado de asimetría y la medida PIN resultante. Más adelante se utilizará, entre otras medidas, la capacidad para predecir el tamaño de la horquilla como indicador de la aptitud del modelo.

Al margen de un trasfondo de referencia común, los dos modelos parten de consideraciones distintas y explotan estrategias diferentes para determinar la medida PIN. El modelo EKOP se basa en los desequilibrios que se producen entre compras y ventas en un determinado espacio de tiempo derivados de la llegada de nueva información al mercado. El modelo de Nyholm utiliza un planteamiento similar al utilizado en los modelos de descomposición de la horquilla, esto es, cambios permanentes en el precio están asociados a la presencia de información, mientras que cambios transitorios se deben a fricciones reales de carácter no informativo. A continuación, se resumen las principales características de cada metodología, así como su técnica de estimación.

### **2.1. EL MODELO DE EASLEY, KIEFER, O'HARA Y PAPERMAN [EKOP]**

Este procedimiento se basa en un planteamiento probabilístico acerca de cómo llegan eventos informativos al mercado, y la forma en la que los distintos inversores y el oferente de liquidez operan. Se asume que los eventos informativos relevantes para el activo se generan inde-

---

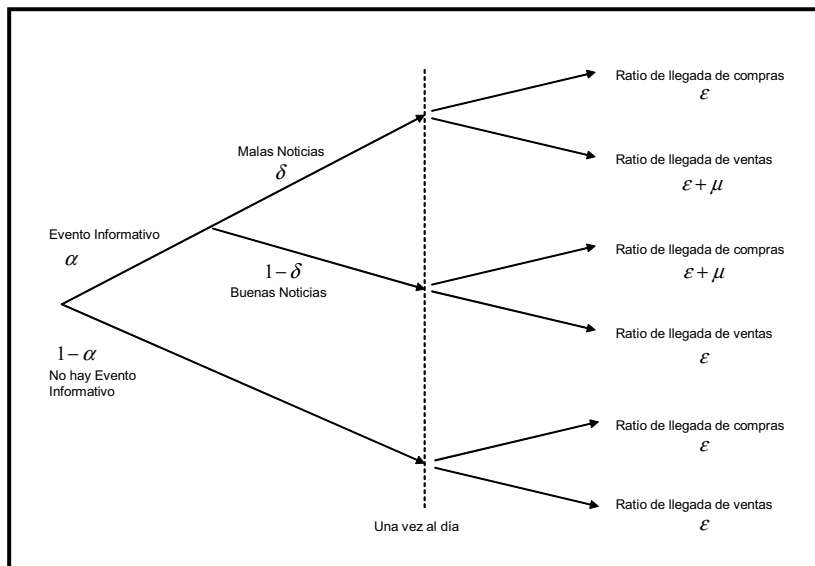
5 Nótese que el planteamiento teórico de los modelos se deduce bajo la figura del creador de mercado como oferente de liquidez. A pesar de que los mercados dirigidos por órdenes no cuentan con creadores de mercado con esta obligación, la oferta de liquidez se realiza de forma voluntaria por aquellos agentes que introducen órdenes límite. El carácter voluntario u obligatorio de la oferta de liquidez no tiene ningún tipo de repercusión sobre el riesgo de selección adversa, al que seguirá haciendo frente cualquier agente interesado en esta actividad.

pendientemente entre sí y a lo largo de los días de negociación,  $i=1, \dots, N$ , y que éstos ocurren con una cierta probabilidad  $\alpha$ . Los eventos informativos pueden generar valor para el activo con probabilidad  $(1-\delta)$ , o representar malas noticias, con probabilidad  $\delta$ . Por sencillez, estos parámetros se asumen constantes durante el intervalo de tiempo considerado.

La llegada de compradores y vendedores desinformados ocurre independientemente del flujo de información, y se modeliza a través de dos procesos independientes de Poisson. Tanto compradores como vendedores desinformados acuden a negociar al mercado mediante un ratio desconocido,  $\varepsilon$ , definido por intervalo de negociación. En el caso de los agentes informados, la llegada de información condiciona completamente su comportamiento. Estos operan comprando cuando identifican buenas noticias, y vendiendo en caso contrario. La llegada de agentes informados también se modeliza mediante un proceso Poisson con parámetro  $\mu$ , independiente de los anteriores. La Figura 1 resume el proceso completo.

**Figura 1**

El presente diagrama describe el proceso de negociación propuesto por EKOP (1996). En el primer nodo, la naturaleza decide si un evento informativo va a producirse. En caso afirmativo, se determina si se trata de buenas o malas noticias. Los tres nodos (no evento, buenas noticias y malas noticias) antes de la línea de puntos ocurren sólo una vez al día. Entonces, dado el nodo seleccionado por un día, los agentes llegan al mercado de acuerdo con el proceso de *Poisson* relevante. Esto es, en los días de buenas noticias el ratio de llegadas será  $\varepsilon + \mu$  para órdenes de compra y  $\varepsilon$  para las órdenes de venta. En los días de malas noticias el ratio de llegadas será  $\varepsilon$  para las compras y  $\varepsilon + \mu$  para ventas. Finalmente, en los días que no se produce evento informativo, sólo los agentes desinformados están presentes en el mercado y el ratio de llegadas para compras y ventas es  $\varepsilon$ .



La caracterización de la medida PIN requiere estimar el vector de parámetros desconocidos  $\theta = (\alpha, \delta, \varepsilon, \mu)'$  que gobierna las secuencias de eventos. La principal dificultad reside en que la llegada de eventos informativos y las operaciones realizadas por los agentes informados son inobservables. La estimación requiere entonces de un planteamiento basado en la secuencia de información observable: la llegada de órdenes de compra o venta. Este procedimiento encuentra un cierto fundamento teórico en el supuesto de que el oferente de liquidez es un agente Bayesiano que modifica sus expectativas atendiendo a la secuencia de órdenes, dado que ésta obedece a la estructura de información subyacente, con más compras en los días que hay eventos positivos y más ventas en los días con eventos negativos. Similarmente, en los días sin eventos informativos no habrá agentes informados en el mercado, y el número de transacciones será relativamente menor. Resulta posible por ello diseñar un modelo probabilístico basado en las distribuciones asociados a estos tres eventos básicos que permita identificar  $\theta$ .

Consideremos, en primer lugar, que en el  $i$ -ésimo día se produce una señal informativa negativa. Las órdenes de venta llegarán con el ratio  $(\mu + \varepsilon)$ , reflejando que tanto agentes desinformados como informados querrán vender, aunque sólo los inversores informados saben que están operando en el sentido correcto. También se observarán llegadas de órdenes de compra, al ratio  $\varepsilon$ , ya que únicamente los agentes desinformados comprarán en un entorno de malas noticias. Bajo la distribución de Poisson, la probabilidad conjunta condicionada de observar un cierto número agregado de órdenes de compras y ventas,  $\{C_i, V_i\}$ , durante ese día viene dada por:

$$\exp(-\varepsilon T) \frac{(\varepsilon T)^{C_i}}{C_i!} \exp(-(\mu + \varepsilon)T) \frac{((\mu + \varepsilon)T)^{V_i}}{V_i!} \quad (1)$$

donde  $T$  representa el número de intervalos de negociación dentro de la sesión. Nótese que la hipótesis de independencia permite factorizar las distribuciones marginales asociadas a compras y ventas. Por un razonamiento completamente similar, en un día con eventos positivos la verosimilitud condicionada de  $\{C_i, V_i\}$  viene dada por:

$$\exp(-(\mu + \varepsilon)T) \frac{((\mu + \varepsilon)T)^{C_i}}{C_i!} \exp(-\varepsilon T) \frac{(\varepsilon T)^{V_i}}{V_i!} \quad (2)$$

y, por último, si no se genera información privada, la probabilidad de una secuencia de compras y ventas viene dada por las operaciones de agentes desinformados, luego:

$$\exp(-2\varepsilon T) \frac{(\varepsilon T)^{C_i+V_i}}{C_i!V_i!} \quad (3)$$

En definitiva, la verosimilitud conjunta de observar la realización  $\{C_i, V_i\}$  se determina al ponderar (1)-(3) por la respectiva probabilidad de ocurrencia de cada suceso ("malas noticias", "buenas noticias", y "ausencia de información"), esto es,  $\alpha\delta$ ,  $\alpha(1-\delta)$  y  $(1-\alpha)$ :

$$\begin{aligned}
 L(C_i, V_i | \theta) = & \alpha \delta \left[ \exp(-\varepsilon T) \frac{(\varepsilon T)^{C_i}}{C_i!} \exp(-(\mu + \varepsilon)T) \frac{((\mu + \varepsilon)T)^{V_i}}{V_i!} \right] + \\
 & + \alpha(1 - \delta) \left[ \exp(-(\mu + \varepsilon)T) \frac{((\mu + \varepsilon)T)^{C_i}}{C_i!} \exp(-\varepsilon T) \frac{(\varepsilon T)^{V_i}}{V_i!} \right] + \\
 & + (1 - \alpha) \left[ \exp(-2\varepsilon T) \frac{(\varepsilon T)^{C_i + V_i}}{C_i! V_i!} \right]
 \end{aligned} \tag{4}$$

y, dado el supuesto de independencia para cada día de negociación, la función de verosimilitud conjunta para el horizonte de  $N$  días es,

$$L(\theta) = \prod_{i=1}^N L(C_i, V_i | \theta) \tag{5}$$

lo que permite identificar  $\theta$  bajo el principio de máxima verosimilitud.

La expresión de la probabilidad implícita de negociación informada se obtiene a partir del análisis del proceso de revisión bayesiano del oferente de liquidez, quién va condicionando sus cotizaciones al flujo de órdenes observable.<sup>6</sup> La medida PIN inferida se obtiene sustituyendo los parámetros poblacionales por sus estimaciones siguiendo la expresión:

$$PIN = \frac{\alpha \delta}{\alpha \mu + 2\varepsilon} \tag{6}$$

De forma intuitiva, la probabilidad de negociación informada en la expresión anterior refleja la proporción de agentes que negocian motivados por información privada ( $\alpha\delta$ ), sobre el total de los que actúan en el mercado ( $\alpha\delta + 2\varepsilon$ ). Este sencillo procedimiento se ha utilizado en gran variedad de aplicaciones empíricas, gozando de general aceptación. Entre otros estudios, el modelo se ha aplicado para extraer el contenido informativo del volumen de negociación y comprobar la precisión de distintos modelos de microestructura (Easley, Kiefer y O'Hara, 1997a y 1997b); analizar el efecto de las asimetrías informativas en el seguimiento de las recomendaciones de analistas financieros (Easley, O'Hara y Paperman, 1998); comprobar si los agentes informados prefieren negociar en el mercado de contado o en el de opciones (Easley, O'Hara y Srinivas, 1998); en el análisis de las operaciones de split (Easley, O'Hara y Saar, 2001); el análisis de la información generada en mercados paralelos (Grammig, Schiereck y Theissen, 2001), o el impacto de la negociación informada en la valoración de activos (Easley, Hvidkjaer y O'Hara, 2002).

<sup>6</sup> Para mayor detalle, consultar el trabajo original de Easley *et al.* (1996). En el mismo se deriva la existencia de la horquilla de precios por motivos de asimetrías de información, sin que sea necesario suponer otro tipo de fricción adicional (de manera similar a lo obtenido en Copeland y Galai, 1983 y Glosten y Milgrom, 1985).



## 2.2. EL MODELO DE RÉGIMEN CAMBIANTE DE NYHOLM

Este modelo no se fundamenta en un planteamiento económico, sino que se basa en una relación heurística de cómo actúa el oferente de liquidez fundamentada en el modelo de variable indicadora. La principal motivación de la extensión es permitir obtener una medida PIN transacción a transacción.

El modelo de variable indicadora asume que las revisiones en el precio del activo siguen el sentido de la transacción, de manera que compras (ventas) provocarán aumentos (disminuciones) en el precio atendiendo a una cierta fracción constante de horquilla observada. Además, la llegada de información pública genera fluctuaciones aleatorias en el precio. El modelo recoge estos elementos en la formulación,

$$\Delta M_t = \alpha \frac{S_{t-1}}{2} Q_{t-1} + \varepsilon_t; \quad \varepsilon_t \sim iid(0, \sigma^2); \quad t = 1, \dots, \tau \quad (7)$$

donde  $\Delta M_t$  representa la modificación en el valor subyacente (aproximada por el medio punto cotizado),  $S_{t-1}$  denota la horquilla cotizada y la variable indicadora  $Q_{t-1}$  señala si la transacción supuso una compra ( $Q_{t-1} = 1$ ) o venta ( $Q_{t-1} = -1$ ). El término de error  $\varepsilon_t$  recoge las fluctuaciones aleatorias, modelizadas como un proceso de ruido blanco con volatilidad  $\sigma$ . El parámetro  $\alpha$  representa el ajuste medio en los precios tras la negociación. Nótese que a pesar de que este modelo asume un comportamiento lineal muy restrictivo y con toda seguridad irrealista, es uno de los elementos básicos en gran parte de la literatura que ha tratado la descomposición de la horquilla de precios. Nótese, finalmente, que en este tipo de modelos el subíndice temporal hace referencia a la ocurrencia de las transacciones, que están irregularmente espaciadas.

Nyholm (2002) permite que las revisiones en el precio se condicionen, en principio, a la llegada de información privada. El modelo (7) se generaliza asociando el tamaño de la revisión a una variable latente inobservable, denotada  $\alpha$ , que toma valor 0 cuando la transacción  $t$  proviene de un agente desinformado, y 1 cuando la negociación tiene origen en una agente informado. El modelo resultante es una extensión de (7) que reconoce una relación no-lineal en la revisión de precios. La expresión resultante es:

$$\Delta M_t = \begin{cases} \alpha_0 \frac{S_{t-1}}{2} Q_{t-1} + \varepsilon_t & , si \quad I_{t-1} = 0 \\ (\alpha_0 + \alpha_1) \frac{S_{t-1}}{2} Q_{t-1} + \varepsilon_t & , si \quad I_{t-1} = 1 \end{cases} \quad (8)$$

donde  $\alpha_0, \alpha_1 \geq 0$  y  $\varepsilon_t \sim iid(0, \sigma^2)$ . Bajo las premisas del modelo, la revisión característica bajo el estado excitado  $(\alpha_0 + \alpha_1)$  es mayor que la observable en el estado normal, dado que el creador de mercado intentará cubrirse así de la mayor pérdida esperada, por lo que  $\alpha_1 \sim \alpha_0$ .

El modelo asume que la variable latente  $I_t$  sigue un proceso estocástico de régimen cambiante Markoviano de primer orden, de manera que la probabilidad de que la variable se encuentre en cada estado en un momento del tiempo sólo depende del estado en el momento anterior. La dinámica temporal del proceso se caracteriza entonces por una matriz de transición  $2 \times 2$ ,

denotada  $\mathbf{P} = \{p_{ij}\}$ , que gobierna el proceso y determina la probabilidad de que la variable se sitúe en un estado u otro en cada periodo de tiempo condicionado al estado en el que se situó en el momento anterior,

$$p_{ii} = \Pr(I_t = i | I_{t-1} = i); \quad i = \{0,1\} \quad (9)$$

y donde

$$\mathbf{P} = \begin{pmatrix} p_{00} & 1 - p_{11} \\ 1 - p_{00} & p_{11} \end{pmatrix} \quad (10)$$

El vector de los parámetros desconocidos  $\theta = (\alpha, \sigma, p_{00}, p_{11})'$  se estima mediante casi-máxima verosimilitud (QML) bajo el supuesto de normalidad en  $\varepsilon_t$ . El supuesto de normalidad en este contexto no es realista debido al comportamiento discreto de la variación de precios, pero el método QML proporciona estimaciones consistentes, aunque ineficientes, si la especificación funcional se realiza correctamente.

Las estimaciones resultantes permiten inferir la dinámica de la cadena oculta de Markov, caracterizando la llegada de información privada transacción a transacción. Ésta se determina mediante la técnica de suavizado (Hamilton, 1994; Kim y Nelson, 1998), infiriendo la probabilidad de observar el estado excitado en cada momento del tiempo dado el conjunto total de información. Con ello se obtiene la medida  $PIN_t = \Pr(I_t | \theta; \psi_\tau)$ , donde  $\psi_\tau$  denota el conjunto total de información disponible para el activo. Nyholm (2002) sugiere también una medida representativa del PIN promedio a lo largo del periodo analizado, similar a la de EKOP, calculada como la media de la secuencia estimada  $\hat{PIN}_t$ . Nótese que bajo el supuesto implícito de ergodicidad en la cadena de Markov, esta medida es representativa de la probabilidad incondicional de que el proceso esté en estado excitado,  $\Pr(I_t = 1)$ , por lo que puede calcularse simplemente como  $\hat{p}_{10} / (\hat{p}_{01} + \hat{p}_{10})$ .

Con la finalidad de comparar las estimaciones procedentes de las dos metodologías, nos centraremos en las probabilidades incondicionales generadas en los dos procedimientos.

### 3. DESCRIPCIÓN DEL MERCADO Y COMPOSICIÓN MUESTRAL

El presente estudio se centra en los activos que son negociados en el mercado español a través del Sistema de Interconexión Bursátil (SIBE). El SIBE es una plataforma electrónica y centralizada que conectó las cuatro plazas tradicionales del mercado bursátil español (Madrid, Barcelona, Bilbao y Valencia) en un único punto de liquidez, siguiendo un diseño similar al del sistema CATS empleado en otros mercados, como Paris y Toronto. El sistema se articula a través del denominado libro de órdenes límite, que registra y acumula las propuestas de compra y venta de todos los participantes. No existen creadores de mercado cotizando continuamente pre-

---

cios y cantidades, sino que son los propios participantes del mercado quienes ofrecen o demandan liquidez a través del uso de varios tipos de órdenes (básicamente, órdenes de mercado, órdenes por lo mejor y órdenes límite) que especifican diferentes condiciones referentes al precio, volumen, y tiempo de expiración.

La contratación para la gran mayoría de activos se realiza en tres fases claramente diferenciadas: subasta de apertura (de 8:30h a 9:00h), sesión abierta (de 9:00h a 17:30h) y subasta de cierre (de 17:30h a 17:35h). La apertura y cierre están basados en una subasta *call* donde se pueden introducir, modificar y cancelar propuestas, pero sin mediar ejecución alguna. Al final del periodo fijado se resuelve la subasta y se intercambian activos a un único precio que maximiza el número de títulos negociados y que representa el precio de referencia para la apertura y cierre de la sesión. La sesión principal está organizada como una contratación continuada, produciéndose una transacción cuando dos propuestas de distinto signo se cruzan. Las propuestas son introducidas a través de los miembros del mercado (agencias y sociedades de valores y entidades financieras autorizadas) y deben venir expresadas en euros (si especifican precio), existiendo dos variaciones mínimas o *ticks* para los mismos: 0.01 € para precios por debajo de 50 € y 0.05 € para precios superiores a 50 €. Finalmente, con el fin de evitar situaciones de alarma, el sistema realiza una serie de paradas automáticas cuando el precio al que se intercambia un título supera unos determinados rangos prefijados periódicamente por Sociedad de Bolsas, S.A. atendiendo a la volatilidad del título en cuestión. Cuando esto ocurre se activa un proceso de subasta de 5 minutos de duración, tras la cual se reanuda la negociación continua.

La distribución de la información relevante de negociación tiene lugar en tiempo real. En sesión abierta, el sistema ofrece información sobre los cinco mejores niveles de cada lado del libro (*pre-trade transparency*) y sobre las transacciones realizadas (*post-trade transparency*). Cuando se produce una subasta (de cierre, de apertura o por superar alguno de los rangos establecidos) el sistema es algo más opaco, ya que sólo se ofrece el precio de equilibrio potencial y el volumen que se intercambiaría a dicho precio. Toda la información es actualizada de manera instantánea en las pantallas de los miembros del mercado, de manera que cualquier introducción, cancelación o modificación de una propuesta queda reflejada inmediatamente.

### 3.1. SELECCIÓN MUESTRAL

Los datos intradiarios necesarios para llevar a cabo el análisis se recopilaron a partir de los ficheros SM procedentes de Sociedad de Bolsas, S.A. Estos ficheros contienen información sobre cualquier evento que afecta al primer nivel del libro de órdenes, como son la introducción y cancelación de órdenes límite, así como las distintas transacciones llevadas a cabo. A diferencia de los mercados dirigidos por precios, cada transacción puede ser calificada como iniciada por un comprador o un vendedor con exactitud, sin necesidad de recurrir a un algoritmo de clasificación. Esto es importante porque permite eliminar la conocida imprecisión de dichos algoritmos y el sesgo que podría derivarse en las estimaciones de la medida PIN (Theissen, 2001). El periodo muestral considerado abarca desde el 1 de septiembre al 29 de diciembre del año 2000, incluyendo 80 días de negociación.

El procedimiento de selección muestral sigue la metodología desarrollada en Easley *et al.* (1996). Todos los activos negociados a través del SIBE son ordenados atendiendo al volumen de negociación durante el año 2000. Los activos son agrupados en siete carteras con el mismo número de activos, donde la primera incluye los activos con mayor frecuencia de negociación.

**Tabla 1**  
**Estadísticos Descriptivos**

La tabla presenta el nombre y abreviatura de las empresas incluidas en nuestra muestra. El número de observaciones (filas de los ficheros SM) y los valores medios de distintas medidas de horquilla, profundidad, actividad, precio y volatilidad aparecen en las distintas columnas. La *Horquilla relativa*, la *Horquilla sobre tick*, la *Profundidad bid € y ask €* y el *Medio punto cotizado* son medias ponderadas por el tiempo que permanecen en el libro calculada sobre los 80 días del periodo muestral. La variable *volumen* es el montante monetario total intercambiado en el periodo considerado en millones de €. *% Cambio en el precio* muestra la rentabilidad obtenida calculada con el primer precio de transacción el día 1 y el último correspondiente al día 80. Finalmente, la *volatilidad* está medida como la desviación estándar de los precios de transacción en el periodo considerado. Las medias de estas variables para cada grupo también son presentadas.

Ticker	Nombre de la Compañía	Observ	Horq. relativa	Horq sobre tick	Prof. Bid €	Prof. Ask €	Volum. Mill €	% Cambio Precio	Medio punto cotizado	Volat
<b>CA: Alta Actividad</b>										
ACE	Acesa S.A.	23091	0.0031	2.83	15745	16717	333.84	7.45	9.02	0.0014
ACR	Aceralia SA	27832	0.0034	3.13	14325	12092	189.13	-3.41	9.24	0.0017
ACS	Construcción y Servicios, S.A.	19137	0.0039	10.32	21657	18412	346.38	-9.16	26.59	0.0014
AGS	Aguas de Barcelona, S.A.	19714	0.0037	5.20	13752	11920	248.16	-2.62	14.24	0.0013
ALB	Alba, S.A.	14641	0.0058	16.00	14645	15860	383.98	-9.86	27.74	0.0017
ANA	Acciona, S.A.	22745	0.0033	12.61	18211	20670	466.08	4.34	37.99	0.0012
AUM	Aurea Conc. Infraestructuras, S.A.	6190	0.0049	8.37	20969	15162	168.06	7.43	17.21	0.0019
DRC	Dragados, S.A.	35401	0.0036	3.87	19463	19344	742.28	27.89	10.57	0.0015
MAP	Mapfre, S.A.	8234	0.0066	12.77	16776	15826	266.93	14.48	19.29	0.0022
NHH	NH Hoteles, S.A.	18319	0.0047	6.26	15010	18105	428.27	-5.39	13.29	0.0016
REE	Red Eléctrica de España, S.A.	21349	0.0043	4.52	12513	8525	133.17	-5.93	10.55	0.0017
SOL	Sol Melía, S.A.	21512	0.0045	4.68	12662	11823	261.62	-14.06	10.62	0.0019
VAL	Vallehermoso, S.A.	19716	0.0044	2.98	12161	11876	213.05	-9.10	6.79	0.0018
	<i>Media de Sección Cruzada</i>	19837	0.0043	7.20	15992	15103	321.61	0.16	16.40	0.0016
<b>CM: Actividad Media</b>										
AEA	Azucarera Ebro Agrícola, S.A.	3243	0.0096	12.87	9295	8499	32.13	-8.03	13.51	0.0034
AZC	Asturiana del Zinc, S.A.	5115	0.0076	7.40	10503	8308	48.71	-11.71	9.90	0.0031
AZK	Azkoyen S.A.	4112	0.0093	6.35	64440	6400	22.71	-22.52	6.91	0.0039
CPF	Campofrío Alimentación, S.A.	3129	0.0080	9.84	11425	7475	26.09	6.24	12.25	0.0026
CRI	Cristalería Española, S.A.	2662	0.0075	23.72	11159	11312	37.22	-17.91	31.97	0.0028
ENC	Grupo Empresarial Ence, S.A.	4606	0.0076	13.52	8538	9608	56.68	-15.73	17.84	0.0032
PAS	Banco Pastor, S.A.	1332	0.0048	22.04	23699	23268	16.41	0.11	45.97	0.0006
PQR	Parques Reunidos, S.A.	4368	0.0093	4.59	5059	6638	30.59	-12.04	4.93	0.0037
SOS	Sos Arana, S.A.	2591	0.0107	9.69	8023	8519	17.86	-16.18	9.06	0.0026
TAZ	Transportes Azkar, S.A.	3243	0.0121	8.28	4561	57811	38.02	-33.89	7.18	0.0047
VDR	Portland Valderrivas, S.A.	1264	0.0127	26.59	9601	8222	24.55	-7.04	21.00	0.0024
VIS	Viscofan, S.A.	9709	0.0058	3.56	7535	6254	58.65	-34.51	6.42	0.0028
ZOT	Zardoya Otis, S.A.	5200	0.0051	4.67	10840	13100	60.68	0.88	9.07	0.0021
	<i>Media de Sección Cruzada</i>	3890	0.0085	11.78	9745	9491	36.18	-13.26	15.08	0.0029
<b>CB: Baja Actividad</b>										
ASA	Tavex Algodonera, S.A.	1070	0.0161	3.79	4633	5017	4.34	-24.80	2.37	0.0035
BAM	Bami S.A.	1739	0.0107	3.20	9499	6958	8.61	0.99	3.02	0.0031
CAF	Cons.y Aux. de Ferrocarriles, S.A.	673	0.0173	37.81	8755	6339	4.80	-11.06	21.99	0.0024
DGI	Dogi International Fabrics, S.A.	2228	0.0127	8.18	5035	3827	7.65	-44.44	6.72	0.0044
ENA	Enaco, S.A.	2525	0.0216	10.92	3289	4724	13.60	27.98	5.11	0.0056
IBG	Iberpapel Gestión, S.A.	804	0.0195	20.09	4063	5455	2.48	-10.36	10.37	0.0034
MCM	Miquel y Costas & Miquel, S.A.	789	0.0228	54.29	12438	6440	7.52	11.36	23.62	0.0025
NEA	Nicolás Correa, S.A.	923	0.0172	5.18	8104	3469	2.09	-35.14	3.08	0.0035
PAC	Papeles y Cartones de Europa, S.A.	1122	0.0164	2.50	6184	5137	6.18	-36.71	1.60	0.0032
RIO	Bodegas Rtojanas, S.A.	728	0.0196	17.35	6227	3652	6.85	-2.60	8.94	0.0025
UBS	Urbanizaciones y Transportes, S.A.	938	0.0219	1.63	8342	6229	2.44	-26.14	0.76	0.0033
VWG	Volkswagen Aktiengesellschaft, S.A.	835	0.0244	46.55	8742	6869	10.63	11.58	54.14	0.0046
ZNC	Española del Zinc, S.A.	893	0.0211	5.55	5132	4579	4.06	-73.23	2.92	0.0040
	<i>Media de Sección Cruzada</i>	1174	0.0186	16.70	6957	5284	6.25	-16.35	11.13	0.0035

**Tabla 2**  
**Horquilla y Punto Medio Cotizado por Carteras**

Los estadísticos descriptivos son presentados en este cuadro para los activos incluidos en la muestra agrupada en tres carteras de volumen. La Horquilla Cotizada Absoluta (HCA) es una media ponderada por tiempo calculada a partir de los datos de los ficheros SM de Sociedad de Bolsas. La Horquilla Cotizada Relativa (HCR) corresponde a la medida anterior estandarizada por el punto medio vigente en cada momento (PMC). El test no paramétrico de Kruskal-Wallis es utilizado para comprobar la hipótesis nula de que los valores de las variables descritas de los activos incluidos en cada una de las tres carteras pertenecen a poblaciones idénticas frente a la hipótesis alternativa de que al menos una de ellas genera valores observados más altos que el resto. El test no paramétrico de Mann-Whitney es utilizado para comprobar la hipótesis de que dos muestras proceden de poblaciones idénticas contra la hipótesis alternativa de que una de ellas tiende a ofrecer valores superiores.

<b>Panel A: Estadísticos Horquilla y Punto Medio</b>			
	CA	CM	CB
Activos Incluidos	13	13	13
<i>HCA</i>			
Media	0.072	0.118	0.231
Mediana	0.052	0.097	0.081
Desv. Std.	0.044	0.077	0.360
<i>PMC</i>			
Media	16.396	15.086	11.123
Mediana	13.290	9.904	5.103
Desv. Std.	9.226	11.900	14.824
<i>HCR</i>			
Media	0.004	0.008	0.018
Mediana	0.004	0.008	0.019
Desv. Std.	0.001	0.002	0.004

<b>Panel B: Igualdad conjunta de posicionamiento central (Test de Kruskal-Wallis)</b>		
	Test	[Prob]
<i>HCA</i>	2.524	[0.28]
<i>PMC</i>	6.825	[0.03]
<i>HCR</i>	31.844	[0.00]

<b>Panel C: Igualdad de posicionamiento central dos a dos (Test de Mann-Whitney)</b>						
	CA vs CM	[Prob]	CA vs CB	[Prob]	CM vs CB	[Prob]
<i>HCA</i>	1.629	[0.09]	0.872	[0.38]	0.256	[0.80]
<i>PMC</i>	0.923	[0.36]	2.410	[0.02]	1.795	[0.07]
<i>HCR</i>	3.949	[0.00]	4.308	[0.00]	4.154	[0.00]

Siguiendo a Easley *et al.* (1996), nos centramos en los activos que corresponden a la segunda, cuarta y sexta carteras, aproximando así tres niveles de actividad: "alta", "media" y "baja". Denotaremos CA, CM y CB a cada uno de estos grupos. Se aplicaron varios filtros para excluir activos que pudieran sesgar las estimaciones propuestas. Así, se eliminaron algunos títulos negociados en el Sistema de Fijación de Precios Únicos (*fixing*) y los activos que no cumplían un grado mínimo de negociación diaria. También fueron excluidos varios activos con precios tan reducidos que implicaban *ticks* excesivamente restrictivos. La lista final de los activos seleccionados, así como información estadística referente a distintos indicadores de liquidez, actividad y volatilidad durante el año 2000 se presenta en la Tabla 1.

En la Tabla 2 se muestra un resumen por carteras de la horquilla de precios (absoluta y relativa), y del nivel de precios de los activos incluidos en las tres carteras. En los paneles B y C de esta tabla se llevan a cabo varias pruebas estadísticas sobre dichas variables (contrastes no-paramétricos de igualdad de medianas de Kruskal-Wallis y Mann-Whitney). Los contrastes rechazan la igualdad de valores medios a los niveles usuales para las variables nivel de precio y horquilla relativa, pero no pueden rechazar la hipótesis nula, ni en conjunto, ni por parejas, para el caso de la horquilla absoluta. Este hecho se debe a que existen diferencias en el nivel de precio medio de los activos incluidos en cada una de las carteras, como se desprende del análisis del punto medio cotizado. Los activos en la cartera de menor actividad poseen también menores precios medios. Claramente, es la horquilla en términos relativos la que ofrece un análisis mucho más robusto del progresivo deterioro del grado de liquidez de los activos considerados. Este hecho será tenido en cuenta más adelante a la hora de realizar el análisis comparativo.

## 4. RESULTADOS EMPÍRICOS

### 4.1. MODELO EKOP

La información intradiaria de los ficheros SM para cada activo se procesa para obtener el número de compras y ventas realizado durante cada uno de los días de la muestra. Sobre esta información se infiere los parámetros y la medida PIN resultante, tal y como se describe en la sección segunda.<sup>7</sup> En la Tabla 3 se presentan las estimaciones resultantes para el modelo EKOP sobre los activos de las tres carteras. Por motivos de claridad en la exposición, se presentan únicamente el valor medio, mediana, y la desviación estándar de los parámetros estimados, así como la media de la medida PIN resultante (panel A).<sup>8</sup> Se presentan también los estadísticos acerca de la significatividad entre las diferencias observadas para cada grupo de activos (paneles B y C). Reviste gran interés comentar estos resultados, y comentar las diferencias observables con los resultados evidenciados para el mercado NYSE en Easley *et al.* (1996). Aunque la comparación rigurosa entre los dos mercados no es posible, la referencia (cualitativa) con otros mercados permite obtener una visión simple e intuitiva acerca de la magnitud relativa de las estimaciones y de sus implicaciones.

El grupo de activos con menor frecuencia de negociación es el que tiene una mayor probabilidad media de negociar frente a un inversor informado. Por término medio, esta probabilidad es 17,9% para el grupo de activos más negociados, 21,7% para el grupo intermedio, y 26,8% para los activos menos negociados. Este resultado es el esperado a priori, dado que se acepta ampliamente en la literatura que los activos menos negociados son menos seguidos por los inversores y presentan un mayor grado de asimetría informativa. En este sentido, las estimaciones de las medidas PIN estarían en la línea de lo obtenido en el NYSE por Easley *et al.* (1996), y para otros mercados dirigidos por órdenes, como muestran Brockman y Chung (2000b).

7 La amplitud T en los procesos de Poisson se toma referida a intervalos horarios. Nótese que la elección particular de este término no tiene mayores repercusiones que reescalar los parámetros  $\epsilon$  y  $\mu$ , pero la elección tomada en este trabajo facilita la computación de la función de verosimilitud.

8 Las estimaciones se realizaron tomando una parrilla de valores iniciales para minimizar la posibilidad de convergencia a óptimo local. La convergencia se obtuvo en todos los casos. Los resultados pormenorizados resultantes de la estimación de cada activo se encuentran disponibles bajo petición.

**Tabla 3**  
**Horquilla y Punto Medio Cotizado por Carteras**

Esta tabla presenta la media, mediana y desviación estándar de los parámetros estimados siguiendo el modelo de Easley *et al* (1996) para 39 activos del mercado español divididos en 3 carteras de actividad. El parámetro  $\mu$  representa el ratio de llegada de agentes informados;  $\varepsilon$ , es el ratio de llegada de agentes desinformados;  $\alpha$ , representa la probabilidad de que ocurra un evento informativo en un día concreto y  $\delta$  es la probabilidad de que la nueva información esté representada por malas noticias. El *PIN* es una variable obtenida siguiendo la expresión:  $PIN = \alpha\mu / (\alpha\mu + 2\varepsilon)$  y mide la probabilidad de negociación basada en información. El test no paramétrico de Kruskal-Wallis es utilizado para comprobar la hipótesis nula de que los valores de los parámetros estimados de las tres carteras de volumen pertenecen a poblaciones idénticas, frente a la hipótesis alternativa de que al menos una de ellas genera valores observados más altos que el resto. El test no paramétrico de Mann-Whitney es utilizado para comprobar la hipótesis de que dos muestras proceden de poblaciones idénticas contra la alternativa de que una de ellas tiende a ofrecer valores superiores. Adicionalmente, para la medida *PIN* se ha realizado un t-test de igualdad siguiendo la metodología Boostrat (Panel D).

Panel A: Parámetros Estimados						
	CA	CM	CB			
Activos Incluidos	13	13	13			
$\mu$						
Media	10.948	4.954	2.130			
Mediana	10.795	3.604	1.303			
Desv. Std.	2.932	4.112	2.402			
$\varepsilon$						
Media	10.832	2.058	0.568			
Mediana	10.892	1.822	0.485			
Desv. Std.	4.342	1.315	0.294			
$\alpha$						
Media	0.404	0.306	0.272			
Mediana	0.380	0.242	0.208			
Desv. Std.	0.082	0.183	0.204			
$\delta$						
Media	0.654	0.680	0.726			
Mediana	0.663	0.725	0.771			
Desv. Std.	0.100	0.218	0.253			
<i>PIN</i>						
Media	0.179	0.217	0.268			
Mediana	0.183	0.190	0.244			
Desv. Std.	0.034	0.079	0.132			
<b>Panel B: Igualdad conjunta de posicionamiento central (Test de Kruskal-Wallis)</b>						
		Test	[Prob]			
$\mu$		23.422	[0.00]			
$\varepsilon$		32.046	[0.00]			
$\alpha$		6.541	[0.04]			
$\delta$		2.597	[0.27]			
<i>PIN</i>		7.588	[0.02]			
<b>Panel C: Igualdad de posicionamiento central dos a dos (Test de Mann-Whitney)</b>						
	CA vs CM	[Prob]	CA vs CB	[Prob]	CM vs CB	[Prob]
$\mu$	3.385	[0.00]	4.154	[0.00]	2.564	[0.01]
$\varepsilon$	4.205	[0.00]	4.205	[0.00]	3.950	[0.00]
$\alpha$	1.487	[0.14]	2.666	[0.01]	0.512	[0.61]
$\delta$	0.718	[0.47]	1.538	[0.12]	0.897	[0.37]
<i>PIN</i>	1.179	[0.24]	2.769	[0.01]	1.384	[0.17]
<b>Panel D: Igualdad de media en PIN dos a dos (Bootstrap t-test)</b>						
	CA vs CM	[Prob]	CA vs CB	[Prob]	CM vs CB	[Prob]
<i>PIN</i>	-1.560	[0.04]	-2.278	[0.00]	-1.139	[0.09]

Un análisis más profundo de las estimaciones permite obtener mayor información sobre el origen de las diferencias observadas entre las distintas carteras, y sus implicaciones para los inversores y el mercado. Si bien es completamente previsible que los ratios de llegada de cada grupo de inversores se reduzcan en los activos menos negociados, como así se observa, estos activos también presentan una menor probabilidad característica de ocurrencia de eventos informativos (esta probabilidad se reduce drásticamente con respecto de los activos más líquidos). Mucho más importante, la relación de agentes informados con respecto de los desinformados es mucho mayor para los activos menos negociados. Por ejemplo, mientras que la magnitud del ratio de llegada de los agentes desinformados ( $\varepsilon = 10,832$ ) es comparable a la de los informados ( $\mu = 10,948$ ) en los activos más negociados, en el caso de los activos menos negociados el ratio de llegada de los agentes desinformados (0,568) sufre una caída dramática con respecto a los agentes informados (2,130). Esta diferencia representa una frecuencia de llegada sensiblemente inferior, que repercute claramente en una mayor medida PIN para estos títulos. Aunque nuestro análisis se centra en un periodo concreto, la relación estadística entre los ratios de llegada poblacionales es muy estable a lo largo del tiempo, como evidencian Easley *et al.*, (2002).

El hecho más significativo en términos de asimetría informativa en el caso de los activos ilíquidos no es tanto el que exista un menor volumen de noticias relevantes, o una menor negociación, si no el que característicamente haya pocos agentes dispuestos a negociar únicamente por motivos de liquidez. De ello se deriva una mayor probabilidad de negociar frente a un inversor informado, y mayores costes esperados derivados del riesgo de selección adversa. Si comparamos el ratio de la llegada de agentes informados en relación a la de los desinformados ( $\hat{\mu}/\hat{\varepsilon}$ ) con la obtenida en el NYSE para cada cartera, observamos que dicha relación es mucho mayor en el caso del mercado español, y que las diferencias entre carteras tienden a ampliarse notablemente al reducir el nivel de actividad.<sup>9</sup> Esto refuerza la idea, que ya se desprendería directamente del primer análisis, de que existe un fuerte grado de aversión al riesgo de selección adversa en el conjunto de inversores del mercado español que lleva a una clara preferencia del inversor desinformado hacia los activos donde ese riesgo es mucho menor. Esta evidencia es importante para los diseñadores de mercado, dado que este fenómeno tiene repercusiones para todas las facetas del mercado, y especialmente, para el éxito de los mecanismos destinados a mejorar la liquidez.

Finalmente, el análisis estadístico presentado en los paneles B y C revela que las diferencias observables entre carteras son significativas. Las diferencias en la medida de probabilidad muestran que, efectivamente, la cartera de menor actividad obtiene un PIN estadísticamente distinto en su comparación con la cartera de actividad alta. Los estadísticos basados en la prueba *bootstrap* rechazan la hipótesis de igualdad de medias, aunque sólo marginalmente en el caso de las carteras representativas de volumen bajo y medio.<sup>10</sup> Tanto los ratios de llegada de los dis-

9 Esta comparación es posible independientemente de la estructura de mercado. Los agentes informados y desinformados utilizarán distintas estrategias dependiendo de las reglas a las que se enfrenten, pero a efectos de este análisis, lo único relevante es su mayor o menor presencia en el mercado. La proporción informado/desinformado obtenida utilizando los valores de la mediana de cada cartera (CA, CM y CB) es (0.77, 1.20 y 1.58) para el NYSE y (0.99, 1.98 y 2.67) para el SIBE.

10 La técnica *bootstrap* o de remuestreo es ahora estándar en aplicaciones empíricas basadas en un número reducido de datos, para las que los resultados asintóticos convencionales podrían ser imprecisos. Efron y Tibshirani (1993) ofrecen una panorámica amplia sobre esta técnica. Este procedimiento permite, bajo condiciones completamente generales, aproximar la distribución exacta de un estadístico. Para ello, dada la muestra de referencia, la idea central consiste en remuestrear tomando un número elevado de submuestras del mismo tamaño que la original, bajo el principio de extracción con reemplazamiento. Sobre cada submuestra, se determina el estadístico de interés. La distribución empírica del estadístico, dado el total de submuestras, infiere la auténtica distribución.



---

tintos grupos de inversores, como la probabilidad de ocurrencia de eventos informativos difieren de manera significativa entre las distintas carteras de negociación. Por último, las diferencias en las estimaciones de  $\delta$  entre las distintas carteras no son significativas. El valor estimado de este parámetro revela una mayor probabilidad de llegada de eventos negativos, que coincide con la fase bajista en la que se situó el mercado durante el periodo analizado.

#### 4.2. MODELO DE RÉGIMEN CAMBIANTE

En el modelo de régimen cambiante, se toma como información de referencia la serie temporal intradiaria formada por los precios generados por todas y cada de una de las transacciones que tuvieron lugar durante el horizonte muestral considerado para cada activo. Los resultados básicos de la estimación del modelo de régimen cambiante se presentan en la Tabla 4 (ver panel A). Las estimaciones de la medida PIN coinciden en señalar un menor componente de asimetría para los activos más líquidos del mercado. La medida de probabilidad incondicional es del 12,7% para la cartera de mayor actividad, 16,5% para la cartera media, y 16,7% para los activos con menor frecuencia de negociación. Estas probabilidades son más reducidas que las evidenciadas por el modelo anterior. El modelo de variable latente revelaría, a través de las probabilidades de transición, que el estado normal es muy persistente (tiene una vida esperada de 10 transacciones), mientras que el estado excitado, vinculado a la llegada de información, tiene una vida esperada muy breve (sólo una transacción).

El valor medio de la revisión en el estado normal está comprendido entre 0,07 para el grupo de mayor actividad, y 0,08 para las otras dos carteras. La revisión en el estado excitado está comprendida entre 1,37 (activos más líquidos) y 1,31 (activos menos negociados). Aunque es tentador hacer una comparación directa de estos coeficientes a lo largo de las distintas carteras, ésta debería realizarse con cierta cautela. El mayor valor de la estimación sugiere que se lleva a cabo un mayor ajuste en los títulos más activos del mercado, pero este coeficiente debe ser aplicado en referencia a la horquilla media que es inferior para estos títulos (ver Tabla 1). Una mayor escala sobre un menor margen podría no suponer una mayor revisión en el precio. Tal y como resultaría previsible, el grupo de activos menos negociados, más volátil, presenta una mayor variabilidad a través del parámetro  $\sigma$  inferido.

Los paneles B y C de la Tabla 4 muestran los resultados de las pruebas de igualdad de parámetros entre las distintas carteras. La prueba de igualdad conjunta es rechazada para todos los parámetros exceptuando  $\alpha_1$  y  $\sigma$  (la igualdad de  $\alpha_0$  es rechazada marginalmente al 10%). Dado que los activos más ilíquidos son más volátiles, las diferencias observadas para  $\sigma$  parecen razonables. Las medidas PIN son distintas para cada cartera debido, fundamentalmente, a los distintos niveles de persistencia de los estados representados por las probabilidades de transición ( $P_{00}$  y  $P_{11}$ ). Estas estimaciones revelan un estado normal muy persistente, y un estado excitado con duración esperada inferior a una o dos transacciones. Por su parte, las pruebas dos a dos ponen de manifiesto que es la cartera de mayor actividad la que presenta un comportamiento claramente distinto al de las otras dos, no encontrando diferencias significativas entre las carteras de negociación media y baja.

Finalmente, Nyholm (2003) considera una extensión del modelo en el que se utiliza el tamaño de la transacción para condicionar la revisión mediante variables binarias. Este modelo también fue estimado, pero al igual que en Nyholm (2003), no se observa ninguna diferencia cuantitativa en la dinámica de la medida PIN respecto a la obtenida en el modelo simple. Dado que

**Tabla 4**  
**Modelo Nyholm (2002): Resumen de los Parámetros Estimados en cada Cartera.**

Esta tabla presenta la media, mediana y desviación estándar de los parámetros estimados siguiendo el modelo propuesto por Nyholm (2002) para 39 activos del mercado español divididos en 3 carteras de actividad. Los parámetros  $\alpha_0$  y  $\alpha_1$  representan la revisión que se produce en la horquilla de precios como consecuencia de la negociación de un agente desinformado e informado, respectivamente;  $P_{00}$  y  $P_{11}$  representa las probabilidades de transición del el proceso de Makov que caracteriza la identificación de agentes con distinta información. La  $PIN$  es la media de las probabilidades estimadas (*smoothed probabilities*) utilizando todas las observaciones para cada uno de los activos. El test no paramétrico de Kruskal-Wallis es utilizado para comprobar la hipótesis nula de que los valores de los parámetros estimados de las tres carteras de volumen pertenecen a poblaciones idénticas frente a la hipótesis alternativa de que al menos una de ellas genera valores observados más altos que el resto. El test no paramétrico de Mann-Whitney es utilizado para comprobar la hipótesis de que dos muestras proceden de poblaciones idénticas contra la alternativa de que una de ellas tiende a ofrecer valores mayores. Adicionalmente, para la medida de  $PIN$  se ha realizado un t-test de igualdad siguiendo la metodología Boostat (Panel D).

Panel A: Parámetros Estimados			
	CA	CM	CB
Activos Incluidos	13	13	13
$\alpha_0$			
Media	0.066	0.081	0.080
Mediana	0.066	0.081	0.085
Desv. Std.	0.008	0.017	0.034
$\alpha_1$			
Media	1.372	1.345	1.316
Mediana	1.384	1.317	1.259
Desv. Std.	0.069	0.096	0.174
$\sigma$			
Media	0.017	0.031	0.059
Mediana	0.014	0.021	0.026
Desv. Std.	0.011	0.020	0.097
$P_{00}$			
Media	0.904	0.877	0.885
Mediana	0.902	0.890	0.879
Desv. Std.	0.013	0.026	0.024
$P_{11}$			
Media	0.342	0.380	0.425
Mediana	0.336	0.387	0.428
Desv. Std.	0.036	0.048	0.058
$PIN$			
Media	0.127	0.165	0.167
Mediana	0.129	0.153	0.174
Desv. Std.	0.019	0.033	0.036

**Tabla 4 cont.**  
**Modelo Nyholm (2002): Resumen de los Parámetros Estimados**  
**en cada Cartera.**

<b>Panel B: Comparación conjunta (Test de Kruskal-Wallis)</b>						
	Test		[Prob]			
$\alpha_0$	5.381		[0.07]			
$\alpha_1$	3.140		[0.21]			
$\sigma$	3.945		[0.14]			
$P_{00}$	13.015		[0.00]			
$P_{11}$	9.903		[0.01]			
<i>PIN</i>	11.133		[0.00]			

<b>Panel C: Comparación dos a dos (Test de Mann-Whitney)</b>						
	CA vs CM	[Prob]	CA vs CB	[Prob]	CM vs CB	[Prob]
$\alpha_0$	2.564	[0.01]	1.333	[0.18]	0.078	[0.94]
$\alpha_1$	1.179	[0.24]	1.436	[0.11]	1.128	[0.26]
$\sigma$	2.077	[0.03]	1.230	[0.22]	0.102	[0.92]
$P_{00}$	2.872	[0.00]	2.410	[0.02]	0.564	[0.57]
$P_{11}$	2.128	[0.03]	3.333	[0.00]	1.846	[0.06]
<i>PIN</i>	3.077	[0.00]	2.667	[0.01]	0.154	[0.88]

<b>Panel D: Comparación PIN dos a dos (Bootstrap t-test)</b>						
	CA vs CM	[Prob]	CA vs CB	[Prob]	CM vs CB	[Prob]
<i>PIN</i>	-3.485	[0.00]	-3.411	[0.00]	-0.196	[0.47]

el objeto de este análisis es la comparación de las medidas PIN, los resultados de este modelo no se presentan, pero se encuentran disponibles bajo petición.

## 5. ANÁLISIS DE CONSISTENCIA

La aplicación de las dos metodologías muestra, a grandes rasgos, evidencia a favor de la existencia de una relación inversa entre el nivel de actividad y el grado de asimetría informativa, de manera que la probabilidad de negociación informada aumenta a medida que la actividad disminuye. Sin embargo, existen diferencias importantes en los resultados evidenciados. En primer lugar, las estimaciones promedio de la medida PIN difieren claramente entre los dos criterios, y, además, muestran patrones diferentes en las carteras. Esto podría constituir una primera evidencia de un comportamiento deficiente en alguna de las dos metodologías.

El hecho de que los dos procedimientos generen estimaciones diferentes no es de por sí concluyente. Al igual que en el caso de los modelos de descomposición de horquilla, los modelos de medida PIN resultarían sensibles a la especificación empleada.<sup>11</sup> Además, dado que ambos modelos se basan en aproximaciones diferentes, pueden observarse diferencias en las estimaciones puntuales. La diferencia en la pauta de probabilidad es visible al representar la distribución de probabilidad empírica de la medida PIN estimada para cada cartera (ver Figura 2). Por ejemplo, la distribución de probabilidad en el modelo EKOP en la cartera de menor negociación se sitúa a la derecha de las otras dos carteras (esto es, se observa una mayor tendencia a valores acumulados más elevados en la cartera de menor volumen). Sin embargo, en el modelo de régimen cambiante es la cartera de mayor actividad la que se sitúa de forma más evidente a la izquierda de las otras dos (con diversos cruces entre ambas).

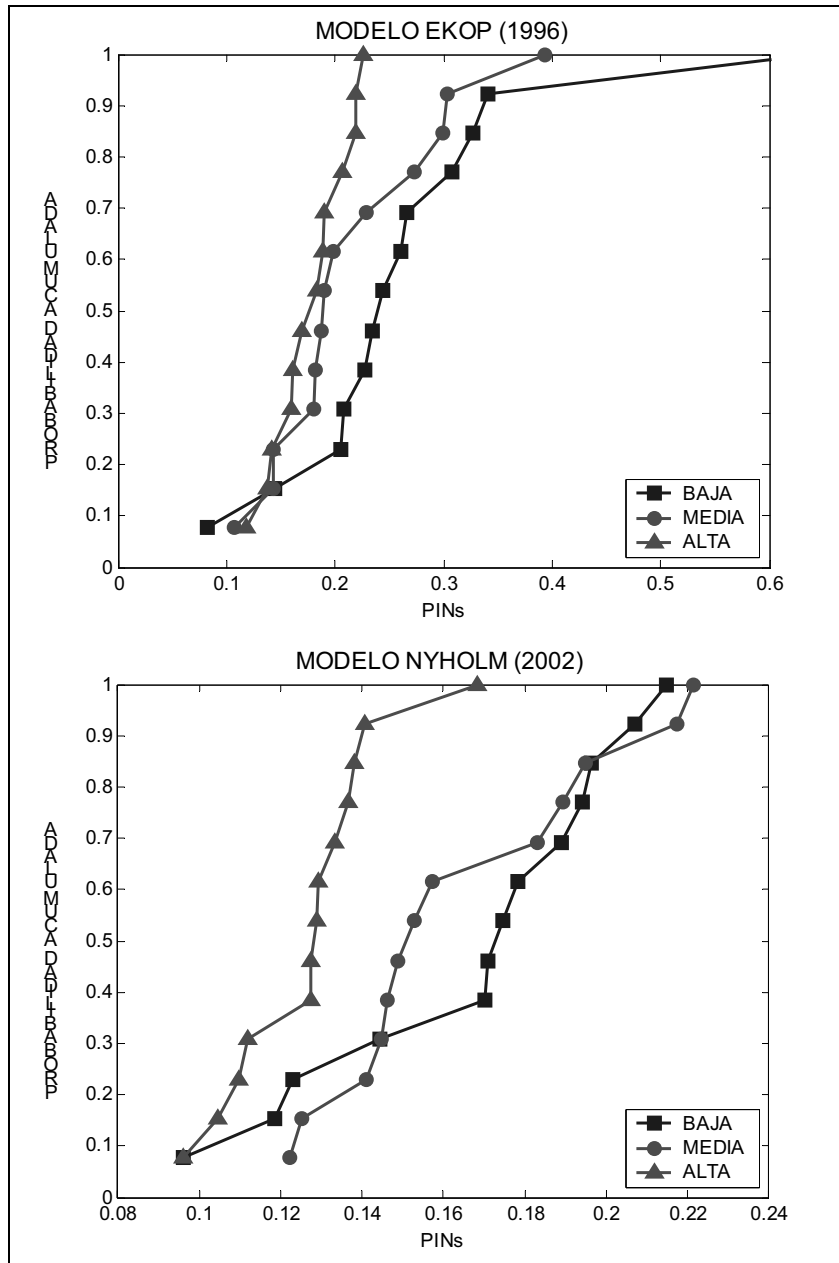
Resulta evidente que las estimaciones deberían mostrar patrones similares, dado que las metodologías de estimación pretenden capturar el mismo fenómeno. Por ello, las estimaciones deberían estar positivamente correlacionadas y tener poder predictivo sobre las variables *proxy* de información asimétrica -recuérdese que la llegada de información es inobservable y que, por tanto, la comparación con la auténtica medida no es posible.<sup>12</sup> Este tipo de análisis tiene la ventaja de ser muy sencillo y permite evaluar si los modelos reúnen ciertos requisitos indispensables. Un análisis basado en la misma filosofía es planteado en Van Ness *et al.* (2001).

La Tabla 5 muestra la matriz de correlaciones en sección cruzada entre las probabilidades estimadas bajo las dos metodologías y la horquilla cotizada en términos absolutos y relativos.<sup>13</sup> El coeficiente de correlación entre las dos medidas de probabilidad estimadas es negativo y significativo (-20%). Este resultado sugiere que las estimaciones resultantes de los modelos no sólo no obedecen a las mismas variaciones, sino que de hecho estarían midiendo dos fenómenos distintos. La ausencia de una relación clara entre estas dos variables se observa claramente en la Figura 3, donde se representan las probabilidades en sección cruzada, activo a activo. Puede observarse como las variaciones en las medidas PIN son en muchas ocasiones de signo opuesto. Sí aceptásemos a priori que la medida EKOP es capaz de aproximar correctamente la llegada de información privada, como se acepta ampliamente desde la literatura de microestructura, el modelo de régimen cambiante no generaría estimaciones con buenas propiedades.<sup>14</sup>

- 11 Por ejemplo, Van Ness *et al.* (2001) muestran en su estudio comparativo estimaciones muy diferentes sobre cinco modelos aplicados al mismo conjunto de datos.
- 12 Por supuesto, esto no impide el análisis comparativo entre distintos modelos, siempre y cuando se considere alguna variable *proxy* de la variable inobservable. Un caso completamente similar en la literatura Financiera ocurre en la vasta literatura de predicción de volatilidad: la imposibilidad de observar la auténtica volatilidad del rendimiento no impide realizar comparaciones entre distintos modelos, o estudiar el comportamiento de un modelo, cuando se considera algún *proxy* de la auténtica volatilidad (habitualmente el cuadrado o el valor absoluto del término de innovación en el rendimiento).
- 13 El tamaño de la horquilla es el *proxy* de información asimétrica más intuitivo y directo, y así se ha aplicado en numerosos estudios. En primer lugar, el propio planteamiento que subyace en el trasfondo común de ambos modelos implica una relación positiva y significativa entre horquilla y grado de asimetría informativa, y, por tanto, entre horquilla y la medida de probabilidad estimada. Esta relación es refrendada y aceptada desde los modelos teóricos (entre otros, Copeland y Galai, 1983; Glosten y Milgrom, 1985).
- 14 Resultaría ingenuo pensar que un planteamiento con el nivel de abstracción del modelo EKOP pueda generar medidas exactas de la auténtica probabilidad de información negociada, pero la literatura de microestructura ha aceptado que esta medida es capaz, al menos, de aproximar la llegada de información privada razonablemente bien. No sólo muestra poder predictivo sobre la horquilla y otros *proxies* de información, sino que en su calidad de medida de iliquidez ha sido utilizada para aplicaciones más avanzadas. Por ejemplo, Easley *et al.* (2002) encuentran que el mercado retribuye el riesgo de iliquidez cuando este factor de riesgo se aproxima mediante la medida EKOP, lo que no ocurre, por ejemplo, si se utiliza la variable volumen.

**Figura 2**

La siguiente Figura muestra las funciones de distribución empíricas correspondientes a las probabilidades de negociación informada (PINs) estimadas para cada uno de los activos de las 3 carteras de actividad, siguiendo las metodologías propuestas por EKOP (1996) y Nyholm (2002) [probabilidades incondicionales].



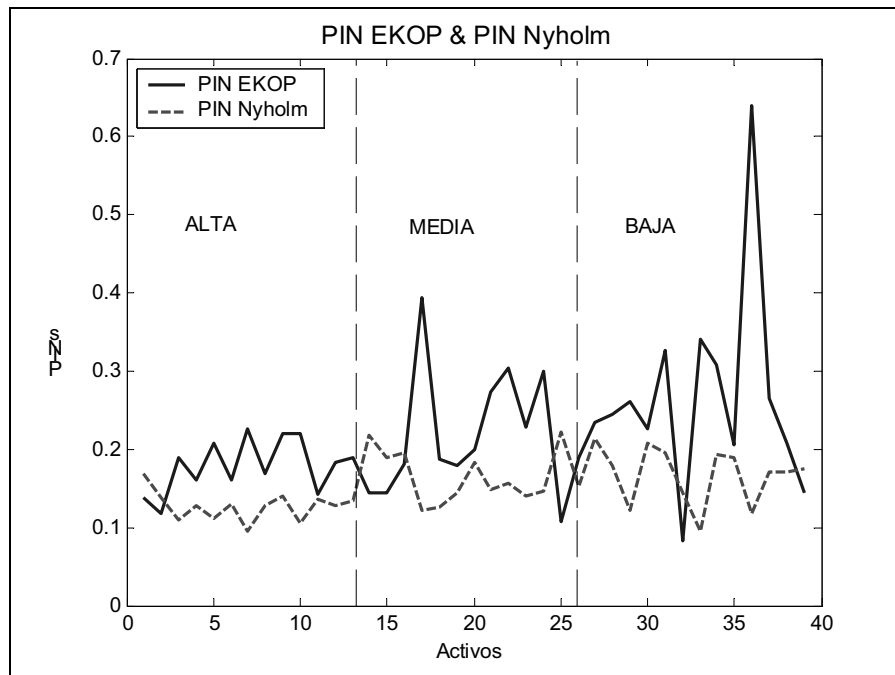
**Tabla 5**  
**Modelo Nyholm (2002): Resumen de los Parámetros Estimados en cada Cartera.**

La presente Tabla muestra la matriz de correlación (sección cruzada) entre las Probabilidades de Negociación Informada (PINs) estimadas con la metodología propuesta por EKOP (1996) y Nyholm (2002) y las medidas de horquilla en términos absolutos y relativos (HCA y HCR).

	<i>HCA</i>	<i>HCR</i>	<i>PIN<sub>EKOP</sub></i>	<i>PIN<sub>Nyholm</sub></i>
<i>HCA</i>	1	0.470	0.118	-0.101
<i>HCR</i>	0.470	1	0.427	0.258
<i>PIN<sub>EKOP</sub></i>	0.118	0.427	1	-0.205
<i>PIN<sub>Nyholm</sub></i>	-0.101	0.258	-0.205	1

**Figura 3**

La siguiente Figura representa las estimaciones de la probabilidad de negociación informada (PIN) para cada uno de los 39 incluidos en nuestra muestra. La línea azul corresponde a las estimaciones siguiendo la metodología propuesta por EKOP (1996), mientras que la línea verde corresponde a las estimaciones siguiendo el trabajo propuesto por Nyholm (2002).



El comportamiento poco satisfactorio del modelo de régimen cambiante se observa además al comparar la correlación con los *proxies* de información. La medida EKOP ofrece unos resultados acorde con lo esperado, con una correlación del 43% con la horquilla relativa y del 12% con la horquilla absoluta. Los resultados del modelo de régimen cambiante son, en el mejor de los casos, mucho más modestos: la correlación con la horquilla relativa se reduce al 26%, mientras que la correlación con la horquilla en términos absolutos es negativa, alrededor del -10%. Al margen del comportamiento con respecto al modelo EKOP, este resultado anómalo no es con-

**Tabla 6**  
**PIN EKOP: Análisis de Regresión.**

Esta tabla presenta los resultados de la estimación de las regresiones lineales dadas por:

$$HCA = \beta_0 + \beta_1 * PMC * PIN + \beta_2 * VOL + \eta$$

$$HCR = \beta_0 + \beta_1 * PIN + \beta_2 * VOL + \eta$$

La variable dependiente en la primera regresión (Panel A) es la horquilla cotizada absoluta calculada como media diaria ponderada por el tiempo. Como variables explicativas se incluye el PIN estimado multiplicado o corregido por el precio del activo medido con el punto medio. Vol representa el volumen. En la segunda regresión (Panel B) la variable dependiente es la horquilla relativa, esto es, corregida por el punto medio, por lo que las variables explicativas quedan representadas por el PIN y VOL.

<b>Panel A: Resultados con HCA</b>			
	Modelo General	Restricción $\beta_1 = 0$	Restricción $\beta_2 = 0$
Constante	-0.017 (-0.46)	0.171 (3.98)	-0.049 (-1.39)
PMC*PIN	0.062 (7.40)	-	0.062 (7.04)
VOL	-2.21E-8 (-2.09)	-2.01E-8 (-1.20)	-
Adj.R <sup>2</sup>	0.596	0.012	0.561
F-Value	29.205	1.475	49.549

<b>Panel A: Resultados con HCR</b>			
	Modelo General	Restricción $\beta_1 = 0$	Restricción $\beta_2 = 0$
Constante	0.009 (3.74)	0.013 (12.81)	0.004 (1.61)
PIN	0.018 (2.09)	-	0.029 (2.87)
VOL	-1.66E-9 (-4.15)	-1.90E-9 (-4.75)	-
Adj.R <sup>2</sup>	0.415	0.362	0.1598
F-Value	14.510	22.564	8.233

sistente con la premisa de que el modelo de régimen cambiante aproxime la llegada de información privada y no puede explicarse en los términos económicos que fundamentan el modelo. Nótese, en contraste, que el análisis de Van Ness *et al.* (2001) muestra que todos los modelos de horquilla analizados están positiva y fuertemente correlacionados entre sí, y correlacionados además con el tamaño de la horquilla.

Para analizar el poder predictivo de las estimaciones sobre las variables *proxy* de asimetría informativa consideramos, en primer lugar y siguiendo a Easley *et al.* (1996), un modelo de regresión lineal donde la medida PIN estimada para cada activo multiplicada por el nivel de precios ( $PIN_i * PMC_i$ ) se utiliza para explicar la horquilla cotizada en términos absolutos,  $HCA_i$ . Como variable de control se considera el volumen de negociación en términos monetarios,  $VOL_i$ :

$$HCA_i = \beta_0 + \beta_1 (PIN_i * PMC_i) + \beta_2 VOL_i + \eta_i \quad (11)$$

**Tabla 7**  
PIN Nyholm: Análisis de Regresión.

Esta tabla presenta los resultados de la estimación de las dos regresiones lineales dadas por:

$$HCA = \beta_0 + \beta_1 * PMC * PIN + \beta_2 * VOL + \eta$$

$$HCR = \beta_0 + \beta_1 * PIN + \beta_2 * VOL + \eta$$

La variable dependiente en la primera regresión (Panel A) es la horquilla cotizada absoluta calculada como media diaria ponderada por el tiempo. Como variables explicativas se incluye PIN estimado multiplicado o corregido por el precio del activo medido con el punto medio. Vol representa el volumen. En la segunda regresión (Panel B) la variable dependiente es la horquilla relativa, esto es, corregida por el punto medio, por lo que las variables explicativas quedan representadas por el PIN y VOL.

<b>Panel A: Resultados con HCA</b>			
	Modelo General	Restricción $\beta_1 = 0$	Restricción $\beta_2 = 0$
Constante	0.002 (0.065)	0.171 (3.98)	-0.030 (-0.85)
PMC*PIN	0.084 (6.85)	-	0.083 (6.48)
VOL	-2.31E-8 (-2.09)	-2.01E-8 (-1.20)	-
Adj.R <sup>2</sup>	0.559	0.012	0.519
F-Value	25.122	1.475	42.847

<b>Panel A: Resultados con HCR</b>			
	Modelo General	Restricción $\beta_1 = 0$	Restricción $\beta_2 = 0$
Constante	0.014 (3.05)	0.013 (12.81)	0.003 (0.64)
PIN	-0.006 (-0.22)	-	0.048 (1.63)
VOL	-1.94E-9 (-4.26)	-1.90E-9 (-4.75)	-
Adj.R <sup>2</sup>	0.345	0.362	0.042
F-Value	11.012	22.564	2.675

En segundo lugar, consideramos un modelo de regresión donde la variable a explicar es la horquilla relativa,  $HCR_i$ . Este análisis proporciona mayor grado de robustez, dado que se ha argumentado que la variable  $(PIN_i * PMC_i)$  en (11) podría tener poder explicativo simplemente porque los activos con mayores horquillas absolutas poseen también mayores precios. En este caso, el modelo de regresión viene dado por:

$$HCR_i = \beta_0 + \beta_1 PIN_i + \beta_2 VOL_i + \eta_i \tag{12}$$

Los modelos (11) y (12) se estiman por mínimos cuadrados, considerando también las versiones restringidas  $\beta_1 = 0$  y  $\beta_2 = 0$ . Los resultados de estas estimaciones se presentan en las Tablas 6 (modelo EKOP) y 7 (modelo de régimen cambiante). Las estimaciones del modelo EKOP tienen mayor poder predictivo medido a través del coeficiente de determinación. El ajuste es razonablemente bueno, pero presenta grandes diferencias entre los dos modelos de regre-



---

sión, atribuibles al sesgo originado por el nivel de precios. En el caso del modelo EKOP (Nyholm), el ajuste es del 60% (56%) en el modelo de regresión (11) al considerar  $\beta_2 = 0$ , mientras que la capacidad explicativa se reduce al 16% (5%) en la regresión (12). Mientras que en el modelo EKOP la estimación de  $\beta_1$  es siempre positiva y significativa, en el modelo de Nyholm la estimación no es significativa en (12) si se incluye el volumen como variable explicativa, y marginalmente significativamente en caso contrario. La ausencia de poder explicativo es muy desafortunada, y contradice además la relación intuitiva implicada en el planteamiento del propio modelo.

## 6. CONCLUSIONES

En este trabajo se ha pretendido, por un lado, establecer una primera comparación de las metodologías existentes para la caracterización del grado de asimetría informativa a través de la probabilidad de información negociada y, por otro, presentar evidencia específica sobre este fenómeno en el mercado español. El análisis realizado es importante para la investigación en el área de asimetría informativa en éste y otros mercados.

La evidencia obtenida es suficiente para concluir que el modelo de régimen cambiante no proporciona estimaciones claramente consistentes con los efectos observables de la llegada de información privada. El principal problema de este procedimiento es la carencia de un fundamento teórico y econométrico sólido que garantice la correspondencia entre la variable latente y la llegada de información. El planteamiento del modelo remarca la existencia de comportamientos no-lineales en el proceso de revisión inmediata de precios, de manera que es posible hablar de cambios "grandes" y "pequeños" en función de una cierta variable. Sin embargo, estas variaciones no tienen por qué corresponderse con cambios permanentes y transitorios en el precio, que es lo que permitiría detectar la presencia de negociación informada. En otras palabras, el ajuste inmediato en el precio es un fenómeno tan complejo e influenciado por tantas variables y la formulación econométrica tan ingenua, que nada impide que los cambios observados tengan su origen en alguna variable relacionada con el entorno de mercado o con la evolución del precio y que sea capaz de generar respuestas asimétricas. De ser revelación de información, el análisis intuitivo realizado habría bastado para identificar una serie de propiedades estadísticas básicas inherentes al planteamiento del modelo, que sin embargo no son verificadas. El procedimiento de régimen cambiante, por tanto, no parece ser una metodología adecuada para medir asimetrías informativas, al menos en su formulación original. Analizar si sería posible extender el modelo para superar estas deficiencias se analiza en un trabajo paralelo (Abad y Rubia, 2004).

Resulta interesante comentar aquí brevemente algunos de los resultados previos obtenidos en ese estudio, pues ello arroja luz sobre las causas que originan el comportamiento deficiente del modelo de régimen cambiante. En ausencia de restricciones, la variable latente que se infiere a través de la estimación tiende a identificarse con la tipología de órdenes (más o menos agresivas) que subyace en las transacciones. Es sabido que órdenes con diferentes niveles de agresividad generan variaciones característicamente diferentes en el medio punto. La probabilidad de estado excitado tiende así a asociarse así con la probabilidad de llegada de órdenes de mercado que sistemáticamente generan mayores variaciones inmediatas en el medio punto, al margen del posible contenido informado de dichas órdenes. En otras palabras, es la arquitectura de las órdenes, y no su contenido informativo, lo que parece inferir el modelo de Nyholm. Este comporta-

miento persiste incluso si se controla por el efecto de otras variables de mercado, como el tamaño de las órdenes, por lo que resulta dudoso que este modelo pueda capturar correctamente el fenómeno para el que inicialmente está diseñado.

Por otro lado, la evidencia empírica obtenida para el mercado español permite ganar una perspectiva más amplia acerca del comportamiento de los inversores, lo que resulta importante tanto para la futura investigación centrada en este mercado, como para los distintos agentes que participan en el mismo. El modelo de Easley *et al.* (1996) revela la existencia de un fuerte comportamiento de aversión al riesgo de selección adversa que aumenta enormemente a medida que los activos son menos negociados. Existe una preferencia inequívoca de los agentes desinformados por negociar sobre los activos de mayor actividad. Este hecho tiene implicaciones claras para los agentes que ofertan liquidez en el mercado, pero también para los reguladores y administración. Así, este comportamiento plantea cuestiones acerca de la capacidad efectiva de distintas reformas que se han centrado precisamente en el grupo de activos menos negociados. Por ejemplo, desde 1998 opera en paralelo a la contratación continua el sistema *fixing*, que pretende mejorar la liquidez de los títulos menos negociados (ver Abad y Rubia, 2003). Dado el fuerte componente de aversión al riesgo evidenciado en este trabajo, parece importante para los agentes y los diseñadores de mercado el comprobar empíricamente si el sistema ha sido capaz de mejorar la liquidez a través de una reducción de la asimetría informativa. Esta cuestión constituye un interesante tema de investigación futura.

## REFERENCIAS

- Abad, D. y A. Rubia, 2003. Actividad, Volatilidad y Rendimientos en el Sistema Fixing Español: Una Evaluación Empírica, *Moneda y Crédito*, 217, 45-72.
- Abad, D. y A. Rubia, 2004. Estimating the Probability of Informed Trading: Further Evidence from an Order-Driven Market, *Papel de Trabajo*, Universidad de Alicante.
- Acosta, J., M.P. Osorno y M.G. Rodríguez, 2000. Los Costes de Selección Adversa en el Mercado Bursátil Español, VIII Foro de Finanzas, Madrid.
- Ahn, H.J., J. Cai, Y. Hamao y R.Y.K. Ho, 2002. The Components of the Bid-Ask Spread in a Limit Order Market: Evidence from the Tokyo Stock Exchange, *Journal of Empirical Finance*, 9, 399-430.
- Brockman, P. y D.Y.Chung, 2000a. Informed and Uninformed Trading in an Electronic, Order-Driven Environment, *Financial Review*, 35, 125-146.
- Brockman, P. y D.Y.Chung, 2000b. An Empirical Investigation of Trading on Asymmetric Information and Heterogeneous prior Beliefs, *Journal of Empirical Finance*, 7, 417-454.
- Brockman, P. y D.Y.Chung, 2002. The Impact of Informed Trading on Corporate Liquidity, *Journal of Multinational Financial Management*, 12, 239-259.
- Choi, J., D. Salandro y K. Shastri, 1988. On the Estimation of Bid-Ask Spreads: Theory and Evidence, *Journal of Financial and Quantitative Analysis*, 23, 219-230.
- Chung, K.H. y M. Li, 2003. Adverse Selection Costs and the Probability of Information-Based Trading, *Financial Review*, 38, 257-272.

- 
- Copeland, T. y D. Galai, 1983. Information Effects on the Bid-ask Spreads, *Journal of Finance*, 38, 1457-1469.
  - Easley, D., N. Kiefer, M. O'Hara y J. Paperman, 1996. Liquidity, Information, and Infrequently Traded Stocks, *Journal of Finance*, 51, 1405-1436.
  - Easley, D., N. Kiefer y M. O'Hara, 1997a. One Day in the Life of a very Common Stock, *Review of Financial Studies*, 10, 805-835.
  - Easley, D., N. Kiefer y M. O'Hara, 1997b. The Information Content of the Trading Process, *Journal of Empirical Finance*, 4, 159-186.
  - Easley, D., M. O'Hara y J. Paperman, 1998. Financial Analysts and Information-Based Trade, *Journal of Financial Markets*, 1, 175-201.
  - Easley, D., M. O'Hara y P.S. Srinivas, 1998. Option Volume and Stock Prices: Evidence on Where Informed Traders Trade, *Journal of Finance*, 53, 431-465.
  - Easley, D., S. Hvidkjaer y M. O'Hara, 2002. Is Information Risk a Determinant of Asset Returns?, *Journal of Finance*, 62, 2185-2221.
  - Efron, B. y R. Tibshirani, 1993. *An Introduction to the Bootstrap*. Chapman-Hall.
  - Farinós, J.E., C.J. García y A.M. Ibáñez, 1999. Efectos del Anuncio de una OPA en la Liquidez y en la Actividad Negociadora de las Empresas Objetivo, *Revista Europea de Dirección y Economía de la Empresa*, 9, 13-27.
  - George, T., G. Kaul y M. Nimalendran, 1991. Estimation of the Bid-Ask Spread and its Components: A new Approach, *Review of Financial Studies*, 4, 623-656.
  - Gómez, J.C. y J. Yagüe, 2002. Tamaño de Transacciones, Introducción de Órdenes y Preferencias por Precios en los Splits de Acciones, Papeles de Trabajo del Instituto Valenciano de Investigaciones Económicas (IVIE), WP-EC 2002-29.
  - Glosten, L.R. y L.E. Harris, 1998. Estimating the Components of the Bid-Ask Spread, *Journal of Financial Economics*, 21, 123-142.
  - Glosten, L.R. y P.R. Milgrom, 1985. Bid, Ask and Transaction Prices in a Specialist Market with Heterogenously Informed Traders, *Journal of Financial Economics*, 14, 71-100.
  - Gramming, J., D. Schiereck y E. Theissen, 2001. Knowing me, Knowing you: Trader Anonymity and Informed Trading in Parallel Markets, *Journal of Financial Markets*, 4, 385-412.
  - Hamilton, J., 1994. *Time Series Analysis*. Princeton University Press: New Jersey.
  - Handa, P., R. Schwartz y A. Tiwari, 2003. Quote Setting and Price Formation in an Order Driven Market, *Journal of Financial Markets*, 6, 461-489.
  - Hasbrouck, J., 1991a. Measuring the Information Content of Stock Trades, *Journal of Finance*, 46, 179-207.

- Hasbrouck, J., 1991b. The Summary Informativeness of Stock Trades: An Econometric Analysis, *Review of Financial Studies*, 4, 571-595.
- Huang, R.D. y H.R. Stoll, 1994. Market Microstructure and Stock Return Predictions, *Review of Financial Studies*, 7, 179-213.
- Huang, R.D. y H.R. Stoll, 1997. The Components of the Bid-Ask Spread: A General Approach, *Review of Financial Studies*, 10, 995-1034.
- Kim, C.H. y C.R. Nelson, 1998. *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. MIT Press.
- Lin, J., G.C. Sanger y G.G. Booth, 1995. Trade Size and Components of the Bid-Ask Spread, *Review of Financial Studies*, 8, 1153-1183.
- Madhavan, A. y S. Smidt, 1991. A Bayesian Model of Intraday Specialist Pricing, *Journal of Financial Economics*, 30, 99-134.
- Madhavan, A., M. Richardson y M. Roomans, 1997. Why Do Security Prices Change? A Transaction-Level Analysis of NYSE Stocks, *Review of Financial Studies*, 10, 1035-1064.
- Minguez-Vera, A. y J.F. Martin-Ugedo, 2003. Concentración Accionarial y Liquidez del Mercado: Un Análisis con Ecuaciones Simultáneas, *Papeles de Trabajo del Instituto Valenciano de Investigaciones Económicas (IVIE)*, WP-EC 2003-20.
- Nyholm, K., 2002. Estimating the Probability of Informed Trading, *Journal of Financial Research*, 35, 485-505.
- Nyholm, K., 2003. Inferring the Private Information Content of Trades: a Regime-Switching Approach, *Journal of Applied Econometrics*, 18, 457-470.
- Roll, R., 1984. A Simply Implicit Measure of the Effective Bid-Ask Spread in a Efficient Market, *Journal of Finance*, 39, 1127-1139.
- Stoll, H., 1989. Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests, *Journal of Finance*, 44, 115-134.
- Theissen, E., 2001. A Test of the Accuracy of the Lee-Ready Trade Classification Algorithm, *Journal of International Financial Markets, Institutions and Money*, 2, 147-165.
- Van Ness, B.F., R.A. Van Ness y R.S. Warr, 2001. How Well do Adverse Selection Components Measure Adverse Selection?, *Financial Management*, 30, 77-98.