



Fac. CC. Económicas y Empresariales
Universidad de La Laguna



Fac. CC. Económicas y Empresariales
Univ. de Las Palmas de Gran Canaria

Intensidad de la actividad de negociación. El caso del futuro del IBEX 35

Jorge V. Pérez Rodríguez*

DOCUMENTO DE TRABAJO 2002-07

* Universidad de Las Palmas de Gran Canaria. Departamento de Métodos Cuantitativos en Economía y Gestión.

Intensidad de la actividad de negociación. El caso del futuro del Ibex 35

Jorge V. Pérez-Rodríguez
Universidad de Las Palmas de Gran Canaria

Resumen:

En este artículo se estima la duración condicional o intensidad esperada de la actividad de negociación del futuro del Ibex35 negociado en intervalos de tiempo irregularmente espaciados mediante los modelos ACD, teniendo en cuenta la periodicidad intradiaria y diversas funciones de pseudo verosimilitud para la densidad del tiempo de llegada de la nueva información en el cálculo de dichas estimaciones. Además, se modeliza conjuntamente el cambio en los precios y la rentabilidad de negociación del futuro mediante modelos del tipo ISAR-GARCH para datos irregularmente espaciados, distinguiendo entre el contenido informativo y no-informativo del tiempo esperado de la llegada de las transacciones y el volumen de negociación. Los resultados muestran que la duración posee contenido informativo para explicar la intensidad esperada de la actividad de negociación, así como una correlación positiva entre la volatilidad condicional de la rentabilidad y el volumen intradiario, y una correlación negativa con la intensidad esperada, ambas significativas, indicando el contenido informativo del tiempo y el volumen de negociación.

Palabras Clave : Microestructura, modelos ACD, Weibull, Gamma Generalizada, ISAR-GARCH.

JEL: C22, C45, C52.

1. Introducción

Desde la perspectiva del análisis de la microestructura del mercado de activos, los movimientos en los precios están causados por la llegada de nueva información, así como por el proceso que incorpora tal información dentro de los precios de mercado. El aumento del flujo y la velocidad de la información en los mercados financieros ha posibilitado, por un lado, que los agentes respondan casi automáticamente a la llegada de la nueva información, pública o privada, de forma simultánea y en todos los mercados; que, además, puedan existir diferencias de opinión entre los agentes sobre la información que sea común; o que los *dealers* ajusten sus posiciones dinámicamente para controlar sus inventarios¹; y, por otro lado, también ha permitido que los agentes inversores obtengan importantes beneficios económicos².

Un aspecto importante en el trabajo empírico de los mercados dirigidos por órdenes es la importancia relativa de las piezas individuales de información que caracterizan la actividad de negociación y las transacciones. Muchos son los trabajos que han analizado estos hechos, no solo modelizando la duración de las transacciones (de acuerdo a diferentes funciones de densidad o diferentes modelos para la duración esperada), como por ejemplo, Engle y Russell (1997, 1998), Zhang, Russell y Tsay (2000), Bauwens y Giot (2000) o Grammig y Maurer (2000), sino que también modelizando la interdependencia de la volatilidad y la duración intradía [véase Grammig y Wellner (2002)].

El propósito de este trabajo es contribuir al análisis de la dinámica de los precios de mercado en España, analizando el futuro del Ibex35 y utilizando datos de transacciones intradiarias que se producen en intervalos irregulares de tiempo. Para ello, la investigación realizada posee dos objetivos. En primer lugar, comparar diferentes funciones de pseudo verosimilitud en la que la densidad subyacente de los tiempos de

¹ No obstante, esta taxonomía puede ampliarse simplemente considerando el potencial *feedback* entre los anteriores factores. Por ejemplo, algunas noticias pueden utilizarse para alcanzar un particular impacto en la volatilidad o los *dealers* pueden ajustar sus inventarios cuando anticipan su propia desventaja relativa frente a los agentes informados.

² En cuanto a los mercados de futuros, podemos decir que la existencia de reducidos costes de transacción han hecho más atrayente la negociación en los mercados de futuros que la realizada para el activo subyacente. Por ello, tales oportunidades de negociación más baratas y flexibles atraen a nuevos y diferentes inversores informados al mercado, con el consiguiente incremento de la actividad de negociación. Asimismo, tal incremento, junto con el arbitraje entre el activo subyacente y el futuro, implican un incremento de la eficiencia del mercado.

las llegadas de las noticias (es decir, tiempos de negociación en los que se produce cada transacción), se supone que se distribuye de forma Exponencial, Weibull o Gamma Generalizada, las cuales se utilizan para computar las estimaciones cuasi máximo verosímiles de los parámetros. A su vez, estas funciones permitirán analizar empíricamente la dinámica de la duración condicional entre dos transacciones consecutivas y modelar el proceso estocástico que la caracteriza a través de los modelos de duración condicional autorregresiva (ACD), propuestos por Engle y Russel (1997, 1998).

En segundo lugar, se propone un modelo conjunto que considera como la distribución conjunta de los tiempos de llegada de las transacciones y el cambio de los precios (de una transacción a otra) está asociado a ésta llegada. En concreto, se modelizan, tanto al rentabilidad como la volatilidad de tales cambios, a través de procesos de tipo AR y GARCH para datos irregularmente espaciados, incluyendo como variables exógenas de la volatilidad condicional a la duración esperada de las operaciones y al volumen de negociación. La finalidad de esta inclusión es analizar si estas variables predicen elevadas o bajas volatilidades en el futuro, en la idea de que éstas afecten la distribución condicional del cambio en los precios.

El trabajo se organiza como sigue. En la sección 2, se consideran algunos modelos alternativos para la duración condicional en términos de la funciones de densidad de las llegadas. En la sección 3, se realiza un análisis empírico, considerando tres aspectos: la descripción de los datos de duración, la estimación de la intensidad condicional de la actividad de negociación y el estudio de la relación entre volatilidad condicional de los precios, la duración esperada de las transacciones y el volumen de negociación. Finalmente, en la sección 4 se establecen las conclusiones más relevantes.

2. Modelización de la intensidad de la actividad de negociación

2.1. El modelo ACD

Una de las más importantes características de los datos de alta frecuencia es el espaciamiento irregular. La llegada de la información se produce en intervalos irregulares de tiempo durante el día. En el análisis econométrico estándar el intervalo de

tiempo es fijo, y por lo tanto, la variable “tiempo” está generalmente omitida del análisis.

Existen muchas maneras de modelizar la duración condicional o intensidad condicional de las transacciones [véase Engle (1996) o Ghysels, Gouriéroux y Jasiak (1998)]. Sin embargo, usaremos el modelo de duración condicional autorregresivo (ACD) propuesto por Engle y Russell (1997, 1998). Este modelo permite una especificación precisa de la dinámica autónoma de la duración, y sirve para predecir la intensidad de la actividad de negociación. Además, dicha predicción puede ser relevante en la explicación de la volatilidad de los precios³. Engle y Russell (1998) desarrollan un método que está directamente relacionado con el espaciado irregular de los datos. Su inspiración es la teoría del “*point process*”, y extienden el análisis estadístico de los datos de tiempo de fallo acelerado de Kalbfleisch y Prentice (1980). En su modelo, se relaciona el tiempo de llegada de información de acuerdo con alguna ley de probabilidad, entendiéndose que dicho proceso es un tipo de proceso dependiente de Poisson.

Supongamos que el tiempo de llegada de nueva información o momento en que ocurre la transacción i -ésima, t_i , es una variable aleatoria que sigue un “*point process*”, y que tiene asociadas algunas variables aleatorias denominadas “*marks*”, que denominaremos por y_i . En este sentido, podría escribirse: $\{(t_i, y_i), i = 1, 2, \dots, N\}$, siendo N el número de *ticks*, puntos temporales o tamaño muestral de datos irregularmente espaciados. La i -ésima observación tiene una densidad condicional conjunta para ambas variables aleatorias condicionadas a la información disponible en el periodo anterior, tal que: $(t_i, y_i) / \Psi_{i-1} \sim f(t_i, y_i / t^{(i-1)}, y^{(i-1)}; \mathbf{q})$, donde $t^{(i-1)} = (t_{i-1}, t_{i-2}, \dots, t_1)$ y $y^{(i-1)} = (y_{i-1}, y_{i-2}, \dots, y_1)$; Ψ_{i-1} representa el conjunto de información disponible hasta el momento t_{i-1} ; y \mathbf{q} es un vector de parámetros desconocidos.

La función f puede factorizarse como un producto entre la función de densidad condicional de los “*marks*” y la distribución marginal de los tiempos de llegada, tal que:

$$f(t_i, y_i / t^{(i-1)}, y^{(i-1)}; \mathbf{q}) = p(y_i / t^{(i)}, y^{(i-1)}; \mathbf{q}_1) q(t_i / t^{(i-1)}, y^{(i-1)}; \mathbf{q}_2) \quad [1]$$

siendo \mathbf{q}_1 y \mathbf{q}_2 dos vectores distintos de parámetros desconocidos. Y la función del logaritmo de verosimilitud resultante es:

$$\log \prod_{i=1}^N f(\cdot) = \sum_{i=1}^N \log f(\cdot) = \sum_{i=1}^N [\log p(y_i/t^{(i)}, y^{(i-1)}; \mathbf{q}_1) + \log q(t_i/t^{(i-1)}, y^{(i-1)}; \mathbf{q}_2)]$$

Atendiendo a la función $q(\cdot)$, y si deseamos analizar la duración entre dos transacciones, resultará conveniente escribir el argumento de dicha función en términos de la misma. Así, considerando $d_i = t_i - t_{i-1}$ como la duración entre dos transacciones consecutivas, puede definirse $q(\cdot)$ mediante la densidad del tiempo de espera entre t_{i-1} y t_i condicionada a y y d . Con ello, Engle y Russell (1997) suponen que existe una dependencia autorregresiva en la densidad condicional de las duraciones, y de esta forma, puede escribirse que la distribución marginal del tiempo de llegadas es: $q(d_i/d^{(i-1)}, y^{(i-1)}; \mathbf{q}_2) = q(d_i/d^{(i-1)}; \mathbf{q}_2)$.

Pues bien, atendiendo a este supuesto, tales autores proponen una especificación para $q(\cdot)$ que denominan modelo de duración condicional autorregresivo (ACD), el cual puede fácilmente formularse en términos de la duración condicional esperada o duración media condicionada de la i -ésima negociación, esto es: $\mathbf{y}_i = E[d_i/d_{i-1}, \dots, d_1; \mathbf{q}_2]$, y, además, por el supuesto básico de que la duración normalizada es independiente e idénticamente distribuida para todo i , es decir, $\frac{d_i}{\mathbf{y}_i} \equiv \mathbf{e}_i \sim iid \mathfrak{D}(\mathbf{q}_2)$, siendo \mathfrak{D} una distribución general definida en el intervalo $(0, \infty)$ con media igual a 1 y vector de parámetros \mathbf{q}_2 .

Bajo el supuesto de las duraciones son débilmente exógenas para los parámetros de interés, por ejemplo \mathbf{q}_2 , no se requiere de una distribución conjunta para $f(\cdot)$, y los parámetros pueden estimarse maximizando el segundo término.

³ Engle y Russell (1997) encuentran que la información, en la forma de duraciones entre eventos del mercado, también ayuda a explicar la evolución de la conducta de las transacciones y de sus precios.

2.2. Extensiones del modelo ACD y diversas funciones de densidad para la llegada de las transacciones

En este epígrafe, y para computar las estimaciones de los parámetros del modelo ACD, consideramos diferentes funciones del logaritmo de pseudo verosimilitud en el que se supone que la densidad de las llegadas de las transacciones es Exponencial, Weibull o Gamma Generalizada. A continuación, describiremos su forma paramétrica y las funciones de pseudo verosimilitud para estimar los parámetros.

El modelo ACD supone que el término de error es multiplicativo y toda la información pasada afecta a la actual a través de la duración media condicional, \mathbf{y}_i . En este sentido, la flexibilidad para caracterizar \mathcal{D} y \mathbf{y}_i proporciona una mayor riqueza al modelo. En nuestro caso, la expresión que utilizamos para \mathbf{y}_i , considerando un proceso lineal generalizado ACD(p,q), es igual a:

$$\mathbf{y}_i = \mathbf{w} + \sum_{j=1}^p \mathbf{a}_j d_{i-j} + \sum_{j=1}^q \mathbf{b}_j \mathbf{y}_{i-j} \quad [2]$$

Como puede verse, este modelo tiene una enorme similitud con el GARCH(p,q) estándar⁴.

Por otro lado, suponiendo que las duraciones son débilmente exógenas para los parámetros de interés, \mathbf{q} , entonces la estimación conjunta de la función f , y los parámetros pueden estimarse sin pérdida de eficiencia mediante la maximización del segundo término definido en [1]. En este estudio, consideramos tres expresiones de la distribución de la duración normalizada, \mathcal{D} . En primer lugar, si la densidad condicional es exponencial, la función cuasi-máximo verosímil Q puede escribirse como:

$$Q(d_i, \mathbf{q}) = \sum_{i=1}^N l_i = - \sum_{i=1}^N \left(\log \mathbf{y}_i + \frac{d_i}{\mathbf{y}_i} \right)$$

siendo \log el logaritmo natural. Bajo este supuesto, el modelo se denomina EACD.

⁴ Véase Engle y Russell (1998). En el caso de que tuviese memoria limitada, sólo las p duraciones más recientes influyen la duración condicional, restringiéndose a cero los parámetros \mathbf{b}_j . También, podrían especificarse otros modelos no lineales, como por ejemplo, el aplicado por Vechiatio (1999) dado el mejor ajuste muestral de las duraciones: $\log \mathbf{y}_i = \mathbf{w} + \mathbf{a}_1 \left(\frac{d_i}{\mathbf{y}_i} - 1 \right) + \mathbf{b}_1 \log \mathbf{y}_{i-1}$, y que es un modelo similar al

EGARCH propuesto por Nelson (1991). O, el recientemente propuesto por Zhang, Russell y Tsay (2000) que es una especificación no lineal denominada TACD o modelo ACD umbral que no emplearemos en este estudio. Por otro lado, estas especificaciones pueden generalizarse considerando variables exógenas que explican la frecuencia de los negocios, o simplemente eligiendo otras expresiones familiares de la literatura ARCH.

En segundo lugar, si utilizamos la distribución Weibull, es decir, $\frac{d_i}{\mathbf{y}_i} \equiv \mathbf{e}_i \sim W(\mathbf{g})$,

entonces la función de densidad es igual a:

$$f(d_i/\mathbf{g}, \mathbf{f}) = \frac{\mathbf{g}}{d_i} \left(\frac{d_i}{\mathbf{f}_i} \right)^{\mathbf{g}} \exp \left[- \left(\frac{d_i}{\mathbf{f}_i} \right)^{\mathbf{g}} \right], \quad \mathbf{f}_i = \frac{\mathbf{y}_i}{\Gamma \left(1 + \frac{1}{\mathbf{g}} \right)}$$

y la función del logaritmo de pseudo verosimilitud para el i -ésimo periodo de negociación es igual a:

$$l_i = \log \left(\frac{\mathbf{g}}{d_i} \right) + \mathbf{g} \log \left[\frac{\Gamma \left(1 + \frac{1}{\mathbf{g}} \right) d_i}{\mathbf{y}_i} \right] - \left(\frac{\Gamma \left(1 + \frac{1}{\mathbf{g}} \right) d_i}{\mathbf{y}_i} \right)^{\mathbf{g}}.$$

Este modelo recibe el nombre de WACD. Cuando $\mathbf{g} = 1$, la función de verosimilitud es similar a la que corresponde al modelo exponencial o EACD.

En tercer lugar, si consideramos la distribución Gamma generalizada estandarizada, tal que: $\frac{d_i}{\mathbf{y}_i} \equiv \mathbf{e}_i \sim G(\mathbf{k}, \mathbf{g})$, su función de densidad es igual a:

$$f(d/\mathbf{k}, \mathbf{g}, \mathbf{l}) = \frac{\mathbf{g} d^{\mathbf{k}\mathbf{g}-1}}{\mathbf{l}^{\mathbf{k}\mathbf{g}} \Gamma(\mathbf{k})} \exp \left[- \left(\frac{d}{\mathbf{l}} \right)^{\mathbf{g}} \right], \quad \mathbf{l} = \frac{\Gamma(\mathbf{k})}{\Gamma \left(\mathbf{k} + \frac{1}{\mathbf{g}} \right)}.$$

Por lo tanto, el vector de parámetros a estimar para \mathfrak{D} es $\mathbf{q}_s = (\mathbf{k}, \mathbf{g})$, mientras que el vector de parámetros contenido en \mathbf{y}_i es igual a $\mathbf{q}_t = (\mathbf{w}\mathbf{a}_1 \dots \mathbf{a}_p, \mathbf{h}_1 \dots \mathbf{b}_q)$. La estimación de los parámetros puede realizarse bajo estos supuestos, considerando la siguiente función del logaritmo de pseudo verosimilitud correspondiente a la i -ésima transacción, la cual es igual a:

$$l_i = \log(\mathbf{g}) + (\mathbf{k}\mathbf{g} - 1) \log \left(\frac{d_i}{\mathbf{l}_i} \right) - \log[(\mathbf{l}_i) \Gamma(\mathbf{k})] - \left(\frac{d_i}{\mathbf{l}_i} \right)^{\mathbf{g}}, \quad \mathbf{l}_i = \mathbf{y}_i \frac{\Gamma(\mathbf{k})}{\Gamma \left(\mathbf{k} + \frac{1}{\mathbf{g}} \right)}$$

Este modelo podríamos denominarlo GACD(p,q).

La estimación del modelo ACD se realiza por cuasi máxima verosimilitud basándose en la función de pseudo verosimilitud en la que la densidad de las llegadas se supone distribuida de forma exponencial (EACD), Weibull (WACD) o Gamma generalizada estandarizada (GACD). La estimación cuasi-máximo verosímil de los

parámetros es consistente para los parámetros y posee una matriz de covarianzas asintótica perfectamente definida.

3. Análisis empírico

Los datos utilizados sobre transacciones son *tick a tick* y corresponden a un contrato del futuro del Ibex35 en el que la información llega en intervalos irregulares de tiempo. Estos datos son descritos por dos tipos de variables: por un lado, el tiempo de llegada de la nueva información, t_i , o tiempo en el que tiene lugar la i -ésima negociación; y por otro lado, un vector de variables observables en aquel periodo o instante, y_i , que son variables denominadas *marks* pues describen los eventos ocurridos. Pues bien, en el contexto financiero, el tiempo de llegada es el punto del tiempo en el que se realiza un contrato sobre un número determinado de activos (en nuestro caso sobre el futuro del Ibex35), mientras que los *marks* son, por ejemplo, los precios *bid-ask*, el *spread*, el volumen del contrato y muchas otras características producidas en las transacciones que han ocurrido, las cuales son de interés en el estudio de la microestructura del mercado, como, por ejemplo, el número de operaciones cruzadas entre oferta y demanda.

En nuestro estudio, disponemos de la información correspondiente al momento en que se realizan las transacciones ($t_i, i=1,2,\dots,N$ periodos espaciados irregularmente), el número de operaciones cruzadas en t_i (NT_{t_i} o NT_i), el *spread* del precio (p_{t_i} o p_i); y el volumen negociado en t_i (V_{t_i} o V_i). Por otro lado, consideramos que la intensidad de la actividad de negociación es igual a $d_i = t_i - t_{i-1}$, o duración entre dos transacciones consecutivas, la cual es siempre positiva y está expresada en segundos.

Cuadro 1

Fechas, días y número de periodos en los que se realizan operaciones cruzadas para el Futuro del Ibex35

Fechas	Día de la semana	Nº de periodos	
		Total de periodos diarios	Eliminando duraciones iguales a cero *
28/09/98	Lunes	8,257	5,032
29/09/98	Martes	7,676	4,792
30/09/98	Miércoles	9,199	5,394
1/10/98	Jueves	10,801	6,115
2/10/98	Viernes	11,747	6,790
5/10/98	Lunes	10,378	6,068
6/10/98	Martes	8,024	4,658
7/10/98	Miércoles	10,871	6,270
8/10/98	Jueves	9,939	5,752
9/10/98	Viernes	7,103	4,274
13/10/98	Martes	7,072	4,351
14/10/98	Miércoles	8,345	4,945
15/10/98	Jueves	8,611	5,147
16/10/98	Viernes	2,803	1,882
		120,826	71,470

Nota: (*) Esta columna se refiere solamente a aquellas operaciones cruzadas una vez se han eliminado las duraciones que eran iguales a cero, esto es, aquellas $s-1$ operaciones en las que los tiempos eran idénticos (siendo s el número de tiempos (hora:minuto), t_i , que son iguales).

El cuadro 1 muestra las fechas, días y número periodos en los que se realizan las operaciones cruzadas que se han analizado. Concretamente, el contrato analizado es el *Ibex Plus Futures*, con código de fecha de expiración: *XJ98*, y observado diariamente durante el periodo que abarca desde 20/7/98 hasta 16/10/98, tercer viernes del mes de Octubre y fecha de expiración del contrato. No obstante, se han elegido los datos del día 16/10/98 en el contrato *XK98*, con vencimiento en Diciembre de 1998, dado que el número de operaciones en la fecha de vencimiento del anterior contrato era obviamente reducido. Los datos son obtenidos de *MEFF Renta Fija*. El horario del mercado está comprendido entre las 10:00 y 17:14 horas. Se han tomado las variables precio, número de operaciones cruzadas, volumen de negociación y los tiempos en que se realizó cada transacción.

Además, en este estudio se han tenido en cuenta otras consideraciones adicionales. En primer lugar, no se consideran las operaciones que tienen lugar en el mismo instante de tiempo; ya que puede darse el caso que existan transacciones que coinciden en el mismo periodo, si bien pueden tener un diferente número de operaciones cruzadas, precio y volumen negociado. Y, en segundo lugar, se crean diversas variables ficticias que especifican el intervalo horario al que pertenece la transacción. En total, se han

creado 8 variables ficticias horarias, reflejando los distintos intervalos horarios que abarcan desde las 10:00 – 10:59 horas hasta las 17:00 - 17:14 horas.

Finalmente, cabría destacar que el periodo elegido coincide con un periodo de crisis financiera (concretamente, el entorno de la crisis bancaria japonesa de 30 de Septiembre de 1998)⁵, lo cual, además, permite inducir el efecto que existe sobre las duraciones de las transacciones en un periodo de crisis.

3.1. Descripción de la duración observada

La descripción de la duración de tiempo entre dos subastas consecutivas, se realiza mediante la metodología estadística del análisis de supervivencia. En este tipo de análisis, la función *hazard* adquiere especial relevancia para el proceso de negociación pues es una función de la intensidad. Esta caracteriza la tasa instantánea de salida por unidad de tiempo [$I(d_i)$, siendo d_i la duración correspondiente]⁶. En el apéndice I y II pueden observarse las frecuencias ajustadas no paramétricamente mediante la densidad *kernel* utilizando el núcleo de Epanechnikov y la ventana espectral h . Estas poseen un comportamiento claramente decreciente, típico de distribuciones de tipo exponencial. Por otro lado, el cuadro 2 muestra las estimaciones no paramétricas de la función *hazard* para diez intervalos en los que se dividen los segundos de duración entre dos transacciones consecutivas, utilizando las estimaciones de Kaplan-Meier. El valor estimado de la función *hazard* puede interpretarse, por ejemplo en el caso del primer intervalo perteneciente a la fecha 28/09/98, como la probabilidad de un 13.48% de que la duración entre dos transacciones consecutivas sea de 12 segundos, habiendo durado 11 segundos, aproximadamente.

En general, podríamos advertir una serie de características para la duración observada. En primer lugar, la existencia de diferentes rangos para los intervalos

⁵ La crisis mexicana de 1994-1995 y asiáticas de 1997-1998 son crisis financieras recientes que han afectado a múltiples países. La crisis de 1997 demostró una fácil y rápida contaminación a otras economías. La crisis de septiembre-octubre de 1998, coincide con la crisis bancaria de Japón y la crisis del fondo de cobertura de riesgos *Long-Term Capital Management (LTCM)*, liderado por los insignes académicos Merton y Scholes.

⁶ La interpretación de la función *hazard*, o de razón de fallo, puede entenderse mejor con el siguiente ejemplo clásico. Así, analizando datos de duración del paro, podríamos decir que la función *hazard* representa la probabilidad de que un individuo salga de una situación de desempleo en un intervalo de tiempo corto, dado que ha estado desempleado en el periodo inmediatamente anterior. Dicho de otra manera, por ejemplo, si un individuo ha estado en paro nueve semanas, ¿cuál es la probabilidad de que en la décima semana esté empleado?.

temporales de cada día, sugiere que la intensidad de la actividad de negociación es claramente distinta porque la nueva información fluye a ritmo diferente haciendo la negociación más o menos frecuente. En segundo lugar, la mayor tasa *hazard* en los días de la crisis. En concreto, si se considera el primer intervalo para cada día, podríamos observar cómo el 30 de Septiembre es el 15%, el 1 de Octubre es el 16.20% y el 2 de Octubre es el 18.92%. Esto implica una mayor actividad de negociación pues las duraciones entre periodos de transacción son más pequeñas. En tercer lugar, teniendo en cuenta la duración de un minuto entre transacciones, podemos observar que la tasa *hazard* o de razón de fallo acumulada para el 28/09/98 es de 37.68%, mientras que para el 2/10/98 es aproximadamente un 74.51%, para descender al 19.75% en la fecha de 15/10/98. Ambos aspectos demuestran que la intensidad con la que se negoció esos días fue elevada. En cuarto lugar, podemos decir que los equivalentes en minutos máximos que aparecen en el cuadro (véanse los últimos intervalos en dicho cuadro), y para cada día de negociación analizado, son: 2, 3.71, 1.78, 1.72, 1.43, 1.93, 2.73, 1.73, 2.35, 2.51, 2.62, 4.82, 3.4 y 5.54, respectivamente. Es decir, en los días de la crisis la frecuencia máxima de duraciones observadas se encuentra entre 1.43 y 1.93 minutos (que representan 86 segundos y 116 segundos, respectivamente). Asimismo, y en todo la muestra, podemos decir que aproximadamente el 75.29% de las duraciones de las transacciones se encuentran por debajo de 5 segundos, mientras que sólo las superan el 24.51%. Esto indica un alto porcentaje de frecuencia de negociación.

Sin embargo, estas estimaciones no informan sobre la posible conducta “estacional” de las duraciones (que sí puede comprobarse en el gráfico de las mismas en el epígrafe II). Para comprobar el efecto estacional en el entorno de los modelos de duración clásicos que emplean covariables, se han creado ocho variables ficticias horarias que reflejan los diferentes intervalos horarios: I_1 (entre 10:00-10:59), I_2 (entre 11:00-11:59), I_3 (entre 12:00-12:59), I_4 (entre 13:00-13:59), I_5 (entre 14:00-14:59), I_6 (entre 15:00-15:59), I_7 (entre 16:00-16:59) e I_8 (entre 17:00-17:14).

Cuadro 2

Estimaciones no paramétricas de la duración entre las transacciones para cada día de negociación

Este cuadro contiene las estimaciones de Kaplan-Meier para todas aquellas operaciones cruzadas en las que las duraciones no son cero.

Intervalo	Hazard	Intervalo	Hazard	Intervalo	Hazard	Intervalo	Hazard	Intervalo	Hazard	Intervalo	Hazard	Intervalo	Hazard
28/09/98		29/09/98		30/09/98		01/10/98		02/10/98		05/10/98		06/10/98	
.0- 12.0	.1348	.0- 22.3	.0826	.0- 10.7	.1500	.0- 10.3	.1620	.0- 8.6	.1892	.0- 11.6	.1476	.0- 16.4	.1060
12.0- 24.0	.0869	22.3- 44.6	.0613	10.7- 21.4	.0956	10.3- 20.6	.1031	8.6- 17.2	.1359	11.6- 23.2	.0941	16.4- 32.8	.0594
24.0- 36.0	.0615	44.6- 66.9	.0355	21.4- 32.1	.0835	20.6- 30.9	.0827	17.2- 25.8	.0736	23.2- 34.8	.0729	32.8- 49.2	.0664
36.0- 48.0	.0625	66.9- 89.2	.0299	32.1- 42.8	.0567	30.9- 41.2	.0898	25.8- 34.4	.0876	34.8- 46.4	.0885	49.2- 65.6	.0530
48.0- 60.0	.0311	89.2- 111.5	.0408	42.8- 53.5	.0734	41.2- 51.5	.0755	34.4- 43.0	.0751	46.4- 58.0	.0416	65.6- 82.0	.0290
60.0- 72.0	.0619	111.5- 133.8	.0448	53.5- 64.2	.0485	51.5- 61.8	.1110	43.0- 51.6	.1203	58.0- 69.6	.0507	82.0- 98.4	.0281
72.0- 84.0	.0263	133.8- 156.1	.0000	64.2- 74.9	.0801	61.8- 72.1	.0388	51.6- 60.2	.0634	69.6- 81.2	.1232	98.4- 114.8	.0523
84.0- 96.0	.0238	156.1- 178.4	.0000	74.9- 85.6	.0267	72.1- 82.4	.0647	60.2- 68.8	.0332	81.2- 92.8	.0000	114.8- 131.2	.0407
96.0- 108.0	.0152	178.4- 200.7	.0000	85.6- 96.3	.0374	82.4- 92.7	.0000	68.8- 77.4	.1163	92.8- 104.4	.0000	131.2- 147.6	.0000
108.0- 120.	.1667	200.7- 223.0	.0897	96.3- 107.0	.1869	92.7- 103.0	.1942	77.4- 86.0	.2326	104.4- 116.0	.1724	147.6- 164.0	.1220
07/10/98		08/10/98		09/10/98		13/10/98		14/10/98		15/10/98		16/10/98	
.0- 10.4	.1638	.0- 14.1	.1272	.0- 15.1	.1112	.0- 15.7	.1073	.0- 28.9	.0660	.0- 20.4	.0906	.0- 33.4	.0497
10.4- 20.8	.1037	14.1- 28.2	.0882	15.1- 30.2	.0684	15.7- 31.4	.0756	28.9- 57.8	.0450	20.4- 40.8	.0626	33.4- 66.8	.0222
20.8- 31.2	.0699	28.2- 42.3	.0632	30.2- 45.3	.0563	31.4- 47.1	.0632	57.8- 86.7	.0219	40.8- 61.2	.0443	66.8- 100.2	.0156
31.2- 41.6	.0524	42.3- 56.4	.0519	45.3- 60.4	.0371	47.1- 62.8	.0538	86.7- 115.6	.0255	61.2- 81.6	.0302	100.2- 133.6	.0172
41.6- 52.0	.0684	56.4- 70.5	.0751	60.4- 75.5	.0558	62.8- 78.5	.0382	115.6- 144.5	.0231	81.6- 102.0	.0196	133.6- 167.0	.0109
52.0- 62.4	.0291	70.5- 84.6	.0473	75.5- 90.6	.0917	78.5- 94.2	.0510	144.5- 173.4	.0138	102.0- 122.4	.0327	167.0- 200.4	.0200
62.4- 72.8	.0231	84.6- 98.7	.0000	90.6- 105.7	.0442	94.2- 109.9	.0637	173.4- 202.3	.0000	122.4- 142.8	.0490	200.4- 233.8	.0299
72.8- 83.2	.1331	98.7- 112.8	.0473	105.7- 120.8	.0000	109.9- 125.6	.0000	202.3- 231.2	.0231	142.8- 163.2	.0000	233.8- 267.2	.0000
83.2- 93.6	.0641	112.8- 126.9	.0000	120.8- 135.9	.0000	125.6- 141.3	.0000	231.2- 260.1	.0000	163.2- 183.6	.0000	267.2- 300.6	.0120
93.6- 104.0	.1923	126.9- 141.0	.1418	135.9- 151.0	.1325	141.3- 157.0	.1274	260.1- 289.0	.0692	183.6- 204.0	.0980	300.6- 334.0	.0599

Nota: Los intervalos están expresados en segundos. Se ha considerado que el número de intervalos fuese de diez por motivos de espacio.

Cuadro 3
Estimaciones paramétricas de modelos de duración tipo Weibull

	28/9/98	29/9/98	30/9/98	1/10/98	2/10/98	5/10/98	6/10/98	7/10/98	8/10/98	9/10/98	13/10/98	14/10/98	15/10/98	16/10/98
I_1	1.2006 (0.03)	1.2167 (0.03)	1.3101 (0.03)	1.2154 (0.02)	1.0335 (0.027)	1.1442 (0.027)	1.3035 (0.03)	1.2803 (0.028)	1.1222 (0.027)	1.4219 (0.034)	1.5760 (0.035)	1.3224 (0.033)	1.3212 (0.029)	2.4833 (0.06)
I_2	1.4554 (0.03)	0.7029 (0.04)	1.4219 (0.03)	1.4535 (0.03)	1.5674 (0.029)	1.3915 (0.028)	1.5615 (0.033)	1.2421 (0.027)	1.5037 (0.028)	2.0585 (0.042)	1.6798 (0.037)	1.8987 (0.036)	1.4101 (0.03)	3.4938 (0.089)
I_3	1.9014 (0.03)	1.8354 (0.04)	1.4802 (0.03)	1.5094 (0.03)	1.6733 (0.028)	1.5777 (0.031)	2.1911 (0.004)	1.3033 (0.029)	1.6746 (0.03)	2.2104 (0.044)	1.6624 (0.036)	1.5558 (0.033)	1.7337 (0.033)	3.1187 (0.075)
I_4	2.4472 (0.05)	2.4136 (0.04)	1.7833 (0.04)	2.2777 (0.04)	1.3096 (0.026)	2.0714 (0.036)	2.3473 (0.042)	2.2727 (0.037)	2.0557 (0.033)	2.3614 (0.046)	2.1692 (0.045)	2.2233 (0.044)	2.8341 (0.049)	3.7889 (0.104)
I_5	2.4007 (0.04)	1.9793 (0.03)	2.5631 (0.05)	2.1284 (0.04)	1.8433 (0.03)	2.0877 (0.036)	2.3648 (0.051)	2.0999 (0.031)	2.3023 (0.038)	2.1871 (0.043)	2.2366 (0.04)	2.5263 (0.038)	2.0743 (0.035)	3.9394 (0.115)
I_6	1.5686 (0.03)	1.5140 (0.03)	1.8081 (0.03)	1.5767 (0.03)	1.4741 (0.025)	1.4828 (0.027)	1.7727 (0.034)	1.3167 (0.027)	1.5482 (0.028)	1.5638 (0.033)	1.8117 (0.037)	1.6992 (0.032)	1.6433 (0.032)	2.6515 (0.059)
I_7	1.3184 (0.03)	1.4611 (0.03)	1.2259 (0.03)	1.0124 (0.03)	0.9657 (0.026)	1.0705 (0.028)	1.2228 (0.032)	1.1399 (0.028)	1.1346 (0.028)	1.3494 (0.034)	1.5589 (0.036)	1.6793 (0.034)	1.2144 (0.03)	1.5525 (0.042)
I_8	1.1002 (0.06)	1.2572 (0.07)	1.0953 (0.06)	0.7890 (0.05)	1.3357 (0.053)	0.9969 (0.06)	1.1128 (0.057)	1.1967 (0.059)	1.1171 (0.056)	1.3836 (0.077)	1.3296 (0.068)	1.2782 (0.061)	1.2598 (0.06)	1.3154 (0.076)
σ	1.0291 (0.011)	1.0757 (0.01)	1.0322 (0.01)	0.9490 (0.009)	0.9673 (0.0091)	0.9991 (0.009)	1.0489 (0.01)	0.9839 (0.01)	0.9763 (0.009)	1.0455 (0.012)	1.0526 (0.012)	1.0469 (0.012)	1.0046 (0.011)	1.0991 (0.023)
λ	0.2147 (0.004)	0.2033 (0.004)	0.2245 (0.004)	0.2486 (0.004)	0.2672 (0.004)	0.2478 (0.004)	0.2022 (0.004)	0.2538 (0.004)	0.2340 (0.004)	0.1809 (0.003)	0.1771 (0.003)	0.2142 (0.004)	0.2158 (0.004)	0.1097 (0.03)
p	0.9717 (0.011)	0.9296 (0.01)	0.9688 (0.01)	1.0536 (0.011)	1.0338 (0.009)	1.0008 (0.009)	0.9533 (0.011)	1.0163 (0.01)	1.0242 (0.01)	0.9565 (0.012)	0.9500 (0.011)	0.9552 (0.01)	0.9954 (0.01)	0.9098 (0.019)
Mediana	3.19	3.32	3.05	2.84	2.62	2.80	3.37	2.75	2.99	3.77	3.84	3.18	3.21	6.09
Log L	-7632.7	-7469.6	-8186.1	-8778.1	-9811.2	-8981.1	-7150.2	-9201.1	-8436.1	-6581.5	-6729.5	-7564.7	-7689.0	-3037.7

Nota: Este cuadro contiene las estimaciones de los modelos paramétricos clásicos en los que la función de densidad es Weibull. Se utilizan como covariables a ocho variables ficticias que determinan el intervalo horario de negociación. Entre paréntesis aparecen los errores estándar. En este caso, la función hazard es $I_p(I_d)^{p-1}$, y dado que existen covariables, los valores de λ son iguales a: $I_i = \exp(-x_i \mathbf{d})$ siendo \mathbf{d} el vector de parámetros desconocidos; mientras que $p = 1/s$ y σ es la desviación. En general, la estimación de los parámetros de interés requiere de la presentación del modelo en términos lineales. De esta forma, la variable duración es transformada en logaritmos. La función de logaritmo de verosimilitud para las T observaciones es: $LnL = \sum_{i=1}^T [c_i(-\ln \mathbf{s} + \ln h(w_i)) + \ln S(w_i)]$, siendo $w_i = (\ln d_i - x_i \mathbf{d}) / \mathbf{s}$; $h(w_i) = g(w_i) / S(w_i)$ es la función hazard, g es la función de densidad, S es la función de supervivencia y c_i es un indicador que describe si la observación i -ésima está censurada a la derecha (=0) o no está censurada (=1).

Pues bien, suponiendo que d_i está generada por un proceso de Poisson de parámetro λ_i , se pueden utilizar algunos modelos clásicos en los que la duración de un fenómeno (como por ejemplo, el tiempo entre transacciones) puede modelizarse de forma bastante simple mediante la densidad de la función Weibull⁷. Concretamente, se emplea un modelo paramétrico de tipo Weibull con covariables, es decir, con variables exógenas que en nuestro caso indican si la transacción i -ésima pertenece al intervalo de negociación horario $I_j, j=1,2,\dots,8$. Los resultados obtenidos aparecen en el cuadro 3. En general, éstos muestran la importancia del intervalo temporal de apertura y cierre, en términos de los parámetros de las variables I_1 e I_2 o I_7 e I_8 , pues éstos son los que presentan un menor valor positivo y por tanto un menor efecto sobre la duración media. En este sentido, los intervalos de apertura y final de la sesión contribuyen en menor medida al incremento de las duraciones medias que los intervalos que corresponden al mediodía⁸.

Finalmente, podemos resaltar que una estimación promedio de la duración en segundos para cada intervalo de transacción, en el total de datos analizados en la muestra, indica que la duración promedio en segundos es igual a: 3.73, 5.03, 5.75, 8.98, 9.61, 5.20, 3.41, y 3.24, respectivamente. Puede verse, cómo comparando estos resultados con los del cuadro 3, los intervalos I_3, I_4 e I_5 son los que poseen una mayor duración promedio entre transacciones.

3.2. Estimación de la intensidad de la actividad de negociación

A continuación, se presentan los resultados de la estimación de los parámetros en el modelo de duración condicional, ACD. El cuadro 4 muestra las estimaciones de los modelos ACD para diferentes distribuciones \mathcal{D} . Además, se recogen las estimaciones de aquellos modelos que han sido filtrados por el componente “estacional” estimado mediante el *spline* cúbico. Sólo se considera una expresión para \mathbf{y}_i , la cual es de tipo GARCH.

⁷ En el caso de los datos sobre transacciones, éstos no están censurados. Por lo tanto, $c_i=1$.

⁸ No obstante, la interpretación de los coeficientes estimados para las covariables no es sencilla. En el caso de los modelos Weibull y Exponencial, $E[d/x_i] = e^{x_i} \Gamma[(1/p)+1]$.

En dicho cuadro se distinguen dos especificaciones para cada distribución que adopta \mathcal{D} . Así, SE indica que el modelo no está corregido por el componente “estacional” estimado; mientras que E indica que el modelo está corregido por el componente “estacional” estimado, es decir, consideramos que la duración está “desestacionalizada”. Engle y Russell (1998) emplean un *spline* cúbico, considerando el siguiente modelo: $E[d_i/t_{i-1}] = s(t_{i-1}) + \mathbf{e}_i$, $\mathbf{e}_i \sim iid(0, \mathbf{S}_e^2)$ que puede estimarse por mínimos cuadrados ordinarios. La especificación del modelo es:

$$s(t_{i-1}) = \sum_{j=1}^k I_j \left[c_j + b_{1,j}(t_{i-1} - k_{j-1}) + b_{2,j}(t_{i-1} - k_{j-1})^2 + b_{3,j}(t_{i-1} - k_{j-1})^3 \right],$$

donde, $I_j = 1$ si $k_{j-1} \leq t_{i-1} \leq k_j$, e $I_j = 0$ en caso contrario⁹. Además, para $j > 1$, c_j y $b_{1,j}$ son restringidos a cero mediante las condiciones de diferenciabilidad habituales, es decir, se suponen iguales a cero. En este sentido, la ecuación [2] se reespecificará sustituyendo d_{i-1} por $\tilde{d}_{i-1} = \frac{d_{i-1}}{\hat{s}(t_{i-1})}$, y estimando los parámetros de dicha ecuación por máxima verosimilitud.

Se han omitido las estimaciones de los coeficientes del *spline* cúbico por motivos de espacio. Por otro lado, se ha añadido a la expresión de la ecuación de la duración condicional una variable ficticia D_i , que toma el valor 1 cuando las duraciones observadas son superiores a 5 desviaciones estándar de la duración realizada; y cero en caso contrario. Con ello, se recogen a todas aquellas observaciones anómalas que puedan provocar imprecisiones en la estimación del modelo.

De los resultados de dicho cuadro, se pueden extraer las siguientes conclusiones. En primer lugar, los modelos estimados para \mathbf{y}_i son EACD(1,1), WACD(1,1) y GACD(1,1). Los órdenes de dichos modelos han sido contrastados mediante criterios de información estándar, como los de Akaike y Schwartz, encontrándose como suficientes los de orden uno.

En segundo lugar, y en cuanto a las estimaciones de los parámetros de \mathbf{y}_i , cabe resaltar que todos son significativos, son estimados de forma robusta para todas las especificaciones consideradas para \mathcal{D} , y cumplen con las condiciones de estacionariedad del modelo de duración condicional. Así, por ejemplo, la persistencia ($\mathbf{a}_1 + \mathbf{b}_1$) es igual, en los modelos SE , a 0.9312, 0.9345 y 0.9657; mientras que en los

modelos E es igual a 0.8299, 0.8303 y 0.8473, respectivamente. De esta manera, podemos decir que, dado que es inferior a la unidad, existe una escasa persistencia en la duración condicional estimada; la cual es todavía menos importante en el caso de los modelos “desestacionalizados”. Esto demuestra que la estacionalidad puede provocar un acercamiento hacia la persistencia en la duración condicional si no es tenida en cuenta.

Cuadro 4

Estimaciones cuasi máximo verosímiles del modelo ACD(1,1) bajo diferentes funciones de densidad del tiempo entre llegadas de las transacciones

Parámetros	Exponencial		Weibull		Gamma Generalizada	
	<i>SE</i>	<i>E</i>	<i>SE</i>	<i>E</i>	<i>SE</i>	<i>E</i>
ω	0.2779 (10.33)	0.4624 (4.32)	0.2665 (10.15)	0.4616 (4.67)	0.2776 (10.64)	0.5002 (3.66)
α_1	0.1089 (19.52)	0.0028 (5.18)	0.1082 (19.52)	0.0029 (5.36)	0.1263 (20.26)	0.0032 (4.05)
β_1	0.8223 (70.83)	0.8271 (23.14)	0.8263 (73.07)	0.8274 (24.76)	0.8394 (85.57)	0.8441 (21.70)
μ_1	12.1961 (15.53)	17.4500 (7.09)	12.0647 (15.23)	17.4372 (8.29)	13.9121 (16.98)	19.849 (6.11)
γ	--	--	1.0506 (430.5)	1.0021 (386.5)	1.2328 (296.2)	1.1514 (237.6)
$t_{g=1}$	--	--	20.73	0.8099	55.93	31.24
κ	--	--	--	--	1.1404 (322.8)	1.2135 (259.3)
Log L	-176739.4	-179410.8	-176574.8	-178410.4	-153414.8	-157318.4
\mathcal{Y}	5.2278 [0.019]	5.1227 [0.017]	5.2566 [0.019]	5.1238 [0.017]	6.4198 [0.024]	6.2520 [0.02]
\mathcal{S}_y	5.2849	4.5667	5.3385	4.5680	6.6070	5.5917
AN(0,1)	-0.101	-0.113	-0.095	-0.088	-0.086	-0.072

Nota: Este cuadro contiene las estimaciones de los coeficientes y sus errores estándar ajustados por la estimación CMV para el modelo de la forma: $\mathbf{y}_i = \mathbf{w} + \mathbf{a}_1 d_{i-1} + \mathbf{b} \mathbf{y}_{i-1} + \mathbf{m} D_i$, donde la i -ésima duración entre eventos es d_i , \mathcal{Y}_i es la duración condicional y D_i es una variable ficticia. El parámetro μ_1 se corresponde con una variable ficticia que vale 1 cuando la duración es mayor que cinco desviaciones típicas ($5\mathcal{S}_d$). Se distinguen dos modelos: a) modelos sin filtrar los datos mediante el componente estacional de la duración observada (*SE*) y b) modelos que filtran la duración mediante el componente estacional estimado por el *spline* cúbico (*E*). La estimación CMV está basada en distintas funciones de pseudo verosimilitud en la que la densidad de las llegadas (\mathcal{D}) es Exponencial, Weibull y Gamma Generalizada. Por otro lado, $t_{g=1}$ es el valor del estadístico *t*-Student para la hipótesis nula de que $\gamma=1$. \mathcal{Y} y \mathcal{S}_y representan la media de las duraciones estimadas y la desviación estándar de las mismas. Entre paréntesis aparece el valor *t*-Student robusto para la hipótesis nula de no significación individual del coeficiente estimado. Entre corchetes aparece el error estándar.

⁹ En este trabajo, el número de intervalos considerados (y, por lo tanto, variables ficticias) es $k=8$.

En tercer lugar, los coeficientes que caracterizan la distribución Weibull (\mathbf{g} y Gamma Generalizada (\mathbf{g} y \mathbf{k}) son significativos en todos los casos estimados; y atendiendo al contraste $t_{g=1}$ en el modelo Weibull, es diferente a la unidad en el caso E , si bien para SE no se rechaza dicha hipótesis, pudiéndose asumir que la tasa *hazard* es constante pues no se rechaza que la distribución sea Exponencial. En el caso del modelo GACD(1,1), no sólo se rechaza individualmente que el coeficiente $\mathbf{g}=1$ sino que no se rechaza que se $\mathbf{k}=0$. Por otro lado, un contraste de tipo Wald para la hipótesis conjunta muestra el rechazo de la misma, y por lo tanto, supone la preferibilidad de dicha distribución para describir la duración condicional y computar las estimaciones del modelo ACD(1,1).

En cuarto lugar, dado que los parámetros \mathbf{a}_i y \mathbf{b}_i son significativos, podemos decir que el modelo de duración condicional posee contenido informativo, esto es, se rechaza que $\mathbf{y}_i = \mathbf{w}$, y por lo tanto, que la duración observada y condicional retardada no posean cierto contenido informativo para explicar la duración condicional o intensidad esperada. Con esto, puede decirse que el proceso de llegadas no sigue una distribución de Poisson, pues la tasa de llegadas no es constante, tal y como sugeriría el hecho de que encontrásemos que $\mathbf{y}_i = \mathbf{w}$.

En quinto lugar, y dado que las verosimilitudes muestrales son distintas al considerar diferentes modelos, el diagnóstico o validación final de los modelos estimados puede realizarse analizando las duraciones estandarizadas. Por ejemplo, en el caso de que éstas sean filtradas, la expresión es $\mathbf{e}_i = d_i / (\mathbf{y}_i \hat{s}(t_{i-1}))$, tal y como hacen Engle y Russell (1997). Estos residuos deben comportarse *i.i.d.*, puesto que el proceso ACD supone que las transformaciones estocásticas de los datos son *i.i.d.*. En este sentido, un contraste sencillo de no existencia de excesos de dispersión puede realizarse mediante el siguiente cociente bajo la hipótesis nula: $\sqrt{T} \left[\frac{(\mathbf{s}_e^2 - 1)}{\mathbf{s}_v} \right] \sim AN(0,1)$, donde \mathbf{s}_v corresponde a la variable $(\mathbf{e}_i - 1)^2$ y \mathbf{s}_e^2 es la varianza de \mathbf{e}_i . En nuestro caso, consideramos las duraciones estandarizadas tanto para el modelo SE como E . Según los resultados obtenidos, podemos decir que éstas son *i.i.d.* a los niveles de significación habituales.

Pues bien, atendiendo a estos resultados, podemos utilizar las duraciones estimadas corregidas de la estacionalidad para caracterizar la duración predicha como variable que

permita explicar la volatilidad condicional de la rentabilidad del futuro del Ibex35. Nótese que las estimaciones de la intensidad condicional medida por \mathbf{y}_i , muestran que la media de las estimaciones de la duración esperada en segundos ($\bar{\mathbf{y}}$) y su desviación (\mathbf{s}_y) son muy parecidas en cada uno de los modelos considerados, mostrándose un incremento de ambas cuando pasamos de la distribución Exponencial a la Gamma Generalizada. Dada la dificultad que existe para comparar los resultados de distintos modelos, pues poseen distintas funciones de pseudo verosimilitud, encontramos que la única forma de evaluar las estimaciones es a través de la significación estadística individual de los parámetros estimados de cada modelo.

Finalmente, cabría destacar que el modelo estimado no ha incluido otras variables observables por el agente (y que, por ejemplo, pudiesen construirse a partir del libro de órdenes), y sólo se ha considerado que la información está resumida por la duración observada y esperada que están retardadas por un periodo, ante la falta de información más precisa sobre la negociación.

3.3. Rentabilidad y volatilidad condicional con datos de transacciones

En finanzas, se han propuesto modelos que contemplan el hecho de la deformación temporal o relación entre el tiempo lógico de los acontecimientos y el tiempo real, considerando una estructura que está formulada en términos de transacciones, pero también donde se especifica la distribución de los eventos en tiempo real. Los trabajos de Mandelbrot y Taylor (1967) o Clark (1973) introdujeron el concepto de “deformación del tiempo” para modelizar la visión subjetiva que del tiempo poseen los agentes o participantes de mercado. En estos modelos, la velocidad del tiempo aumenta cuando aumenta el flujo de información. Este flujo de información puede medirse por: a) la tasa de interés, el volumen de transacción (Mandelbrot y Taylor) o la volatilidad

¹⁰ Otras características son que el activo que se negocia tiene un valor de liquidación que es aleatorio; el conjunto de precios es aquel en el que el beneficio esperado para cualquier negocio es nulo; tales precios competitivos implican que los precios son iguales a la expectativa condicional del valor de liquidación del activo, dado el tipo de negocio que ocurre (de esta forma, la comprensión de la conducta de los precios implica el análisis de los cambios en las creencias de los creadores de mercado); se supone que el aprendizaje es bayesiano (el creador de mercado conoce que el flujo de órdenes está correlacionado con el valor de los activos, pero no conoce cual es su valor cierto); cada agente tiene una creencia a priori sobre el valor de activo; utilizando un aprendizaje bayesiano, el proceso de actualización tiene buenas propiedades de convergencia. En modelos análogos a GM, el precio al final de cualquier activo negociado; por ejemplo, su valor de liquidación, se supone que es una variable aleatoria cuya ley de probabilidad está caracterizada por momentos cuyos parámetros son exógenos.

(Clark)¹¹. También el interés por la deformación del tiempo aparece en Hausman, Lo y MacKinlay (1992), o en Engle y Russell (1997 y 1999) quienes proponen un modelo que tiene en cuenta explícitamente el espaciado irregular de los datos de transacciones, además de medir y predecir la intensidad con la que se realizan las mismas (modelo analizado en la sección 2).

La idea de este epígrafe es la especificación conjunta de $p(\cdot)$ y $q(\cdot)$. Para ello, y en un sentido similar a Vecchiato (1999), se completa la expresión [1] no sólo considerando la duración condicional sino modelizando conjuntamente la rentabilidad del futuro del Ibex35 y su volatilidad condicional, mediante una especificación que considera el espaciado irregular de los datos de transacciones.

El tipo de modelos que emplearemos introduce como nuevas variables exógenas a la duración observada, para recoger el efecto del espaciado irregular de los datos.

Siguiendo a Pai, Polasek y Kozumi (1995), consideremos un modelo para datos irregularmente espaciados en forma de un $AR(p)$ [en adelante $ISAR(p)$] con funciones de respuesta $\Phi_j[\Delta^j t_i]$, $j=1, \dots, p$; $i=1, \dots, N$; siendo $\Delta^j = 1 - L^j$ donde L^j es el operador de retardos de j periodos. Las funciones de respuesta, Φ_j , son funciones de $\Delta^j t_i$ y generalmente tienen diferentes grados de decaimiento y permiten ponderar las observaciones de las variables retardadas utilizando pesos que dependen de la duración de tiempo. Nótese que la duración entre dos transacciones consecutivas es Δt_i (o d_i), mientras que la distancia entre dos transacciones que están separadas j ticks puede escribirse como $\Delta^j t_i$. Considerando que las rentabilidades observadas en el periodo t_i , tienen en cuenta la longitud del intervalo de duración, podemos escribir que:

$r_i \equiv \frac{(\log p_i - \log p_{i-1})}{\Delta t_i}$. De esta forma, también se permite eliminar la estacionalidad de

los datos.

El modelo $ISAR(p)$ con errores gaussianos y condicionalmente heterocedástico puede escribirse como:

¹¹ Estas nuevas escalas temporales se mostraron adecuadas a las propiedades de las series temporales. Bastante más que el tiempo físico, ellas son la escala de tiempo natural o relevante en la que se mide el proceso de generación de precios. Stock (1988) mostró que en tales escalas de tiempo los precios podrían estar linealmente relacionados a los fundamentos subyacentes. Stock también demostró que esta escala de tiempo podría capturar la heterocedasticidad condicional exhibida por las series de tiempo. Con el cambio de volumen de las transacciones en las frecuencias intradía, el uso de tales escalas de tiempo parece ser lo más relevante. Así, las transformaciones de escala temporal son una alternativa interesante a la clase de modelos ARCH. Estas transformaciones son: a) tiempo físico; b) theta (ϑ)-desestacionalizado y c) tiempo intrínseco. Por ejemplo, véase para un mayor detalle Dacorogna, Müller, Nagler, Olsen y Pictet (1993).

$$r_i = \mathbf{f}_0 + \sum_{j=1}^p \Phi_j[\Delta^j t_i] r_{i-j} + v_i; \quad v_i / \Psi_{i-1} \sim N(0, \mathbf{s}_i^2); \quad i = 1, \dots, N \quad [3]$$

donde existen N puntos irregularmente espaciados; se supone que el error condicionado a la información existente en $i-1$, Ψ_{i-1} , se distribuye Normal con media nula y varianza variable, \mathbf{s}_i^2 (que es la varianza del cambio en los precios sobre la i -ésima transacción condicionada al pasado y presente de las duraciones entre transacciones). En este estudio consideramos tres tipos de funciones de respuesta, Φ_j : una para datos igualmente espaciados y dos para datos irregularmente espaciados. Así, podemos citar: la función constante, $\Phi_j[\Delta^j t_i] = \mathbf{f}_j$, que corresponde al caso de datos igualmente espaciados; la función exponencial, $\Phi_j[\Delta^j t_i] = \mathbf{f}_{0j} + \mathbf{f}_{1j} \exp(-\Delta^j t_i)$; y, la función recíproca, $\Phi_j[\Delta^j t_i] = \mathbf{f}_{0j} + \mathbf{f}_{1j} (1/\Delta^j t_i)$.

Para distintos valores de \mathbf{f}_{0j} y \mathbf{f}_{1j} se tienen diferentes funciones decrecientes. El vector de parámetros es $\Phi = (\mathbf{f}_0, \mathbf{f}_{01}, \mathbf{f}_{11}, \dots, \mathbf{f}_{0p}, \mathbf{f}_{1p})'$ ⁽¹²⁾. La condición de estacionariedad de [3] depende del espacio de parámetros y de la distribución de d_i . En este sentido, por ejemplo, si suponemos que $d_i \sim G(\mathbf{k}, \mathbf{g})$ y la duración es independiente de los errores, el valor esperado $E[\Phi_j[\Delta^j t_i]] = E[\mathbf{f}_{0j} + \mathbf{f}_{1j} \exp(-\Delta^j t_i)] = \mathbf{f}_{0j} + \mathbf{f}_{1j} (\mathbf{g} + 1)^{-k_j}$, $\mathbf{k} > 1$. Así, denominando a dicho valor esperado por \mathbf{f}_j , la condición de estacionariedad en un proceso ISAR(1) con $j=1$ es equivalente a analizar las raíces del polinomio $1 - \mathbf{f}_1 L = 0$, las cuales deben caer fuera del círculo unitario.

La estimación del modelo ISAR(p) podría realizarse por MCO siempre que el término de perturbación fuese homocedástico y no correlacionado serialmente. Sin embargo, dado que se ha supuesto que este término de perturbación es condicionalmente heterocedástico en el conjunto de información Ψ_{i-1} , se ha considerado que la dinámica del proceso para la varianza condicional puede formularse en términos de un modelo de tipo GARCH.

¹² En este caso, la matriz de regresores está formada por $N \times (2p+1)$ elementos, formados por las variables retardadas sin ponderar, las variables ponderadas en las que los pesos dependen de la duración del tiempo, y una constante.

En este sentido, y dada la naturaleza de los datos irregularmente espaciados, también podrían plantearse diferentes funciones de respuesta en los modelos de tipo ARCH. Extendiendo el modelo de Pai, Polasek y Kozumi (1995) al caso GARCH(k, q), tenemos que:

$$\mathbf{s}_i^2 = \mathbf{w} + \sum_{j=1}^q A_j [\Delta^j t_i] v_{i-j}^2 + \sum_{r=1}^k \mathbf{b}_r \mathbf{s}_{i-r}^2 \quad [4]$$

donde las funciones $A_j [\Delta^j t_i]$ pueden ser equivalentes a las descritas en anteriormente. Así, $A_j [\Delta^j t_i] = \mathbf{a}_{0j}$ en el caso constante, $A_j [\Delta^j t_i] = \mathbf{a}_{0j} + \mathbf{a}_j \exp(-\Delta^j t_i)$ en el caso exponencial, y $A_j [\Delta^j t_i] = \mathbf{a}_{0j} + \mathbf{a}_j (1/\Delta^j t_i)$ en el caso recíproco. Los parámetros $\mathbf{a}_{0j}, \mathbf{a}_j, j = 1, \dots, q$ y $\mathbf{b}_r, r = 1, \dots, k$, deben cumplir las condiciones de estricta no negatividad y estacionariedad habituales en los modelos GARCH, pero además teniendo en cuenta la distribución de d_i , como en el caso ISAR(p).

Al considerar como variable exógena la duración observada, \mathbf{s}_i^2 es la varianza por segundo del proceso de precios condicionado al conjunto de información de las duraciones pasadas y presentes. De esta forma, las expresiones [3] y [4] caracterizan conjuntamente a la rentabilidad y volatilidad en datos irregularmente espaciados, y que denominaremos como procesos ISAR-GARCH.

3.3.1. La intensidad y el volumen de negociación en la volatilidad condicional.

3.3.1.1. Una justificación

¿Por qué considerar el tiempo esperado y el volumen de negociación de las transacciones en la volatilidad condicional?. En cuanto al primer aspecto, podemos argumentar que la importancia del tiempo de los negocios en la valoración de los precios es una cuestión más bien empírica. Generalmente, en los modelos que explican la microestructura del mercado¹³, el tiempo de negociación no juega un papel importante puesto que se considera que no es exógeno al proceso de ajuste de los precios, y puede decirse que es irrelevante para la conducta del precio porque, en sí mismo, no posee contenido informativo. Sin embargo, si el tiempo fuese exógeno al

¹³ En este sentido, podemos aludir a que la importancia de la microestructura de los mercados financieros es fundamental, pues permite analizar la evolución de los precios (*bid/ask* o su *spread*), los requerimientos de los márgenes, el grado de competencia en el mercado, el comportamiento de los creadores de mercado (*market makers*), la frecuencia en la ejecución de las órdenes o la volatilidad intradía.

proceso de precios, entonces podría tener sentido su análisis. Por ejemplo, si el tiempo estuviese correlacionado con cualquier factor vinculado con la valoración de los precios, entonces la presencia o ausencia de negociación proveería de información a los participantes del mercado. Por lo tanto, si los participantes del mercado pueden aprender observando el tiempo de negociación, entonces el ajuste de los precios a la información disponible también dependería del tiempo [véase O'Hara (1995, pág.169)]. En esta última línea, algunos modelos recientes en finanzas consideran que el tiempo entre las transacciones contiene información importante en la descripción de la dinámica de los precios. Kyle (1985) y Admati y Pfleiderer (1988), entre otros, han sugerido que la frecuencia de las transacciones incorpora información sobre el mercado, es decir, el tiempo de negociación está relacionado con la existencia de nueva información. Sin embargo, la noción del tiempo como una señal que advierte de la presencia o ausencia de negociación y provee información a los participantes del mercado, se desarrolla con la investigación de Diamond y Verrecchia (1987)¹⁴, aunque también es tratada por Easley y O'Hara (1992), y Hausman, Lo y MacKinlay (1992), quienes han sugerido que el tiempo puede afectar al proceso de ajuste de los precios, aunque también comentan que no está todavía claro cómo.

En cuanto al segundo aspecto, mientras que la conexión empírica entre los movimientos en los precios y el volumen es clara [véanse Epps y Epps (1976), Rogalski (1978), Tauchen y Pitts (1983) y Sharma, Mougoue y Kamath (1996), entre otros], no es tan obvio porqué esto se produce o es así. Investigadores empíricos han establecido algunos patrones asimétricos para el volumen y la dirección de los cambios en los precios, si bien la generalidad de estos resultados está sujeta a debate. Por ejemplo, se dice que el volumen es mayor cuando los precios crecen que cuando disminuyen, y que esto puede estar relacionado con las restricciones sobre ventas al descubierto. Por otro lado, la relación precio-volumen en el mercado de futuros no es la misma que en el mercado del subyacente [véase Karpoff (1987)]. Por ejemplo, Kim y Verrecchia (1991) han modelizado la conexión entre la información pública de los anuncios y el volumen, encontrando que el cambio en el volumen es proporcional a la precisión de la señal de

¹⁴ Estos autores tienen la intuición de que si los agentes negociadores son incapaces de negociar en ciertos momentos, entonces observando la ausencia de negociación puede indicarse el estado subyacente. Concretamente, estos autores aluden a tres razones: 1) que el agente seleccionado para negociar simplemente no desea hacerlo; 2) que un agente desinformado sea incapaz de vender al descubierto; o 3) que un agente informado de las malas noticias pueda ser incapaz de negociar si las ventas al descubierto están prohibidas.

información pública, y es decreciente con la cantidad de pre-anuncios públicos y de la información privada.

No obstante, y en general, la investigación teórica no proporciona una respuesta definitiva. En la literatura sobre la microestructura del mercado de activos la conexión entre precio-volumen es menos clara. Por ejemplo, Kyle (1985) no considera que el volumen de negociación sea un factor en el proceso de ajuste de los precios. La razón es que los agentes informados siempre ajustan sus cantidades de órdenes para mantener igual su proporción relativa de negocio. Una de las razones de porqué es difícil evaluar la conexión precio-volumen es que no está claro qué cantidad de información proporciona el volumen a los participantes en el mercado. Aún así, numerosos investigadores han examinado la posible relación en un entorno de expectativas racionales, como por ejemplo Pflleiderer (1984), Campbell, Wang y Grossman (1991) y Blume, Easley y O'Hara (1994), entre otros. En general, existen dos aproximaciones. Una primera que analiza el volumen cuando existen agentes que transaccionan con diferentes señales de información [por ejemplo, véase Wang (1994)]. Una segunda, determina que en la información inherente al volumen estadístico, los agentes pueden aprender de su observación [véase Blume, Easley y O'Hara (1994), entre otros]. En este sentido, podemos decir que el papel del volumen en el proceso de ajuste de los precios es facilitar el aprendizaje de la incertidumbre que es subyacente a dicho proceso.

3.3.1.2. *Un modelo GARCH con variables exógenas*

Suponiendo que \mathbf{s}_i^2 es un proceso GARCH(1,1) con variables exógenas como:

$$\mathbf{s}_i^2 = \mathbf{w} + A[\Delta t_i]v_{i-1}^2 + \mathbf{b}\mathbf{s}_{i-1}^2 + \mathbf{h}\mathbf{y}_i + \mathbf{h}_2V_i \quad [5]$$

donde $j=1$; los parámetros son \mathbf{w} , \mathbf{a}_{01} , \mathbf{a}_{02} , \mathbf{b}_1 , \mathbf{h}_1 , \mathbf{h}_2 y donde, por ejemplo en el caso exponencial, $\mathbf{a}_{01} + \mathbf{a}_{11} \exp(-d_i) + \mathbf{b}_1 < 1$ es la condición de existencia de la varianza incondicional, que depende no sólo del espacio de parámetros sino de la distribución de d_i , como se ha visto. Este modelo es claramente simétrico, es decir, la volatilidad es la misma tanto cuando los *shocks* son negativos como cuando son positivos pero de igual magnitud.

Los modelos propuestos [3] y [5] permiten estudiar la importancia de la duración condicional (\mathbf{y}_i) y el volumen de negociación de las transacciones (V_i) en el proceso de ajuste de los precios, pues consideramos que incorporan información relevante sobre el mercado en la evolución de la volatilidad condicional de las rentabilidades. Si el

efecto sobre los precios es debido a la ausencia de negocio, esto significaría que el tiempo no es exógeno al proceso de ajuste de los precios. Las implicaciones del papel del tiempo en la conducta de los precios sugiere que si, por ejemplo, los precios son martingalas, el conocimiento de la secuencia de precios no es un estadístico suficiente para toda la información del mercado. Sin embargo, si un investigador examina únicamente los precios de las transacciones, entonces el cómo los precios se mueven hacia $t+1$ no es independiente de todas aquellas transacciones precedentes. Ya que los agentes informados son más comunes que los desinformados, la negociación es también más común cuando ha ocurrido algún evento informativo. De esta forma, cuando existe nueva información el proceso de valoración es mayor. Engle y Russell (1997) demuestran que la volatilidad de los precios por unidad de tiempo está inversamente relacionada con la duración de las transacciones¹⁵. Diamond y Verrecchia (1987) sugieren que las restricciones de ventas al descubierto pueden inducir una correlación negativa entre la frecuencia de la negociación y el movimiento de los precios. En este sentido, cabe esperar que el parámetro h_1 sea negativo. Por otro lado, cabe resaltar que una importante implicación empírica del modelo AP es que los agentes privados provocan una más elevada volatilidad del precio durante los periodos en que existe una alta actividad de negociación. Sus resultados teóricos son a menudo citados en conexión con el volumen observado y el agrupamiento o *clustering* de volatilidad en los mercados financieros¹⁶. Aunque la asociación entre volatilidad y actividad de mercado es bien conocida, existen pocos trabajos que la cuantifiquen o modelos que la relacionen formalmente. Easley y O'Hara (1992) sugieren que las tasas de transacciones más frecuentes están debidas a una gran fracción de agentes informados. En un modelo de

¹⁵ Engle y Russell (1997) muestran que la relación entre las duraciones de precios y la volatilidad es aproximadamente igual a $S^2 = c^2/Y_t$, siendo c una constante. Esto significa que la varianza por unidad de tiempo es inversamente proporcional a la duración.

¹⁶ El modelo AP no se dirige específicamente hacia los patrones sistemáticos en la llegada de noticias, o al impacto potencial de las noticias sobre la distribución de información privada en el tiempo. De acuerdo a esto, el patrón de los ajustes de precios y la volatilidad observada en un mercado particular está determinada conjuntamente por los patrones de llegada de información discutida en AP. Por ejemplo, cuando los agentes no tienen una desventaja informativa, los precios se ajustan a las noticias sobre su publicación/divulgación. Con la información asimétrica, los precios se ajustan a la información privada antes de que las noticias sean anunciadas, completan su ajuste antes de que las noticias lleguen y en la ausencia de más noticias en periodos subsecuentes de negociación, la volatilidad declina. La anterior discusión sugiere una estrategia para la valoración de las relaciones empíricas entre la llegada de información pública y privada y la volatilidad de mercado. Un aspecto de la estrategia es contrastar el resultado del modelo AP relacionando la actividad de mercado al *clustering* de la volatilidad: *ceteris paribus*, la volatilidad tiende a variar directamente con la actividad de mercado como una consecuencia de que existen agentes informados.

expectativas racionales, los especialistas conocen esto y harán que los precios sean más sensibles al flujo de órdenes cuando las transacciones sean más frecuentes. En este sentido, cabe esperar que $h_2 > 0$.

La estimación se realiza por máxima verosimilitud, donde la función de densidad normal condicional es:

$$f(v_i/\Psi_{i-1}) = (2\mathbf{ps}_i)^{-\frac{1}{2}} \exp\left(-\frac{v_i^2}{2\mathbf{s}_i^2}\right),$$

siendo Ψ_{i-1} el conjunto de información en $i-1$; y la función del logaritmo de verosimilitud bajo normalidad condicional del error es igual a:

$$\text{Log } L(\mathbf{q}) = -\frac{1}{2} \sum_{i=1}^N \left[\log(\mathbf{s}_i^2) + \frac{v_i^2}{\mathbf{s}_i^2} \right],$$

siendo \mathbf{q} el vector de parámetros y N el tamaño muestral. En general, el algoritmo numérico de optimización utilizado para la estimación de los parámetros es Berndt, Hall, Hall y Hausman [BHHH, 1974]. Sin embargo, a veces es difícil justificar el supuesto de normalidad condicional, entonces la función $L(\mathbf{q})$ puede estar mal especificada. En este sentido, las estimaciones asintóticas de \mathbf{q} son cuasi-máximo verosímiles, y los parámetros serán consistentemente estimados incluso si es violado el supuesto de normalidad condicional. El método empleado es el de Bollerslev y Wooldridge, quienes aplican el procedimiento cuasi-máximo verosímil para procesos ARCH, obteniendo una estimación consistente y robusta asintóticamente de la matriz de varianzas y covarianzas de los parámetros del proceso¹⁷.

El cuadro 5 muestra las estimaciones del modelo ISAR(1)-GARCH(1,1) para las rentabilidades intradiarias del futuro del Ibex35. En él, se consideran tres tipos de modelos. El Modelo I, considera que ni la intensidad condicional ni el volumen de negociación tienen un carácter informativo sobre los cambios que experimentan los precios y que medimos por la volatilidad condicional. En el Modelo II, sólo se considera el carácter informativo de la intensidad esperada de la negociación; mientras que en el Modelo III se considera el contenido informativo tanto de la intensidad condicional como del volumen de negociación de todas las transacciones realizadas en cada periodo.

¹⁷ Esta escasa habilidad del modelo ARCH (o de sus extensiones), para pasar un contraste de diagnóstico simple, ha conducido a la utilización de distribuciones condicionadas más generales que la distribución Normal. Así, por ejemplo, con el fin de eliminar el alto grado de leptocurtosis de la distribución del error, Bollerslev (1987) propuso la distribución t -Student condicional, que posee colas más gruesas que la distribución normal, donde los grados de libertad tienden a ∞ , incluyendo la distribución Normal como caso límite.

Por otro lado, dentro de cada uno de los modelos I, II y III, se consideran diferentes combinaciones de las funciones de respuesta. Concretamente, tres para el modelo ISAR ($\Phi_j[\Delta^j t_i]$) y tres correspondientes al modelo GARCH ($A_j[\Delta^j t_i]$). Por motivos de espacio y dado que los resultados obtenidos muestran que dichas especificaciones no cumplen las condiciones de estacionariedad de la varianza condicional, se han omitido las estimaciones de la función de respuesta constante y sus combinaciones, así como las combinaciones de funciones de respuesta exponencial-exponencial. Por lo tanto, se han incluido solamente los resultados de las combinaciones exponencial-recíproco (que denominamos *Exp-Rec*) y recíproco-recíproco (que denominamos *Rec-Rec*).

Los parámetros han sido estimados por cuasi máxima verosimilitud, aunque bajo el supuesto de normalidad condicional, mediante el procedimiento de Bollerslev y Wooldridge. El periodo total ha sido dividido en dos partes. Una primera submuestra de estimación que contiene 71369 observaciones, y una segunda submuestra para evaluar la predicción de la rentabilidad que contiene las últimas 100 observaciones.

En cuanto a los resultados de las estimaciones, pueden apreciarse las siguientes cuestiones. En primer lugar, los contrastes habituales de selección de los órdenes de los modelos, como los criterios de información de Akaike y Schwartz, muestran que existe una escasa estructura regular en las rentabilidades y también en su varianza condicional, pudiéndose representar por órdenes unitarios para ambos polinomios de retardos. Por otro lado, se considera que el valor j en las funciones de respuesta es unitario, considerando que la duración entre transacciones es consecutiva e igual a 1 *tick*. En segundo lugar, y dada la existencia de *outliers*, se han introducido dos variables ficticias para que controlasen la elevada variabilidad de algunas rentabilidades intradía. Estas variables son: D_{1i} , que vale 1 si la rentabilidad es superior a $4\mathbf{s}_r$ (siendo \mathbf{s}_r la desviación estándar histórica de la rentabilidad) y cero en caso contrario; y D_{2i} que vale 1 si la rentabilidad es menor a $-4\mathbf{s}_r$ y cero en caso contrario. Ambos coeficientes son significativos indicando un importante efecto de las observaciones anómalas o rentabilidades atípicas en la muestra, tanto las que son de signo positivo como las que son de signo negativo. En tercer lugar, los coeficientes estimados son bastante estables entre las diferentes combinaciones de modelos y, en el caso de la varianza condicional, cumplen con las condiciones de signo.

Cuadro 5

Estimaciones cuasi máximo verosímiles del modelo ISAR(p)-GARCH(k,q)

	Modelo I		Modelo II		Modelo III	
	<i>Exp-Rec</i>	<i>Rec-Rec</i>	<i>Exp-Rec</i>	<i>Rec-Rec</i>	<i>Exp-Rec</i>	<i>Rec-Rec</i>
ϕ_0	0.00029 (1.38)	0.00029 (1.72)	0.00029 (1.77)	0.000297 (1.59)	0.00029 (1.50)	0.00029 (1.74)
μ_1	-0.3408 (-7.01)	-0.3408 (-7.12)	0.341 (-8.56)	-0.341 (-7.46)	-0.341 (-7.66)	-0.341 (-6.89)
μ_2	0.2753 (5.18)	0.2753 (5.58)	0.275 (5.39)	0.276 (6.79)	0.276 (5.49)	0.275 (5.67)
ϕ_{01}	-0.3121 (-19.9)	-0.3478 (-25.78)	-0.312 (-35.9)	-0.348 (-19.5)	-0.313 (-28.2)	-0.348 (-33.1)
ϕ_{11}	0.4861 (6.97)	0.2182 (11.62)	0.486 (9.35)	0.218 (8.80)	0.486 (8.57)	0.218 (12.4)
ω	0.00029 (5.54)	0.000296 (5.93)	0.000304 (5.14)	0.0003047 (20.1)	0.000308 (5.49)	0.000306 (5.07)
α_{01}	0.1796 (4.48)	0.1784 (4.51)	0.182 (4.00)	0.182 (7.06)	0.184 (4.21)	0.183 (4.63)
α_{11}	0.1313 (6.98)	0.1323 (2.31)	0.126 (1.98)	0.127 (1.90)	0.124 (1.89)	0.125 (2.92)
β_1	0.6695 (16.8)	0.6711 (18.49)	0.670 (15.7)	0.67 (54.9)	0.666 (16.35)	0.667 (16.1)
η_1	--	--	-0.0000014 (-1.64)	-0.00000145 (-1.43)	-0.00000837 (-2.14)	-0.0000083 (-1.96)
η_2	--	--	--	--	0.00000759 (2.55)	0.0000076 (2.45)
\mathbf{s}_v	0.05584	0.05581	0.05584	0.05581	0.05584	0.055809
Log L	184974.7	184992.6	184976.8	184994.6	184994.7	185012.5
<i>Asimetría</i>	0.1636	0.1549	0.1689	0.1741	0.1943	0.1914
<i>No linealidad positiva</i>	-1.7002	-1.6989	-1.7144	-1.6951	-1.7007	-1.6991
<i>No linealidad negativa</i>	1.3282	1.3388	1.3294	1.3163	1.2989	1.309
Evaluación de la capacidad predictiva						
EAM	0.045428	0.045531	0.045428	0.045515	0.045428	0.045516
EPAM	118.7233	118.9250	118.7230	118.8591	118.7229	118.8620
RECM	0.065235	0.065074	0.065235	0.065358	0.065236	0.065359
U-Theil	0.863029	0.869289	0.863031	0.864883	0.863044	0.864795
DA	-5.765993	-5.765993	-5.765993	-5.765993	-5.765993	-5.765993

Nota: Este cuadro contiene las estimaciones de los modelos:

$$r_i = \mathbf{f}_0 + \sum_{s=1}^2 \mathbf{m}D_s + \sum_{j=1}^p \Phi_j [\Delta^j t_i] r_{i-j} + v_i; \quad v_i / \Psi_{i-1} \sim N(0, \mathbf{s}_i^2); \quad i = 1, \dots, N.$$

$$\mathbf{s}_i^2 = \mathbf{w} + \sum_{j=1}^q A_j [\Delta^j t_i] v_{i-j}^2 + \sum_{r=1}^k \mathbf{b}_r \mathbf{s}_{i-r}^2 + \mathbf{h}y_i + \mathbf{h}_2 V_i.$$

Las estimaciones de los modelos se han realizado considerando que $p=1, j=1$ y $q=k=1$, en todos los casos. El periodo de predicción contiene 71369 observaciones, mientras que el periodo de predicción comprende las 100 últimas observaciones. D_{1i} y D_{2i} representan variables ficticias binarias. Estas se han añadido por conveniencia para recoger el efecto de las observaciones atípicas que poseen un signo positivo (D_1) y las que poseen signo negativo (D_2). El criterio que utilizamos para decidir si una observación es atípica es ± 4 desviaciones estándar (σ) de la rentabilidad realizada. Se utiliza el método de Bollerslev y Wooldridge para obtener estimaciones robustas de los errores estándar. Entre paréntesis aparecen los valores t -Student robustos.

Así, por ejemplo, en términos de los coeficientes de la función $\Phi[\Delta t_i]$ para el modelo ISAR(1), f_{01} es siempre negativo, indicando la existencia de correlación serial negativa para los datos de alta frecuencia, y siendo consistente con el “hecho estilizado” encontrado en los datos de elevada frecuencia que muestra como están negativamente correlacionados. Por otro lado, el coeficiente de decaimiento f_{11} es siempre positivo tanto para la función exponencial como recíproca, indicando un rápido decrecimiento de dicha función puesto que el valor es inferior a 0.5.

Por otro lado, y en términos del modelo GARCH(1,1), la función $A[\Delta t_i]$ posee coeficientes positivos para a_{01} y el parámetro de decaimiento a_{11} , cuyos valores estimados son pequeños. En este sentido, también la duración afecta a la volatilidad en idéntico sentido que a la rentabilidad. En todos los casos del cuadro 4, la condición de estacionariedad debe cumplirse que $a_{01} + a_{11}d_i^{-1} + b_1 < 1$. Esto implica que existe una dependencia del espacio de los parámetros y también respecto de d_i . Nótese que, dado que se han eliminado las duraciones iguales a cero, la condición de estacionariedad se cumpliría para toda observación i -ésima, pues $d_i \geq 1$ en nuestra muestra.

En cuarto lugar, el diagnóstico y evaluación intramuestral de los modelos puede realizarse mediante varios contrastes. Por un lado, comparando los diferentes modelos *Exp-Rec* mediante el contraste de razón de verosimilitudes: $RV = -2[\log L_0 - \log L_1]$, siendo $\log L_0$ el logaritmo de verosimilitud bajo la hipótesis nula y $\log L_1$ el correspondiente al modelo bajo la hipótesis alternativa, y distribuyéndose como χ^2_q , siendo q el número de grados de libertad. Así, el modelo que incluye conjuntamente la duración condicional y el volumen de negociación (es decir, el modelo III) es preferible a quien no lo incluye (modelo I) o a quien sólo incluye la duración condicional (modelo II). Los resultados se pueden resumir en el cuadro 6:

Cuadro 6. Ratios de verosimilitud.

Modelo	RV₁	RV₂
<i>Exp-rec</i>	40	35.8
<i>Rec-rec</i>	39.8	35.8
q	2	1
$\chi^2_{5\%,q}$	5.99	3.84

Como hipótesis alternativa se considera al modelo III. RV_1 es el ratio de verosimilitudes entre el modelo III y el modelo I; mientras que RV_2 es la razón de verosimilitudes entre el modelo III y el modelo II. q es el número de restricciones y $\mathbf{C}_{5\%,q}^2$ son los valores críticos para q .

La conclusión es idéntica para los modelos *Rec-Rec*. Por otro lado, mediante el error estándar de la regresión observamos ciertas similitudes entre los modelos. Finalmente, se añaden los estadísticos propuestos por Engle y Ng (1993), basados en la curva de impacto de las noticias, los cuales tienen, según estos autores una mayor potencia en la detección de especificaciones erróneas relacionadas con los efectos asimétricos. Estos contrastes son denominados de asimetría, no linealidad positiva y no linealidad negativa. En el cuadro 4 aparecen los coeficientes *t*-Student del cociente de la regresión entre los residuos estandarizados (\hat{v}_i/\mathbf{S}_i) sobre la variable cualitativa $S_{i-1}^- = 1$ si $v_i < 0$; sobre el producto $(1 - S_{i-1}^-)v_{i-1}$, y sobre el producto $S_{i-1}^-v_{i-1}$. Los resultados de los contrastes parecen indicar que los modelos están correctamente especificados al nivel de significación del 5%, si bien al 10%, y sólo en términos del contraste de no linealidad positiva, no resultan correctamente especificados los modelos. Por otro lado, y aunque no se presentan en este cuadro, podemos añadir que los excesos de curtosis en los residuos estandarizados son menores que los residuos estimados, lo que indicaría igualmente una especificación correcta del modelo de varianza condicional.

En quinto lugar, la interpretación de los resultados de los coeficientes \mathbf{h}_1 y \mathbf{h}_2 es coherente con las consideraciones hechas en el epígrafe 3.1.. Así, por ejemplo, considerando el contenido informativo del tiempo esperado y volumen de transacciones (modelo II y III, respectivamente), podemos decir que, en todos los casos, \mathbf{h}_1 es negativo, es decir, \mathbf{S}_i^2 (o varianza por segundo del cambio en los precios) está inversamente relacionada con la duración condicional estimada \mathbf{y}_i (o intensidad esperada de la actividad de negociación), si bien es sólo significativo en el modelo III al 5%, y al 10% en el modelo II. Este hecho es consistente con las implicaciones que se derivan de Diamond y Verrecchia (1987) y Engle y Russell (1997), justificando su contenido informativo. Por otro lado, y considerando el modelo III, observamos como \mathbf{h}_2 es positivo, lo que indica que cuando aumenta el volumen aumenta la volatilidad de la rentabilidad, y viceversa. Por lo tanto, la volatilidad tiende a variar con la actividad de mercado, aspecto éste que liga fácilmente con los resultados obtenidos en múltiples

trabajos sobre la relación entre volumen y precios en línea con Clark(1973). Por otro lado, podemos destacar que un ajuste de regresión entre el volumen de negociación y las rentabilidades positivas ($r_i > 0$) y negativas ($r_i < 0$), construidas como variables ficticias, las rentabilidades negativas incrementan más el volumen de negociación que las rentabilidades positivas¹⁸. Sin embargo, esta asimetría en el comportamiento del volumen no resultó significativa cuando se incluyó en el modelo de GARCH(1,1). Además, cabe resaltar que tampoco resultó significativo estadísticamente el efecto asimétrico sobre la volatilidad de la existencia de rentabilidades negativas y positivas, cuando se consideró la estimación de un modelo ISAR(1)-GJR(1,1), siendo GJR acrónimo del modelo de Glosten, Jagannathan y Runkle (1993).

Finalmente, y en sexto lugar, observando como el modelo III parece preferible al resto de modelos - a tenor de los resultados obtenidos en el periodo de estimación – comparamos todos los modelos, evaluando la predicción de las rentabilidades. En concreto, realizamos una comparación competitiva de las predicciones de los modelos estimados. El periodo de predicción elegido consta de los 100 últimos periodos de la muestra. Las predicciones se obtienen con horizonte de un periodo (*one-step-ahead*), asumiendo que se conoce la duración observada de las transacciones. Para evaluar la capacidad predictiva de los distintos modelos empleamos varios estadísticos de la bondad de la predicción. Las medidas de bondad que utilizamos en este trabajo se basan en $h=1, \dots, 100$ periodos de predicción para r_i , que denominaremos por \hat{r}_i . En primer

lugar, los estadísticos del error absoluto medio: $EAM = \frac{1}{100} \sum_{h=1}^{100} |r_h - \hat{r}_h|$; el error

porcentual absoluto medio: $EPAM = \frac{1}{100} \sum_{h=1}^{100} |r_h - \hat{r}_h| / r_h \times 100$; la raíz cuadrada del error

cuadrático medio: $RECM = \sqrt{\frac{1}{100} \sum_{h=1}^{100} (r_h - \hat{r}_h)^2}$; el estadístico:

$U - Theil = RECM / \left[\sqrt{\frac{1}{100} \sum_{h=1}^{100} r_h^2} + \sqrt{\frac{1}{100} \sum_{h=1}^{100} \hat{r}_h^2} \right]$. Por otro lado, destacamos el contraste

de predicción direccional de los cambios (o de independencia de las predicciones y de

¹⁸ Es decir, cuando las rentabilidades son negativas, la variación esperada del volumen es mayor. Además, cuando las rentabilidades muestrales son negativas (49206 periodos), la variación del volumen es negativa en 32438 (es decir, un 65.92% de las veces). Por otro lado, cuando la rentabilidad es positiva (22263 periodos), la variación negativa del volumen sólo existe en 14683 (es decir, un 65.95%). Por lo tanto, las proporciones muestrales de la variación negativa del volumen y la rentabilidad –positiva o negativa -son casi idénticas.

los valores actuales de los precios) de Pesaran y Timmermann (DA, 1992). Bajo la hipótesis nula los valores reales y predichos son independientes. En este caso, la distribución del estadístico es $N(0,1)$. El procedimiento de contraste es el siguiente:

$$DA = [\text{var}(SR) - \text{var}(SRI)]^{-0.5}(SR - SRI), \quad \text{donde} \quad SR = 100^{-1} \sum_{h=1}^{100} I_i[y_h \cdot \hat{y}_h > 0],$$

$$SRI = p\hat{p} + (1-p)(1-\hat{p}), \text{ siendo } SRI \text{ es el ratio de éxito en el caso de independencia}$$

entre los valores reales y los predichos bajo la hipótesis nula. El resto de elementos son iguales a: $p = 100^{-1} \sum_{h=1}^{100} I_i[y_h > 0]$, $\hat{p} = 100^{-1} \sum_{h=1}^{100} I_i[\hat{y}_h > 0]$, $\text{var}(SR) = 100^{-1}[SRI(1-SRI)]$ y $\text{var}(SRI) = 100^{-2}[H(2p-1)^2 p(1-p) + (2\hat{p}-1)^2 \hat{p}(1-\hat{p}) + 4p\hat{p}(1-p)(1-\hat{p})]$.

Los resultados no muestran una preferencia clara de algún modelo atendiendo a los criterios analizados: *EAM*, *EPAM*, *RECM*, *U* de Theil y contraste *DA*. Por ejemplo, en términos de los modelos denominados *Exp-Rec*, tenemos que los criterios *EAM* permanecen iguales en todos los modelos, para *EPAM* es preferible el modelo III, para *RECM* es preferible el modelo I y II, y para el contraste *DA* se rechaza la hipótesis nula en todos los casos. En el caso del modelo *Rec-Rec*, las preferencias son diferentes. Por ejemplo, según *EAM* y *EPAM* es preferible el modelo II; para *RECM* y *U* de Theil, es el modelo III y, finalmente, para el contraste *DA*, también se rechaza la hipótesis nula. De todo ello, se deduce que existe un conflicto importante para destacar cual es el mejor modelo en términos predictivos, lo que sugiere en última instancia el carácter aleatorio del movimiento de precios y la escasa predecibilidad de las rentabilidades utilizando esta clase de modelos.

4. Conclusiones.

En este trabajo se ha analizado la dinámica subyacente en el tiempo entre transacciones consecutivas, así como su importancia para explicar los cambios en los precios utilizando datos intradiarios espaciados irregularmente del futuro del Ibex35.

En primer lugar, los resultados muestran que existe un modelo que permite caracterizar la dinámica de la duración condicional o intensidad de la actividad de negociación del futuro del Ibex35. Este modelo es un modelo de duración condicional autorregresivo. Además, bajo la distinta consideración de la densidad del tiempo entre llegadas, se comparan los resultados de las estimaciones usando diferentes funciones de

pseudo verosimilitud que tratan de recoger el carácter de los datos muestrales y estimar, de forma cuasi máximo verosímil, los parámetros de dichos modelos. A pesar de que no se comparan directamente las funciones de pseudo logaritmo verosimilitud, un modelo significativo que anida al caso Exponencial y Weibull, como es la función de densidad Gamma Generalizada, parece adaptarse estadísticamente a los datos de duración observados, dada la significación de los coeficientes.

En segundo, y último lugar, usando los modelos empíricos para datos irregularmente espaciados de tipo ISAR-GARCH, que conjuntamente modelizan las rentabilidades y volatilidades condicionales del futuro del Ibex35, permiten concluir que tanto la intensidad esperada de la actividad de negociación como el volumen de las transacciones realizadas poseen un contenido informativo en la explicación de la dinámica de la varianza condicionada por segundo de las rentabilidades de dicho activo.

Referencias bibliográficas

- Admati, A. and Pfleiderer, P. (1988): "A theory of Intraday Patterns: Volume and Price Variability". *The Review of Financial Studies*, 1, 3-40.
- Bauwens, L. and Giot, P. (2000): "The Logarithmic ACD Model: An Application to the Bid-Ask Quote Process of Three NYSE Stocks". *Annales d'Economie et de Statistique*, 60, 117-149.
- Black, F. (1976): "Studies in Stock Price Volatility". American Statistical Association. Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section, 177-181.
- Blume, L.E., Easley, D. and O'Hara, M. (1994): "Market Statistics and Technical Analysis: The Role of Volume". *Journal of Finance*, 49, 153-182.
- Bollerslev, T. (1987): "A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return". *Review of Economics and Statistics*, 69, 542-547.
- Campbell, J., Grossman, S.J. and Wang, J. (1991): "Trading Volume and Serial Correlation in Stock Returns". *Working Paper*, Princeton University.
- Clark, P.K. (1973): "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices". *Econometrica*, 41, 135-156.
- Dacorogna, M., Müller, U., Nagler, R., Olsen, R. and Pictet, O. (1993): "A Geographical Model for the Daily and Weekly Seasonal Volatility in the FX market". *Journal of International Money and Finance*, 12(4), 413-438.
- Diamond, D. and Verrecchia, R. (1987): "Constraints on Short-Selling and Asset Price Adjustments to Private Information". *Journal of Financial Economics*, 18, 277-311.
- Easley, D. and O'Hara, M. (1992): "Time and the Process of Security Price Adjustment". *Journal of Finance*, 47, 577-606.
- Engle, R. (1996): "The Econometrics of Ultra-high Frequency Data". DP 96-15. University of California, San Diego. También, como documento NBER, Working Paper, 5816.

- Engle, R.F. and Ng, V. (1993): "Measuring and Testing the impact of News on Volatility". *Journal of Finance*, 48, 1749-1778.
- Engle, R. and Russell, J. (1997): "Forecasting the Frequency of Changes on Quoted Foreign Exchange Prices with the Autoregressive Conditional Duration Model". *Journal of Empirical Finance*, 4, 187, 212.
- Engle, R. and Russell, J. (1998): "Autoregressive Conditional Durations: A New Model for Irregularly Spaced Data". *Econometrica*, 66(5), 1127-1162.
- Epps, T. W. and Epps, M. L. (1976): "The Stochastic Dependence of Security Price Change and Transaction Volumes: Implications for the Mixture-of-Distributions Hypothesis". *Econometrica*, Vol. 44, pp. 305-322.
- Foster, F. and Viswanathan, S. (1993): "Variations in Trading Volume, Return Volatility, and Trading Costs: Evidence of Recent Price Formation Models". *Journal of Finance*, 48, 187-211.
- Ghysels, E. and Jasiak, J. (1994): "Garch for Irregularly Spaced Financial Data: The ACD-GARCH Model". C.R.D.E., Université de Montreal, unpublished manuscript.
- Ghysels, E., Gouriéroux, C and Jasiak, J. (1998): "Stochastic Volatility Duration Models". *Working Paper*, Pennsylvania State University.
- Glosten, L. and Milgrom, P. (1985): "Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneous Informed Traders". *Journal of Financial Economics*, 13, 71-100.
- Glosten, L, Jagannathan, R. and Runkle, D. (1993): "Relationship Between the Expected Value and the Volatility of the Nominal Excess Return of Stocks". *Journal of Finance*, 48, 1779-1801.
- Godhart, C. and O'Hara, M. (1997): "High Frequency Data in Financial Markets Issues and Application". *Journal of Empirical Finance*, 4, 73-114.
- Grammig, J. and Maurer, K. (2000): "Non-Monotonic Hazard Functions and the Autoregressive Conditional Duration Model". *Econometrics Journal*, 3(1), 16-39.
- Grammig, J. and Wellner, M. (2002): "Modeling the Interdependence of Volatility and Inter-Transaction Duration Processes". *Journal of Econometrics*, 106(2), 369-401.
- Hausman, J., Lo, A. and MacKinlay, C. (1992): "An Ordered Probit Analysis of Transaction Stock Prices". *Journal of Financial Economics*, 31, 319-379.
- Holden, C. and Subrahmanyam, A. (1992): "Long-lived Private Information and Imperfect Competition". *Journal of Finance*, 47, 247-270.
- Jackson, M. (1991): "Equilibrium Price Formation, and the Value of Private Information". *Review of Financial Studies*, 4, 1-16.
- Kalbfleisch, J. and Prentice, R. (1990): *The Statistical Analysis of Failure Time Data*. New York: John Wiley & Sons.
- Karpoff, J. (1987): "The Relation Between Price Change and Trading Volume. A Survey". *Journal of Financial and Quantitative Analysis*, 22, 109-126.
- Kim, O. and Verrecchia, R. (1991): "Market Reactions to Anticipated Announcements". *Journal of Financial Analysis*, 30, 273-310.
- Kyle, A. (1985): "Continuous Auctions and Insider Trading". *Econometrica*, 53, 1315-1336.
- Kyle, A. (1989): "Informed Speculation with Imperfect Competition". *Review of Economics Studies*, 56, 317-355.
- Mandelbrot, B. and Taylor, H. (1967): "On the Distribution of Stock Prices Differences". *Operation Research*, 15, 1057-1062.

- O'Hara, M. (1995): *Market Microstructure Theory*. Blackwell Publishers, Massachusetts.
- Pai, J., Polasek, W. and Kozumi, H. (1995): "Irregularly Spaced AR and ARCH (ISAR-ARCH) Models". WWZ-Discussion Papers, N° 9509. University of Basel.
- Pesaran, M.H. and Timmermann, P. (1992): "A Simple Nonparametric Test of Predictive Performance". *Journal of Business and Economic Statistics*, 10, 461-465.
- Pfleiderer, P. (1984): "The Volume of Trade and the Variability of Prices: A Framework for Analysis in Noisy Rational Expectations Equilibria". *Working Paper*, Stanford University.
- Polasek, W. and Pai, J. (1996): "ISAR-ARCH(p,q) Processes for Irregularly Spaced Time Series". *Working Paper*, University of Basel.
- Rochet, J. and Vila, J. (1994): "Insider Tradings and Market Manipulations: Existence and Uniqueness of Equilibrium". *Review of Economic Studies*, 61, 131-152.
- Rogalski, R.J. (1978): "The Dependence of Prices and Volume". *Review of Economics and Statistics*, 60, 268-274.
- Russell, J. and Engle, R. (1998): "Econometrics Analysis of Discrete-Valued Irregularly-Spaced Financial Transactions Data Using a New Autoregressive Conditional Multinomial Model". Working Paper Universidad de Chicago.
- Sharma, J.L., Mougoue, M. and Kamath, R. (1996): "Heteroscedasticity in Stock Market Indicator Return Data: Volume versus GARCH Effects". *Applied Financial Economics*, 6, 337-342.
- Stock, J. (1998): "Estimating Continuous Time Process Subject to Time Deformation". *Journal of the American Statistical Association*, 83, 77-85.
- Tauchen, G. and Pitts, M. (1983): "The Price Variability-Volume Relationship on Speculative Markets". *Econometrica*, 51, 485-505.
- Vechiatto, W. (1999): "New Models for Irregularly Spaced Time Series Analysis with Application to High Frequency Financial Data". *Working Paper Banca Intesa*, n° 1.
- Wang, J. (1994): "A Model of Competitive Stock Trading Volume". *Journal of Political Economy*, 102, 127-168.
- Zhang, M., Russell, J. and Tsay, R. (2001): "A Nonlinear Autoregressive Conditional Duration Model with Applications to Financial Transaction Data". *Journal of Econometrics*, 104, 179-207.

Abstract:

In this paper we analyze the intensity of trading activity and trading volume using data on index future Ibex35 that arrives at irregular intervals of time. We contribute to the analysis in two aspects. First, we consider a conditional duration model on past arrival times, and use deseasonalized durations and some different pseudo likelihood functions to compute quasi-maximum likelihood estimates of the parameters. The results show that duration holds information about past and it explains the intensity of trading activity for future index. Second, we specify a model for change prices and future index returns, distinguishing non-informative and informative aspects of expected intensity and trading volume by ISAR-GARCH models for irregularly spaced transaction data. In this case, the results show a negative correlation between conditional volatility and expected durations, and a positive correlation with trading volume, indicating an important informative characteristic from both variables in the dynamic structure for volatility of future index returns.

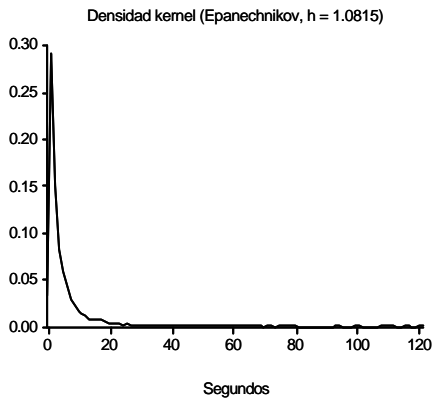
Keywords: Microstructure, ACD models, Weibull, Gamma Generalized, ISAR-GARCH.

JEL: C22, C45, C52.

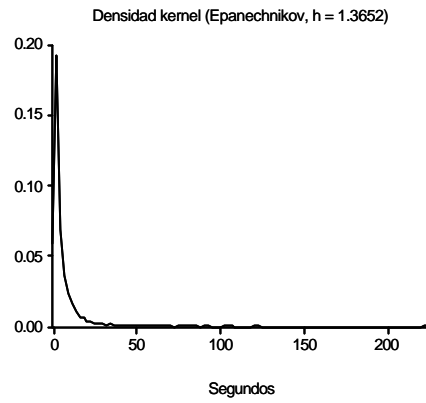
APENDICE I

FRECUENCIAS ESTIMADAS DE LAS DURACIONES PARA CADA DÍA DE NEGOCIACION (28/9/98 HASTA 16/10/98)

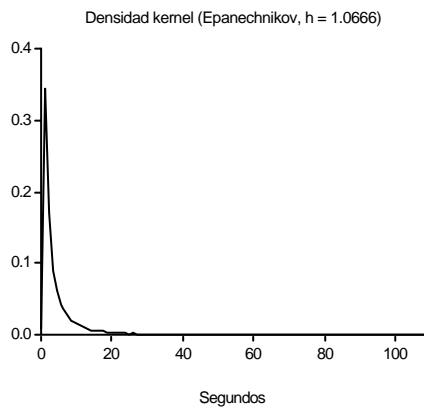
(i) Día 28/9/98



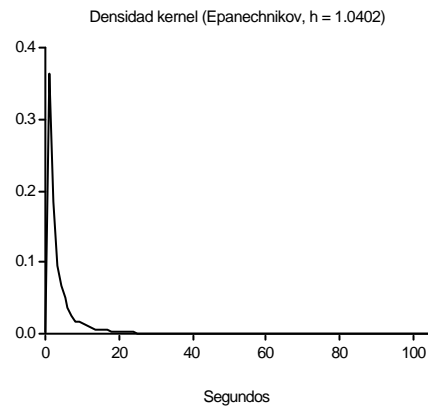
(ii) Día 29/9/98



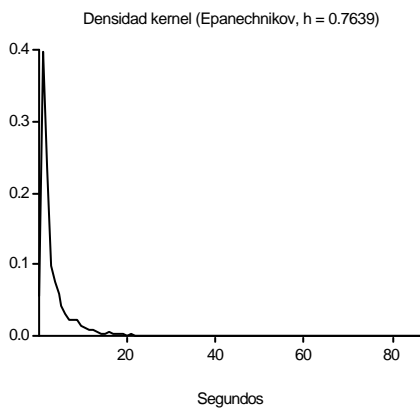
(iii) Día 30/9/98



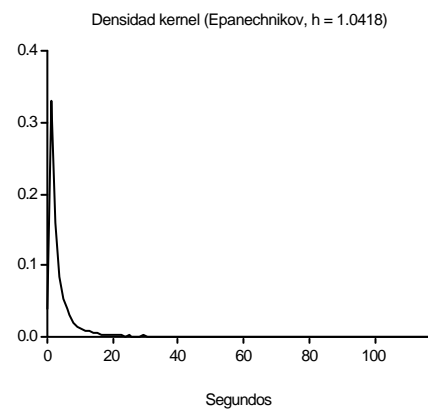
(iv) Día 1/10/98



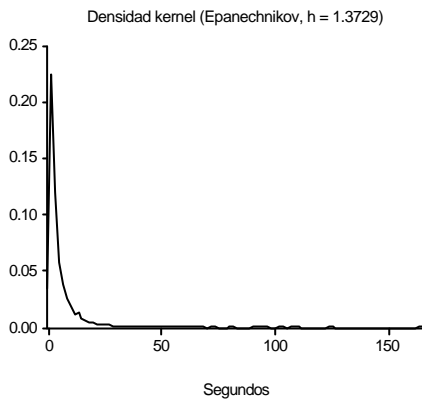
(v) Día 2/10/98



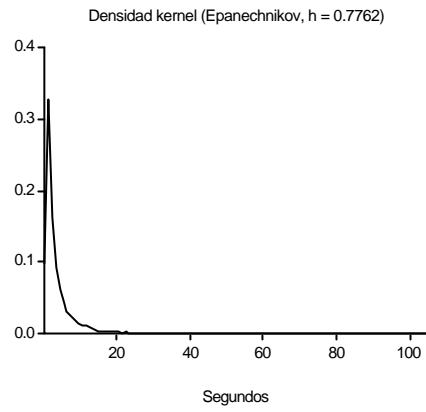
(vi) Día 5/10/98



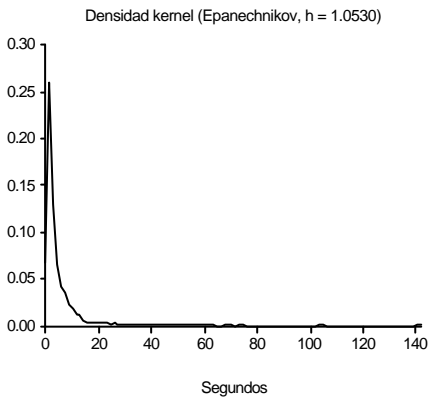
(vii) Día 6/10/98



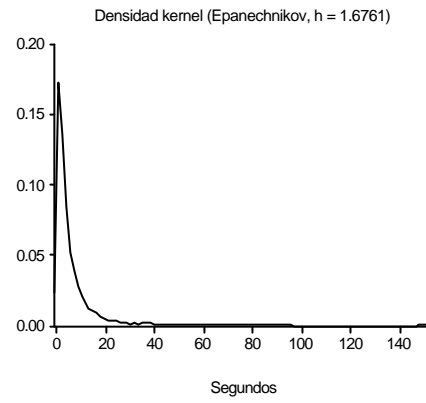
(viii) Día 7/10/98



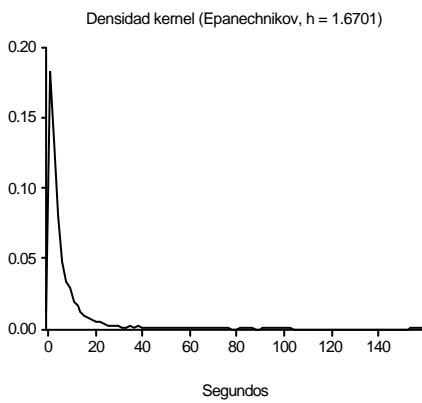
(ix) Día 8/10/98



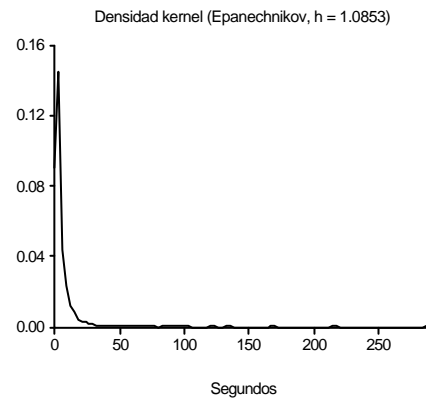
(x) Día 9/10/98



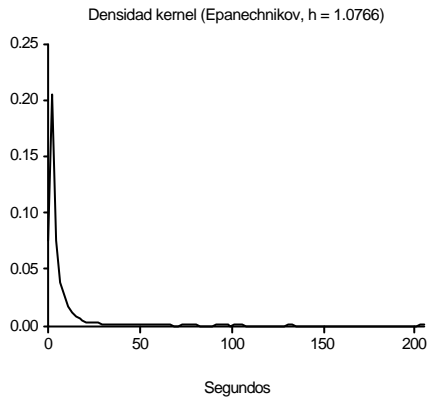
(xi) Día 13/10/98



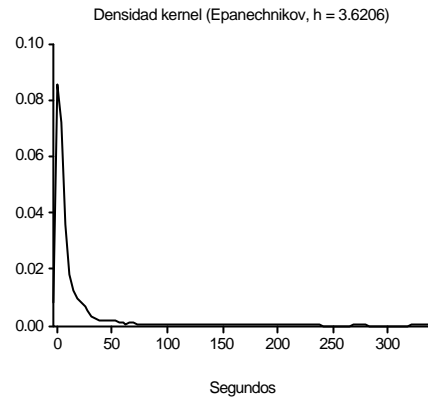
(xii) Día 14/10/98



(xiii) Día 15/10/98



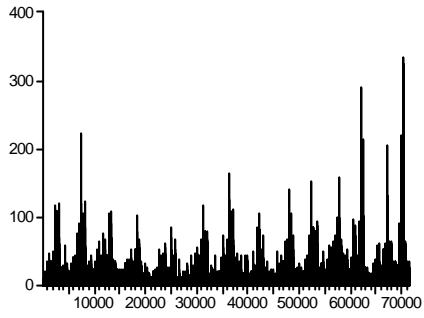
(xiv) Día 16/10/98



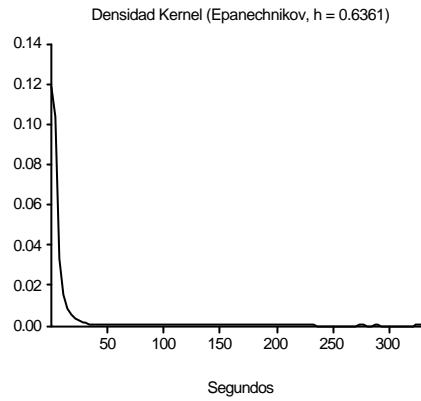
APENDICE II

DURACION, FRECUENCIAS ESTIMADAS Y EVOLUCIÓN DEL INDICE, VOLUMEN Y RENTABILIDAD DEL FUTURO DEL IBEX35

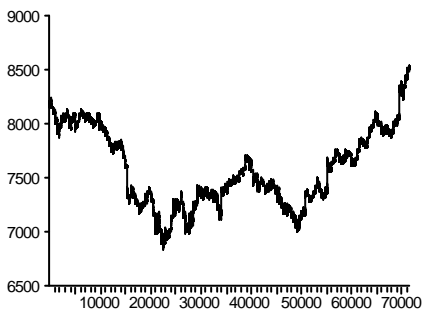
(i) Duración Observada



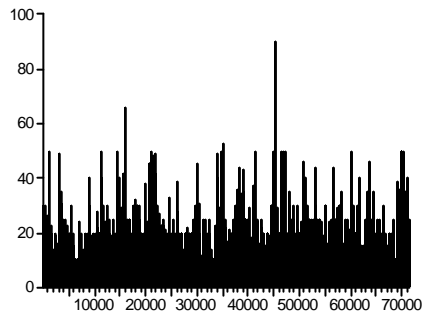
(ii) Densidad de la duración observada



(iii) Evolución del futuro del índice Ibex35



(iv) Evolución del volumen de negociación



(v) Evolución de la rentabilidad

