



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

ARBITRAJE EN MERCADOS FINANCIEROS FRAGMENTADOS

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

PEDRO IGNACIO CONCHA VILLEGAS

PROFESOR GUÍA:  
MARCELA VALENZUELA BRAVO

MIEMBROS DE LA COMISIÓN:  
ALEJANDRO BERNALES SILVA  
PATRICIO VALENZUELA AROS

Este trabajo ha sido parcialmente financiado por Instituto Milenio para la investigación en Imperfecciones de Mercados y Políticas Públicas, ICM IS130002

SANTIAGO DE CHILE

2017



RESUMEN DE LA MEMORIA PARA OPTAR  
AL TÍTULO DE MAGÍSTER EN ECONOMÍA APLICADA  
Y AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL  
POR: PEDRO IGNACIO CONCHA VILLEGAS  
FECHA: 2017  
PROF. GUÍA: SR. MARCELA VALENZUELA BRAVO

## ARBITRAJE EN MERCADOS FINANCIEROS FRAGMENTADOS

Los grandes cambios que se han producido en las últimas décadas en cuanto a la forma de tranzar lo activos financieros ha generado nuevas preguntas, y una gran cantidad de investigación asociada a las consecuencias de estos fenómenos, investigación que está lejos de ser concluyente en la mayoría de los temas abordados y que es de gran interés a la hora de generar políticas o diseños de mercado que favorezcan la mayor calidad de los mismos.

En la presente investigación se estudia el arbitraje que se produce a causa de la fragmentación de los mercados financieros, fenómeno reciente, con datos de alta frecuencia en mercados donde se evidencia la presencia de los llamados inversionistas de alta frecuencia, otro fenómeno de gran relevancia y estudio en las últimas décadas.

Para lo anterior, se utilizan datos de 22 acciones pertenecientes al índice AEX (Amsterdam Exchange Index) durante todo el año 2014, que son tranzadas simultáneamente en más de un mercado. Específicamente, se utilizan los libros de órdenes y transacciones y los libros de órdenes límites (LOB) de los mercados Bats Chi X y Euronext, para las acciones estudiadas.

La metodología utilizada consiste en primero identificar los instantes de arbitraje que se dan entre estos mercados, luego utilizar el tradicional modelo logit, donde en este caso la variable dependiente es igual a uno para aquellos días o intervalos en que existe alguna potencial oportunidad de arbitraje. La variable dependiente es regresionada con un conjunto de medidas asociadas a alguna dimensión del libro de órdenes límites, es decir, con ciertas variables que miden alguna característica de microestructura de mercado.

Los resultados encontrados indican que efectivamente existen ciertas características o ‘ruido de microestructura’ como lo denomina Marshall et al. (2013) [29] que inciden o que predominan cuando ocurren los instantes de potencial arbitraje.

Dentro de los principales resultados, se puede mencionar que la liquidez tiene una importante relación sobre estos eventos, siendo la profundidad la dimensión más importante y que presenta una clara relación negativa con la presencia de estos eventos. Además, no sólo es importante la liquidez en la cima del libro sino también la forma en que la liquidez se distribuye a lo largo del libro de órdenes límites. La volatilidad presente en el libro y la actividad transaccional observada tienen una relación positiva y significativa con la presencia de las oportunidades de potencial arbitraje.



*Con mucho amor y agradecimiento, para mi mamá y el recuerdo de mi mamá.  
Muchas gracias a estas grandes mujeres, por tanto.*



# Agradecimientos

Al finalizar este paso por la Universidad, no puedo sino agradecer a tantas personas que han ayudado a hacer de esta una muy bonita experiencia.

Primero agradecerle a mi mamá, sin su constante esfuerzo, rigurosidad, compañía y amor incondicional, jamás hubiese cumplido mi sueño de estudiar en la Universidad de Chile. Agradecerle a mi mamá, me hubiese encantado que me hubiese visto terminando esta etapa, pero le agradeceré siempre por tanto amor y por influir tanto en quien soy. Agradecerle a mi tata, mis tíos/as y primas por su cariño y por creer en mí, y a la Mía por la compañía e incondicionalidad.

Me siento muy feliz por la gente que pude conocer en la Universidad y agradecido de las experiencias que viví, sin embargo, no puedo dejar de agradecer a esos amigos que conozco hace ya más de 15 años en algunos casos. Muchas gracias a aquellos miembros de Altura09, que a pesar de los caminos distintos que tomamos, de la dificultad de hacer calzar los tiempos, no he dejado de sentir genuina preocupación por el otro, gracias por haberme acompañado en mis momentos de estrés y angustia en la Universidad y por construir un hermoso grupo de amigos. Muchas gracias también a todos aquellos grandes amigos que mantengo del Colegio Santa Cruz.

En mis primeros años en la Universidad me hice de muy buenos amigos, como olvidar esas noches de estudio en San Diego y esas ollas de tallarines, donde claramente sin ustedes esos años de estrés no hubiesen sido recordado hoy como grandes años, gracias a su compañía, muchas gracias plan comuneros. En tercero tuve un pequeño pero importante paso por Minas, y la importancia radica en las personas que conocí ahí, amigos que pese a mi cambio de carrera me siguieron haciendo parte de sus actividades, y hasta el día de hoy conservo muy buenos amigos mineros. También agradezco por la oportunidad que tuve de hacer mi última práctica en Indonesia, y agradezco a los amigos que conocí en una de las mejores experiencias que he vivido, especialmente muchas gracias al Warung.

Finalmente, pero no menos importante, muchas gracias a aquellos amigos que conocí en Industrias. A aquellos amigos que hice con Taller de Industrias, muchas gracias a LOSLA <3 por siempre hacer más alegre mis días en la Universidad y muchas gracias al Gym y el Ñam, sin ustedes los últimos años en la Universidad no hubiesen sido lo mismo. Agradecer a los amigos que conocí en el Magíster, gracias por el constante ánimo que nos dimos y por valorarnos los unos a los otros. También agradecer a todos quienes conocí en el equipo docente de algún ramo, sé que probablemente se me olviden personas, pero realmente a todos quienes estuvieron en algún momento de alguna u otra forma estoy muy agradecido, ya que hicieron de mi paso por la Universidad una excelente experiencia.

También agradezco profundamente a mi profesora guía, Marcela Valenzuela, por su dedicación, tiempo y consejos. Realmente fue una gran guía en este cierre de ciclo. Adicionalmente, agradecer al profesor Patricio Valenzuela, porque debido a su motivación en clases pude re encantarme con la carrera, en un momento que lo necesitaba. Muchas gracias también a quienes trabajan en el Centro de Finanzas, por la disposición y ayuda entregada, y al Instituto Milenio para la investigación en Imperfecciones de Mercados y Políticas Públicas por su financiamiento parcial (ICM IS130002).





# Tabla de Contenido

<b>Introducción</b>	<b>1</b>
<b>1. Literatura relacionada</b>	<b>5</b>
<b>2. Mercados y Datos</b>	<b>9</b>
<b>3. Arbitraje</b>	<b>13</b>
<b>4. Análisis Empírico</b>	<b>20</b>
4.1. Objetivo . . . . .	20
4.2. Dimensiones del libro a considerar . . . . .	21
4.2.1. Liquidez . . . . .	21
4.2.2. Incertidumbre . . . . .	28
4.2.3. Actividad transaccional . . . . .	30
4.2.4. Otros . . . . .	32
4.3. Metodología econométrica . . . . .	33
<b>5. Resultados</b>	<b>35</b>
5.1. Modelo diario . . . . .	35
5.2. Modelo predictivo . . . . .	41
<b>6. Análisis de sensibilidad</b>	<b>46</b>
<b>Conclusión</b>	<b>48</b>
<b>Bibliografía</b>	<b>51</b>
<b>7. Anexos</b>	<b>54</b>

# Índice de Tablas

2.1. Resumen de valores promedios diarios transaccionales de las acciones analizadas.	12
3.1. Duración de oportunidades de arbitraje . . . . .	16
3.2. Potenciales ganancias porcentuales asociadas a las oportunidades de arbitraje	17
3.3. Depth asociado a las oportunidades de arbitraje . . . . .	18
3.4. Máxima ganancia asociada a las oportunidades de arbitraje . . . . .	18
7.1. Valores promedios diarios transaccionales de las acciones analizadas de Bats Chi X. . . . .	54
7.2. Valores promedios diarios transaccionales de las acciones analizadas de Euronext.	55
7.3. Depth como único regresor, logit diario. . . . .	56
7.4. Spread como único regresor, logit diario. . . . .	57
7.5. Distribución de liquidez como único regresor, logit diario. . . . .	58
7.6. Volatilidad como único regresor, logit diario. . . . .	59
7.7. Actividad transaccional como único regresor, logit diario. . . . .	60
7.8. Otras medidas como único regresor, logit diario. . . . .	61
7.9. Logit diario, múltiples controles. . . . .	62
7.10. Depth como único regresor, logit predictivo. . . . .	63
7.11. Spread como único regresor, logit predictivo. . . . .	64
7.12. Distribución de liquidez como único regresor, logit predictivo. . . . .	65
7.13. Volatilidad como único regresor, logit predictivo. . . . .	66
7.14. Actividad transaccional como único regresor, logit predictivo. . . . .	67
7.15. Otras medidas como único regresor, logit predictivo. . . . .	68
7.16. Logit predictivo, múltiples controles. . . . .	69
7.17. Logit diario, análisis de sensibilidad. . . . .	70
7.18. Logit predictivo, análisis de sensibilidad. . . . .	71

# Índice de Ilustraciones

2.1. Volumen y valor nominal de AEX, según mercado . . . . .	11
3.1. Duración oportunidades de arbitraje . . . . .	15



# Introducción

En las últimas décadas se han producido grandes cambios en la forma de tranzar los activos financieros; se ha visto un incremento sustancial en cuanto a las plataformas disponibles a la hora de tranzar un mismo activo financiero, fenómeno llamada fragmentación.

Otro importante cambio ha sido el incremento sustancial de la tecnología utilizada por los agentes involucrados en los distintos mercados financieros.

Paralelamente, los reguladores han respondido a estos cambios implementando nuevas normativas que han buscado alentar la proliferación de nuevas plataformas para tranzar los activos financieros y de esta forma aumentar la competencia. En Estados Unidos fue la normativa RegNMS (Sistema nacional de regulación de mercados) en 2005 y en el caso de Europa fue la normativa MiFID (Directiva sobre mercados financieros) en 2007, las que buscaron adaptarse a la nueva realidad imperante.

La serie de cambios, ha alentado fuertemente el desarrollo de investigación en torno a estos temas relacionados a la microestructura de mercado, investigación que hasta el momento no es concluyente, pues siguen habiendo preguntas sin cerrar en torno a los beneficios o daños a causa de los cambios del último tiempo.

En el marco de la presente investigación se estudia el arbitraje en mercados financieros fragmentados, específicamente se busca entender que características son las que prevalecen o podrían dar paso a la existencia de estos eventos.

Estudiar un fenómeno tan relevante en la economía financiera como lo es el arbitraje es de gran importancia. Lo anterior, debido a que la aparición de eventos de potencial arbitraje está asociado a instantes en que los activos financieros se encuentran mal valorados, lo que puede dar paso a que ciertos agentes del mercado se beneficien a expensas de otros agentes.

Adicionalmente, existe evidencia de que las rentas de arbitraje pueden producir ciertos efectos negativos y no deseados para los mercados financieros. Foucault, Kozhan y Tham (2016) [15] señalan que las oportunidades de arbitraje de corta duración incrementan el costo del riesgo de selección adversa para ciertos inversionistas, por otra parte, Budish, Cramton y Shim (2015) [8] muestran que las rentas de arbitraje perjudican la provisión de liquidez y generan una carrera socialmente ineficiente por aumentar la velocidad de ejecución de las órdenes.

La utilidad derivada de entender el arbitraje en mercados fragmentados va por una parte

en beneficio de los reguladores y también hacia los inversionistas.

Los reguladores, quienes deben generar políticas y diseños de mercado que favorezcan la calidad de los mismos, al entender que características dan paso a las oportunidades de arbitraje pueden generar diseños de mercado tendientes a evitar los efectos negativos asociados a las rentas de arbitraje.

Los inversionistas por su parte al conocer que características dan paso a la aparición de estos eventos, pueden tomar decisiones más informadas a la hora de decidir invertir en un activo financiero o en otro.

Adicionalmente, si se pone en contexto las grandes cantidades de dinero que son tranzadas diariamente en los mercados financieros y la presencia de múltiples intereses globales, puede comprenderse que estas rentas de arbitraje pueden producir un gran daño a la economía global.

Para investigar este fenómeno de arbitraje en mercados fragmentados, se lleva a cabo un estudio empírico en que se utilizan datos de libros de órdenes límites (LOB) y libros de órdenes y transacciones para acciones pertenecientes al índice AEX, índice compuesto de los 25 principales valores tranzados de forma primaria en Euronext Amsterdam, los cuales son tranzadas simultáneamente en más de un mercado. Específicamente, se utilizan 22 acciones pertenecientes al índice AEX, con datos de alta frecuencia que corresponden al año 2014 y tomando en consideración los mercados Bats Chi X y Euronext.

La metodología utilizada para llevar adelante la investigación consiste en una primera etapa identificar las oportunidades de arbitraje que se dan en las acciones estudiadas entre los mercados de Bats Chi X y Euronext. A grandes rasgos, estos instantes ocurren cuando el primer mejor precio de compra en un libro de órdenes límites es mayor al primer mejor precio de venta de la misma acción en el otro mercado.

Luego de identificadas las oportunidades de arbitraje se utiliza el tradicional modelo logit, donde en este caso la variable dependiente es igual a uno para aquellos días o intervalos de tiempo en que ocurren una potencial oportunidad de arbitraje y cero en caso contrario.

Esta variable dependiente es regresionada con un conjunto de medidas, las cuales detectan el estado del libro de órdenes límites con respecto a alguna dimensión del libro en particular.

Para detectar las oportunidades de arbitraje y para medir las variables independientes, fue necesario primero realizar ciertos filtros sobre los libros de órdenes límites para de esta forma limpiar los datos de posibles errores.

Una vez realizada la limpieza en los datos, se construyó un libro de órdenes límites con una frecuencia de un segundo, para de esta forma medir todas las variables con la regularidad antes señalada.

Los resultados encontrados validan las hipótesis que son desarrolladas en el presente trabajo.

Por una parte, se concluye la relevancia que tiene la liquidez presente en el libro de órdenes límites sobre la presencia de oportunidades de arbitraje, siendo la profundidad la dimensión más relevante y que presenta una clara relación negativa con la aparición de estos eventos.

Otro aspecto a destacar, es que la liquidez no sólo es importante en la cima del libro, sino que también es relevante como esta se distribuye a lo largo del libro de órdenes límites. Lo anterior, puede detectarse a través de la gran significancia que tiene la variable Relative Liquidity, y también a través de la mayor importancia que tiene el Depth Beyond por sobre el Depth Best.

Los resultados indican que la aparición de los eventos de potencial arbitraje tiene una relación positiva y significativa con la incertidumbre presente en el libro de órdenes límites y con la actividad transaccional observada.

Adicionalmente, se pudo identificar que existe cierto grado de asimetría en relación a la información revelada en los distintos lados de mercado del libro de órdenes límites, siendo en algunos casos más relevante el lado de compra por sobre el lado de venta.

Lo anterior pues para una misma variable medida de manera aislada para el lado de compra y para el lado de venta, se pudo ver en distintas configuraciones econométricas que la variable en el lado de la compra mantiene una alta significancia, mientras que la significancia de la variable del lado de la venta del libro desaparece en algunos casos.

Las oportunidades de arbitraje encontradas pueden ser muy bien caracterizadas a través de la descripción realizada por Foucault, Kozhan y Tham (2016) [15] para las oportunidades de arbitraje de alta velocidad. Los autores antes mencionados las caracterizan como eventos; i) frecuentes, ii) de muy corta duración, iii) más eficientemente explotadas por máquinas que por humanos y iv) con una ganancia muy leve por oportunidad de arbitraje. Características que se ajustan de muy buena manera a los eventos encontrados.

Específicamente, en el tramo horario comprendido entre las 8:00 y las 16:00 horas durante todo el año 2014 con las acciones antes mencionadas se detectaron 4,496 instantes de potencial arbitraje. La duración promedio de las oportunidades encontradas es de 1.739 segundos y con una tasa porcentual de potencial ganancia asociada de 0.03428. Además, en el 42.71 % de los casos el lado de compra asociado a la potencial oportunidad es de Bats Chi X y en el restante 57.29 % de los casos corresponde a Euronext.

La presente investigación está estructurada de la siguiente forma; en el capítulo 1 se expone la investigación más pertinente desarrollada hasta la fecha en torno al arbitraje en mercados fragmentados, en el capítulo 2 se presentan los datos con los que se trabaja y se realiza una breve caracterización de los mercados involucrados y su funcionamiento, en el capítulo 3 se caracteriza la forma en que se obtienen los instantes de potencial arbitraje y se realiza una caracterización de los eventos encontrados, en el capítulo 4 se desarrollan las hipótesis en torno al arbitraje en mercados fragmentados y se presenta la metodología econométrica que se utilizará para validar las hipótesis, en el capítulo 5 se exponen los resultados obtenidos para luego desarrollar las principales conclusiones del trabajo, y finalmente en el capítulo 6 se exponen ciertas tablas que son pertinentes destacar en la presente investigación.





# Capítulo 1

## Literatura relacionada

El arbitraje es uno de los fenómenos que emerge debido a la fragmentación de los mercados financieros, fenómeno que no ha sido tan estudiado como si ha ocurrido con los efectos de la fragmentación o de la competencia entre mercados, se entiende por arbitraje “la simultánea compra y venta de un activo financiero equivalente en dos mercados diferentes con el objetivo de obtener una ganancia, que viene de la discrepancia en los precios” (Bodie, Kane y Marcus, 2009, p. 319 [6]).

En la presente sección, se presenta a grandes rasgos la investigación más relevante y pertinente desarrollada a la fecha en torno al arbitraje, sin embargo, en el desarrollo de la totalidad del presente trabajo se profundiza en ciertos estudios realizados, cuando esto sea pertinente mencionar.

Gagnon y Karolyi (2010) [19] estudian los ADR's y otras acciones que están listadas en Estados Unidos y en algún otro mercado local, y buscan discrepancias de precios y retornos diarios. Se incluyó el retorno como variable pues en un mercado accionario sin fricciones e integrado, se debería asegurar no sólo que la diferencia en precios fuera cero sino también que la diferencia en los retornos diarios fuera igual a cero.

Los autores antes mencionados, utilizan una muestra de 506 parejas de acciones de 35 países en el período comprendido entre 1993 y 2004. Muestran que la desviación de la paridad de precios entre las parejas de acciones es en promedio alrededor de 4.9 basis points (como % del precio de la acción en el mercado local) y con una desviación estándar de 1.4 % para una pareja típica. Estas desviaciones en los precios pueden alcanzar valores que van desde un descuento de -40.4 % hasta un premio de 127.4 %.

El análisis se realizó por firma-trimestre, y se realizó un análisis cros seccional mediante series de tiempo de panel por firma-trimestre. En total, se usaron 12,572 firmas-trimestre.

Las medidas que se utilizaron para medir el arbitraje fueron: el logaritmo natural del cociente entre el precio en Estados Unidos y el precio de la acción en el mercado doméstico. Adicionalmente, se computaron los cambios diarios de la variable anterior, es decir, la dife-

rencia en el retorno diario.

Se realizaron distintos análisis regresionales de manera cros seccional y dentro de las mayores conclusiones que se obtuvo fue, que la diferencia logarítmica de los precios y que la diferencia de los retornos están relacionadas positivamente con los costos de tenencia. La capitalización de mercado tiene un efecto negativo y significativo sobre ambas diferencias.

Se muestra que el poder explicativo de los proxies para los costos de tenencia permanece fuerte, aun controlando por costos transaccionales. A nivel de países, se muestra una fuerte y negativa relación entre la magnitud de la disparidad de precios y los proxies para el grado de desarrollo económico y financiero.

La disparidad de precios tiende a ser mucho menor cuando no hay restricciones sobre las ventas cortas en el mercado local y cuando los inversionistas tienen un mayor grado de conocimiento del mercado.

Por el contrario, la disparidad de precios es mayor cuando los inversionistas extranjeros tienen acceso limitado al mercado local, cuando los costos transaccionales son altos, y bajo condiciones de mercado más volátiles.

Estudiar el arbitraje, producto de la existencia de múltiples plataformas para tranzar un mismo activo, tiene importancia pues existe evidencia de que los inversionistas concentran su actividad en las horas de solapamiento entre distintos mercados. Menkveld (2008) [30] encontró un aumento en el volumen, la volatilidad y una ligera mayor oferta de liquidez para las horas superpuestas versus aquellas no superpuestas.

Schultz y Shive (2010) [37] dicen ser los primeros en examinar la microestructura de como situaciones de precio erróneo ('mispricing') son creadas y resueltas. Para lo anterior, examinan dos clases de acciones que ofrecen derechos sobre el mismo flujo de caja y que son emitidas por la misma compañía, cada pareja de acciones tiene los mismos derechos sobre los flujos de cajas, pero distintos derechos de voto.

Uno de los resultados encontrados que relaciona la liquidez con el arbitraje, y que es distinto al que muestran la mayoría de las investigaciones es que la causa más común de las brechas que crean las situaciones de arbitraje es la presión de precios que mueve a la acción más activa (nonvoting) a que su precio se salga de línea. Lo anterior, es contra intuitivo, pues se esperaría que la acción menos activa fuera la que estuviese asociada a precios erróneos.

Marshall, Nguyen y Visaltanachoti (2013) [29] realizan un estudio empírico y son unos de los primeros en estudiar qué características de mercado prevalecen cuando se presentan las oportunidades de arbitraje, utilizando datos de alta frecuencia.

Para detectar cómo afectan las características de los mercados sobre la presencia de oportunidades de arbitraje se utilizaron datos 'intraday' de dos ETFs. Los ETFs son fondos de inversión que buscan replicar el comportamiento de algún índice en particular.

Se estudiaron instantes de desajustes de precio entre los ETFs, SPY y IVV, entre sí y con el

índice que buscan replicar, es decir, el índice S&P500. Los datos van desde febrero del 2001 hasta agosto del 2010.

Las valoraciones erróneas o instantes de potencial arbitraje en los datos no son eventos frecuentes, con la barrera para las desviaciones de precios de 0.2% se encuentran 183 instantes en el período 2001-2010. Estos instantes desaparecen rápidamente.

La rápida desaparición de las oportunidades de arbitraje es ligeramente diferente a otros resultados encontrados en el mercado de acciones. Por ejemplo, Froot y Dabora (1999) [18] muestran que los desvíos en los precios en acciones listadas en diferentes lugares pueden durar sobre 4 años.

Para estudiar los determinantes en la presencia de oportunidades de arbitraje se estimó una regresión logit, donde la variable dependiente binaria era igual a uno en los días que se creaban oportunidades de arbitraje y cero en los otros casos, esta variable fue regresionada sobre los spreads, el valor tranzado, diferencias en el NAV y el VIX. La magnitud de las ganancias asociadas a estas oportunidades fueron regresionadas con estas mismas variables.

Los resultados muestran que las diferencias de precios, cuando son suficientes para generar ganancias de arbitraje, son más probables que ocurran en días que los spreads son más altos. Los resultados muestran que es más probable que se generen oportunidades de arbitraje en los días menos líquidos.

Los autores antes mencionados proponen que existen ciertas características que podrían incidir en la presencia de oportunidades de arbitraje, sin embargo, existen otros autores que mencionan que las oportunidades de arbitraje se deben a otras razones.

Budish, Cramton y Shim (2015) [8] desarrollan un modelo teórico en que hay un activo X que se transa en un mercado con un diseño CLOB y existe una señal pública sobre el valor de X, Y. El valor fundamental de X está perfectamente correlacionado a la señal pública Y.

Se concluye con el modelo que las oportunidades de arbitraje son construidas en esencia por el diseño CLOB (Mercados de órdenes límites continuos) imperante actualmente en los mercados financieros, dado que el procesamiento en serie en tiempo continuo crea rentas de arbitraje incluso cuando la información es pública simétricamente. Estas rentas de arbitraje argumentan perjudican la provisión de liquidez e inducen a una carrera armamentista interminable y socialmente excesiva por la velocidad.

Este modelo teórico es acompañado de un análisis empírico, utilizando datos de dos índices que replican el comportamiento del índice S&P500.

Dentro del período analizado, que va desde el 2005 hasta el 2011 encuentran que la duración de las oportunidades de arbitraje disminuye dramáticamente, desde una mediana de 97 ms en 2005 a 7ms en 2011. Lo que argumentan está muy relacionado con la fuerte inversión de las firmas transaccionales de alta frecuencia en velocidad durante este período.

Encuentran también que la rentabilidad de estas oportunidades permanece relativamente

constante durante este período. La frecuencia de las oportunidades de arbitraje varía en el tiempo, pero esta variación está conducida casi enteramente por la variación en la volatilidad de mercado.

Se menciona que el diseño de mercado actual predominante (CLOB) conduce a oportunidades de arbitraje mecánicas y que la competencia no ha afectado el tamaño ni la frecuencia de las oportunidades de arbitraje, esto sólo ha aumentado, argumentan, el riesgo de cuán rápido uno puede ser capturado por un agente que posea una velocidad en su ejecución de órdenes mayor.

Foucault, Kozhan y Tham (2016) [15] presentan el desarrollo de un modelo de arbitraje en mercados cruzados, muestran ciertas proposiciones que se derivan del modelo y se testean empíricamente.

La parte empírica es realizada con datos de tres parejas de monedas (dólar-euro, dólar-libra y libra-euro), entre estas parejas se producen en ciertos instantes arbitraje triangular. Se usa data de Reuters D-3000 desde el 2 de enero del 2003 al 30 de diciembre del 2004.

El arbitraje puede ser dañino o beneficioso para los inversionistas, dependiendo cual sea la causa de esta oportunidad de arbitraje. Cuando la oportunidad es debido a shocks de demanda u oferta transientes (presión de precios), los arbitrades actúan implícitamente como proveedores de liquidez al explotar esta oportunidad. En este caso, tanto el arbitrador como la contraparte se benefician.

Cuando las oportunidades son debido al ajuste asincrónico en los precios de los activos luego de un arribo de información, la ganancia de los arbitrades se hace a expensas de la contraparte, lo cual constituye un arbitraje tóxico. Una transacción tóxica se define, según Easley et al. (2012) [13] de la siguiente forma: “Un flujo de órdenes es considerado como tóxico cuando se selecciona adversamente a un ‘market maker’ que puede no ser consciente de que está proveyendo liquidez a expensas de pérdidas”.

Los resultados del paper sugieren que la eficiencia ganada en los precios por los traders de alta frecuencia viene con el incremento en el costo del riesgo de selección adversa.

# Capítulo 2

## Mercados y Datos

La proliferación de nuevas plataformas o mercados que permiten tranzar un activo financiero en distintos lugares es un fenómeno de gran expansión durante las últimas décadas. Este fenómeno de fragmentación en los mercados financieros es bastante fuerte en Estados Unidos, y con un cierto retraso y menos estudiado en Europa.

La fragmentación ha potenciado el desarrollo de nuevas y mejores tecnologías en los sistemas transaccionales, lo cual ha ayudado a disminuir los tiempos de ejecución de órdenes y ha significado que las diferentes plataformas hayan intensificado la competencia entre sí, beneficiando a los inversionistas, por ejemplo, a través de la reducción de las cuotas o comisiones cobradas.

Lo anterior, ha sido producto no sólo de la fragmentación financiera sino también de los potentes cambios tecnológicos que han ocurrido en la forma de tranzar los activos financieros.

La proliferación de los llamados algoritmos transaccionales u operadores de alta frecuencia es un fenómeno creciente de las últimas décadas que ha crecido de forma paralela al fenómeno de la fragmentación.

Un ejemplo de los grandes cambios tecnológicos que ha experimentado la forma de tranzar los activos financieros es el hecho que en 2010 se completó la construcción de un cable de fibra óptica de alta velocidad que conectaba los mercados financieros de New York y Chicago, el resultado de esta inversión fue reducir el tiempo de conexión entre estos dos mercados de 16 a 13 milisegundos.

Sin embargo, no existe un claro consenso sobre los efectos que puede provocar la fragmentación de los mercados financieros. Específicamente, existen quienes por una parte creen que la fragmentación puede dañar la calidad de los mercados a través de la liquidez disponible.

Otros investigadores señalan que el aumento en la velocidad de algunos inversionistas podría perjudicar o crear un costo de selección adversa para aquellos inversionistas que no posean sistemas transaccionales tan sofisticados.

Los cuestionamientos sobre las ventajas o desventajas asociadas a la proliferación de ‘traders’ de alta velocidad tampoco muestran un consenso entre los investigadores. Por lo tanto, generar investigación y encontrar relaciones empíricas con datos de alta frecuencia en mercados fragmentados es de suma relevancia.

Los grandes cambios que se han visto en los mercados financieros han sido recogidos en parte por los reguladores correspondientes y han surgido ciertas normativas que buscan adaptarse a la nueva realidad vigente.

La Unión Europea introdujo el 1 de Noviembre del 2007 el instructivo llamado MiFID, por sus siglas en inglés que significa Directiva sobre Mercados Financieros, el cual entre otros asuntos alienta la competencia entre bolsas locales y fomenta la entrada de nuevos actores al mercado.

Estudiar los fenómenos producidos a causa de la fragmentación es un asunto de gran interés dentro de la investigación desarrollada en la microestructura de mercados y en finanzas en general.

Los datos utilizados son útiles para estudiar este fenómeno y sus consecuencias, pues corresponden a datos de órdenes y transacciones diarias de acciones que son tranzadas de manera simultánea en múltiples mercados en Europa, siendo los dos más importantes; Euronext y Bats Chi X.

Chi X, inició sus operaciones en Europa en Marzo del 2007. Inicialmente sólo operaba una plataforma electrónica transaccional con acciones pertenecientes a índices de Holanda y Alemania. Su crecimiento ha sido muy rápido y significativo, en 2011 se convirtió en el mayor mercado accionario Europeo de acuerdo a las estadísticas transaccionales publicadas por la Federación de Bolsas Europeas. Actualmente tiene el carácter de ‘Recognised Investment Exchange’ (RIE), y es posible tranzar más de 4,500 activos financieros pertenecientes a los 15 mayores mercados europeos.

Euronext también es una bolsa de valores paneuropea que opera desde el 22 de septiembre del 2000, es el producto de la fusión de distintas bolsas locales de Europa, inicialmente la Bolsa de Paris, Ámsterdam y Bruselas.

Tanto Euronext como Bats Chi X presentan mecanismos transaccionales similares, pues ambas permiten tranzar de acuerdo a libros de órdenes electrónicos, es decir, son mercados de órdenes límites continuos y las transacciones son efectuadas en orden de acuerdo a la prioridad de precio y de llegada.

Es importante realizar una breve descripción sobre el diseño de mercado presente tanto en Euronext como en Bats Chi X, este diseño de mercado predomina actualmente en la mayoría de los mercados financieros y se denomina CLOB, por sus siglas en inglés que significa “Libro de órdenes límites continuo”.

El conjunto o bloque de órdenes se construyen a partir de las llamadas ordenes límites. Las ordenes límites indican precio, cantidad de acciones involucradas y si se trata de una orden de compra o de venta. Estas órdenes pueden ser enviadas durante todo el día transac-

cional, y pueden también ser retiradas parcialmente o en su totalidad.

Si bien existen otro tipo de órdenes que pueden enviar los ‘traders’, son las ordenes límites las básica y el resto suelen ser variaciones de las órdenes límites o alguna con ciertas características en particular. Por ejemplo, una orden de mercado es una orden límite, pero con la característica que el precio se determinará de tal forma que la orden sea ejecutada inmediatamente.

Una transacción se realizará cuando una nueva orden límite sea enviada con las siguientes características; ya sea una orden de compra con un precio igual o mayor que el mejor precio de venta disponible en el mercado en ese instante, o una orden de venta con un precio igual o menor que el actual mejor precio de compra.

Los datos diarios utilizados son de alta frecuencia, es decir, existen diferencias entre dos observaciones consecutivas que pueden ser de órdenes de tiempo inferiores al segundo. Se utilizaron acciones pertenecientes al índice AEX.

Este índice está compuesto de los 25 principales valores que se tranzan de manera primaria en Euronext Amsterdam y que forman el índice Amsterdam Exchange Index. El 16 de Febrero del año 2017 AEX tenía un valor nominal consolidado de 3,163,002,813 euros, donde el 55.83 % era tranzado en Euronext y un 22.36 % en Bats Europe.

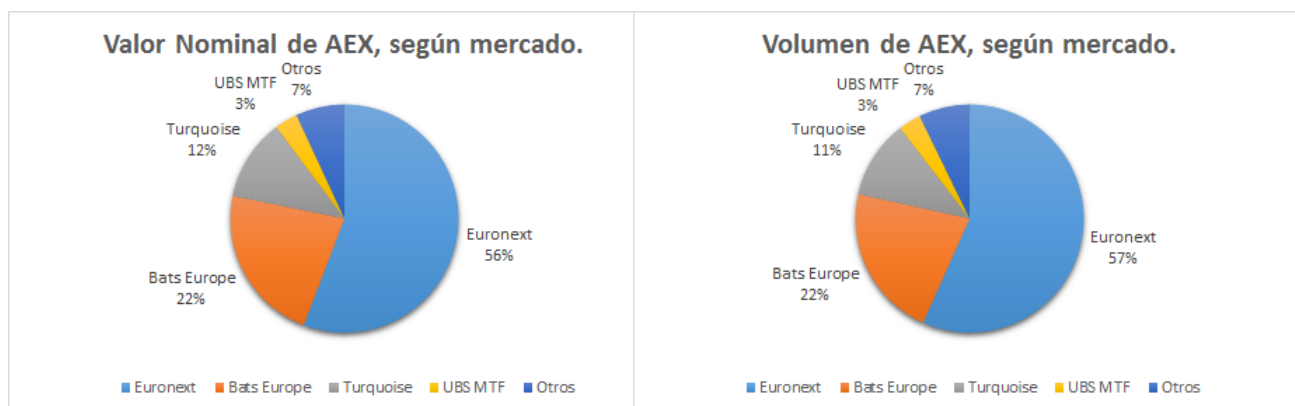


Figura 2.1: Volumen y valor nominal de AEX, según mercado

El volumen comprendido por AEX a la fecha antes mencionada era de 163,472,911, donde el 56.71 % era tranzado en Euronext y un 21.73 % en Bats Europe. En la figura 2.1 se puede apreciar el detalle del valor nominal y del volumen de AEX, según mercados.

	Volumen promedio	Valor promedio	Órdenes de compra promedio	Órdenes de venta promedio
Total Chi X	1,211,370.114	17,025,176.57	34,925,920.85	34,768,155.06
Total Euronext	4,067,840.255	57,456,229	218,685,698.5	210,652,962.7

Tabla 2.1: Resumen de valores promedios diarios transaccionales de las acciones analizadas.

Se analizaron 22 acciones pertenecientes a AEX, que se tranzaron de manera simultánea en Euronext y Bats Chi X, durante todo el año 2014. Como se puede ver en la tabla 2.1, el promedio diario del valor tranzado por acción en Euronext es de 57,456,229 euros y el promedio diario de número de acciones tranzadas en Euronext es de 4,067,840, en el caso de Bats Chi X los valores son 17,025,176 euros diarios y 1,211,370 acciones diarias. En la misma tabla se puede ver el número de órdenes de compra y venta diarias promedio.

En la tabla 7.1, se puede ver el detalle para cada acción analizada de Bats Chi X sobre su valor y volumen promedio diario tranzado, y el número de órdenes de compra y venta diarias promedio. La misma información, pero relativa a las acciones analizadas de Euronext, se puede encontrar en la tabla 7.2.

Para cada acción se cuenta con dos libros, el libro de ordenes límites (LOB) y el libro de ordenes y transacciones. El LOB corresponde a lo que ven los inversionistas en cada momento y es donde se pueden ver las órdenes que quedan en espera por ser ejecutadas de acuerdo a la prioridad de precio, es decir, los precios de compra más altos y los precios de venta más bajos son más probables de ser ejecutados, y son ordenados de mejores a peores precios en espera.

El libro de ordenes y transacciones muestra el detalle de cuando van ingresando las ordenes, ya sean de compra o venta, y cuando se realizan las transacciones, a partir de este libro se construye el libro consolidado, LOB. Sin embargo, existen cierto tipo de ordenes en las que no es público todo el detalle, como lo es el caso de las ordenes ocultas o 'iceberg', por lo que no es posible reconstruir el LOB a partir del libro de ordenes y transacciones por completo. Lo anterior no es un problema, pues no es necesario reconstruir el LOB debido a que ya se dispone de él.

El LOB contiene el id de la acción, fecha y hora a la cual corresponde la información, además de proporcionar los tres mejores precios (quotes), de compra y venta, con la cantidad (depth) que es ofrecida a cada precio. En tanto el libro de ordenes y transacciones proporciona el precio y la cantidad de acciones involucradas en cada orden entrante o transacción realizada, más la información sobre la fecha, hora y acción en cuestión.



# Capítulo 3

## Arbitraje

Uno de los fenómenos que emerge debido a la fragmentación de los mercados es el arbitraje, es decir, tomar ventaja de las diferencias de precios que surgen sobre un mismo activo subyacente que es tranzado simultáneamente en más de un mercado.

El arbitraje es importante dentro de la economía financiera, pues está relacionado a dos temas de gran relevancia como lo son el cumplimiento de la ley de único precio y las hipótesis de los mercados eficientes. Se entiende por arbitraje “la simultánea compra y venta de un activo financiero equivalente en dos mercados diferentes con el objetivo de obtener una ganancia, que viene de la discrepancia en los precios” (Bodie, Kane, and Marcus, 2009, p. 319 [6]). En teoría no se requiere capital y no hay riesgos, sin embargo, existen impedimentos debido a fricciones de mercado e información imperfecta que puede limitar el arbitraje.

Como se vio en el capítulo sobre literatura relacionada, el arbitraje es un fenómeno estudiado de diversas formas y de gran importancia, sin embargo, el estudio con datos de alta frecuencia es un fenómeno relativamente nuevo, como lo señalan Marshall, Nguyen y Vitsaltanachoti (2013) [29] "muy poco se sabe sobre las características de mercado dentro del día que existen cuando las oportunidades de arbitraje son creadas", lo que ellos denominan microestructura de arbitraje.

Como evidencia de lo contemporáneo que es el estudio de este fenómeno se puede mencionar que recién en 2016, Foucault, Kozhan y Tham [15] parecieran ser los primeros en modelar este canal de arbitraje y entregar evidencia sobre su importancia para la liquidez. En ese mismo trabajo, los autores antes mencionados señalan que más investigación es necesaria para establecer robustez sobre las conclusiones que entregan para otro tipo de oportunidades de arbitraje de alta frecuencia, ellos utilizan datos de libros electrónicos sobre monedas.

Con los datos antes mencionados, se procedió a detectar las oportunidades de arbitraje que se daban entre una misma acción que es tranzada simultáneamente en Bats Chi X y en Euronext. Las oportunidades de arbitraje se detectaron en el intervalo de tiempo comprendido entre las 8:00 y las 16:00, período durante el cual ambos mercados operan. Antes de identificar los instantes en que se dan estas oportunidades de arbitrajes es necesario realizar ciertos

filtros sobre los datos, con el objetivo de limpiar los datos de posibles errores.

Exclusivamente se dejaron para cada libro de órdenes límites las observaciones en que al menos existe un mejor precio de venta y un mejor precio de compra mayor a cero, los depths o cantidades ofrecidas a estos precios también deben ser mayores a cero. Adicionalmente, fueron borradas las observaciones en que el mejor precio de compra es mayor al mejor precio de venta para una acción en un mercado en particular, ya que de lo contrario se estaría frente a una situación de arbitraje dentro de un mercado en particular, esta última limpieza de datos es recomendada por Schultz y Shive (2010) [37] y es aplicada también por Marshall, Nguyen y Visaltanachoti (2013) [29] en su trabajo sobre arbitraje.

Luego de realizada la limpieza en los datos, es necesario también obtener el horario de cierre que tiene cada mercado día a día. Debido a que si bien se estudió desde las 8:00 hasta las 16:00 y normalmente los mercados tienen horarios de apertura y cierre constantes, hay días en que cierran antes por eventos especiales como Navidad, por lo cual en esos casos se estudian los mercados hasta el menor horario de cierre entre Bats Chi X y Euronext. Este horario es posible obtenerlo de los libros fácilmente, pues se puede ver cuál es el momento en que hubo una última actualización en los datos.

El siguiente paso es generar día a día y para cada acción un libro de órdenes límites (LOB) que tenga regularidad en el tiempo, dado que en los datos es posible ver instantes en que en un segundo el LOB es actualizado muchas veces, como otras ocasiones en que el LOB no cambia durante segundos.

Se construyó un LOB con una regularidad de un segundo, en los casos en que durante un segundo haya más de una observación se dejó la última, pues corresponde a lo que ven los inversionistas al finalizar ese segundo. Al tener el LOB limpio y con una frecuencia regular para una acción y día en particular, se realizó un cruce con el LOB de la misma acción y día que es tranzada en el otro mercado, luego de haber seguido el mismo procedimiento antes descrito con este libro.

Finalmente, con esta base diaria por acción donde se puede ver el libro consolidado en ambos mercados es que se procedió a detectar las oportunidades de arbitraje.

La detección de las oportunidades de arbitraje se realizó basado en el método descrito por Marshall, Nguyen y Visaltanachoti (2013) [29], pero con ciertas variaciones. La condición para que en un instante exista una oportunidad de arbitraje es:

$$\frac{bid_i}{ask_j} > 1 + gap, j \neq i \quad (3.1)$$

Donde  $j$  e  $i$  hacen referencia a los mercados Euronext y Bats Chi X. Es decir, se estará frente a una oportunidad de arbitraje cada vez que el primer mejor precio de compra de un libro sea mayor al primer mejor precio de venta del libro que se observa en el otro mercado, teniendo como mercados a Bats Chi X y a Euronext.

Este paso sirve para identificar cuáles serán los lados de mercado relevantes, de compra y venta, los cuales se identifican para cada oportunidad de arbitraje, siendo  $i$  el lado de

compra relevante y j el lado de venta relevante, es decir, involucrados en la oportunidad de arbitraje.

En este caso se está frente a una oportunidad de arbitraje debido a que un inversionista que tenga acceso a los dos mercados podría comprar la acción al precio de venta y luego venderlo al precio de compra correspondiente, obteniendo como ganancia la diferencia entre estos precios, esto lo podría repetir con las acciones que estén disponibles a los precios involucrados en la oportunidad de arbitraje.

La variable *gap* presente en la ecuación 3.1 puede ser cualquier valor mayor o igual a cero, este valor actúa como una frontera sobre la ganancia exigida a la hora de considerar una oportunidad de arbitraje, y puede entenderse como la diferencia necesaria para que sea rentable arbitrar y se puedan cubrir los costos que podrían estar asociados a la transacción.

Se consideró *gap* igual a cero pues el principal objetivo de este trabajo es determinar qué factores prevalecen cuando se está en presencia de una oportunidad de arbitraje, mas que determinar cuál podría ser la ganancia potencial real asociada.

El hecho de que *gap* sea igual a cero significa que se considera como una potencial oportunidad de arbitraje cada vez que el primer mejor precio de compra en un libro sea mayor estricto al primer mejor precio de venta presente en el otro libro, sin importar cuan mayor sea.

Con los datos antes mencionados, en el año 2014, entre las 8:00 y las 16:00 y siguiendo el procedimiento descrito para identificar las oportunidades de arbitrajes se detectaron 4,496 instantes en que se está frente a una potencial oportunidad de arbitraje.

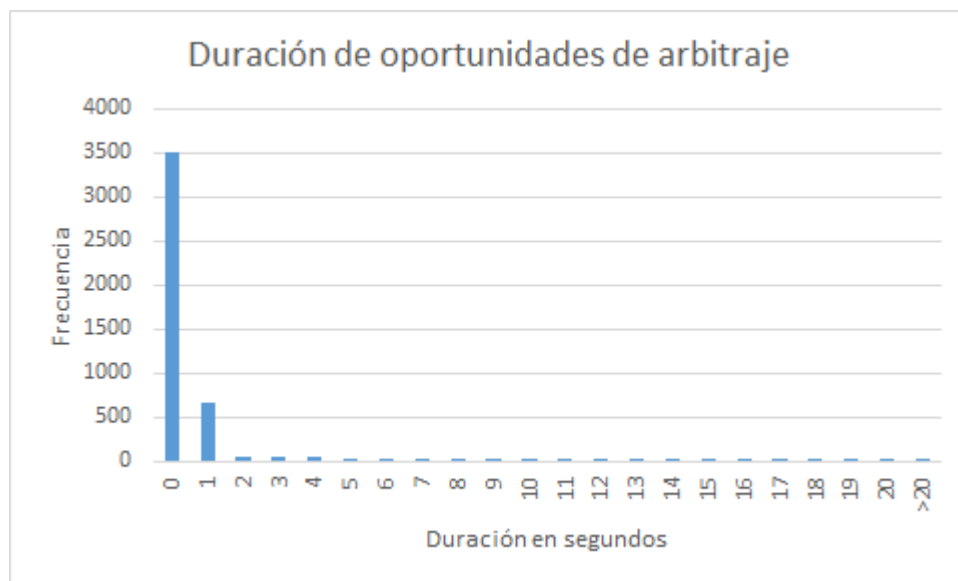


Figura 3.1: Duración oportunidades de arbitraje

Como se puede ver en la figura 3.1 la mayor cantidad de los eventos encontrados desaparecen rápidamente en el tiempo. La duración promedio de estas oportunidades de arbitraje es de 1.739 segundos, con una mediana de 0 segundos y el tercer cuartil también tiene una duración de 0 segundos. La oportunidad de arbitraje encontrada con mayor duración es de 1517 segundos, que es equivalente a una duración de 25 minutos y 17 segundos.

Como se mencionó anteriormente para calcular las oportunidades de arbitraje primero fue necesario obtener el libro de órdenes límites con frecuencia de 1 segundo. Una oportunidad de arbitraje tiene una duración de 0 segundos cuando en un segundo en particular se da la oportunidad de arbitraje y en el segundo inmediatamente siguiente desaparece.

Duración	Frecuencia	Porcentaje
Mayor o igual a 0 segundos	4,496	100 %
Mayor o igual a 1 segundo	979	21.8 %
Mayor o igual a 2 segundos	314	7 %
Mayor o igual a 5 segundos	149	3.3 %
Mayor o igual a 10 segundos	83	1.85 %

Tabla 3.1: Duración de oportunidades de arbitraje

Foucault, Kozhan, Tham (2016) [15] caracterizan las oportunidades de arbitraje de alta velocidad como: i) frecuentes, ii) de muy corta duración, iii) más eficientemente explotadas por máquinas que por humanos y iv) con una ganancia muy leve por oportunidad de arbitraje.

A continuación, se detalla por qué en este caso las oportunidades de arbitraje estudiadas tienen un comportamiento acorde a las características antes mencionadas.

El hecho de que las oportunidades de arbitraje sean frecuentes, los autores antes mencionados lo plantean a escala de mili segundos, en la presente investigación se puede ver que la mayor frecuencia se da mientras menor sea el tiempo de duración a considerar (como se ve en la tabla 3.1). Lo anterior, también valida la corta duración de estos eventos, como puede verse en la tabla 3.1 sólo un 3.3 % del total de las oportunidades de arbitraje detectadas tienen una duración mayor o igual a 5 segundos.

Además, se detectaron 4,496 oportunidades de arbitraje para 22 pares de acciones en el año 2014 que contó con 255 días transaccionales, lo anterior da un promedio de 0.8 oportunidades de arbitraje por día y por pareja de acciones, lo anterior considerando que no se tomó el 100 % del tiempo en que ambos mercados operan conjuntamente.

Lo anterior, permite considerar los eventos estudiados como frecuentes, ya que son eventos que se esperaría no se dieran con tanta regularidad en el tiempo, además al considerar escalas de tiempo de mili segundos es probable que la frecuencia fuese mucho mayor. El hecho de que estas oportunidades sean más eficientemente explotadas por máquinas que por humanos tiene que ver precisamente por lo fugaz de estos eventos.

Para el cálculo del potencial beneficio porcentual asociado a cada oportunidad de arbitraje encontrada se utilizaron los supuestos de Mitchell et al. (2002) [32], que es que se requiere

un margen del 50 % tanto para la posición larga como para la posición corta involucrada al aprovechar la oportunidad de arbitraje. Esta metodología también fue utilizada por Marshall, Nguyen y Visaltanachoti (2013) [29] en su trabajo sobre arbitraje.

Siguiendo a los autores anteriores es que se calculó la ganancia porcentual asociada a cada oportunidad de arbitraje de la siguiente forma:

$$Tasa\ porcentual\ de\ ganancia = 100 \cdot \frac{Bid\ Price - Ask\ Price}{0,5 \cdot (Bid\ Price + Ask\ Price)} \quad (3.2)$$

Para el cálculo de estas tasas se utilizaron los precios que se observan al inicio de la oportunidad de arbitraje, con sus respectivos lados de compra y venta asociados en la oportunidad de arbitraje en cuestión.

	N	% del total	Mín.	25 % Per	Mediana	Media	75 %	Máx.
Todas	4996	100	0.01	0.01194	0.01681	0.03428	0.02487	10.53
Bid Chi X	2131	42.71	0.01	0.012	0.01697	0.0296	0.02496	6.186
Bid Euronext	2365	57.29	0.01	0.0119	0.01673	0.03849	0.02486	10.53

Tabla 3.2: Potenciales ganancias porcentuales asociadas a las oportunidades de arbitraje

Como se puede ver en la tabla 3.2, la ganancia promedio asociada al total de oportunidades de arbitraje es de 0.03428 %, una ganancia promedio asociada a cada oportunidad de arbitraje muy baja, lo que ratifica la última característica propuesta por Foucault, Kozhan, Tham (2016) [15] sobre las oportunidades de arbitraje de alta velocidad, el hecho de que estas oportunidades tienen asociadas una ganancia muy leve por oportunidad.

Como se puede ver en la misma tabla, el comportamiento de esta tasa de ganancia es similar cuando la oportunidad en cuestión tiene el lado de compra involucrado de Chi X o Euronext. Hasta el tercer quintil no se observan muchas diferencias, la media cuando el lado de compra involucrado es de Euronext es más alta debido a la cola superior, en que se observan unos valores máximos más altos comparado a cuando el lado de la compra involucrado es de Chi X.

En promedio, estas tasas porcentuales de potencial ganancia asociada a las oportunidades de arbitraje son más bajas que las encontradas por Marshall, Nguyen y Visaltanachoti (2013) [29] en su trabajo de arbitraje con ETF's.

Los autores antes mencionados encuentran una tasa promedio de 0.33 % y 0.32 %, según cual sea el activo que está sobrevalorado o subvalorado que ellos estudian. Sin embargo, en el estudio con ETF's se encuentran valores extremos menos dispersos que los encontrados en el presente estudio, en el trabajo con ETF's la tasa máxima de ganancia encontrada es de 1.87 % y la mínima de 0.2 % , en comparación a la de 10.53 % y 0.01 % respectivamente, encontradas en la presente investigación.

Otra forma de entender y caracterizar estas oportunidades de arbitraje es a través de la cantidad de acciones involucradas, depth, en la potencial oportunidad. Siendo el mínimo depth, entre el lado de compra y venta involucrado en la oportunidad, la cantidad relevante,

pues un agente que tenga acceso a los dos mercados podrá explotar esta oportunidad con esa cantidad de acciones como máximo.

	N	% del total	Mín.	25 % Per	Mediana	Media	75 %	Máx.
Todas	4996	100	1	73	196.5	589.2	586.5	21580
Bid Chi X	2131	42.71	1	75	200	595.3	600	21580
Bid Euronext	2365	57.29	1	72	194	583.7	563	17190

Tabla 3.3: Depth asociado a las oportunidades de arbitraje

En la tabla 3.3 se puede apreciar el detalle para esta variable, donde se observa que al iniciarse la oportunidad de arbitraje el mínimo depth promedio entre los dos lados relevantes es de 589.2 acciones.

Con esta cantidad de acciones se puede tranzar y aprovechar de esta forma la oportunidad creada, donde la ganancia en euros asociada estará dada por la diferencia entre el precio de compra y el precio de venta involucrados, por acción.

	N	% del total	Mín.	25 % Per	Mediana	Media	75 %	Máx.
Todas	4996	100	0.001	0.33	0.885	2.808	2.23	900
Bid Chi X	2131	42.71	0.001	0.33	0.885	2.72	2.233	900
Bid Euronext	2365	57.29	0.001	0.327	0.885	2.888	2.22	650

Tabla 3.4: Máxima ganancia asociada a las oportunidades de arbitraje

En la tabla 3.4 se aprecia el detalle de multiplicar esta diferencia entre los precios involucrados en la oportunidad por la cantidad de acciones disponibles para cerrar la oportunidad. En promedio, la ganancia en euros por oportunidad es de 2.808 euros, mediante la metodología recién explicada, donde se asume cero costos para cerrar la oportunidad en cuestión.

Del total de oportunidades de arbitraje detectadas en el presente estudio, en el 57.29 % de los casos el lado de la compra involucrado es de Euronext y en el 42.71 % de los casos es de Chi X.

Del total de oportunidades de arbitraje encontradas, el 29 % se encuentra en el tramo horario [8:00,10:00], mismo porcentaje que ocurre en el tramo [14:00,16:00], más abajo vienen aquellos eventos que ocurren en el tramo [12:00,14:00], alcanzando un 22 % y finalmente el 19 % restante ocurre en el tramo horario [10:00,12:00].

Lo anterior podría colaborar a erradicar la idea de que las oportunidades de arbitraje se dan exclusivamente en algún horario en particular, sin embargo, si existe un porcentaje considerablemente mayor en aquellos horarios más cercanos a la apertura y cierre de las bolsas.

Otro gran cambio que se ha dado en la última década en los mercados financieros es la presencia de inversionista que poseen sistemas de alta velocidad que les permite automatizar sus órdenes y responder dentro de fracciones de segundos.

Budish et al. (2015) [8] señala que los retornos de los arbitradores de alta velocidad son sustanciales dado que las oportunidades de arbitraje son muy frecuentes a escala de milisegundos. La actual literatura ha mostrado que estos tiempos de ejecución pueden ir del orden de los milisegundos (Hasbrouck, Saar, 2013 [23]).

Dicho lo anterior, es que una de las diferencias de la presente investigación con el método planteado por Marshall, Nguyen y Visaltanachoti (2013) [29] para detectar las oportunidades de arbitraje tiene que ver con el supuesto sobre el tiempo de ejecución de las órdenes.

Marshall, Nguyen y Visaltanachoti (2013) [29] utilizan un enfoque conservador en ese sentido, considerando un tiempo de ejecución de 15 segundos, por lo que consideran en su investigación aquellas oportunidades de arbitraje que al menos se mantienen por este lapso de tiempo. Sin embargo, ellos mismos señalan que trabajos actuales muestran tiempos de ejecución inferiores a un segundo (Hendershott, Moulton, 2011 [25]), incluso pudiendo reaccionar algunos ‘traders’ a eventos del mercado dentro de 2 o 3 milisegundos (Hasbrouck, Saar, 2013 [23]).

En lo que sigue de la investigación, para la mayor parte del análisis realizado se consideran aquellas oportunidades que duran al menos un segundo. De esta forma, se es consecuente con las últimas investigaciones relacionadas sobre la velocidad de los agentes.

Además, al no considerar todas las oportunidades de arbitraje (incluyendo las de duración 0 segundos) no se entra en conflicto con lo que plantea Budish et al. (2015) [8]. Los autores recién mencionados plantean que las oportunidades de arbitraje se generan de forma mecánica a escalas de milisegundos dado el diseño de mercado imperante, ellos muestran que en mercados fragmentados, los mercados no se correlacionan a escala de milisegundos.

# Capítulo 4

## Análisis Empírico

### 4.1. Objetivo

El objetivo de la presente investigación es generar evidencia empírica para entender uno de los fenómenos que se produce a causa de los cambios que han experimentado los mercados financieros en el último tiempo, específicamente debido a la fragmentación de los mercados, el fenómeno que se busca caracterizar y entender es el de arbitraje en mercados fragmentados.

Al igual que sobre los efectos que han producido las grandes variaciones de las últimas décadas en los mercados financieros, no existe un consenso entre los investigadores ni entre quienes desarrollan políticas para regular los mercados sobre qué factores influyen en la presencia de oportunidades de arbitraje ni cuales son las características de mercado que predominan durante estos fenómenos.

Este fenómeno de arbitraje ha sido denominado por algunos investigadores como arbitraje tóxico que como se ha mencionado anteriormente se caracteriza porque los 'traders' obtienen una ganancia producto de los precios ofrecidos en diferentes mercados, utilizando un precio de compra en un mercado y un precio de venta en otro mercado. Este fenómeno aparece debido a un retraso en la actualización de los precios por parte de los proveedores de liquidez frente a nueva información.

Han existido distintos acercamientos para tratar de caracterizar este fenómeno. Budish et al. (2015) [8] señalan que el arbitraje que se produce es mecánico y meramente consecuencia del actual diseño de mercado imperante, es decir, libros de órdenes límites continuos. Por otra parte, Marshall et al. (2013) [29] señalan que hay ciertas características imperante en el mercado dentro del día que podrían influir en la aparición de oportunidades de arbitraje, como una disminución en la liquidez o un aumento en el riesgo de liquidez. Foucault et al. (2016) [15] también muestran teórica y empíricamente que existe una relación positiva entre iliquidez y arbitraje tóxico.

A través de la presente investigación se busca generar evidencia empírica a favor o en contra de



alguna de las distintas líneas argumentativas desarrolladas en el tema, y si es posible generar nuevas perspectivas de investigación. Se busca ayudar en la caracterización y entendimiento de este fenómeno recientemente estudiado.

## 4.2. Dimensiones del libro a considerar

### 4.2.1. Liquidez

#### Definición e intuición

Probablemente una de las dimensiones más importantes a tener en cuenta en cuanto al estado del libro de órdenes límites o de un mercado financiero en general es la liquidez.

Si bien la liquidez podría definirse de forma rigurosa de acuerdo a algún modelo en particular, existen ciertos consensos sobre este concepto. La liquidez se entiende como el “nivel al cual una orden puede ser ejecutada dentro de un corto plazo de tiempo a un precio cercano al valor de consenso que hay sobre un activo financiero” (Foucault, Pagano, Roell, 2013 [17]).

La intuición detrás de este concepto no es tan lejana, pues en general todos los mercados tienen cierto grado de liquidez o iliquidez. Un mercado es líquido cuando es relativamente fácil realizar transacciones del activo en cuestión, es decir, cuando se puede convertir el activo en dinero fácilmente.

En el caso del mercado accionario, se dirá que un mercado es líquido cuando los ‘traders’ son capaces de ejecutar grandes cantidades de órdenes de forma rápida y a un bajo costo.

Por ejemplo, si al comprar una gran cantidad de acciones sólo puedo realizarlo a un precio unitario por acción mucho mayor al precio que se observaba previamente, estaremos frente a un mercado para esa acción en particular ilíquido.

En general si en un mercado se mueve mucho dinero en lapsos cortos de tiempo se estará frente a un mercado líquido, ya que es más factible convertir el activo a dinero.

La liquidez no es una constante para un mercado, ya que esta varía a lo largo del tiempo y entre activos que pueden ser tranzados en un mismo mercado. La liquidez puede relacionarse con un concepto económico muy utilizado frecuentemente, la elasticidad. Esto pues un pequeño cambio en la oferta o la demanda, en un mercado líquido, no debería impactar mayormente en un cambio en los precios (Hasbrouck, 2007 [22]).

Se suele decir que la liquidez tiene distintas dimensiones, siendo las más importantes: la profundidad (depth), amplitud (breadth) y la resiliencia (resiliency).

La profundidad se refiere a la cantidad ofrecida del activo subyacente, por ejemplo, si sobre el precio actual existe una gran cantidad incremental del activo subyacente disponible

para la venta se estará frente a un mercado profundo. Lo mismo ocurrirá, si bajo el precio actual hay una gran cantidad incremental de compradores para el activo subyacente. Un mercado es ancho cuando presenta varios participantes, ninguno de los cuales puede ejercer un poder de mercado significativo. Finalmente, el mercado será resiliente cuando el efecto en los precios debido a los procesos transaccionales es pequeño y desaparece rápidamente (Hasbrouck, 2007 [22]).

Dentro de cada dimensión es posible distinguir ciertas variables que miden la liquidez dentro de esa dimensión. Por ejemplo, para ver la profundidad es posible observar la cantidad ofrecida tanto al precio de compra como el de venta (depth). Más adelante se detalla con mayor profundidad las variables que son utilizadas para capturar la liquidez del libro.

Habiendo clarificado lo que se entiende por liquidez, es sensato pensar que mientras mayor sea el impacto de una transacción o mientras más complejo sea realizarla, deberían producirse con mayor regularidad distorsiones en los precios del activo subyacente, lo que incidiría positivamente en la presencia de oportunidades de arbitraje. Es decir, a mayor iliquidez debería ser más frecuente la aparición de estos eventos en cuestión.

## Literatura previa

Si bien poco se sabe sobre las características del libro que prevalecen cuando se dan oportunidades de arbitraje en mercados fragmentados, probablemente sea la liquidez la dimensión más estudiada a la fecha y la que más ha sido relacionada con arbitraje.

Gagnon y Karolyi (2010) [19] estudian las discrepancias o desviaciones de precios y retornos de acciones que se tranzan simultáneamente en distintos mercados alrededor del mundo. Se comparan los precios de distintos ADRs y otro tipo de acciones ‘cross listed’ en Estados Unidos con los precios sincrónicos en el respectivo mercado local, realizando un ajuste de moneda correspondiente. La comparación es hecha diariamente.

Los ADR’s o recibos depositarios americanos por sus siglas en inglés son un mecanismo que tienen las empresas no originarias de Estados Unidos para emitir sus acciones en los mercados locales de Estados Unidos, es decir, son un título financiero donde el activo subyacente es el mismo que el de la acción en el mercado local.

Dentro de las hipótesis que tienen los autores antes mencionados es que las acciones más ilíquidas en uno u otro mercado, o ambos, tienen mayores diferencias de precios o retornos entre las parejas de acciones ‘cross listed’ que estudian. Sin embargo, no encuentran en todos sus resultados que los ‘proxies’ que utilizan para liquidez en el mercado local y en el mercado americano sean significativos.

Pero existen ciertos resultados que dan respaldo a la hipótesis sobre liquidez y arbitraje, si bien no relacionada con características dentro del día transaccional sino más bien con otras formas de explicar la liquidez. En general, los activos financieros que pertenecen a mercados más desarrollados tienden a ser más líquidos. Los autores antes mencionados encuentran que,

a un nivel de país, existe una relación fuerte y negativa entre las discrepancias de precios y ciertos proxies que utilizan para medir el nivel de desarrollo económico y financiero en los mercados.

Si bien el trabajo de Gagnon y Karolyi (2010) [19] muestra ciertas relaciones sobre cuando existe discrepancia de precios, tiene como limitantes que se trabajan con precios diarios y a veces puede darse que no exista un completo solapamiento de las horas transaccionales, siendo irreal obtener una rentabilidad asociada a esta oportunidad de arbitraje.

Adicionalmente, no se detecta el momento exacto en que se produce este fenómeno al trabajar con precios de cierre diario y no hacer una investigación más detallada dentro del día. Lo anterior no ocurre en la presente investigación, donde las oportunidades de arbitraje son detectadas a nivel de segundo y en todo momento se consideran dos mercados que operan sincrónicamente.

Foucault, Kozhan y Tham (2016) [15] en su paper “Toxic arbitrage” realizan la contribución de modelar este canal de arbitraje y entregar evidencia sobre su importancia para la liquidez. Dicho paper, pareciera ser el primero en hacerlo. Ellos presentan un modelo teórico y las hipótesis derivadas del modelo se testean con libros de órdenes de tres parejas de monedas (dólar-euro, dólar-libra y libra-euro), estos datos tienen una exactitud del orden de los 10 mili segundos. Entre estas parejas se producen en ciertos momentos arbitraje triangular.

En este paper es importante lo que denominan el “mix arbitrage”, que tiene que ver con la porción de arbitraje que se debe a un ajuste asincrónico en los precios de los activos luego de un arribo de información (arbitraje tóxico) y aquel debido a shocks de liquidez (no tóxico). Este trabajo estudia el efecto que tiene el “mix arbitrage” sobre la liquidez.

Dentro de las hipótesis testeables más importantes que se derivan del modelo está que la iliquidez debería ser mayor en períodos donde: i) la fracción de oportunidades de arbitraje que son tóxicas es mayor. ii) la probabilidad que una oportunidad de arbitraje tóxico termine con una transacción de arbitraje sea mayor. Adicionalmente, se predice que cuando el arbitraje se hace más tóxico, las oportunidades de arbitraje deberían ser más cortas.

Dentro de los mayores resultados empíricos, se muestra que la probabilidad que una oportunidad de arbitraje tóxico termine con una transacción tiene un efecto positivo sobre la iliquidez. Se encuentra también una relación positiva y significativa entre la fracción diaria de oportunidades de arbitraje que son tóxicas y la iliquidez. El hecho de que la duración de las oportunidades de arbitraje tóxico sean más cortas en los días en que la fracción de oportunidades tóxicas es mayor también se muestra en los datos.

Marshall, Nguyen y Visaltanachoti (2013) [29] presentan un trabajo esencialmente empírico, para detectar como afectan las características de los mercados sobre la presencia de oportunidades de arbitraje. Se utilizaron datos al interior del día transaccional de dos ETFs que replican el comportamiento del índice S&P500.

Los resultados muestran que las diferencias de precios, cuando son suficientes para generar ganancias de arbitraje, son más probables que ocurran en días que los spreads son más

altos. Los resultados muestran que es más probable que se generen oportunidades de arbitraje en los días menos líquidos.

Adicionalmente, realizan un estudio de los eventos de las oportunidades de arbitraje y dentro de los principales resultados obtenidos se muestra una disminución en la liquidez cuando ocurren estos fenómenos.

## Hipótesis

En relación a la liquidez, **“se espera que la presencia de oportunidades de arbitraje sea mayor en aquellos momentos que la liquidez sea menor”**.

## Variables

Antes de detallar aquellas variables relacionadas a la liquidez, es importante mencionar ciertos aspectos considerados en el cálculo de todas las medidas que capturan algún estado del libro de órdenes límites.

Todas las medidas fueron calculadas con una frecuencia de un segundo, luego para obtener la variable  $X_{i,t}$  se calcula el valor promedio que tiene la medida  $X$  en el intervalo de tiempo  $t$  para la pareja de acciones  $i$ , lo anterior se hace de manera separada para Chi X y Euronext. Finalmente, como esta medida tiene que ver con la pareja  $i$ , es decir, la misma acción siendo tranzada en dos mercados, es que para obtener la variable  $X_{i,t}$  final es que se calcula el promedio de la variable observada en Chi X con lo observado en Euronext.

Una de las variables consideradas, para detectar en cada período de tiempo y para cada pareja de acciones utilizadas la liquidez es el depth, esta medida está relacionada específicamente con la profundidad del libro de órdenes límites, que cabe recordar es una de las dimensiones de la liquidez.

El depth tiene que ver con la cantidad de acciones ofrecidas en el libro de órdenes límites a algún precio en particular. Las distintas variaciones del depth que se calcularon son las siguientes:

- Depth beyond: Corresponde a la cantidad de acciones disponibles en el LOB al segundo y tercer mejor precio, tanto de compra como de venta.
- Depth best: Corresponde a la cantidad de acciones disponibles en el LOB al primer mejor precio, tanto de compra como de venta.
- Depth bid: Corresponde a la cantidad de acciones disponibles en el LOB al primer mejor precio, en el lado de la compra.
- Depth ask: Corresponde a la cantidad de acciones disponibles en el LOB al primer mejor precio, en el lado de la venta.

Para el cálculo de todas las medidas asociadas al depth, las variables son dividida por 100.000.

Existen ciertas medidas que son llamadas 'crossed', para el cálculo de estas variables se obtienen los mejores precios existentes entre Chi X y Euronext, es decir, como si se tratará de un único mercado o puede entenderse como el consolidado resultante entre estos dos mercados.

Matemáticamente, la forma de obtener estos mejores precios consolidados es la siguiente:

$$Ask\ quote\ crossed = \min\{Ask\ quote\ Chi\ X, Ask\ quote\ Euronext\} \quad (4.1)$$

$$Bid\ quote\ crossed = \max\{Bid\ quote\ Chi\ X, Bid\ quote\ Euronext\} \quad (4.2)$$

Luego de obtener los mejores precios, se consideran exclusivamente esos lados del LOB a la hora de calcular alguna de las medidas llamada 'crossed'. Esos lados pueden corresponder a Chi X o Euronext, y en algunos casos pueden ser ambos, cuando, por ejemplo, el mejor precio sea el mismo en Chi X y Euronext, por lo que en esos casos se considerará la suma entre Chi X y Euronext para ese lado del libro.

Adicionalmente a las variables del depth antes mencionadas, se calculó el Depth best crossed, Depth bid crossed y Depth ask crossed. Estas medidas son similares del Depth best, Depth bid y Depth ask antes descritas pero calculados a través del mecanismo para las variables crossed antes mencionado.

Otro conjunto de variables relacionadas a la liquidez tiene que ver con el spread. Dentro de las medidas relacionadas al spread se encuentran aquellas relacionadas al bid ask spread y otro conjunto relacionado al effective spread.

Las diferentes versiones del bid ask spread consideradas como covariables son las siguientes:

- Bid ask spread average: Es la diferencia entre el primer mejor precio de venta y el primer mejor precio de compra.
- Bid ask spread crossed: Es similar al bid ask spread average, pero esta vez calculado a través del procedimiento descrito para las medidas 'crossed'.
- Bid ask spread stand: Corresponde al cociente entre el bid ask spread average y el midquote promedio que se observa en ese mismo período de tiempo t.

Adicionalmente a las medidas relacionadas al bid ask spread antes descritas, se calculó el effective spread y sus respectivas componentes. Estas medidas se calcularon utilizando como guía el trabajo de Hendershott, Jones y Menkveld (2011) [24].

El effective spread corresponde a:

$$Effective\ spread_{i,t} = \frac{q_{i,t}(p_{i,t} - m_{i,t})}{m_{i,t}} \quad (4.3)$$

Donde  $q_{i,t}$  es igual a 1 cuando una transacción es iniciada por un comprador e igual a -1 cuando es iniciada por un vendedor,  $p_{i,t}$  corresponde al precio al que se tranza la acción i en

el período  $t$  y  $m_{i,t}$  es el midquote observado en el período  $t$  para la acción  $i$ .

El effective spread puede descomponerse en la suma del realized spread y el adverse selection.

El realized spread corresponde a:

$$\text{Realized spread}_{i,t} = \frac{q_{i,t}(p_{i,t} - m_{i,t+\Delta})}{m_{i,t}} \quad (4.4)$$

El adverse selection corresponde a:

$$\text{Adverse selection}_{i,t} = \frac{q_{i,t}(m_{i,t+\Delta} - m_{i,t})}{m_{i,t}} \quad (4.5)$$

De la definición de estas dos últimas variables, es evidente ver que su suma corresponde a la definición dada para el effective spread.

El effective spread y sus componentes fueron calculadas considerando un  $\Delta$  igual a 30 segundos.

Otra medida que fue calculada relacionada a la liquidez es la medida llamada relative liquidity (RLIQ), la cual es desarrollada por Valenzuela, Zer, Fryzlewicz y Rheinländer (2015) [38]. Esta medida captura la distribución de la liquidez disponible a lo largo del libro de órdenes límites.

La metodología para calcular el relative liquidity presenta algunas diferencias con la desarrollada por los autores antes mencionados, sin embargo, esta variable sigue capturando en esencia lo mismo.

Cuando esta variable es más grande quiere decir que la liquidez disponible se concentra mayormente en torno al primer mejor precio, lo cual puede ocurrir ya sea porque la mayor parte del depth ofrecido está en los mejores precios o porque los siguientes precios son muy cercanos al primer mejor precio.

La metodología usada para el cálculo de esta variable es la siguiente:

1. Al igual que con el resto de las variables, se obtuvo el LOB cada un segundo para cada acción para proceder a calcular la variable. Los siguientes pasos son aplicados con una frecuencia de un segundo.
2. Se calculó la distancia entre precios (ajustándola con el tick) para cada orden límite relativa a los mejores precios, matemáticamente:

$$\Delta_{i,t}^{buy} = (p_t^B - p_i^{buy})/tick \quad (4.6)$$

$$\Delta_{i,t}^{sell} = (p_i^{sell} - p_t^A)/tick \quad (4.7)$$

Donde  $p_t^B$  ( $p_t^A$ ) es el mejor precio de compra (venta) en el segundo  $t$  y  $p_i^{buy}$  ( $p_i^{sell}$ ) es el  $i$ -ésimo mejor precio.

3. Luego, se calcula el porcentaje acumulado del depth ofrecido al  $i$ -ésimo mejor precio.

4. Se calculan los pesos que se le asignarán a cada porcentaje acumulado de depth con la siguiente función  $g(\lambda, \Delta)$ :

$$g(\lambda, \Delta) = \frac{\exp(-\lambda\Delta)}{\sum_{\Delta} \exp(-\lambda\Delta)} \quad (4.8)$$

Lo anterior, se realizó con una función de decaimiento exponencial de parámetro 0.366.

5. Finalmente, se calcula el relative liquidity cada un segundo de la siguiente forma:

$$RLIQ_{i,t}^{buy} = \sum_{\Delta} F_{i,t}^{buy} g(\lambda, \Delta) \quad (4.9)$$

Donde  $F_{i,t}^{buy}$  corresponde a la función de distribución acumulada del depth para el lado de la compra y  $g(\lambda, \Delta)$  es la función de pesos descrita en la ecuación 4.8. El procedimiento es similar para el lado de la venta.

Se calculó el relative liquidity para el lado de la compra, para el lado de la venta y una tercera variable que promedia las dos medidas anteriores. Estas variables se calcularon de manera separada para Chi X y Euronext, y luego se promediaron entre sí, al igual que con el resto de las variables como ya fue mencionado.

Una última medida asociada también a la distribución de la liquidez en el libro de órdenes límites es la variable 'slope'. Se calculó esta medida basado en el trabajo de Foucault, Kozhan y Tham (2016) [15].

La variable 'slope' o pendiente es igual al promedio entre los dos siguientes ratios: i) La diferencia entre el segundo mejor precio de venta y el primer mejor precio de venta, dividido por el número de acciones ofrecidas al primer mejor precio de venta y ii) La diferencia entre el primer mejor precio de compra y el segundo mejor precio de compra, dividido por el número de acciones disponibles al primer mejor precio de compra.

Esta variable captura cuan profundo o delgado es el libro de órdenes límites en los primeros mejores precios, cuando la variable  $slope_{i,t}$  es más grande quiere decir que el LOB es más delgado para ese período t y ese activo i en particular.

La variable slope es mayor cuando el número de acciones ofrecidas al primer mejor precio es menor o cuando los siguientes precios están más distantes del primer mejor precio.

## 4.2.2. Incertidumbre

### Definición e intuición

El concepto de incertidumbre es muy importante en finanzas y economía, pero es un concepto usado en distintas disciplinas de manera muy genérica, aunque con algunas variaciones según sea el caso particular.

Muchas veces el concepto de incertidumbre es usado de manera indistinta con el de riesgo, sin embargo, existen ciertos consensos sobre algunas diferencias tras estas ideas. En general se dice que se está frente a una situación de riesgo cuando una variable puede tomar distintos valores, pero se conoce la probabilidad de ocurrencia de cada caso. Por incertidumbre se entiende cuando sólo existen aproximaciones sobre el posible valor que tomará una variable, pero no hay probabilidades definidas con exactitud.

En el caso particular de esta investigación y en general en finanzas, la incertidumbre está dada porque los precios de los activos como las acciones siguen procesos aleatorios, los cuales pueden modelarse, pero no con exactitud.

Dado este desconocimiento sobre el valor futuro que pueden tomar los precios, por ende, los retornos y todas las otras medidas que se derivan de los precios es que existe una constante preocupación por parte de los agentes del mercado sobre qué tan alta es la incertidumbre existente y las posibles variaciones que pueden tener las variables de interés en el futuro.

Una mayor incertidumbre está asociada a mayores variaciones futuras con respecto al escenario observado actualmente. Dicho lo anterior es que si en el futuro las variables de interés como los precios se alejan dramática e inesperadamente de los precios actuales y algunos inversionistas no logran actualizar sus órdenes es muy probable que se generen instantes en que los activos se encuentren mal valorados, lo que podría dar origen a instantes donde sea posible arbitrar los mercados.

### Literatura previa

En relación a la incertidumbre dentro del libro de órdenes límites, nuevamente Marshall, Nguyen y Visaltanachoti (2013) [29] son unos de los pocos en aportar lo que ellos denominan ‘ruido de microestructura’ y de relacionarlo con arbitraje, concluyendo que pareciera jugar un rol importante dentro del problema de valoración que conduce finalmente a las oportunidades de arbitraje. Dentro de este ‘ruido de microestructura’ están contenidas ciertas variables relacionadas a la incertidumbre.

Los autores antes mencionados, en su trabajo sobre evidencia de arbitraje con ETF's, muestran empíricamente que la desviación estándar del spread aumenta considerablemente durante las oportunidades de arbitraje. Este aumento es entre un 280 % y un 300 % durante la ventana de arbitraje versus 20 días transaccionales atrás al evento en cuestión. Además, durante las ventanas de arbitraje la desviación estándar del spread aumenta versus una ventana in-



mediatamente previa a la aparición de estos eventos. Adicionalmente, se encuentra que la desviación estándar de los retornos aumenta cuando aparecen las oportunidades de arbitraje.

Budish, Cramton y Shim (2015) [8] muestran que la frecuencia de las oportunidades de arbitraje ha variado en el tiempo, y señalan que esta variación está conducida casi enteramente por la variación en la volatilidad de mercado. Luego, a pesar de que el principal argumento de estos autores es que las oportunidades de arbitrajes se dan de forma mecánica dado el diseño de mercado imperante, le asignan un rol importante a la volatilidad y por ende a la incertidumbre presente en los mercados.

## **Hipótesis**

En cuanto a la incertidumbre, **“se espera que a mayor incertidumbre en el mercado sea más probable o más frecuente la existencia de oportunidades de arbitraje”**.

## **Variables**

Las variables que se construyeron para medir la incertidumbre dentro del libro de órdenes límites son la volatilidad del midquote y la volatilidad de los retornos. Se calcularon los retornos a partir del midquote, tanto simples como logarítmicos.

### 4.2.3. Actividad transaccional

#### Definición e intuición

Por actividad transaccional se va a entender el nivel de transacciones que se observan con respecto a un activo en particular. Es decir, tiene que ver con el nivel de dificultad por el cual las órdenes de compra y venta son ejecutadas.

La actividad transaccional es una dimensión que puede variar por activo, pero también puede variar para un mismo activo a lo largo del tiempo. Como se verá más adelante con mayor profundidad, la actividad transaccional se ha relacionado con la liquidez disponible en un mercado.

No es tan difícil ver las relaciones que tiene la liquidez y la actividad transaccional, si entendemos la liquidez como el conjunto de condiciones que presenta la oferta y la demanda de un activo financiero en un mercado en particular, se puede ver que la actividad transaccional no es más que la sinergia entre la oferta y la demanda, luego la actividad transaccional va a estar condicionada por la liquidez.

Al notar lo anterior se puede ver que, al aumentar la actividad transaccional, la liquidez disponible en el mercado va sufriendo variaciones. Es decir, una mayor actividad transaccional, podría relacionarse con una mayor volatilidad de la liquidez disponible.

Como se explicó anteriormente una mayor volatilidad podría generar instantes con precios distorsionados, lo cual podría dar pie a la presencia de oportunidades de arbitraje, sobre todo si se trata de la volatilidad de una de las dimensiones más importantes de un mercado, la liquidez.

#### Literatura previa

Marshall, Nguyen y Visaltanachoti (2013) [29] encuentran en su estudio de casos de arbitraje que el valor tranzado aumenta mientras están presentes las oportunidades de arbitraje, lo cual indica un aumento en la volatilidad de la liquidez.

El riesgo de liquidez tiene que ver con el riesgo de que para algún activo en particular no se pueda encontrar una contraparte debido a la ausencia de liquidez en ese mercado, lo que lleva a que los compradores tengan que aumentar sus precios o los vendedores tengan que disminuir sus precios para poder tranzar el activo en cuestión.

Johnson (2008) [27] señala que el volumen representa el segundo momento de los cambios de la liquidez o riesgo de liquidez. Pastor, Stambaugh (2003) [36] y Acharya, Pedersen (2004) [1] destacan la importancia del riesgo de liquidez sobre la fijación de precio de los activos.

Johnson (2008 [27]) señala sobre el riesgo de liquidez, que es importante desde el punto de vista de las políticas, debido al peligro que suponen grandes caídas en la liquidez, lo que

puede provocar distorsiones en los precios, interrupciones en la transferencia de riesgos y posiblemente liquidación ineficaz de las inversiones reales.

El aumento que se produce del valor total de las transacciones, en el trabajo de Marshall, Nguyen y Visaltanachoti (2013) [29], mientras están presentes las oportunidades de arbitraje podrían ser un indicio también de que existen ciertos agentes en el mercado que sacan provecho de estas oportunidades de arbitraje, con lo que se ve aumentada la actividad transaccional en presencia de estos eventos.

## Hipótesis

En relación a la actividad transaccional, **“se espera que a mayor actividad transaccional sea más probable o más frecuente la existencia de oportunidades de arbitraje”**.

## Variables

A diferencia del cálculo de las variables relacionadas a las dimensiones antes descritas, para las presentes variables no se consideró el LOB cada un segundo. Las variables asociadas a la actividad transaccional son medidas calculadas a partir del libro de órdenes y transacciones, y dependen del nivel de órdenes y transacciones que se observe en el período de tiempo  $t$  en particular.

Las variables que capturan el nivel transaccional son:

- $Volumen_{i,t}$ : Corresponde a la cantidad de acciones  $i$  tranzadas durante el período  $t$ .
- $Valor_{i,t}$ : Corresponde a la cantidad de acciones  $i$  multiplicado por el respectivo precio al que son tranzadas durante el período  $t$ .
- $Órdenes\ de\ compra_{i,t}$ : Corresponde a la cantidad de órdenes de compra entrantes para la acción  $i$  durante el período de tiempo  $t$ .
- $Órdenes\ de\ venta_{i,t}$ : Corresponde a la cantidad de órdenes de venta entrantes para la acción  $i$  durante el período de tiempo  $t$ .

## 4.2.4. Otros

### Definición e intuición

En la presente sección se agregan otros controles, que buscan identificar algún estado del libro de órdenes límites que no pueda ser clasificado directamente dentro de las principales dimensiones antes descritas, es decir, variables que no tengan una directa relación con la liquidez, la actividad transaccional ni la volatilidad.

Esta dimensión es más generalizada por lo que no se plantean ni desarrollan hipótesis. El principal objetivo de agregar esta sección es incluir otros controles que ayuden a aislar el efecto de las variables relacionadas a las dimensiones anteriores, y de esta forma darles mayor robustez a las conclusiones obtenidas.

### VARIABLES Y LITERATURA PREVIA

Dentro de los controles considerados están:

- **Distbid:** Corresponde a la diferencia entre el primer mejor precio de compra y el tercer mejor precio de compra. Esta variable es un buen control pues toma en consideración de cierta forma las discrepancias que tienen los inversionistas a lo largo del libro de órdenes límites o también puede ser visto como la estrechez que tiene el libro, por lo que al considerar esta covariable podría aislarse el efecto mecánico debido a que los precios a lo largo del libro se encuentran más cercanos.
- **Distask:** En este caso corresponde a la diferencia entre el tercer mejor precio de venta y el primer mejor precio de venta. La utilidad de esta variable es similar a la de la variable descrita anteriormente, pero en este caso asociado al lado de venta del libro de órdenes límites.
- **Diff Midquote:** Esta variable corresponde al valor absoluto entre la diferencia del precio o midquote promedio que se observa en Chi X con el observado en Euronext. De cierto modo esta variable también captura cierta discrepancia existente entre los inversionistas que tranzan un mismo activo, en este caso la diferencia es entre quienes lo hacen en Chi X y quienes lo hacen en Euronext.

Otra variable considerada es el desbalance de órdenes u 'order imbalance'. Marshall, Nguyen y Visaltanachoti (2013) [29] en su trabajo con ETF's muestran que durante los eventos de arbitraje que ellos estudian el desbalance de órdenes aumenta, si bien es importante mencionar que ellos calculan el desbalance de órdenes como la diferencia entre el valor absoluto de las transacciones iniciadas por compradores y las transacciones iniciadas por vendedores, dividida por la suma de las dos, a diferencia de lo realizado en este trabajo en que el desbalance de órdenes es igual al valor absoluto de la diferencia entre el depth ofrecido al mejor precio de compra y el depth ofrecido al mejor precio de venta. Se calculó el desbalance de órdenes antes descrito y su versión 'crossed'.

### 4.3. Metodología econométrica

Para testear las hipótesis antes descritas se utiliza el tradicional modelo logit, donde la variable dependiente es igual a 1 para aquellos períodos y parejas de acciones en que hay presencia de oportunidades de arbitraje y 0 en caso contrario. Es decir, la variable dependiente es de la forma:

$$y_{i,t} = \begin{cases} 1, & \text{hay oportunidad de arbitraje para la pareja de acciones } i \text{ en el periodo } t \\ 0, & \sim \end{cases} \quad (4.10)$$

Esta variable dependiente <sup>1</sup> es regresionada con un conjunto de medidas que indican alguna dimensión o estado del libro de órdenes límites para la pareja de acciones  $i$  en el período  $t$  o  $t-1$ . Las covariables utilizadas son las variables antes descritas. Las variables tanto independientes como dependientes se han construido de tal forma que se está trabajando con datos de panel.

En este caso los individuos están determinados por las distintas parejas de acciones  $i$ , son la misma acción, pero se habla de pareja de acciones pues es la misma acción, pero tranzada en dos mercados distintos, Chi X y Euronext. En cada período utilizado es posible distinguir las características de distintas parejas de acciones  $i$ , y para cada pareja de acciones  $i$  es posible ver las características que prevalecen en los libros de órdenes límites en distintos períodos temporales  $t$ .

Por tratarse de datos de panel, es que al realizar las distintas configuraciones del modelo logit se consideran efectos fijos tanto por individuos como por períodos de tiempo, pues el objetivo es determinar cómo influyen las características del libro sobre las oportunidades de arbitraje, por lo que es necesario limpiar el efecto que podría deberse exclusivamente a alguna pareja de acciones en particular o a algún período de tiempo específico y no al estado del libro de órdenes límites (LOB) ni al libro de órdenes y transacciones.

Se utilizaron principalmente dos configuraciones temporales al estimar los distintos modelos logit.

El primero tiene como objetivo determinar cuáles son las características del libro que más prevalecen cuando ocurren las oportunidades de arbitraje, en este modelo cada período  $t$  de tiempo representa un día transaccional en específico. Luego, la variable dependiente será igual a 1 si en el *día*  $t$  para la *pareja*  $i$  de acciones se observa una oportunidad de arbitraje entre Chi X y Euronext. Lo que se busca con este modelo es conocer cuáles son las características del libro que se observan en los días que hay oportunidades de arbitraje.

El segundo modelo tiene como objetivo determinar si existen variables que tengan algún efecto predictivo sobre la posibilidad de que ocurra una oportunidad de arbitraje. Para este modelo, el período  $t$  está dado por intervalos de cierta duración, concretamente se utilizan intervalos de dos horas.

---

<sup>1</sup>sólo se consideraron aquellas oportunidades de arbitraje de duración mayor o igual a 1 segundo

Luego, como las oportunidades de arbitraje se detectaron entre las 8:00 y las 16:00, para cada día existen 4 períodos de duración dos horas. Sin embargo, para cada día sólo se tomaron como variable dependiente los últimos 3 períodos de dos horas. No se considera el primer intervalo pues las variables dependientes son regresionadas con variables independientes que se observan en el intervalo inmediatamente anterior, por lo que como con en el primer intervalo no se tiene información previa dentro del mismo día transaccional, no se consideró como variable dependiente.

Las variables dentro de cada período de dos horas son calculadas en sub intervalos de 30 minutos, luego para obtener la variable independiente final se promedia considerando los intervalos de 30 minutos existentes dentro del período de dos horas correspondiente y que estén dentro del horario transaccional de ese día en particular.

El cálculo es hecho cada 30 minutos y no directamente cada 2 horas debido a que existen días en que la actividad transaccional termina antes de las 16:00 horas y sólo se consideran sub intervalos en que su duración completa este dentro del horario transaccional.

Si se hubiesen calculado las variables cada dos horas y siguiendo el criterio anterior, se hubiesen dejado fuera muchos intervalos, debido a esto y sumado al hecho de que las conclusiones no deberían variar drásticamente es que se siguió el procedimiento antes descrito.

# Capítulo 5

## Resultados

### 5.1. Modelo diario

A continuación, se hace un análisis de los principales resultados obtenidos a partir del modelo diario, el cual como se indicó anteriormente consiste en un modelo logit en que la variable dependiente binaria es igual a 1 para aquellos días y parejas de acciones en que se presentó una oportunidad de arbitraje, y cero en caso contrario.

Las variables independientes consideran algún estado en particular del libro, el *día*  $t$  y para esa *pareja*  $i$  de acciones en cuestión. Es decir, es el modelo que busca entender que características prevalecen aquellos días en que se presenta alguna oportunidad de arbitraje.

Las variables independientes consideradas que buscan capturar la liquidez de una acción en el libro de órdenes límites son principalmente el spread y el depth, con sus respectivas variaciones. Aquellas relacionadas a la distribución de la liquidez son la variable ‘relative liquidity’ y ‘slope’.

En el modelo diario, al observar los efectos que tienen estas medidas de liquidez de manera aislada, es decir, solo regresionando la variable dependiente binaria con una medida de liquidez como variable independiente<sup>1</sup>, se observa que el depth y sus distintas variaciones tienen un efecto significativo sobre la probabilidad de ocurrencia de un evento de arbitraje en un día en particular para una pareja de acciones específica.

En la tabla 7.3 se ven los resultados al considerar como única covariable el depth en sus distintas versiones. El efecto del depth es significativo de manera aislada y con el signo esperado, pues en todas las variantes del depth se ve que a mayor depth disponible en el libro de órdenes límites menor es la probabilidad de que exista una oportunidad de arbitraje en un *día*  $t$  en particular, todas las versiones calculadas del depth tienen una significancia muy alta pues tienen un p-value menor a 0.01.

---

<sup>1</sup>todos los modelos consideran efectos fijos por individuo y período de tiempo

A mayor depth, el mercado es más líquido en cuanto a la profundidad disponible, luego a mayor liquidez menor es la probabilidad de que ocurra en un *día t* una oportunidad de arbitraje como era lo esperado.

En relación al spread y sus distintas variaciones en el modelo diario, se puede ver que las conclusiones no son tan categóricas como en el caso del depth. Cuando el spread es más grande, se está en una situación de mayor iliquidez, luego según las hipótesis antes desarrolladas se esperaría que a mayor spread mayor probabilidad de ocurrencia de un evento de arbitraje.

En la tabla 7.4 se presentan los resultados al considerar como única covariable el spread en sus distintas versiones. Se puede notar que al considerar el bid ask spread promedio y el bid ask spread dividido por el midquote promedio (bid ask spread estandarizado) como únicos regresores se tiene el efecto contrario al esperado, sin embargo, su efecto sobre la probabilidad de ocurrencia de arbitraje en un *día t* no presenta significancia estadística.

Existen autores que consideran como una mejor medida de liquidez el effective spread, esta medida podría reflejar de mejor forma la liquidez del mercado ya que considera la distancia entre el verdadero precio al cual se transa y el midquote. Al considerar esta variable como único regresor, calculada a 30 segundos, no se ve ningún grado de significancia.

El effective spread puede descomponerse en el realized spread y el price impact o adverse selection, cabe destacar que al considerar como único regresor el adverse selection a 30 segundos se obtiene el signo esperado y con una significancia máxima. Esta medida se relaciona con el impacto en los precios que tienen las transacciones o también puede entenderse como la pérdida bruta de los demandantes de liquidez debido a selección adversa (Hendershott, Jones, Menkveld 2011 [24]).

Es decir, esta última variable es probablemente la que se encuentra más relacionada en estricto rigor con la liquidez, específicamente con la dimensión de resiliencia. Un mercado será más líquido en el sentido de resiliencia cuando el impacto en los precios debido a una transacción sea lo más bajo posible.

Los resultados que se ven en la tabla 7.4 muestran que a mayor impacto en los precios a corto plazo (30 segundo) mayor es la probabilidad de ocurrencia de arbitraje, lo cual es coherente con la hipótesis de que a menor liquidez disponible mayor es la probabilidad de ocurrencia de un evento de arbitraje.

Los resultados obtenidos de utilizar múltiples covariables se muestran en la tabla 7.9, donde ya no se regresa la variable dependiente con una sola covariable sino que con un conjunto de covariables que miden distintas dimensiones y estados del libro.

Se puede ver en la tabla 7.9, que el efecto del depth se mantiene al agregar el efecto de otras variables, lo cual le da gran robustez a esta variable y consolida la relación negativa existente entre la profundidad de un mercado y la presencia de oportunidades de arbitraje en un día transaccional.

Al probar con distintas configuraciones, la medida relacionada con el depth que perdura



frente a las otras es el *depth beyond*, que en este caso representa el *depth* disponible al segundo y tercer mejor precio, tanto de compra como de venta. Esta variable tiene una gran significancia en distintas configuraciones econométricas.

Lo anterior se puede ver en el modelo 4 presentado en la tabla 7.9, donde al considerar en un mismo modelo el *depth best* y el *depth beyond* como covariables, se aprecia que es esta última versión del *depth* la que sobrevive en cuanto a significancia. Adicionalmente, en la misma tabla se aprecia que en todas las configuraciones con múltiples covariables, cuando se utiliza el *depth beyond*, siempre esta variable presenta una significancia muy alta ya que presenta un *p-value* menor a 0.01.

Esto es algo relevante puesto que existe cierto consenso de que el *depth* a los mejores precios es la información más relevante y más considerada por los inversionistas, a través de la presente investigación se puede ver que el *depth* a los mejores precios si es relevante en cuanto a la presencia de oportunidades de arbitraje, pero pareciera ser más informativo el *depth beyond*.

Lo anterior, pudiera deberse a que el *depth beyond* en este caso solo considera el segundo y tercer mejor precio, pues no se tiene un detalle mayor del libro de órdenes límites, por lo cual de todas formas se está tomando en cuenta el *depth* ofrecido a precios probablemente muy cercanos a los primeros mejores precios y no tan distantes como podría entenderse el *depth beyond*.

Otro aspecto relevante a destacar en relación al *depth*, es que al considerar las características que predominan en los días que hay oportunidades de arbitraje el lado de la compra es más significativo e informativo que el lado de la venta.

Por ejemplo, en el modelo 3 de la tabla 7.9 se consideran simultáneamente el *depth bid* y el *depth ask* consolidado o ‘crossed’, es la primera variable la que mantiene su significancia. En la misma tabla, se puede apreciar el modelo 7 donde se considera como covariables al mismo tiempo el *depth bid* y *depth ask* promedio, donde nuevamente es el lado de la compra el que presenta una mayor significancia comparativamente al lado de la venta.

Lo anterior se destaca pues existen distintos investigadores que señalan que efectivamente es el lado de compra el lado más informativo dentro de la configuración del libro de órdenes límites, es decir, no es simétrica la información revelada en el libro de órdenes límites. Por ejemplo, Valenzuela, Zer, Fryzlewicz y Rheinländer (2015) [38] destacan esta asimetría existente entre el lado de la compra y el lado de venta del libro de órdenes límites, estos autores señalan específicamente que es el lado de la compra el más informativo de la volatilidad futura.

En los modelos que consideran múltiples variables independiente se encontró que la mayoría de las versiones del *spread* no poseen significancia estadística.

El *realized spread*, que como se mencionó anteriormente es una componente del *effective spread*, es la única variable del *spread* que frente a múltiples covariables mantiene su efecto negativo y significativo en algunas configuraciones. Como se mencionó, esto no contradice las hipótesis desarrolladas, pues es la componente llamada ‘adverse selection’ o de impacto

en los precios la medida mayormente asociada a la liquidez, la cual no presenta significancia. Además, podría deberse sólo a un efecto mecánico la significancia del realized spread.

Otra posible explicación a el hecho de que al considerar algunas variantes del spread no sean significativas y no tengan el signo esperado, que, por ejemplo, si encuentra en su estudio Marshall, Nguyen y Visaltanachoti (2013) [29], es lo que señalan muchos investigadores de que la mayor velocidad de los arbitradores puede dañar la calidad de los mercados, dado que los dealers pueden ampliar el bid ask spread para cubrirse del riesgo de no poder actualizar los precios (Copeland y Galai, 1983 [11]).

Es decir, dada la gran velocidad de algunos inversionistas, quienes no poseen esta velocidad tienen un costo de selección adversa por lo que podrían ser más precavidos con la liquidez que ofrecen en el mercado, ampliando a priori el spread. Por lo que habría otro fenómeno que explicaría los movimientos del spread.

En la tabla 7.5 se aprecian los resultados de utilizar como únicos regresor aquellas variables relacionadas a la forma en qué se distribuye la liquidez a lo largo del libro de órdenes límites.

En la tabla antes mencionada se puede notar que no sólo influye la liquidez sobre la presencia de oportunidades de arbitraje, sino también cómo esta se distribuye a lo largo del libro de órdenes límites. Lo anterior se puede notar dada la alta significancia que presentan las distintas versiones del ‘relative liquidity’ y la significancia que también presenta la variable ‘slope’ de manera aislada.

Cuando el ‘relative liquidity’ es más alto quiere decir que la liquidez está más concentrada en torno al mejor precio, lo cual puede ocurrir ya sea porque los otros precios son muy cercanos al mejor precio o porque el depth ofrecido se concentra principalmente en los primeros mejores precios.

Como ya se ha mencionado, esta medida puede interpretarse como una medida de distribución de liquidez. Dada la alta significancia estadística y la relación positiva encontrada entre esta variable y la presencia de oportunidades de arbitraje, es que se puede pensar que es conveniente para el mercado que la liquidez no sólo se concentre en los mejores precios, sino que esté disponible a lo largo del libro de órdenes límites para de esta forma evitar la presencia de oportunidades de arbitraje.

Lo anterior debido a que, al consumirse la liquidez ofrecida a los mejores precios mediante una gran transacción, por ejemplo, si no existe liquidez suficiente en el remanente del LOB pueden generarse situaciones de distorsiones de precios que colaboren en la formación de oportunidades de arbitraje.

La importancia de la variable ‘relative liquidity’ no sólo se puede apreciar en la tabla 7.5 de manera aislada, sino que también al considerar múltiples controles en la tabla 7.9. En todos los modelos en que se consideró la variable ‘relative liquidity’ en la tabla 7.9 se aprecia un p value menor a 0.01 y con una relación positiva con la presencia de oportunidades de arbitraje.

Otro aspecto relevante a destacar en torno a la variable ‘relative liquidity’ tiene que ver nuevamente con la asimetría existente entre el lado de compra y el de venta del LOB.

En el modelo 5 presente en la tabla 7.9 se consideró como variante del ‘relative liquidity’ esta medida, pero sólo considerando el lado de la compra y en el modelo 6 presente en la misma tabla se consideró la variable, pero sólo con el lado de la venta, se puede ver que si bien en ambos casos el efecto es positivo sobre la presencia de oportunidades de arbitraje, pero que cuando se considera la versión del lado de la compra la variable tiene un p value menor a 0.01 y al considerar el lado de la venta hay un p value sólo menor a 0.1 para esta variable.

Nuevamente, se aprecia una significancia mayor para una variable cuando se considera el lado de la compra comparativamente al lado de la venta.

En relación a la variable ‘slope’ se puede ver en la tabla 7.9 que en la mayoría de las versiones en que se consideró esta variable, se mantuvo la alta significancia y el efecto positivo sobre la presencia de oportunidades de arbitraje.

Sólo en el modelo 2 presente en la tabla 7.9 perdió su significancia, lo cual puede deberse probablemente a que se consideró en ese modelo simultáneamente la variable ‘relative liquidity’ y la variable ‘slope’, ambas relacionadas a la distribución de la liquidez, y donde la variable que mantuvo su significancia fue la llamada ‘relative liquidity’.

Sin embargo, el análisis de la variable slope es distinto al de la variable relative liquidity, puesto que la variable slope toma en cuenta mayormente la distribución presente en los mejores precios del LOB, en cambio la variable ‘relative liquidity’ puede considerar precios más lejanos.

Mientras mayor es ‘slope’ quiere decir que el libro es más delgado en los mejores precios, lo cual puede deberse a que haya una menor cantidad de acciones ofrecidas a esos precios o a que los otros precios estén más distantes.

Los resultados muestran que a mayor slope mayor probabilidad de que exista un evento de arbitraje, lo cual ratifica la importancia del depth en los mejores precios y de la forma en que se distribuye la liquidez a lo largo del LOB.

En la tabla 7.6 se pueden ver los resultados obtenidos al considerar cómo única covariable alguna medida relacionada a la incertidumbre presente en el libro de órdenes límites, ya sea la volatilidad del midquote o de los retornos, en su versión simple o logarítmica.

Se puede ver que todas las variables calculadas tienen una alta significancia y una relación positiva con la presencia de oportunidades de arbitraje, lo cual muestra que los resultados son coherentes de acuerdo a las hipótesis desarrolladas con anterioridad de que a mayor incertidumbre más probable es la aparición de eventos de potencial arbitraje.

Es decir, la probabilidad de que en un *día t* para una *pareja i* de acciones exista un instante de arbitraje es mayor mientras mayor sea la volatilidad observada en ese *día t* para esa *pareja i* de acciones.

Estos resultados muestran una gran robustez, puesto que el efecto positivo y la alta significancia se mantienen en los modelos que consideran múltiples estados del libro, lo que se puede apreciar en mayor detalle en la tabla 7.9. Lo anterior ocurre al considerar cualquiera de las medidas asociadas a la incertidumbre presente en el LOB.

Con respecto a cómo se relaciona la actividad transaccional de manera aislada y la presencia de oportunidades de arbitraje se pueden ver los resultados en la tabla 7.7.

Los resultados indican que a mayor actividad transaccional más probable es la ocurrencia de algún instante de arbitraje. Tanto al medir la actividad transaccional a través del volumen tranzado o del valor tranzado en un día en particular de manera aislada, se encuentra que tiene una relación positiva y significativa (p value menor a 0.01) con la presencia de eventos de arbitraje.

El volumen es el número de acciones tranzadas en un día en particular y el valor es el número de acciones tranzadas multiplicado por el precio respectivo al cual se tranzan, claramente existe un problema de colinealidad al considerar ambas variables juntas en un modelo.

Sin embargo, se puede ver que en el modelo diario es el volumen el que sigue teniendo significancia al considerar también como covariable el valor tranzado, lo que se puede observar en el modelo 3 presente en la tabla 7.9.

Se utilizó como covariables el logaritmo natural del volumen y del valor tranzado con el objetivo de suavizar el efecto de estas variables, dada la alta magnitud que presentan estas medidas comparativamente con las otras variables.

En todas las configuraciones con múltiples variables, a excepción de cuando van juntas volumen y valor, el efecto del volumen tiene una alta significancia, teniendo siempre un p value menor a 0.01. Se puede concluir que el volumen es una variable explicativa con gran robustez que explica que características predominan en los días que se presentan instantes de arbitraje.

Luego, se puede concluir que a mayor actividad transaccional en un *día t* en particular más probable o frecuente es la presencia de oportunidades de arbitraje.

Los resultados de utilizar como único regresor las variables que no se encuentran relacionadas directamente con ninguna de las dimensiones del libro antes descritas, y que están presentes en la categoría otros, son presentados en la tabla 7.8.

En la tabla antes mencionada se puede ver que el desbalance de órdenes, la diferencia de miduqote y el desbalance de órdenes consolidado o ‘crossed’ si presentan un efecto significativo al considerar estas variables de manera aislada sobre la presencia de oportunidades de arbitraje, pero con un efecto marginal muy bajo en cuanto a magnitud.

Si bien al considerarlas de manera aisladas estas variables presentan un alto nivel de significancia, en ninguna configuración con múltiples controles son significativos sus efectos como se puede ver en la tabla 7.9, por lo que puede concluirse que los modelos aislados sólo

muestran el efecto producido por otras variables no consideradas.

A modo de resumen, al considerar las variables que predominan y tienen un efecto significativo en los días que hay presencia de oportunidades de arbitraje, se puede mencionar que, en los días con mayor liquidez en cuanto a profundidad disponible es menos probable que exista una oportunidad de arbitraje.

También, en los días en que la liquidez está muy concentrada en torno a los mejores precios comparativamente con el resto del libro de órdenes límites es más probable que haya una oportunidad de arbitraje.

Finalmente, aquellos días más volátiles y con una mayor actividad transaccional presentan una mayor probabilidad de ocurrencia de arbitraje.

## 5.2. Modelo predictivo

Los resultados anteriores tienen que ver con las características que predominaban los días que hay presencia de alguna oportunidad de arbitraje. A continuación, se muestran los principales resultados obtenidos en relación a aquellas variables de estado del libro de órdenes límites que podrían ser consideradas predictores de algún evento de arbitraje futuro.

En relación a aquellas variables contemporáneas del LOB que ayudan a predecir la presencia de oportunidades de arbitraje en un período de tiempo futuro se puede ver que el depth también tiene una alta significancia predictiva y con un comportamiento similar al visto en los modelos diarios.

En la tabla 7.10 se presentan los resultados del modelo predictivo con intervalos de dos horas dentro del día transaccional, en que se utiliza cómo único regresor del logit la variable depth en sus distintas versiones.

A mayor depth disponible actual en el libro de órdenes límites, menor es la probabilidad de un evento de arbitraje en un intervalo de tiempo inmediatamente sucesivo, lo cual puede verse de manera aislada en la tabla 7.10 o considerando múltiples efectos en la tabla 7.16.

En el modelo 3 presente en la tabla 7.16 se puede ver nuevamente que al utilizar en un mismo modelo el depth best y el depth beyond, es esta última variable la que presenta una mayor significancia estadística.

Por lo que nuevamente, es el depth beyond la variable más importante dentro de las distintas versiones del depth calculadas. En todos los modelos en que se utilizó como covariable el depth beyond, a excepción del modelo 3, esta variable siempre presenta un p value menor a 0.01.

En el modelo predictivo no se observa la asimetría en la información revelada por el lado de la compra y la revelada por el lado de la venta con respecto al depth, hecho que si ocurría

en el modelo diario.

Como se puede ver en el modelo 5 presente en la tabla 7.16 al considerar el depth bid crossed y el depth ask crossed, ambas covariables mantienen un alto nivel de significancia estadística, lo mismo ocurre en el modelo 6 presente en la misma tabla al considerar simultáneamente el depth bid promedio y el depth ask promedio.

En la tabla 7.11 se aprecian los resultados del modelo predictivo de intervalos de dos horas en que se utiliza como único regresor del logit la variable spread, en sus distintas versiones.

En esta tabla se puede ver que al igual que en el modelo diario la componente del effective spread llamada adverse selection tiene un efecto positivo y significativo sobre la presencia de oportunidades de arbitraje futuras, y que la componente llamada realized spread tiene un efecto negativo y ligeramente menos significativo que la componente antes mencionada.

Nuevamente, el effective spread no tiene significancia, pero en este caso, las distintas versiones del bid ask spread presentan una alta significancia estadística y una relación negativa con la presencia de oportunidades de arbitraje futuras.

Al considerar múltiples covariables, el bid ask spread en sus distintas versiones y el realized spread, mantienen su significancia estadística como se puede ver en la tabla 7.16.

La forma en que influyen predictivamente estas variables sobre la presencia de oportunidades de arbitraje llama la atención pues, los resultados muestran que a mayor spread menor es la probabilidad de una oportunidad de arbitraje futuro, lo que indica que mientras mayor sea la iliquidez medida a través del spread actualmente, menos probable es que en el futuro haya una oportunidad de arbitraje.

Sin embargo, cuando se considera el effective spread que como se indicó anteriormente es una mejor medida de liquidez relacionada al spread, desaparece la significancia predictiva de esta variable.

Lo anterior podría indicar, que los resultados que se pueden ver en la tabla 7.11 de manera aislada o en la tabla 7.16 considerando múltiples controles, relacionados al bid ask spread están más bien mostrando un efecto mecánico debido a la mayor estrechez existente entre los precios, que hace más fácil que sólo pequeñas variaciones en la liquidez consumida hagan posible una distorsión en los precios y no reflejen un estado real de la liquidez del libro.

Además, recordando que el adverse selection es probablemente la medida más relacionada a una de las dimensiones de liquidez en el mercado, resiliencia, se ve que de manera aislada también tiene un poder predictivo importante y consecuente con las hipótesis planteadas. Sin embargo, no se puede decir que este poder predictivo es robusto puesto que, al considerar esta variable en modelos con múltiples controles, su efecto desaparece.

Los resultados de considerar como único regresor para el modelo predictivo de intervalos de dos horas, alguna variable relativa a la distribución de la liquidez, se pueden ver en la tabla 7.12.

Nuevamente, la variable *relative liquidity* tiene un efecto positivo y muy significativo sobre la presencia de oportunidades de arbitraje, en este caso en un intervalo de tiempo futuro.

La significancia de la variable *slope* desaparece en el modelo predictivo, lo que indicaría que es más importante la distribución de la liquidez a lo largo del LOB y no sólo con respecto a los primeros precios.

En la tabla 7.16 se puede apreciar que en todos los modelos que se utiliza como regresor a la variable *relative liquidity*, esta variable presenta siempre un efecto positivo y con un *p* value menor a 0.01.

En el modelo 7 presente en la misma tabla, se aprecia que nuevamente el *relative liquidity* del lado de la compra presenta una mayor significancia estadística comparativamente al lado de la venta.

En el modelo predictivo nuevamente aparece la importancia de la volatilidad de los precios y de los retornos, para en este caso predecir la aparición de eventos de arbitraje en un intervalo inmediatamente próximo.

En la tabla 7.13 se pueden ver los resultados de considerar como único regresor en el modelo predictivo con intervalos de dos horas, alguna variable relacionada a la incertidumbre presente en el libro de órdenes límites.

Se puede concluir nuevamente que una mayor incertidumbre actual aumenta la probabilidad de un evento de arbitraje, en este caso en un intervalo inmediatamente sucesivo.

Sin embargo, en la tabla 7.16 se puede ver que si bien en la mayoría de las configuraciones, en que se utiliza alguna variable de incertidumbre esta tiene un efecto positivo y significativo, pero, presenta una significancia menor al modelo diario.

Al tomar en cuenta la actividad transaccional, se puede ver que en el modelo predictivo nuevamente queda claro que a mayor actividad transaccional más probable es la aparición de un evento de arbitraje.

En la tabla 7.14 se pueden ver los resultados de utilizar como único regrsor en el modelo predictivo con intervalos de dos horas, alguna variable asociada a la actividad transaccional.

En la tabla antes mencionada, se observa un comportamiento similar al modelo diario, donde el logaritmo natural del volumen y del valor tranzado tienen un efecto positivo y muy significativo sobre la presencia de oportunidades de arbitraje, y donde las órdenes de compra y venta no tienen ninguna incidencia sobre la presencia de oportunidades de arbitraje en un intervalo inmediatamente próximo.

En la tabla 7.16 se puede apreciar que, al utilizar la variable logaritmo natural del volumen tranzado, está siempre tiene un *p* value menor a 0.01 en todas las configuraciones y presenta una relación positiva sobre la presencia de oportunidades de arbitraje futuras.

Nuevamente, el volumen tranzado tiene una significancia mayor al valor tranzado, lo cual se puede ver en el modelo 2 presente en la tabla recién mencionada, donde al utilizar estas variable juntas en un mismo modelo, es el volumen el que mantiene su alta significancia.

Los resultados de utilizar como único regresor en el modelo predictivo de intervalos de dos horas, las variables que no se encuentran relacionadas directamente con ninguna de las dimensiones del libro antes descritas, y que están presentes en la categoría otros, son presentados en la tabla 7.15.

La única variable que frente a otros controles mantiene una significancia estadística es la variable llamada order imbalance, la cual como se puede ver en los modelos 4, 5 y 6 presentes en la tabla 7.16 presenta significancia estadística y con una relación positiva sobre la presencia de oportunidades de arbitraje.

Lo anterior, tiene sentido pues cuando la liquidez está cargada a un lado del libro, ya sea el de compra o venta, es más fácil que alguna transacción haga que se consuma la poca liquidez comparativa de un lado y de esta forma el libro quede muy cargado hacia un lado, ya sea por compradores o vendedores y la ausencia de lo opuesto.

Este hecho podría provocar distorsiones en los precios debido a una ausencia de la contraparte, pudiendo generar eventos de arbitraje futuro, lo cual nuevamente da cuenta de la importancia no solo de la liquidez sino también de que esta esté distribuida a lo largo del libro y no esté tan concentrada.

En este caso al hablar de libro, se hace referencia al libro consolidado entre Chi X y Euronext, lo cual también es relevante de destacar ya que se puede ver que no sólo las características propias de un libro pueden influir en la presencia de oportunidades de arbitraje, sino que también las características que resultan a la hora de considerar a ambos mercados como uno.

Sin embargo, no se puede concluir con gran robustez sobre este fenómeno, debido a que por una parte al regresionar esta variable de manera única en el logit predictivo, presenta un efecto contrario y significativo. Además, la magnitud del coeficiente asociado es muy bajo comparativamente al resto de estimadores.

El efecto que parecieran tener las variables *distbid* y *distask*, que se pueden ver en la tabla 7.15, desaparece al considerar múltiples controles, como se aprecia en la tabla 7.16. Luego, estas variables no tienen un poder explicativo sobre la presencia de oportunidades de arbitraje futuras.

A modo de resumen, se puede destacar la importancia que nuevamente tiene en el modelo predictivo de oportunidades de arbitraje las variables *depth*, las distintas medidas de volatilidad, la actividad transaccional medida a través del volumen tranzado y la variable *relative liquidity*, que captura la distribución de liquidez en el libro de órdenes límites.

Lo cual les da robustez a los resultados encontrados en torno a estas respectivas dimensiones del libro, que también se pueden ver en el modelo diario y que consolidan las conclusiones



relativas a las hipótesis desarrolladas en el presente trabajo.

# Capítulo 6

## Análisis de sensibilidad

En el presente capítulo se realiza un análisis de sensibilidad sobre los resultados obtenidos en relación a cómo afectan las características de micro estructura sobre la presencia de oportunidades de arbitraje en un mercado bursátil fragmentado.

Para llevar a cabo lo anterior se realizan nuevas estimaciones de los modelos diarios y predictivos.

Una de las principales diferencias incluidas en los nuevos modelos econométricos tiene que ver con que para todas las estimaciones realizadas en este capítulo de sensibilidad se incorporan todas las oportunidades de arbitraje encontradas, es decir, no se consideran sólo aquellas con una duración mayor o igual a 1 segundo como se hace en el capítulo de resultados sino todos los eventos de duración mayor a 0 segundos.

Lo anterior, se realiza debido a que es interesante saber si los efectos descritos previamente se mantienen para aquellas oportunidades de arbitraje de más corta duración. Esto es muy relevante, ya que como se ha mencionado previamente hay una fuerte evidencia de que los inversionistas poseen cada vez sistemas más sofisticados a la hora de tranzar, los que se caracterizan por tener tiempos de ejecución de órdenes de magnitudes inferiores al segundo.

La segunda principal variación incluida en este capítulo tiene que ver con la variable *gap* definida en la ecuación 3.1, se estimaron los modelos diarios y predictivos con  $gap=0$ ,  $gap=0.005$  y  $gap=0.01$ .

Vale decir, incluyendo todos los eventos de arbitraje, independiente de su duración temporal, se estimaron los modelos con los valores de *gap* antes mencionados. Esta última variación sobre cuando se considera hay una oportunidad de arbitraje tiene el objetivo de corroborar como influyen las variables estudiadas al considerar una diferencia mínima exigida entre los precios involucrados en la oportunidad de arbitraje, tal como queda determinado por la ecuación 3.1.

Los valores antes mencionados para la variable *gap* fueron elegidos de acuerdo a la distribución observada para la diferencia existente entre el precio de compra y el precio de venta

involucrados en las oportunidades de arbitraje encontradas. Esta diferencia tiene una mediana igual a 0.005 y el tercer cuartil toma un valor de 0.01, valores considerados a la hora de estudiar el arbitraje con un gap mayor a cero.

Adicionalmente, para darle mayor robustez a los resultados relativos al poder predictivo que tienen las variables de micro estructura sobre las oportunidades de arbitraje, se estima un modelo predictivo con las mismas características antes explicadas, pero con intervalos de una hora de duración en lugar de dos horas.

Finalmente, en todas las nuevas estimaciones se agregó un nuevo control para darle mayor robustez a los resultados. Este control es la medida de iliquidez de Amihud (2002) [2], este es un control muy utilizado, por ejemplo, Valenzuela, Zer, Fryzlewicz y Rheinländer (2015) [38] en su trabajo que presentan la medida ‘relative liquidity’ lo consideran como un control en sus modelos.

Esta medida es igual al valor absoluto del retorno observado en un período de tiempo dividido por el valor tranzado en ese mismo período de tiempo para algún activo en particular. La medida de iliquidez de Amihud es una medida de iliquidez debido a que está relacionada al impacto que tienen las transacciones de un activo sobre los precios del mismo, es decir, asociada a la dimensión de la liquidez llamada resiliencia.

En la tabla 7.17 se ven los principales resultados del modelo logit diario con las variaciones explicadas en el presente capítulo.

Se puede ver que las variables que presentan un efecto que no desaparece al considerar un *gap* mayor a cero son la volatilidad de los retornos, el logaritmo del volumen tranzado y la variable slope crossed.

Estas variables mantienen el mismo efecto que ya ha sido explicado anteriormente, por una parte, se ve que los días de mayor incertidumbre en cuanto a los precios observados para los activos financieros, es más probable la ocurrencia de los eventos de arbitraje.

Por otra parte, se ve que los días de mayor actividad transaccional es más probable la ocurrencia de estos eventos, lo cual como ya ha sido mencionado se ha relacionado con la volatilidad de la liquidez.

Finalmente, cuando el libro es más estrecho en la parte superior de este, más probable es la ocurrencia de los eventos de arbitraje, lo que se puede ver con el efecto positivo que presenta la variable slope crossed al considerar distintos valores de *gap*.

Además, se ve que el depth beyond crossed tiene un efecto negativo y significativo con  $gap=0$  y  $gap=0.005$ , pero esta significancia desaparece al considerar  $gap=0.01$ . Adicionalmente, se puede ver que nuevamente el depth beyond crossed al ir en un modelo junto al depth best crossed, es la primera variable la que destaca por su significancia.

Finalmente, en cuanto al efecto de la variable relative liquidity crossed se ve que, al considerar todos los eventos, independiente de su duración temporal, esta medida mantiene su

efecto positivo y significativo. Sin embargo, este efecto desaparece al considerar valores positivos para la variable *gap*.

En la tabla 7.18 se pueden ver los principales resultados de los modelos logit predictivos, considerando intervalos de una y dos horas, con las variaciones explicadas en el presente capítulo.

Nuevamente las variables más robustas, debido a que mantienen su efecto y significancia al considerar un *gap* mayor a cero, son la volatilidad de los retornos y el logaritmo del volumen tranzado. Variables que como se puede ver en la tabla 7.18 mantienen un efecto positivo sobre la probabilidad de ocurrencia de arbitraje en un intervalo de tiempo inmediatamente posterior, lo cual ocurre al considerar intervalos de dos o una hora.

En relación al *depth beyond crossed* y al *relative liquidity crossed*, se ven los mismos efectos antes explicados, sin embargo, la significancia de estos efectos desaparece al considerar un *gap* mayor a cero.

En los modelos predictivos se puede ver nuevamente la preponderancia del *depth beyond crossed* por sobre el *depth best crossed*.

El hecho de que los efectos predictivos de las variables de micro estructura se mantengan al considerar intervalos de una hora, le entrega mayor robustez a las conclusiones del presente trabajo, ya que se puede concluir que existen variables que pueden ayudar a predecir la aparición de estos eventos en el futuro, y que esto no está restringido a algún intervalo de tiempo en específico.

Si bien algunas variables pierdan significancia al considerar *gap* mayor a cero puede deberse a que para esos casos hay muy pocas observaciones en las que hay un evento de arbitraje, por lo que podría ser relevante en un estudio futuro ampliar la muestra y corroborar los efectos cuando hay un *gap* mayor a cero.

# Conclusión

A través de la presente investigación se ha podido estudiar un fenómeno que emerge debido a la fragmentación de los mercados financieros, el fenómeno de arbitraje en mercados fragmentados.

Una correcta forma de caracterizar las oportunidades de arbitraje encontradas al analizar 22 acciones pertenecientes al índice AEX durante el año 2014, es utilizando la descripción utilizada por Foucault, Kozhan, Tham (2016) [15] para las oportunidades de arbitraje de alta velocidad, es decir, como eventos; i) frecuentes, ii) de muy corta duración, iii) más eficientemente explotadas por máquinas que por humanos y iv) con una ganancia muy leve por oportunidad de arbitraje.

A través del análisis econométrico realizado, se ha podido comprobar que existen ciertas características en el libro de órdenes límites y en el libro de órdenes y transacciones que prevalecen cuando ocurren los eventos de arbitraje, habiendo incluso ciertas variables que podrían tener un poder predictivo frente a estos eventos.

Se puede concluir con respecto a las hipótesis desarrolladas que los eventos de potencial arbitraje son más frecuentes cuando la liquidez es menor, la dimensión de la liquidez que pareciera tener mayor incidencia sobre la presencia de estos eventos es la profundidad del mercado.

Adicionalmente, no sólo la liquidez es relevante, sino también la forma en que esta se distribuye a lo largo del libro de órdenes límites. No sólo es relevante que haya liquidez en la cima del libro, sino también a lo largo de este. Lo anterior, se puede destacar dada la gran relevancia que tiene la variable *depth beyond*, por sobre el *depth* a los mejores precios, y por la alta significancia que muestra la variable *relative liquidity* en las distintas configuraciones econométricas.

Con respecto a la incertidumbre en el libro de órdenes límites, se puede concluir que efectivamente la presencia de eventos de potencial arbitraje es mayor cuando la incertidumbre es mayor. Adicionalmente, la presencia de estos eventos tiene una relación positiva y significativa con la actividad transaccional.

Uno de los objetivos de la presente investigación es generar evidencia empírica que pueda relacionarse con las distintas ideas actuales desarrolladas en torno al arbitraje en mercados financieros fragmentados.

Por una parte, Marshall et al. (2013) [29] plantean que existen ciertas características de microestructura de mercado que influyen en la presencia de oportunidades de arbitraje, y por otro lado Budish et al. (2015) [8] plantean que la presencia de oportunidades de arbitraje está determinado mecánicamente por el diseño de mercado actual imperante, es decir, del estilo CLOB.

Ambas hipótesis pueden no ser contrarias entre sí, ya que por una parte se evidencia que efectivamente hay ciertas características del LOB o ruido de microestructura, como lo definen Marshall et al. (2013) [29], que influyen sobre la presencia de oportunidades de arbitraje.

Sin embargo, estas características de microestructura no son independientes al diseño de mercado imperante, por lo cual las propuestas de Budish et al. (2015) [8] sobre modificar el diseño actual y transitar hacia uno que elimine las rentas de arbitraje, y con esto todos los posibles costos de selección adversa que sufren los agentes del mercado, podría ser comprensivo.

En relación a los fenómenos derivados de la fragmentación de los mercados financieros y los distintos cambios que se han producido a nivel mundial en la forma de tranzar los activos financieros queda mucha investigación por realizar.

En futuras investigaciones sería pertinente medir la presencia de los llamados algoritmos transaccionales de alta frecuencia, para de esta forma aislar el efecto que podría generar la mayor velocidad a la hora de ejecutar las órdenes y de esta forma estudiar la idea relativa al costo de selección adversa que sufren ciertos agentes del mercado, idea que ha tomado gran fuerza en el último tiempo por distintos investigadores.

Adicionalmente, sería pertinente acompañar la investigación empírica con el desarrollo de modelos matemáticos que expliquen el arbitraje en mercados fragmentados, y que consideren en su desarrollo la no independencia de las variables de microestructura con el diseño de mercado.

Finalmente, estudiar en más detalle la evolución de estos eventos de potencial arbitraje, específicamente estudiar si efectivamente están siendo aprovechadas estas oportunidades de arbitraje o cómo estos instantes en que los activos están mal valorados desaparecen.

# Bibliografía

- [1] ACHARYA, V. V., AND PEDERSEN, L. H. Liquidity risk and asset pricing. *Journal of Financial Economics* 77 (2004), 375–410.
- [2] AMIHUD, Y. Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets* 5, 1 (2002), 31–56.
- [3] AMIHUD, Y., LAUTERBACH, B., AND MENDELSON, H. The value of trading consolidation: evidence from the exercise of warrants. *Journal of Financial and Quantitative Analysis* 38, 04 (2003), 829–846.
- [4] BARUCH, S., KAROLYI, G. A., AND LEMMON, M. Multimarket trading and liquidity: theory and evidence. *The Journal of Finance* 62, 05 (2007), 2169–2200.
- [5] BERNALES, A. How fast can you trade? high-frequency trading in dynamic limit order markets, 2013. Working paper Banque de France.
- [6] BODIE, Z., KANE, A., AND MARCUS, A. J. *Investments, Eight Edition*. International Edition, 2009.
- [7] BOEHMER, B., AND BOEHMER, E. Trading your neighbor’s etfs: Competition or fragmentation? *Journal of Banking and Finance* 27, 9 (2003), 1667–1703.
- [8] BUDISH, E., CRAMTON, P., AND SHIM, J. The high-frequency trading arms race: Frequent batch auctions as a market design response. *The Quarterly Journal of Economics* 130, 4 (2015), 1547–1621.
- [9] CHAKRAVARTY, S., CHIYACHANTANA, C., AND JIANG, C. The choice of trading venue and relative price impact of institutional trading: Adrs versus the underlying securities in their local markets. *Journal of Financial Research* 34, 4 (2011), 537–567.
- [10] CHLISTALLA, M., AND LUTAT, M. Competition in securities markets: the impact on liquidity. *Financial Markets and Portfolio Management* 25, 2 (2011), 149–172.
- [11] COPELAND, T. E., AND GALAI, D. Information effects on the bid ask spread. *the Journal of Finance* 38, 5 (1983), 1457–1469.
- [12] DRAUS, S., AND ACHTER, M. V. Circuit breakers and market runs, 2016.

- [13] EASLEY, D., DE PRADO, M. M. L., AND O'HARA, M. Flow toxicity and liquidity in a high-frequency world. *Review of Financial Studies* 25, 5 (2012), 1457–1493.
- [14] FOUCAULT, T., AND GEHRIG, T. Stock price informativeness, cross-listings, and investment decisions. *Journal of Financial Economics* 88, 1 (2008), 146–168.
- [15] FOUCAULT, T., KOZHAN, R., AND THAM, W. W. Toxic arbitrage. *Review of Financial Studies* (2016), hhw103.
- [16] FOUCAULT, T., AND MENKVELD, A. J. Competition for order flow and smart order routing systems. *The Journal of Finance* 63, 1 (2008), 119–158.
- [17] FOUCAULT, T., PAGANO, M., ROELL, A., AND RÖELL, A. *Market liquidity: theory, evidence, and policy*. Oxford University Press, 2013.
- [18] FROOT, K. A., AND DABORA, E. M. How are stock prices affected by the location of trade? *Journal of Financial Economics* 53, 2 (1999), 189–216.
- [19] GAGNON, L., AND KAROLYI, G. A. Multi-market trading and arbitrage. *Journal of Financial Economics* 97, 1 (2010), 53–80.
- [20] GAJEWSKI, J. F., AND GRESSE, C. Centralised order books versus hybrid order books: A paired comparison of trading costs on nsc (euronext paris) and sets (london stock exchange). *Journal of Banking & Finance* 31, 9 (2007), 2906–2924.
- [21] HALLING, M., MOULTON, P. C., AND PANAYIDES, M. Volume dynamics and multi-market trading. *Journal of Financial and Quantitative Analysis* 48, 02 (2013), 489–518.
- [22] HASBROUCK, J. *Empirical market microstructure: The institutions, economics, and econometrics of securities trading*. Oxford University Press, 2007.
- [23] HASBROUCK, J., AND SAAR, G. Low-latency trading. *Journal of Financial Markets* 16, 4 (2013), 646–679.
- [24] HENDERSHOTT, T., JONES, C. M., AND MENKVELD, A. J. Does algorithmic trading improve liquidity? *The Journal of Finance* 66, 1 (2011), 1–33.
- [25] HENDERSHOTT, T., AND MOULTON, P. C. Automation, speed, and stock market quality: The nyse's hybrid. *Journal of Financial Markets* 14, 4 (2011), 568–604.
- [26] HENGELBROCK, J., AND THEISSEN, E. Fourteen at one blow: The market entry of turquoise. *Review of Finance* 10, 1 (2009), 99–152.
- [27] JOHNSON, T. C. Volume, liquidity, and liquidity risk. *Journal of Financial Economics* 87 (2008), 388–417.
- [28] KOZHAN, R., AND THAM, W. W. Execution risk in high-frequency arbitrage. *Management Science* 58, 11 (2012), 2131–2149.
- [29] MARSHALL, B. R., NGUYEN, N. H., AND VISALTANACHOTI, N. Etf arbitrage: Intraday



- evidence. *Journal of Banking & Finance* 37, 9 (2013), 3486–3498.
- [30] MENKVELD, A. J. Splitting orders in overlapping markets: A study of cross-listed stocks. *Journal of Financial Intermediation* 17, 2 (2008), 145–174.
- [31] MENKVELD, A. J. High frequency trading and the new market makers. *Journal of Financial Markets* 16, 4 (2013), 712–740.
- [32] MITCHELL, M., PULVINO, T., AND STAFFORD, E. Limited arbitrage in equity markets. *Journal of Finance* 57, 2 (2002), 551–584.
- [33] O’HARA, M., AND YE, M. Is market fragmentation harming market quality? *Journal of Financial Economics* 100, 3 (2011), 459–474.
- [34] PARLOUR, C., AND SEPPI, D. J. *Limit Order Markets: A Survey*. North-Holland, 2008, ch. 2, pp. 61–93.
- [35] PARLOUR, C. A., AND SEPPI, D. J. Liquidity-based competition for order flow. *Review of Financial Studies* 16, 2 (2003), 301–343.
- [36] PASTOR, L., AND STAMBAUGH, R. F. Liquidity risk and expected stock returns. *Journal of Political Economy* 111 (2003), 642–685.
- [37] SCHULTZ, P., AND SHIVE, S. Mispricing of dual-class shares: Profit opportunities, arbitrage, and trading. *Journal of Financial Economics* 98, 3 (2010), 524–549.
- [38] VALENZUELA, M., ZER, I., FRYZLEWICZ, P., AND RHEINLÄNDER, T. Relative liquidity and future volatility. *Journal of Financial Markets* 24 (2015), 25–48.
- [39] VAN KERVEL, V. Competition for order flow with fast and slow traders. *Review of Financial Studies* (2015), hhv023.

# Capítulo 7

## Anexos

### 1. Anexo A

Stock	Average volume	Average value	Average bid orders	Average ask orders
AGNa.CHI	2436213.443	15456578.91	66032845.99	64939922.2
AHa.CHI	943291.4196	12614885.54	22496881.23	22614629.49
AFp.CHI	1123544.314	9711005.457	20134460.35	20117924.82
AKZAa.CHI	277872.3059	15137175.53	4992170.592	4943410.957
MTa.CHI	2624983.153	28698701.06	96780310.87	96807624.91
ASMLa.CHI	444242.3843	31008121.59	7891178.451	7629215.537
CORAa.CHI	72524.8902	2611841.717	1189087.361	1322645.341
FURa.CHI	256779.3765	6467353.269	3394196.29	3383322.298
UNAA.CHI	1481866.835	45052739.05	74179151.79	74330921.34
HEIAa.CHI	309260.5059	16663997.12	5181300.192	5146146.898
WKLla.CHI	276242.7686	5889615.158	5551243.886	5908834.247
INGAa.CHI	4547204.337	47806014.63	143435381.6	143047854.3
DSMa.CHI	251918.4157	12623593.25	5474276.2	5475716.941
KPNa.CHI	5824405.557	14937619.02	126230590.3	124808874.6
PHIAa.CHI	1242083.506	29664179.64	36597715.39	37358480.1
PNLa.CHI	987530.698	3464291.524	15034531.53	15216465.83
RANDa.CHI	229694.2078	9191396.859	4104756.714	4186795.875
RENa.CHI	608672.3137	10166136.18	22435558.69	22635095.93
RDSAa.CHI	1576105.737	43773002.49	88774145.02	86215823.54
SBMOa.CHI	522186.902	5922506.067	9243122.345	9212967.471
TNTEa.CHI	494859.5098	2847022.289	7281844.012	7464823.114
BOKAa.CHI	118659.9216	4846108.236	1935509.831	2131915.482
TOTAL	1211370.114	17025176.57	34925920.85	34768155.06

Tabla 7.1: Valores promedios diarios transaccionales de las acciones analizadas de Bats Chi X.

La tabla muestra el volumen y valor promedio diario tranzado, y el número de órdenes de compra y venta promedio diario, para todas las acciones utilizadas tranzadas en Bats Chi X.

## 2. Anexo B

Stock	Average volume	Average value	Average bid orders	Average ask orders
AEGN.AS	7213307.431	45896285.98	348875287.6	334279109.9
AHLN.AS	3356057.765	44955853.92	210161590.7	204881074
AIRF.PA	3744479.835	32212486.49	80848997.04	81247728.71
AKZO.AS	673140.4275	36888135.38	34720433.85	34067156.21
ISPA.AS	9643441.122	105423925.5	623302740.9	603600968.3
ASML.AS	1436314.133	99841883.29	60233154.02	59185490.35
COR.AS	360701.8902	13141702.93	14472786.69	14338236.71
FUGRc.AS	1155652.337	27520244.99	25194775.95	25085628.2
UNc.AS	4697838.49	143184810.1	322060804.3	312283219.6
HEIN.AS	865210.8588	46145951.96	36904945.23	35625906.7
WLSNc.AS	780965.149	16616774.73	44524672.35	43018717.63
ING.AS	18562574.8	195327691.5	1239986199	1159478913
DSMN.AS	801894.8863	40155962.16	32141009.98	31591542.96
KPN.AS	15883840.13	40674821.26	672727815.2	662697099.1
PHG.AS	3576472.427	85735751.97	211683024.8	210920170
PTNL.AS	3916969.925	13802863.84	115433710.5	116641442.7
RAND.AS	669260.5647	26695632.66	26063005.85	24583269.75
ELSN.AS	2008890.749	33533987.09	154161584.1	149237118.3
RDSa.AS	6047970.518	168740007.9	398195730.4	381252719.8
SBMO.AS	1977741.294	22688559.61	65844392.09	62084895.07
TNTE.AS	1765280.271	10503537.31	79361041.24	74114071.6
BOSN.AS	354480.6078	14350167.54	14187664.99	14150700.3
TOTAL	4067840.255	57456229	218685698.5	210652962.7

Tabla 7.2: Valores promedios diarios transaccionales de las acciones analizadas de Euronext. La tabla muestra el volumen y valor promedio diario tranzado, y el número de órdenes de compra y venta promedio diario, para todas las acciones utilizadas tranzadas en Euronext.

### 3. Anexo C

VARIABLES	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Depth beyond average	-5.962*** (0.556)						
Depth best average		-25.10*** (2.678)					
Depth bid average			-44.31*** (5.274)				
Depth ask average				-44.31*** (5.461)			
Depth best crossed					-12.83*** (1.374)		
Depth bid crossed						-22.90*** (2.727)	
Depth ask crossed							-22.32*** (2.770)
Constant	0.190 (0.613)	0.579 (0.603)	0.436 (0.594)	0.519 (0.612)	0.384 (0.597)	0.302 (0.591)	0.309 (0.604)
Observations	4,994	4,994	4,994	4,994	4,994	4,994	4,994
Stock FE	YES	YES	YES	YES	YES	YES	YES
Days FE	YES	YES	YES	YES	YES	YES	YES
Pseudo R-squared	0.176	0.165	0.162	0.162	0.166	0.164	0.163
Marginal effects							
Depth beyond average	-0.501***						
Depth best average		-2.177***					
Depth bid average			-3.869***				
Depth ask average				-3.858***			
Depth best crossed					-1.110***		
Depth bid crossed						-1.987***	
Depth ask crossed							-1.945***

Tabla 7.3: Depth como único regresor, logit diario.

La tabla presenta los resultados obtenidos de estimar el modelo logit diario descrito en la metodología econométrica considerando una única covariable. En este caso la covariable utilizada es el depth en sus distintas versiones. (Errores estándares robustos en paréntesis. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ )

4. Anexo D

VARIABLES	Model 1	Model 2	Model 3	Model 4	Model 5
Realized spread (30s)	-2,173*				
	(1,296)				
Adverse selection (30s)		12,965***			
		(3,149)			
Effective spread (30s)			-987.1		
			(1,010)		
Bid ask spread average				-43.70	
				(53.71)	
Bid ask spread standardized					-926.0
					(977.3)
Constant	0.515	-0.626	0.107	0.192	0.697
	(0.764)	(0.615)	(0.719)	(0.838)	(1.200)
Observations	4,994	4,994	4,994	4,994	4,994
Stock FE	YES	YES	YES	YES	YES
Days FE	YES	YES	YES	YES	YES
Pseudo R-squared	0.140	0.143	0.139	0.142	0.141
Marginal Effects					
Realized spread (30s)	-203.7*				
Adverse selection (30s)		1,208***			
Effective spread (30s)			-92.72		
Bid ask spread average				-4.072	
Bid ask spread standardized					-86.63

Tabla 7.4: Spread como único regresor, logit diario.

La tabla presenta los resultados obtenidos de estimar el modelo logit diario descrito en la metodología econométrica considerando una única covariable. En este caso la covariable utilizada es el spread en sus distintas versiones. (Errores estándares robustos en paréntesis. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ )

5. Anexo E

VARIABLES	Model 1	Model 2	Model 3	Model 4
Relative liquidity average	10.46*** (1.285)			
Relative liquidity ask side average		8.273*** (1.033)		
Relative liquidity bid side average			10.76*** (1.234)	
Slope				309.3*** (95.68)
Constant	-5.372*** (0.902)	-4.300*** (0.798)	-5.485*** (0.888)	-0.295 (0.594)
Observations	4,994	4,994	4,994	4,994
Stock FE	YES	YES	YES	YES
Days FE	YES	YES	YES	YES
Pseudo R-squared	0.165	0.159	0.168	0.140
Marginal Effects				
Relative liquidity average	0.923***			
Relative liquidity ask side average		0.743***		
Relative liquidity bid side average			0.942***	
Slope				29.05***

Tabla 7.5: Distribución de liquidez como único regresor, logit diario.

La tabla presenta los resultados obtenidos de estimar el modelo logit diario descrito en la metodología econométrica considerando una única covariable. En este caso la covariable utilizada es relativa a la distribución de liquidez en el LOB. (Errores estándares robustos en paréntesis. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ )

## 6. Anexo F

VARIABLES	Model 1	Model 2	Model 3
Midquote volatility	3.701*** (0.598)		
Return volatility		10,274*** (1,333)	
Log return volatility			10,258*** (1,328)
Constant	-0.447 (0.582)	-1.353** (0.599)	-1.352** (0.599)
Observations	4,994	4,994	4,994
Stock FE	YES	YES	YES
Days FE	YES	YES	YES
Pseudo R-squared	0.150	0.157	0.157
Marginal Effects			
Midquote volatility	0.344***		
Return volatility		942.5***	
Log return volatility			940.7***

Tabla 7.6: Volatilidad como único regresor, logit diario.

La tabla presenta los resultados obtenidos de estimar el modelo logit diario descrito en la metodología econométrica considerando una única covariable. En este caso la covariable utilizada es la volatilidad en sus distintas versiones. (Errores estándares robustos en paréntesis. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ )

7. Anexo G

VARIABLES	Model 1	Model 2	Model 3	Model 4
Log volumen	1.341*** (0.106)			
Log valor		1.380*** (0.113)		
Órdenes de venta			-4.91e-10 (7.98e-10)	
Órdenes de compra				-7.29e-10 (8.70e-10)
Constant	-19.23*** (1.591)	-22.87*** (1.921)	-0.269 (0.595)	-0.269 (0.594)
Observations	4,994	4,994	4,994	4,994
Stock FE	YES	YES	YES	YES
Days FE	YES	YES	YES	YES
Pseudo R-squared	0.178	0.174	0.139	0.139
Marginal effects				
Log volumen	0.115***			
Log valor		0.120***		
Órdenes de venta			0	
Órdenes de compra				-6.84e-11

Tabla 7.7: Actividad transaccional como único regresor, logit diario.

La tabla presenta los resultados obtenidos de estimar el modelo logit diario descrito en la metodología econométrica considerando una única covariable. En este caso la covariable utilizada es alguna medida asociada a la actividad transaccional. (Errores estándares robustos en paréntesis.

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ )



8. Anexo H

VARIABLES	Model 1	Model 2	Model 3
Order imbalance	-0.000543*** (7.43e-05)		
Order imbalance crossed		-0.000312*** (4.03e-05)	
Diff. midquote			4.21e-05** (1.67e-05)
Constant	0.658 (0.613)	0.588 (0.605)	-0.281 (0.592)
Observations	4,994	4,994	4,994
Stock FE	YES	YES	YES
Days FE	YES	YES	YES
Pseudo R-squared	0.156	0.159	0.143
Marginal effects			
Order imbalance	-4.75e-05***		
Order imbalance crossed		-2.71e-05***	
Diff. midquote			3.99e-06**

Tabla 7.8: Otras medidas como único regresor, logit diario.

La tabla presenta los resultados obtenidos de estimar el modelo logit diario descrito en la metodología econométrica considerando una única covariable. En este caso la covariable utilizada corresponde a alguna variable comprendida en la dimensión otros. (Errores estándares robustos en paréntesis. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ )

## 9. Anexo I

VARIABLES	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Order imbalance crossed		8.70e-05 (0.000106)	4.60e-05 (0.000100)				
Return volatility	8,544*** (2,010)	3,503** (1,363)		3,999*** (1,299)			3,911*** (1,297)
Log volumen	0.913*** (0.148)	1.080*** (0.138)	0.863** (0.384)	1.105*** (0.143)	1.027*** (0.137)	1.064*** (0.136)	1.069*** (0.134)
Depth best crossed		-13.36*** (4.318)					
Slope		1,712 (2,055)	4,977*** (1,677)	5,271* (3,060)	5,314*** (1,910)	5,795*** (1,853)	
Relative liquidity average	6.318*** (1.104)	6.836*** (1.093)	3.482*** (1.148)				7.103*** (1.094)
Diff. midquote		1.13e-05 (1.17e-05)					
Realized spread (30s)	-3,582*** (1,054)						
Adverse selection (30s)	-6,183 (3,800)						
Order imbalance	-8.02e-05 (9.14e-05)						
Depth beyond average	-3.417*** (0.842)			-3.245*** (1.042)	-3.162*** (0.537)	-3.429*** (0.546)	
Depth bid average							-22.52*** (8.343)
Depth ask average							-18.72** (8.189)
Midquote volatility			1.805*** (0.687)				
Log valor			0.396 (0.420)				
Depth bid crossed			-10.78** (4.433)				
Depth ask crossed			-7.393 (6.380)				
Dist bid			-0.0157*** (0.00277)		-0.0134*** (0.00271)	-0.0152*** (0.00272)	
Depth best average				-3.062 (5.105)			
Dist ask				-0.0163*** (0.00255)	-0.000812 (0.000533)	-0.000730 (0.000449)	
Log return volatility					4,509*** (1,286)	4,214*** (1,255)	
Relative liquidity bid side average					4.137*** (1.159)		
Relative liquidity ask side average						1.899* (1.065)	
Constant	-15.26*** (2.129)	-18.85*** (1.939)	-19.79*** (2.514)	-15.37*** (2.055)	-16.53*** (1.985)	-15.84*** (1.963)	-18.53*** (1.888)
Observations	4,994	4,994	4,994	4,994	4,994	4,994	4,994
Stock FE	YES	YES	YES	YES	YES	YES	YES
Days FE	YES	YES	YES	YES	YES	YES	YES
Pseudo R-squared	0.220	0.216	0.224	0.223	0.227	0.225	0.215

Tabla 7.9: Logit diario, múltiples controles.

La tabla presenta los resultados obtenidos de estimar el modelo logit diario descrito en la metodología econométrica considerando multiples controles.

## 10. Anexo J

VARIABLES	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Depth beyond average	-5.347*** (0.549)						
Depth best average		-22.23*** (2.622)					
Depth ask average			-40.78*** (5.841)				
Depth bid average				-35.97*** (5.216)			
Depth best crossed					-11.48*** (1.340)		
Depth bid crossed						-18.56*** (2.653)	
Depth ask crossed							-21.07*** (2.928)
Constant	-1.651 (1.074)	-1.327 (1.082)	-1.347 (1.083)	-1.515 (1.082)	-1.478 (1.085)	-1.624 (1.081)	-1.487 (1.087)
Observations	8,688	8,688	8,688	8,688	8,688	8,688	8,688
Stock FE	YES	YES	YES	YES	YES	YES	YES
Intervals FE	YES	YES	YES	YES	YES	YES	YES
Pseudo R-squared	0.132	0.124	0.122	0.119	0.125	0.120	0.124
Marginal effects							
Depth beyond average	-0.216***						
Depth best average		-0.927***					
Depth ask average			-1.699***				
Depth bid average				-1.527***			
Depth best crossed					-0.475***		
Depth bid crossed						-0.783***	
Depth ask crossed							-0.869***

Tabla 7.10: Depth como único regresor, logit predictivo.

La tabla presenta los resultados obtenidos de estimar el modelo logit predictivo con intervalos de dos horas descrito en la metodología econométrica considerando una única covariable. En este caso la covariable utilizada es el depth en sus distintas versiones. (Errores estándares robustos en paréntesis. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ )

## 11. Anexo K

VARIABLES	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Realized spread (30s)	-2,195** (1,016)					
Adverse selection (30s)		5,590*** (2,129)				
Effective spread (30s)			-1,408 (933.8)			
Bid ask spread average				-84.99*** (15.37)		
Bid ask spread standardized					-1,896*** (571.8)	
Bid ask spread crossed						-129.1*** (16.56)
Constant	-1.217 (1.138)	-2.256** (1.090)	-1.458 (1.142)	-0.956 (1.082)	0.100 (1.256)	-0.616 (1.074)
Observations	8,688	8,688	8,688	8,688	8,688	8,688
Stock FE	YES	YES	YES	YES	YES	YES
Intervals FE	YES	YES	YES	YES	YES	YES
Pseudo R-squared	0.104	0.103	0.103	0.113	0.109	0.121
Marginal effects						
Realized spread (30s)	-100.8**					
Adverse selection (30s)		257.3***				
Effective spread (30s)			-64.95			
Bid ask spread average				-3.788***		
Bid ask spread standardized					-85.46***	
Bid ask spread crossed						-5.668***

Tabla 7.11: Spread como único regresor, logit predictivo.

La tabla presenta los resultados obtenidos de estimar el modelo logit predictivo con intervalos de dos horas descrito en la metodología econométrica considerando una única covariable. En este caso la covariable utilizada es el spread en sus distintas versiones. (Errores estándares robustos en paréntesis. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ )

## 12. Anexo L

VARIABLES	Model 1	Model 2	Model 3	Model 4
Relative liquidity average	8.729*** (1.080)			
Relative liquidity ask side average		6.314*** (0.838)		
Relative liquidity bid size average			8.913*** (1.018)	
Slope				-765.0 (914.2)
Constant	-6.451*** (1.220)	-5.203*** (1.147)	-6.515*** (1.217)	-2.006* (1.072)
Observations	8,688	8,688	8,688	8,688
Stock FE	YES	YES	YES	YES
Intervals FE	YES	YES	YES	YES
Pseudo R-squared	0.122	0.116	0.125	0.102
Marginal effects				
Relative liquidity average	0.380***			
Relative liquidity ask side average		0.281***		
Relative liquidity bid size average			0.384***	
Slope				-35.40

Tabla 7.12: Distribución de liquidez como único regresor, logit predictivo.

La tabla presenta los resultados obtenidos de estimar el modelo logit predictivo con intervalos de dos horas descrito en la metodología econométrica considerando una única covariable. En este caso la covariable utilizada es relativa a la distribución de liquidez en el LOB. (Errores estándares robustos en paréntesis. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ )

13. Anexo M

VARIABLES	Model 1	Model 2	Model 3
Midquote volatility	11.63*** (2.580)		
Return volatility		7,563*** (1,382)	
Log return volatility			7,512*** (1,379)
Constant	-2.297** (1.093)	-2.831*** (1.084)	-2.825*** (1.084)
Observations	8,688	8,688	8,688
Stock FE	YES	YES	YES
Intervals FE	YES	YES	YES
Pseudo R-squared	0.107	0.110	0.110
Marginal effects			
Midquote volatility	0.533***		
Return volatility		343.1***	
Log return volatility			340.8***

Tabla 7.13: Volatilidad como único regresor, logit predictivo.

La tabla presenta los resultados obtenidos de estimar el modelo logit predictivo con intervalos de dos horas descrito en la metodología econométrica considerando una única covariable. En este caso la covariable utilizada es la volatilidad en sus distintas versiones. (Errores estándares robustos en paréntesis. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ )

## 14. Anexo N

VARIABLES	Model 1	Model 2	Model 3	Model 4
Log volumen	0.895*** (0.0855)			
Log valor		0.901*** (0.0910)		
Órdenes de venta			-7.13e-10 (2.95e-09)	
Órdenes de compra				-1.17e-09 (2.90e-09)
Constant	-13.69*** (1.546)	-15.81*** (1.762)	-2.034* (1.071)	-2.032* (1.071)
Observations	8,688	8,688	8,688	8,688
Stock FE	YES	YES	YES	YES
Intervals FE	YES	YES	YES	YES
Pseudo R-squared	0.127	0.124	0.102	0.102
Marginal effects				
Log volumen	0.0383***			
Log valor		0.0389***		
Órdenes de venta			0	
Órdenes de compra				-5.41e-11

Tabla 7.14: Actividad transaccional como único regresor, logit predictivo.

La tabla presenta los resultados obtenidos de estimar el modelo logit predictivo con intervalos de dos horas descrito en la metodología econométrica considerando una única covariable. En este caso la covariable utilizada es alguna medida asociada a la actividad transaccional. (Errores estándares robustos en paréntesis. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ )

15. Anexo Ñ

VARIABLES	Model 1	Model 2	Model 3	Model 4	Model 5
Order imbalance	-0.000419*** (8.18e-05)				
Order imbalance crossed		-0.000231*** (4.18e-05)			
Distbid			-0.0191*** (0.00279)		
Distask				-0.0184*** (0.00279)	
Diff. midquote					-7.31e-06 (7.08e-06)
Constant	-1.355 (1.088)	-1.448 (1.088)	-0.878 (1.099)	-0.861 (1.093)	-2.036* (1.072)
Observations	8,688	8,688	8,688	8,688	8,688
Stock FE	YES	YES	YES	YES	YES
Intervals FE	YES	YES	YES	YES	YES
Pseudo R-squared	0.113	0.115	0.126	0.124	0.102
Marginal effects					
Order imbalance	-1.80e-05***				
Order imbalance crossed		-9.88e-06***			
Distbid			-0.000831***		
Distask				-0.000779***	
Diff. midquote					-3.38e-07

Tabla 7.15: Otras medidas como único regresor, logit pedictivo.

La tabla presenta los resultados obtenidos de estimar el modelo logit predictivo con intervalos de dos horas descrito en la metodología econométrica considerando una única covariable. En este caso la covariable utilizada corresponde a alguna variable comprendida en la dimensión otros. (Errores estándares robustos en paréntesis. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ )



## 16. Anexo O

VARIABLES	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Realized spread (30s)	-1.821** (838.6)	-2.322** (995.9)					
Adverse selection (30s)	-1.076 (2.075)	-1.521 (2.164)					
Return volatility	4.071** (1.721)				2.892** (1.340)		3.028** (1.362)
Log volumen	0.669*** (0.103)	1.248*** (0.416)	0.570*** (0.108)	0.800*** (0.0968)	0.737*** (0.101)	0.768*** (0.101)	0.698*** (0.101)
Order imbalance	-1.13e-05 (7.73e-05)	-1.50e-05 (7.82e-05)					
Depth beyond average	-3.606*** (0.798)	-3.542*** (0.797)	-1.856** (0.941)				-3.136*** (0.551)
Relative liquidity average	4.391*** (0.966)	4.243*** (0.979)	3.087*** (1.100)	4.871*** (0.942)	3.361*** (1.072)	2.997*** (1.069)	
Slope	-65.18 (543.3)						
Log return volatility		4.024** (1.702)					
Log valor		-0.617 (0.429)					
Bid ask spread crossed			-59.41*** (18.83)				
Midquote volatility			11.41*** (3.118)	3.659 (2.573)		7.412*** (2.749)	
Order imbalance crossed			0.000106 (6.51e-05)	0.000192** (7.75e-05)	0.000171** (7.42e-05)	0.000144** (6.78e-05)	
Distbid			-0.00650 (0.00668)		-0.00821 (0.00675)	-0.00934 (0.00749)	-0.00466 (0.00589)
Distask			-0.000707 (0.00647)		-0.000921 (0.00649)	-0.00166 (0.00718)	-0.00513 (0.00601)
Depth best average			-14.41* (8.529)				
Effective spread (30s)				-820.7 (648.8)			
Depth best crossed				-16.28*** (3.918)			
Diff. midquote				-1.62e-05 (4.48e-05)			
Depth bid crossed					-12.77*** (3.995)		
Depth ask crossed					-15.18*** (4.813)		
Midquote average						0.00867 (0.0123)	
Depth bid average						-23.05*** (7.005)	
Depth ask average						-26.94*** (9.156)	
Relative liquidity bid size average							5.781*** (1.431)
Relative liquidity ask size average							-3.014** (1.418)
Constant	-12.41*** (1.740)	-10.23*** (2.375)	-9.894*** (1.903)	-14.35*** (1.693)	-12.89*** (1.776)	-12.74*** (1.815)	-12.03*** (1.772)
Observations	8,688	8,688	8,688	8,688	8,688	8,688	8,688
Stock FE	YES	YES	YES	YES	YES	YES	YES
Intervals FE	YES	YES	YES	YES	YES	YES	YES
Pseudo R-squared	0.158	0.158	0.166	0.159	0.163	0.163	0.164

Tabla 7.16: Logit predictivo, múltiples controles.

La tabla presenta los resultados obtenidos de estimar el modelo logit predictivo con intervalos de dos horas descrito en la metodología econométrica considerando multiples controles. (Errores estándares robustos en paréntesis.\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ )

## 17. Anexo P

VARIABLES	Model 1 (gap=0)	Model 2 (gap=0.005)	Model 3 (gap=0.01)
Return volatility	9,560*** (2,435)	6,767*** (2,176)	5,268** (2,612)
Log volumen	1.139*** (0.137)	1.037*** (0.154)	1.590*** (0.246)
Depth beyond crossed	-2.380*** (0.397)	-1.166** (0.548)	0.0118 (1.245)
Depth best crossed	1.554 (1.954)	3.499 (2.890)	-5.424 (5.103)
Relative liquidity crossed	5.442*** (0.879)	-2.026** (0.989)	-1.615 (2.152)
Slope crossed	6,683*** (2,192)	6,522*** (2,248)	7,408*** (2,607)
Effective spread (30s)	-3,723*** (1,011)	-2,049** (1,026)	-1,481 (1,312)
Amihud's illiquidity measure	-3.923e+07 (3.313e+07)	4.853e+06 (5.251e+07)	4.491e+07 (6.810e+07)
Constant	-16.61*** (1.973)	-15.32*** (2.227)	-26.55*** (3.653)
Observations	5,500	4,972	3,549
Stock FE	YES	YES	YES
Days FE	YES	YES	YES
Pseudo R-squared	0.348	0.397	0.403

Tabla 7.17: Logit diario, análisis de sensibilidad.

La tabla presenta los resultados obtenidos de estimar el modelo logit diario con multiples controles y considerando las variaciones introducidas en el capítulo de análisis de sensibilidad. (Errores estándares robustos en paréntesis.\*\*\* p<0.01, \*\* p<0.05, \* p<0.1)

## 18. Anexo Q

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intervalo	2 horas	2 horas	2 horas	1 hora	1 hora	1 hora
Gap	0	0.005	0.01	0	0.005	0.01
VARIABLES						
Return volatility	1,165 (939.0)	5,183*** (1,439)	11,252*** (3,459)	1,627** (635.0)	1,879** (866.9)	3,732** (1,559)
Log volumen	0.744*** (0.0662)	0.628*** (0.0917)	0.697*** (0.189)	0.734*** (0.0481)	0.773*** (0.0677)	0.812*** (0.132)
Depth beyond crossed	-1.497*** (0.242)	-0.548 (0.536)	-1.046 (1.610)	-1.395*** (0.204)	-0.549 (0.446)	-0.987 (1.469)
Depth best crossed	-0.380 (1.164)	0.00975 (2.272)	-1.815 (8.402)	-0.621 (0.928)	0.193 (1.946)	-2.174 (6.685)
Relative liquidity crossed	4.170*** (0.585)	-0.485 (0.863)	-1.249 (2.205)	4.171*** (0.478)	-0.327 (0.737)	-1.171 (1.740)
Slope crossed	2,234** (983.9)	2,173* (1,240)	6,466*** (2,375)	1,608** (672.3)	1,231 (817.9)	2,563 (1,576)
Amihud's illiquidity measure	7.345e+06 (5.823e+06)	6.084e+06 (1.175e+07)	-4.781e+07* (2.676e+07)	-5.193e+06 (3.785e+06)	-7.453e+06 (6.981e+06)	-5.872e+06 (1.251e+07)
Constant	-12.79*** (1.065)	-12.22*** (1.590)	-29.81*** (2.740)	-13.12*** (0.933)	-13.43*** (1.322)	-15.14*** (2.052)
Observations	13,572	9,679	4,097	25,869	16,858	6,246
Stock FE	YES	YES	YES	YES	YES	YES
Intervals FE	YES	YES	YES	YES	YES	YES
Pseudo R-squared	0.195	0.228	0.299	0.166	0.202	0.286

Tabla 7.18: Logit predictivo, análisis de sensibilidad.

La tabla presenta los resultados obtenidos de estimar los modelos logit predictivos con múltiples controles y considerando las variaciones introducidas en el capítulo de análisis de sensibilidad.

(Errores estándares robustos en paréntesis.\*\*\* p<0.01, \*\* p<0.05, \* p<0.1)