Una nueva metodología para identificar los determinantes de la liquidez accionaria relevante para distintos perfiles de la inversión

María del Pilar Monteagudo

Segundo Premio / Categoría Jóvenes Profesionales

12° Premio de Investigación Económica

"Dr. Raúl Prebisch" 2020

BANCO CENTRAL
DE LA REPÚBLICA ARGENTINA

UNA NUEVA METODOLOGÍA PARA IDENTIFICAR LOS DETERMINANTES DE LA LIQUIDEZ ACCIONARIA RELEVANTES PARA DISTINTOS PERFILES DE LA INVERSIÓN

RESUMEN

A partir de un conjunto variado de indicadores de liquidez desarrollados por la literatura de la microestructura de mercado y un amplio set de variables abordadas por los estudios de *asset pricing*, la presente investigación presenta una nueva metodología que combina técnicas econométricas avanzadas y de *distribution based clustering* conjuntamente con algoritmos de *machine learning* a los efectos de encontrar el conjunto de determinantes que mejor permita explicar la liquidez para distintos grupos de acciones que caracterizan diferentes perfiles de inversión en el mercado local, ello sobre la base de construir índices compuestos por las medidas de liquidez que mejor predicen los retornos al interior de cada clúster. A partir de los resultados obtenidos, este estudio tiene como fin último brindarle al mercado local herramientas que le permitan una gestión más eficiente de la liquidez direccionada por grupo de acciones que comparten entre sí patrones intradiarios de liquidez.

Palabras clave: liquidez, microestructura de mercado, *asset pricing*, acciones, *distribution based clustering, machine learning*, econometría

Códigos JEL: G1, C5, C6

1.-INTRODUCCIÓN

Dentro del amplio campo de la investigación financiera, los procesos de intercambio y formación de precios en los mercados han acaparado la atención de diversos autores, de cuyos estudios surgidos principalmente a partir de las últimas dos décadas del siglo pasado fue brotando el sólido cuerpo teórico que dio vida a la *microestructura de mercado* como subdisciplina de las finanzas, estrechamente ligada y complementaria a las tradicionales teorías de *asset pricing*. Abocada a dimensionar el impacto que tanto los mecanismos como los costos de transacción ejercen sobre el comportamiento de los precios de diferentes activos (Liu, 2010) -en otros términos, la magnitud del desvío de los precios de sus valores de equilibrio (Biais, Glosten & Spatt, 2004)-, y de los actores que participan de los mercados, la microestructura de mercado ha logrado identificar y dimensionar la influencia de un conjunto de imperfecciones de mercado.

De los diversos aspectos que analiza este cuerpo teórico, la magnitud en la cual una menor liquidez –como uno de los principales costos de transacción-, impacta en el retorno de las distintas clases de activos, constituye uno de los aspectos de mayor interés dentro de la literatura especializada (Madhavan, 2000). Dentro de este marco, una multiplicidad de estudios ha debatido respecto a las medidas y modelos cuantitativos más apropiados a fin de mensurar este vínculo, no solo para el caso de los países desarrollados (véase e.g. Amihud y Mendelson, 1986; Chordia, Subrahmanyam y Anushman, 2001; Fama y French, 1993; entre otros), sino también para los mercados de países emergentes (Kearney, 2012; Rouwenhorst, 1999; entre otros).

No obstante, comprender exhaustivamente el fenómeno de la liquidez no solamente requiere analizar su impacto en los retornos, sino también identificar el conjunto de los factores que mejor expliquen su comportamiento al interior de cada mercado. Al respecto, un conjunto de estudios han trabajo sobre esto último empleado diferentes enfoques (Chordia, Sarkar & Subrahmanyam, 2005; Stoll, 2000; entre otros). En lo particular, un análisis exhaustivo se vuelve todavía más relevante en mercados poco profundos, donde la falta de liquidez y con ello el exceso de volatilidad constituyen los dos riesgos más críticos y disuasivos para los inversores. Desde un punto de vista del alcance práctico, la identificación de los principales focos causales de la liquidez constituye un insumo sumamente relevante a la hora de otorgar efectividad al diseño de políticas que permitan ir superando paulatinamente las fricciones al interior de los mercados y de las firmas, y

generen de este modo señales más eficientes. No obstante, en mercados concentrados y volátiles, con pocas firmas cotizando y con patrones de operación intra-diaria muy disímiles (i.e., poca profundidad, amplitud y flexibilidad, siguiendo la definición de liquidez de Hasbrouck, 2007), evaluar la liquidez desde un punto de vista general (y no segmentado) puede ocasionar sesgos en la interpretabilidad del ajuste de los modelos y la significancia de las variables. En otras palabras, las diferentes medidas de liquidez abordadas en la literatura de la microestructura de mercado permiten captar las diversas dimensiones del fenómeno, pero que no siempre pueden resultar simultáneamente significativas para grupos de acciones que exhiben perfiles de riesgo de la inversión marcadamente diferentes.

En atención a lo expuesto, y siendo este estudio el primero para Argentina en abordar en profundidad el fenómeno de la liquidez accionaria, se propone una nueva metodología que combina el uso de técnicas de econometría avanzadas, *distribution based clustering* y algoritmos de *machine learning* a los efectos de encontrar el conjunto de determinantes que permita explicar la liquidez para distintos grupos de acciones, ello partiendo de las medidas de liquidez que mejor permitan predecir los retornos esperados al interior de cada clúster. El acceso a datos intradiarios de la negociación de cada acción en el mercado permitirá la construcción de clústeres por aproximación de distribuciones de liquidez, empleando para ello una base de más de setecientos cincuenta mil registros (750.000). El análisis será efectuado para todo el conjunto de acciones que operaron en el mercado argentino durante el período 2009q4-2019q2. Una vez obtenidos los resultados a partir de llevar a cabo un análisis pormenorizado con técnicas cuantitativas avanzadas, el fin último de esta investigación consiste en proporcionar al mercado local herramientas útiles para la gestión de la liquidez accionaria partiendo de un enfoque de clústeres que definen diferentes perfiles de la inversión de acuerdo con patrones intra diarios de las operaciones.

Dando cuenta de lo expuesto, el resto del trabajo se estructura como sigue. En la Sección 2 se desarrolla una revisión de la literatura sobre la relación entre liquidez y retornos tanto para países desarrollados y emergentes, así como también las diferentes medidas de liquidez (y las dimensiones que capturan) y sus posibles determinantes. A continuación, en la Sección 3, se abordan los aspectos metodológicos adoptados en el presente estudio y los modelos a contrastar empíricamente. Por su parte, en la Sección 4 se relevan los

resultados empíricos obtenidos, mientras que en la Sección 5 se abordan las conclusiones principales de esta investigación.

2.- MARCO TEÓRICO

2.1- Microestructura de mercado y asset pricing: el rol de la liquidez

La complejidad y el alcance de los factores que afectan a los retornos esperados de las diferentes clases de activos han ocupado la atención tanto de académicos como prácticos. En este marco, aquellos relacionados con los mecanismos y costos de transacción, como la liquidez, han sido abordados principalmente por la microestructura de los mercados (Easley & O´Hara ,2003, Madhavan, 2000; Biais, Glosten & Spatt, 2005). De acuerdo con Hasbrouck (2007), la liquidez ha sido definida también como sinónimo de "profundidad, amplitud y resilencia (flexibilidad)"- basándose en Kyle (1985)-. En efecto, en un mercado con profundidad existe oferta y demanda suficiente para concretar una operación de compra o venta por encima y por debajo del precio de mercado, contando con un número suficiente de órdenes de compra y venta. En mercados flexibles, los efectos sobre el precio vinculados al proceso de negociación son pequeños y se diluyen en el cortísimo plazo. En un mercado con amplitud, se cuenta con una cantidad suficiente de volumen negociado, lo que permite que se satisfagan las preferencias de los demandantes y oferentes de forma constante. De este modo, cuanto más líquido el mercado menor la distancia entre los precios y sus valores fundamentales.

De la categorización brindada por Amihud, Mendelson & Pedersen (2005), pueden distinguirse diversas fuentes de (i) liquidez. En referencia, una primera fuente de iliquidez puede encontrarse en los costos de transacción exógenos como e.g., gastos de corretaje, costos de procesamiento de pedidos o impuestos sobre transacciones. Igualmente, se pueden identificar también las presiones de demanda y el riesgo de inventario. En otras palabras, cuando los agentes mantienen una presencia que no es permanente en el mercado (o bien no hay el suficiente volumen), los oferentes deben vender la porción deseada de su posición a un hacedor de mercado o *market maker* (cuya función es precisamente proveer inmediatez), que probablemente compre anticipadamente, para luego deshacer su posición (reventa). Esta maniobra, no obstante, llevará a este último a exponerse al riesgo de un cambio en el precio todo tiempo que mantenga el activo en su poder, hasta que logre venderlo, situación que se traducirá en un costo para el vendedor.

Asimismo, es probable que se generen transacciones entre partes con desigual información, no solo en términos del valor fundamental del activo sino también de la dinámica de los futuros flujos de órdenes, evento que se traducirá en mayores costos.

2.2.- Liquidez y retornos esperados en mercados accionarios desarrollados y emergentes

La controversia respecto al grado en que la liquidez constituye un factor de injerencia en el precio de los activos financieros ha sido plasmada en un vasto número de investigaciones especializadas. En este sentido, estudios teóricos como el de Merton (1987) afirman que la liquidez debe ser considerada como un factor idiosincrático, a diferencia del tradicional CAPM. Ahondando en la diversidad de este conjunto de estudios empíricos, Amihud y Mendelson (1986) y Eleswarapu (1997) han encontrado el efecto positivo del diferencial de precio comprador-vendedor (*bid-ask spread*) sobre los retornos de los activos ajustados por riesgos, evidenciando que el retorno de los activos es una función creciente y cóncava del spread. Por su parte, Easley, Hvidkjaer y O'Hara (2002), evidenciaron que la liquidez tiene un efecto significativo y positivo sobre los retornos. Dentro de este mismo orden de ideas, Brennan, Chordia y Subrahmanyam (1998) evaluaron la significatividad, para el período de 1966-1995, de la liquidez medida como el volumen negociado en un modelo de asset pricing multifactorial, controlando por tamaño, ratio valor libro/valor de mercado de cada compañía, rentabilidad por dividendos y retornos pasados. El estudio realizado sobre las acciones del NYSE de Brennan et al. (2012) constituye una contribución en la comprensión del impacto sobre los retornos de la liquidez asociada a las órdenes de compra y de venta.

Otra de las contribuciones seminales se apoya en el trabajo de Datar, Naik y Radcliffe (1998), quienes propusieron a la rotación como nueva medida de liquidez para evaluar el impacto sobre los retornos del NYSE en el período 1962-1991, controlando por el valor contable/valor de mercado, tamaño y beta de la firma (siguiendo a Fama y French, 1993) y el Efecto-Enero (Eleswarapu y Reinganum, 1993), asociando de este modo la liquidez a la frecuencia de las operaciones. Este resultado fue corroborado, e.g., por Chordia, Subrahmanyam y Anshuman (2001). Otras investigaciones en la materia incluyen los trabajos de Amihud, Mendelsonn y Lauterbach (1997) y Amihud (2002), quienes hallaron a la liquidez como predictor de los retornos a partir de replicar la medida de impacto precio de Kyle y capturando también los costos de transacción mediante los días con

retornos cero - controlando por tamaño, volatilidad, rentabilidad por dividendos y retornos pasados para un conjunto de acciones del NYSE-.

Luego del impacto que estas investigaciones tuvieron en el mundo desarrollado, el estudio de la relación entre retornos y liquidez en mercados accionarios también se convirtió en una inquietud de los países emergentes, en especial porque abundan las fricciones en esta clase de mercados financieros (Batten & VinhVo, 2014). Una de las investigaciones preliminares en la materia, la de Rouwenhorts (1999), testeó para 20 país emergentes los determinantes de los retornos, encontrando resultados similares a los hallados en países desarrollados respecto a los factores de French y Fama, pero para la rotación como medida de liquidez. No obstante, las conclusiones derivadas de los estudios de Claessens, Dasgupta y Glen (1998) y Jun, Marathe y Shawky (2003) encontraron una relación sorprendentemente positiva entre los retornos esperados y la rotación, contradiciendo los hallazgos típicos. De un modo no esperado, Lischewski y Voronkova (2012) ratificaron la ausencia de relación significativamente estadística entre retornos ajustados por riesgo y liquidez. De acuerdo con Bakaert, Harvey y Lundblad (2007), para un conjunto de 19 países emergentes de Latinoamérica y el Sudeste Asiático, la proporción de retornos diarios cero resulta la medida apropiada para predecir los retornos.

2.3.- Alternativas para medir la liquidez

Como puede desprenderse de lo tratado en la subsección previa, a lo largo de los distintos estudios han sido abordadas diversas clases de medidas de liquidez, siendo este aspecto por demás crítico a la hora de mensurar su impacto sobre el exceso de retornos. Siguiendo la clasificación brindada por Lesmond (2005), una clase de estimadores de la liquidez buscan captar el costo directo en que se puede incurrir cuando se lleva a cabo una determinada transacción. Dentro de este grupo, el indicador más representativo y básico es el *spread bid-ask* cotizado y sus derivados. Asimismo, también se pueden encontrar medidas como el coeficiente λ de impacto precio desarrollado por Kyle (1985) y adoptada luego por numerosos estudios. Sin embargo, no siempre se encuentra disponible la información para construir dichas medidas que requieren de una frecuencia intra-diaria.

En este marco, Goyenko et al. (2009) ha puesto de manifiesto que un buen número de medidas de liquidez construidas a partir de datos de baja frecuencia también resultan útiles para reflejar los costos de transacción que captan de forma directa estimadores

como los ya vistos. En este marco, otra clase de indicadores denominados reflexivos, asocian la liquidez con características de la operatoria, como e.g., el volumen, ya sean las clásicas de rotación, medida de Amihud (Amihud, 2002), etc. Por último, existe otra clase de medidas de liquidez que, como señala el autor, reflejan indirectamente los costos de transacción a partir del comportamiento del precio. Dentro de este conjunto se encuentran el estimador de Roll (1984)- véase Butler, Grullon & Weston (2005)-, la medida de Lesmond et al. (1999), entre otras.

Para la primera categoría de medidas definidas por Lesmond (2005), i.e, las que captan de forma directa el costo de transacción se pueden definir:

(1) Spread Bid-Ask Cotizado:

$$QS = A_{i,t}^{l} - B_{i,t}^{h}$$

, es la medida más simple y se computa como la diferencia entre el precio ask más bajo ($A_{i,t}^{l}$) y el precio bid más alto $B_{i,t}^{h}$ (Butler, Grullon & Weston, 2005). Si, a su vez, se afecta esta última medida dividiéndola por el punto medio entre ambos precios (promedio) o bien por el precio de la operación, se obtiene el *spread bid-ask* relativo. No obstante, este tipo de medidas no contemplan el precio al que fue ejecutada la operación.

(2) Spread Bid-ask efectivo:

Esta medida refiere al costo de ejecución inmediato, y se computa como la diferencia entre el precio de transacción y el punto medio cotizado:

$$ES = 2\ abs\left(P_{i,t} - M_{i,t}\right) = \ 2\ IN_{i,t}(P_{i,t} - M_{i,t})$$

, donde $P_{i,t}$ es el precio de transacción efectivo de la acción i para la operación t,$M_{i,t}$ se calcula como $M_{i,t} = \frac{(B_{i,t} + A_{i,t})}{2}$, y $IN_{i,t}$ representa un indicador que vale 1 si la negociación se efectuó como compra o -1 si se inició como venta; en otros términos, es dos veces el costo efectivo de la operación (Lee, Lin & Tsao, 2006). Esta medida se apoya en el supuesto que las operaciones por encima del punto medio son compras y las que están por debajo del promedio constituyen ventas. Al igual que para el spread cotizado, el efectivo es conveniente verlo en términos relativos.

(3) Impacto precio de Kyle (1985):

El impacto precio estima el coeficiente λ resultante de la regresión univariada definida,

$$R_{i,t} = \theta + \lambda OIB_{i,t} + \varepsilon_{i,t}$$

, siendo $R_{i,t}$ el retorno en t de la acción i, y $OIB_{i,t}$ la diferencia entre el volumen de transacción de las operaciones iniciadas como compra y venta. Este coeficiente busca captar la sensibilidad del precio o flexibilidad del mercado al flujo de órdenes.

No obstante, como puede apreciarse, estas medidas de liquidez requieren de información intra diaria en mayor o menor nivel de detalle, datos que no son de acceso público. Es por ello que, a fin de poder hacer extensiva la investigación en este campo, diversos autores han reversionado estas medidas sin perder de vista las limitaciones propias de acceso en que incurren los investigadores. En este marco, medidas como el volumen negociado o la proporción de días con volumen positivo (Panigo, Grandes y Pasquini, 2006) capturan la liquidez a partir de la frecuencia de las operaciones. Asimismo, hay medidas que asocian la liquidez a la frecuencia de las operaciones como la rotación,

(4) Rotación:

$$R_{i,t} = \frac{Vol_{i,t}}{Accionesencirculación_{i,t-1}}$$

, siendo $Vol_{i,t}$ el volumen negociado y las acciones en circulación vigentes al inicio del período según datos del balance de cada una de las firmas (Datar, Naik & Radcli¤e, 1998). Sin embargo, por la forma en que está construida, esta medida puede no representar la caída en la liquidez del mercado, e.g., en momentos de crisis de liquidez donde hay una escasez cada vez mayor de papeles disponibles para la negociación (Froot et al., 2001). Tomando esto en consideración, Tang, K., & Wang, C. (2011) definen la rotación negociable. No obstante, y como indica Armitage et al. (2014), esta medida no contempla la variabilidad de costos de transacción entre acciones, como sí lo refleja la proporción de días con retorno cero definida por Lesmond et al. (1999). En otras palabras, un papel con altos costos de transacción tenderá a tener movimientos de precios con menor frecuencia y por lo tanto más días con retorno cero que una acción con costos más bajos.

Intentando replicar el efecto impacto precio de Kyle y la proporción de días con retornos cero, Amihud (2002) desarrolló una medida de liquidez:

(5) Iliquidez de Amihud:

$$Amihud_{i,t} = \frac{1}{d}\sum_{t=1}^{n}\frac{\left|R_{i,t,d}\right|}{DVol_{i,d,t}}$$

, siendo d el número de días de negociación en un período,$\left|R_{i,t}\right|$ el valor absoluto de los retornos diarios promedio del activo i y $DVol_{i,t}$ el volumen negociado del activo promedio en el mismo período *t.* En particular, Brennan et al. (2012) proponen una modificación a la medida de Amihud, ajustando el retorno absoluto a la rotación.

Por último, otra clase de estimadores menos utilizados emplean los precios como insumo principal, entre las que se distinguen la medida de liquidez de Roll (1984), que define implícitamente al spread efectivo a partir de las covarianzas de precios – aunque no siempre queda definida cuando la covarianza es negativa-, o el gamma impacto precio de Pastor y Stambaugh (2003), que busca replicar el impacto precio desarrollado por Kyle y extendido por Amihud, pero a través de un modelo multifactorial de baja frecuencia.

2.4.- Explorando los determinantes de la liquidez en los mercados accionarios: una revisión de las principales variables de interés

Sobre la base de las numerosas investigaciones aquí consideradas, los resultados revelan que, en líneas generales, los determinantes de la liquidez accionaria puede clasificarse en cuatro grupos: (i) aspectos idiosincráticos como los fundamentales y la gestión interna de las firmas; (ii) variables macroeconómicas, institucionales y legales dentro de los países, (iii) factores no diversificables y estacionales y (iv) aspectos de la negociación, i.e., mecanismos y variables representativas de la operación de cada firma en los mercados y otras características institucionales, regulatorias y estructurales. A continuación, se exhibe un cuadro comparativo para un conjunto amplio de estudios.

Cuadro 1.- Presencia de los distintos conjuntos de regresores exógenos de la liquidez accionaria en la literatura especializada

Estudio	Aspectos fundamentales de la firma	Variables Macroeconómicas e institucionales	Factores no diversificables y estacionales	Aspectos de la negociación de la firma
De Cesari, Espenlaub & Khurshed (2011)			x	x
Chordia, Sarkar & Subrahmanyam (2005)			x	x
Al Salamat (2016)	x	x		x
Jain (2003)	x			x
Stoll (2000)	x			x
Dicle, M. F. (2008).	x	x	x	x
Huang & Chang (2015)	x	x		
Chan, K., Hameed, A., & Kang, W. (2013).	x		x	
Tang, K., & Wang, C. (2011).	x			x
Jain, P., Jiang, C., & Mekhaimer, M. (2016)	x		x	x
Lee, J. H., Lin, S. Y., Lee, W. C., &Tsao, C. Y. (2006)			x	x
Lee, K. H., Sapriza, H., &Wu, Y. (2016).	x	x	x	x
Rehman, I. U., Mahdzan, N. S., & Zainudin, R. (2016).		x		
Chung, K. H., & Van Ness, R. A. (2001)				x
Gopalan, Kadan & Pevzner (2012)	x			x
Hillert et al, 2016	x			

Dentro del conjunto de determinantes asociados a aspectos fundamentales de la firma, De Cesari et al. (2011) pone en evidencia cómo la recompra o venta de acciones propias incrementa la liquidez, en especial de aquellos papeles más ilíquidos, permitiendo disminuir los costos de transacción por información asimétrica y el costo del patrimonio cuando un papel está sub o sobre valorado (análogamente Hillert, A., Maug, E., & Obernberger, S. (2016) encuentran que este efecto se acreciente aún más en períodos de crisis). Asimismo, Charoenwong et al. (2011) señala que aquellas firmas con proporción de activos líquidos en balance ganan en liquidez accionaria en tanto acarrean menores costos de información por "incertidumbre de valuación", incluso por encima del efecto "incertidumbre de inversión" planteado por Myers & Rajan (1998) y que asocia activos líquidos con financiamiento de proyectos ilíquidos, distribución de disponibilidades o riesgo a expropiación de fondos. Otros estudios como el de Jain et al. (2016) convalidan la hipótesis de valuación para la proporción de I+D, o la importancia de los activos líquidos en firmas pequeñas con menores oportunidades de crecimiento, restricciones financieras y mala gestión interna (Gopalan et al., 2012). Asimismo, otros ratios de

gestión como los pasivos de corto plazo sobre activos (Choi & Cook, 2006), volumen y crecimiento de activos, entre otros, contribuyen a determinar el nivel de liquidez accionaria (véase también Chordia et al., 2005). Medidas de rentabilidad interna como el Valor Libro a Valor Mercado – que refleja el ROE potencial-, el tamaño de la firma (como medida de riesgo), y el apalancamiento han sido evaluadas en conjunto, en tanto e.g., un valor libro más alto con un incremento de la deuda podría reflejar un crecimiento con mayor rentabilidad para el accionista (Lipson y Mortal, 2009). Por otra parte, el apalancamiento no siempre es percibido como un factor de riesgo, en tanto puede mejorar la liquidez al reducir el costo de agencia entre gerentes y accionistas por las mejores decisiones gerenciales en compañías con altos niveles de deuda (Goyenko et al., 2009). Más aún, factores como los ingresos por acción (Brennan & Subrahmanyam, 1996), el ratio precio-beneficio (Liem & Basana, 2012) y la política de dividendos (Banerjee et al., 2007) han demostrado tener una injerencia para explicar la variabilidad de la liquidez. De acuerdo con Griffin (2010), las acciones que pagan dividendos satisfacen mejor la necesidad de liquidez de los inversores, volviéndose más atractivas en especial si se negocian con poca frecuencia. Sin embargo, esto no pareciera verificarse para las firmas cuyas acciones son más líquidas, que al tener un menor costo de capital tienden a financiar un mayor número de proyectos con valor presente positivo, lo cual se ve desalentado por la distribución de dividendos.

No obstante, en la medida en que los factores idiosincráticos juegan un papel crucial en la determinación de la liquidez accionaria, los aspectos macroeconómicos también han sido estudiados desde una óptica similar. En este marco, Al Salamat (2016) identifica que la tasa de interés y el crecimiento del producto bruto interno tienen un efecto positivo sobre la liquidez, mientras que la inflación contribuye negativamente. Rehman et al. (2016), determinan que variables como la base monetaria, tasa de interés, el ciclo de negocios y el flujo de inversiones extranjeras tienen un impacto rezagado y significativo en el corto plazo, aunque en el largo plazo son la producción industrial y el flujo neto de divisas los que resultan significativos. Dentro de este conjunto, las tasas de corto plazo se vuelven un factor relevante asociado al costo de financiamiento de los inventarios de los *market makers* y los *traders* durante la jornada de negociación. Como señalan Dicle (2008), en períodos donde los portafolios están desbalanceados el riesgo de financiar las operaciones se vuelve más elevado en períodos de tasas crecientes. Cabe considerar, por otra parte, factores tales como el rating de deuda soberana, que afecta a la liquidez

negativamente a través del rebalanceo de portafolio, la selección adversa y la liquidación de posiciones por pérdidas colaterales (Lee et al., 2016).

Como se ha mencionado al inicio de la subsección, existen factores no diversificables o sistémicos que tienen impacto en la liquidez. En tal sentido, tanto la liquidez como la volatilidad del mercado (e.g. índice VIX, véase Huang & Chang, 2015) constituyen determinantes a nivel individual. En este marco, Coughenour y Saad (2004) demuestran cómo la liquidez individual se ve afectada por la variación de la liquidez del portafolio de mercado, y más precisamente Lee et al. (2006) identifican la presencia de covariación con la liquidez de la industria, en especial en las firmas más pequeñas. Por su parte, y como medida clásica de riesgo no diversificable, Chan et al. (2013) determina que a mayor beta de una acción- i.e., mayor volatilidad sistémica-más liquidez. En efecto, y a diferencia de lo expuesto por Dicle (2008), como la información de mercado es más fácilmente observable que la específica de la acción, podría argumentarse que cuando una acción tiene una gran correlación con el mercado, el movimiento esperado del papel es más claro y por lo tanto menos sensible a órdenes particulares, capturando así un mayor volumen. Por otra parte, es usual reconocer en estudios variables categóricas que contribuyen a captar el efecto "día de la semana", "mes del año", "vacaciones", momentos de crisis, eventos de anuncios macroeconómicos relevantes, entre otros (Chordia et al., 2005).

Finalmente, y en cuanto a los mecanismos y características de la negociación, regulación y estructura al interior de los mercados, diversos estudios asocian una mayor volatilidad de los retornos de la firma con un impacto negativo en liquidez por los mayores costos de inventario (Stoll, 2000). Asimismo, también ha sido utilizado el precio de cierre en tanto precios más bajos suelen ser más riesgosos (Jain, 2003). En lo que refiere particularmente a la regulación y estructura de negociación, uno de los pocos análisis integrales realizado por Jain (2003) pone de relieve cómo las características del diseño dentro de 51 mercado accionarios afectan su nivel de liquidez, dentro de las cuales se distinguen: (i) forma de negociación; (ii) presencia de *market makers*; (iii) incremento del tamaño del *tick* (véase e.g. Chakravarty, Harris, & Wood, 2001), entre otros. Otras investigaciones analizan movimientos en la liquidez por cambios normativos en el manejo de órdenes (Chung & Van Ness, 2001), o de las operaciones de venta corta (Ko & Lim, 2006).

3.- METODOLOGÍA

3.1.- Indicadores de liquidez relevantes según el perfil de la inversión: modelización por clúster

La imposibilidad de aproximar la liquidez a través de una única medida capaz de captar todas las dimensiones del fenómeno ya ha sido tratada en estudios precedentes. En este sentido, trabajos como los de Lam, Tam y Dong (2019) han propuesto la construcción de una única medida de liquidez a partir de técnicas como el Análisis de Componentes Principales (PCA) que capte las diferentes dimensiones, i.e., cantidad negociada, frecuencia de operación, el impacto precio y los costos de transacción. Sin embargo, estas dimensiones pueden resultar más o menos relevantes para diferentes perfiles de la inversión. Es posible argumentar que inversores con un capital pequeño busquen minimizar el costo de entrada y salida (spread bid-ask) a partir de una mayor profundidad del libro de órdenes e inversores más grandes tengan capacidad de relegar costos e invertir en acciones menos líquidas con un perfil más riesgoso en busca de oportunidades.

En este marco de hipótesis, es requerido estimar el modelo general de exceso de retornos (véase 3.2) para distintos grupos de acciones clasificadas en base a datos de liquidez con información intradiaria. Se procederá a agrupar acciones mediante el método de *distribution based clustering*, i.e., cada clúster incluirá las acciones que más se aproximen en la distribución intra diaria. Para ello se procederá en dos pasos: (1) *data dimensionality reduction* y (2) *clustering*. En primera instancia, se reducirán las observaciones por acción para el spread bid-ask relativo (al precio, para la comparabilidad) a los cuatro momentos de la distribución (i.e. media, desvío estándar, simetría y curtosis), luego estandarizadas. Una vez ello, se segmentará a partir de dichos parámetros tomando una medida de distancia euclediana cuadrática entre distribuciones. La muestra empleada será de 3 meses (máxima disponible) durante el período 2019m9 a 2019m12.

Al respecto del primer paso, la medida de distancia Wasserstein- Mallows (véase Ni, Bresson, Chan & Esedoglu, 2009) permite reducir y comparar distribuciones a partir de sus cuatro momentos. Tomando las funciones de distribución empíricas para la variable de interés es posible comparar las distribuciones acumuladas inversas $F^{-1}(t) = Y$ de dos histogramas cualesquiera a partir de esta métrica, que resulta de la familia de las medidas

de distancia euclidianas aplicada a las distribuciones de probabilidad. En este sentido, se puede definir para $p = 1$ y $p = 2$ la distancia de Wasserstein-Mallows como:

$$dW^p(H_i, H_j) = \sqrt[p]{\int_0^1 \left|F_i^{-1}(t) - F_j^{-1}(t)\right|^p dt}$$

Las distribuciones son comparadas a partir de los diferentes momentos de la distribución, i.e., ubicación, escala, y forma (Rueshendorff, 2001). La descomposición de la dW^p con $p = 2$ para dos distribuciones resulta en:

$$dW^2(Y_i, Y_j) = (\mu_i - \mu_j)^2 + (\sigma_i - \sigma_j)^2 + 2\,\sigma_i\sigma_j(1 - \rho(Y_i, Y_j))$$

Debido a la gran cantidad de datos, es necesario comparar la distancia entre dos distribuciones reduciendo los parámetros de la forma al coeficiente de asimetría y curtosis.

A continuación, se utilizará el algoritmo K-means (K-medias), que permite el agrupamiento de un conjunto de observaciones en el que cada una de ellas pertenece al clúster cuyo valor medio de cada parámetro (conjuntamente) más se aproxime. Siguiendo a Makles (2012), se computarán diez soluciones - para un número máximo establecido de *k* (k=1 ... 10) grupos por cada métrica empleada y luego se compararán. Debido a que el procedimiento es sensible a los centroides iniciales, se replicarán las soluciones para 50 posiciones iniciales aleatorias, promediando luego los valores obtenidos. El número óptimo de clústeres se evaluará a partir de computar, graficar y observar los quiebres pronunciados del WSS (suma de las distancias al cuadrado de cada observación al centroide del clúster), el coeficiente Eta Cuadrado η^2, que cuantifica la reducción proporcional del WSS por cada k comparada con la suma total de los cuadrados (TSS), el coeficiente de Reducción Proporcional del Error (PRE) que muestra la reducción proporcional de WSS para la solución del clúster k en comparación con k-1 y el índice de Calinski-Harabasz, que evalúa la validez de un k a partir del promedio de la suma de cuadrados inter e intra-clúster. La interpretabilidad y consistencia combinada de todas las medidas permitirá encontrar el número de k óptimos que minimicen las distancias (eucledianas) intra-clúster y las maximice inter-clúster.

3.2.- *Dimensionality reduction technique* y formas funcionales flexibles para la selección del mejor modelo por clúster

En primer lugar, sólo mediante el uso de datos de panel es posible observar a las mismas unidades o grupos de unidades en diferentes momentos del tiempo, requerido para nuestro modelo de retornos (exceso). Como es bien sabido, no es posible suponer que las observaciones dentro de un panel sean necesariamente independientes, i.e., la presencia de factores omitidos que afectan a una variable (pero no a otra) en un momento del tiempo también podrán hacerlo al siguiente. No obstante, como refiere Baltagi (1995), es posible mejorar la consistencia de las estimaciones suponiendo que esta heterogeneidad individual no observable deriva fundamentalmente de efectos individuales específicos por unidad de corte transversal, aunque constantes en el tiempo (véase también Elosegui et al., 2006).

Asimismo, la inclusión conjunta de todas las variables de liquidez disponibles además de las variables de control potencialmente relevantes no necesariamente conlleva a la mejor estimación del verdadero proceso generador de los datos. Es bien sabido que existe al respecto un *trade-off* entre sesgo y varianza, más comúnmente asociado a la problemática de *overfitting* o sobreajuste. Por otra parte, no resulta teóricamente evidente que la relación entre las variables examinadas sea estrictamente lineal ni contemporánea. Para este primer problema, el método de abordaje más sencillo consiste en utilizar el *principal component analysis* (PCA por sus siglas, ver Lam & Chan, 2019). Sin embargo, este método radica en la escasa capacidad de interpretación económica de los componentes obtenidos, que combinan distintas proporciones de varianza de las variables subyacentes. Para el segundo problema, por su parte, la aproximación más difundida recae en la utilización de redes neuronales artificiales con distintos *layers* (capas) complementarios que incluyan funciones de activación no lineales diversas. Aunque la utilización de este tipo de algoritmos suele ser particularmente eficiente para obtener predicciones robustas, resulta mucho más deficiente para su utilización en análisis de inferencia causal, ello debido a la compleja interpretación de las relaciones subyacentes.

Por los motivos previamente mencionados, en la presente investigación utilizaremos un enfoque alternativo, que combina la potencialidad de los algoritmos de *machine learning* con la interpretabilidad de econometría, ello a los efectos de encontrar una forma funcional para la ecuación de exceso de retornos por clúster, que combine buen poder

predictivo y niveles de interpretabilidad robustos y asequibles para los resultados de inferencia causal. Esta metodología se basa en cuatro etapas complementarias:

1. *Feature extraction*: Generación de nuevas variables explicativas a partir de las variables originales, construyendo para cada una de ellas los primeros rezagos. Si bien la generación de otras transformaciones puede aumentar el poder predictivo, su incorporación es deficiente en términos de la interpretabilidad financiera.
2. *Regularization* (preselección): Reducción preliminar del espacio de variables relevantes a través del algoritmo LASSO (*least absolute shrinkage & selection operator*) - (Tibshirani; 1996)-. Este algoritmo, si bien no garantiza la identificación del verdadero proceso generador de los datos, es una herramienta flexible y poderosa para eliminar de manera apropiada las variables menos relevantes para la ecuación examinada.
3. *All-subset-regression*: Tomando como insumo los resultados de la preselección vía LASSO (15 variables), el algoritmo econométrico GSREG (véase Gluzmann y Panigo, 2013) permite obtener el mejor modelo explicativo final (para la muestra completa, dejando afuera cierta cantidad de observaciones para la confirmación *out-of-sample*) sobre la base de diversos criterios informativos (Akaike, Bayesiano, Mallow, etc.) y test de residuos (para poder garantizar inferencia estadística).
4. *Optimal Model Validation:* los métodos de *resampling* como el *cross-validation* se encuentran ampliamente difundidos en el campo de la estimación de modelos y selección de variables, ya que permiten iterar los modelos múltiples veces excluyendo cada vez un set diferente de observaciones de entrenamiento (o de validación) calculando una tasa de error para asegurar la mejor predicción de nuevas observaciones donde no se conoce la variable predicha. Siguiendo a Jung & Hu (2015) se utilizará una versión que permite promediar entre los mejores modelos obtenidos en (3) basado en la técnica de *K-fold validation* y asegurando la adyacencia temporal entre ambos sets (no se emplean muestras aleatorias). Tomando como base la selección mediante GSREG (3), se determinan los K modelos óptimos que minimizan el error cuadrático medio de la predicción (RMSE, por sus siglas en inglés) *out-of-sample.* Los coeficientes resultan los valores promedio ponderados (por el criterio) entre los K modelos óptimos. A fin de disminuir el trade-off entre sesgo y varianza se emplearán K=10 grupos, dejando un set de validación de 8 trimestres. Este procedimiento asegura estimadores más estables y eficientes.

Siguiendo esta metodología en cuatro etapas, y para cada clúster, se estimará un modelo de exceso de retornos dinámico multifactorial[2], como se detalla a continuación. Tanto la variable dependiente (retornos) como los regresores serán analizados sobre una base trimestral -recopilados de la plataforma financiera Thomson Reuters Eikon-, en pesos argentinos y construidos para todas las acciones que hayan cotizado de forma continua durante el período 2009q4-2019q2. El modelo se estructura como sigue:

$$(1) \quad ER^{e}{}_{i,t} = \alpha + \sum_{j=1}^{k} \beta_j liq *_{i,t} + \sum_{j=1}^{k} \gamma_j X_{j,i,t} + \vartheta ER_{m_{i,t}} + \omega ADR_{i,t} + \sum_{j=\underline{j}}^{j} \theta_j Year_{j,t} + \sum_{j=\underline{j}}^{j} \delta_j Sector_{j,t} + \varphi \sigma Fx_{j,t} + \tau Ln * + \psi ER_{i,t-1} + \varepsilon_{i,t}$$

$i = 1, \dots, N$(Unidades de corte transversal); $t = 1, \dots, T$ (Observaciones de serie de tiempo)

, donde $ER^{e}{}_{i,t}$ es la variable dependiente que representa los retornos (exceso) para la firma *i* en el período *t,* $liq *_{i,t}$ es un set de *=*1,...,n* medidas de liquidez para la firma *i* en *t* desarrolladas en el marco teórico, $X_{j,i,t}$ desde $X_{1,i,t}$ ... $X_{k,i,t}$ corresponde al set de *k=n* variables de control específicas de la firma, $ER_{m_{i,t}}$ es el retorno del mercado (exceso) que busca captar la parte de la varianza de la variable dependiente explicada por el riesgo sistémico (no idiosincrático), $ADR_{i,t}$ una variable *dummy* que identifica la tenencia de ADR por parte de la firma *i* (las firmas que cotizan en mercados del exterior se suponen más diversificadas, integradas, y con más posibilidades de ser negociadas), $Sector_{j,t}$ una variable *dummy* que buscan captar el efecto diferencial por sector sobre el exceso de retornos no explicado por factores idiosincráticos (se utilizará la clasificación de Thompson Reuters), $\sigma Fx_{j,t}$ es el tipo de cambio ARS/USD y pretende corregir los efectos de la pérdida de valor de la moneda sobre el comportamiento de los retornos que no hayan sido capturados por los movimientos del índice del mercado ni las heterogeneidades estructurales (representa un efecto sustitución de portafolio), $Ln *$ es cada uno de los regresores rezagados n=1 períodos, $ER_{i,t-1}$ la variable dependiente rezagada un período que busca captar el efecto *momentum o contrarian* y $\varepsilon_{i,t}$un componente aleatorio con media cero y varianza σ_ε. Por su parte, α denota el efecto no observado y fijo específicos de la firma (no incluido en el *panel regression*), β_j son los $\beta_1 \dots \beta_n$los coeficientes de regresión que determinan el impacto de cada medida de liquidez sobre los retornos (exceso) esperados, ϑ el factor sistémico de los retornos, γ_j de $\gamma_1 \dots \gamma_n$ los estimadores de

2 véase Panigo, Grandes y Pasquini (2006).

los regresores independientes de control, φ el coeficiente de regresión del tipo de cambio, ω δ_j estimadores para las variables independientes *dummy*, τ representa los estimadores de las variables independientes rezagadas y ψ el coeficiente de regresión del primer rezago de los retornos. Dentro del conjunto de variables de control, se emplearán el tamaño de la firma (compañías más grandes son menos riesgosas y exigen menos retornos diferenciales), el ratio valor mercado a libro (un bajo ratio es un activo subvaluado que para recomponer su valor requerirá un exceso de retornos), el apalancamiento, que implica un riesgo de desvalorización del patrimonio, el ROA y su volatilidad como indicadores financieros del crecimiento genuino, y el circulante como aproximación *floating* disponible - y refleja el riesgo asociado a la operación en los mercados (siguiendo lo propuesto por Wong et al, 2006). A continuación, se presenta el set de variables de liquidez a emplear en este modelo:

Identificador de variable endógena y liquidez	Descripción	Identificador de variables de control	Descripción
retornos	Retornos de cada acción	apalancamiento	Deuda total/ activos corrientes
spread_cotizado_medio	Spread cotizado relativo (punto medio)	tamaño_negocio	Logaritmo natural de los Activos totales
spread_efectivo_medio	Spread Efectivo relativo (punto medio)	circulante	Acciones en circulación
ln_volumen	Volumen operado (logaritmo natural)	ROA	Beneficio neto sobre activos
rotacion	Nominales operados por cada 1000 títulos en circulación	mv_bv	Precio de Mercado / Valor Libro
amihud_rotacion	Medida de Amihud ajustada por rotación	ROA_desvio	Desvío Estándar 3q ROA
amihud	Medida de Amihud	FX	Tipo de Cambio Nominal (en pesos)
d_volpositivo	Proporción de días con volumen positivo	indice_global	Índice de Mercado

Nota: los spreads se calculan a partir de los precios más altos y bajos del día.

3.3.- Determinantes de la liquidez: un modelo por clúster multifactorial

Las medidas de liquidez que resulten predictores óptimos para el exceso de retornos serán empleadas para construir un índice de liquidez compuesto representativo por clúster[3]. A partir de los cuatro grupos de determinantes de la liquidez revisados en la sección 2.3, se abordará un modelo dinámico de múltiples dimensiones. En este marco, el procedimiento de selección de variables óptimas empleando *machine learning* debe complementarse con

[3] El índice se construye como una suma producto de los coeficientes por las variables.

el poder explicativo de los estimadores. En efecto, el principal efecto que genera estimaciones sesgadas radica en la autocorrelación de los errores dentro de los paneles, potenciado fundamentalmente por la evidencia de endogeneidad entre liquidez y retornos (véase Arellano, 1990). Para ello, es necesaria una estimación simultánea de los mejores modelos de retornos y liquidez por clúster, y la técnica más apropiada para ello es el método de 3SLS (*Three-Stage Least Squares*). Como señalan Crowe et al. (2010), este método también es robusto a la autocorrelación de los residuos, generando así estimadores insesgados. No obstante, esta consistencia será contrastada a partir de la prueba general de autocorrelación para datos de panel de Cumby-Huizinga (Cumby y Huizinga, 1992). A partir de la metodología expuesta, el modelo de determinantes de la liquidez general a estimar para cada clúster se estructura como sigue:

$$(2) \qquad IL_{i,t} = \alpha + \sum_{j=1}^{k} \gamma_j X_{j,i,t} + \tau Ln * + \psi IL_{i,t-n} + \varepsilon_{i,t}$$

$i = 1, \dots, N$(Unidades de corte transversal); $t = 1, \dots, T$ (Observaciones de serie de tiempo)

, donde $IL_{i,t}$ es la variable dependiente que representa el índice de liquidez para la firma *i* en el período *t,* $X_{j,i,t}$ desde $X_{1,i,t}$ ... $X_{k,i,t}$ corresponde al set de *k=n* variables de interés analizadas por la literatura especializada, $Ln*$ es cada uno de los regresores rezagados n=1...2 períodos y $IL_{i,t-1}$ la variable dependiente rezagada n=1...2 períodos y $\varepsilon_{i,t}$un componente aleatorio con media cero y varianza σ_ε. Por su parte, α denota el efecto no observado y fijo específicos de la firma (no incluido en el *panel regression*), $\gamma_j$de $\gamma_1 \dots \gamma_n$los coeficientes de regresión que determinan el impacto de los regresores independientes sobre la liquidez, τ representa los estimadores de las variables independientes rezagadas y ψ el coeficiente de regresión del primer y segundo rezado de la liquidez. En la tabla seguidamente se detalla el conjunto de regresores utilizados en el modelo:

Identificador	Descripción	Identificador	Descripción
liquidez	Índice de (I) Liquidez	deuda_corriente	Pasivos de corto plazo/Activos totales
tamaño_negocio	Logaritmo natural de los Activos totales	prop_activosliq	Activos Líquidos/Activos totales
ROA	Beneficio neto sobre activos	EMBI_pp	Rating de deuda soberana EMBI Argentina (pp.)
ROA_desvio3q	Desvío Estándar 3q ROA	tasa_interes_pp	Tasa de interés bono soberano 2Y USA (pp.)
MV_BV	Precio de Mercado / valor libro	PBI_pp	Tasa de crecimiento Producto Bruto Interno (%YOY, pp.)

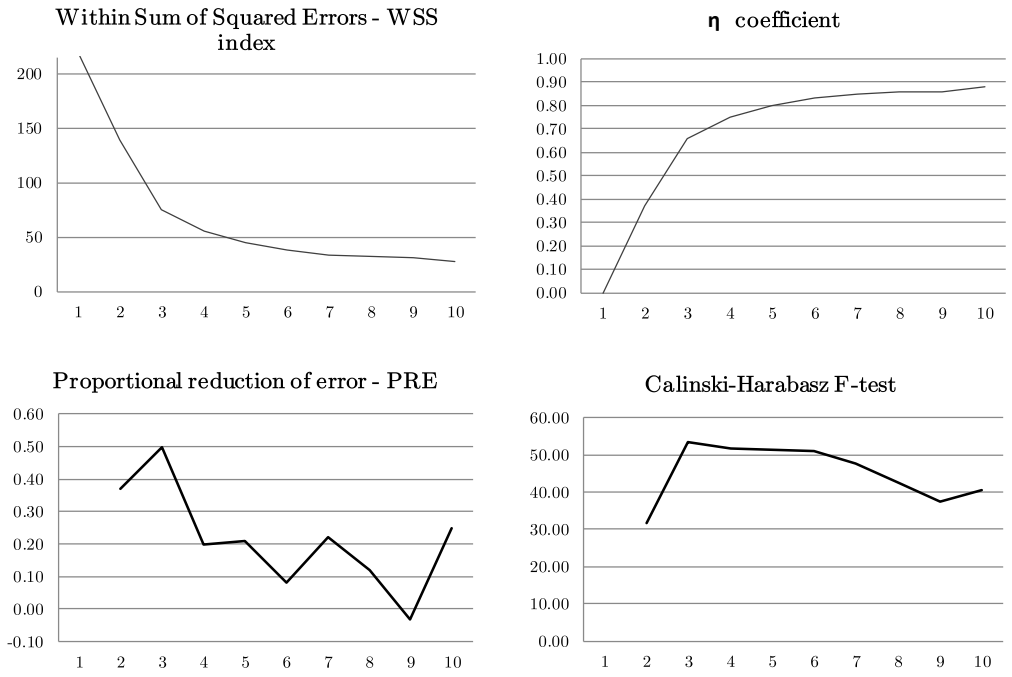
EPS	(Utilidad neta-dividendos en acciones preferenciales) / Prom. Acciones en circul.	FX_desviod	Desvío estándar diario del tipo de cambio nominal ARS/USD
crecimiento_activos	Tasa de crecimiento 1 período de los Activos del Balance	FX	Tipo de cambio nominal ARS/USD
apalancamiento	Deuda total/ activos corrientes	indice_global_desviod	Volatilidad (diaria) del índice de mercado
PE_ttm	Precio corriente de la acción/EPS últimos 12 meses	Beta	Coeficiente de Riesgo Sistémico Beta
PE_norm	PE con ganancias normalizadas	mercado_liquidez	Índice de Liquidez para el mercado
tasa_dividendos	Dividendo por acción/precio	industria_liquidez	Índice de Liquidez por Industria
pagodiv_dum	Dummy=1 si hay pago de dividendos	retornos_desviod	Desvío estándar diario de los retornos
recompra_dum	Dummy=1 si hay aumento de acciones en cartera propia	precio_cierre	Precio de cierre por acción
deuda_acida	Pasivos de corto plazo/Activos líquidos	retornos	Retornos por acción

4.- RESULTADOS EMPÍRICOS

4.1.- Clusterización óptima para la muestra de acciones del panel general

El método *distribution based clustering* fue aplicado a los parámetros de la distribución del spread *bid-ask* relativo para cada firma. A continuación, se exhiben los resultados obtenidos para la muestra de datos intradiarios.

Gráfico 1.- Soluciones de K-Means para los *k* grupos (*k = 1,...,10*)



Como se desprende para el conjunto de los cuatro gráficos, los resultados indican que la segmentación en K=3 clústeres es la solución óptima, ello a partir de obtener el mismo número de clústeres para cada una de las medidas abordadas. En términos más específicos, para el WSS se aprecia un quiebre marcado para K=3. Dicho de otro modo, añadir un clúster más no minimiza aún más la varianza intra-clúster (ni la maximiza inter-clúster). De la misma forma, el coeficiente Eta, exhibe la máxima reducción del WSS (aproximadamente 70%) para la misma solución. Considerando la reducción proporcional del error (PRE), se aprecia que el pico de la función se encuentra en K=3, reflejando una disminución del alrededor del 40% de la varianza al interior de cada clúster (y maximización inter-clúster) en relación con segmentar la muestra en 2. Por último, el quiebre del índice Calinski- Harabasz revela que tres clústeres es la segmentación óptima que permite minimizar las distancias intra-clúster y maximizarlas inter-clúster. Concomitantemente, fue verificado que las firmas para j=1,...,50 en cada métrica computada para la solución k=3 efectivamente se mantenían dentro del mismo clúster[4]. En la Tabla 1 a continuación se presenta una vista comparativa de los momentos de la distribución (estandarizados) de la liquidez intra-diaria por clúster.

Tabla 1.- Momentos de la distribución estandarizada (promedio) por clúster del spread bid-ask relativo

Clúster	N acciones	Media	Desvío Estándar	Simetría	Curtosis
1	8	1.96286	1.52673	-0.54733	-0.30782
2	43	-0.36488	0.30389	-0.12003	-0.19231
3	5	-0.00262	0.17069	1.90796	2.1464

Nota: Se consideraron sólo las acciones ordinarias que hayan cotizado de forma continua durante el período 2009q4-2019q2. De la muestra original de 68 papeles se consideraron 56. **Clúster 1:** DYCA, FIPL, ROSE, MORI, FERR, PATA, COME, TXAR; **Clúster 2:** BMA, GGAL, TECO2, PAMP, BRIO, CEPU, TGSU2, HARG, IRCP, CTIO, BPAT, IRSA, CRES, ESME, CAPX, MOLI, INDU, CECO2, EDN, MIRG, BHIP, BOLT, LEDE, CELU, OEST, GCLA, AUSO, CADO, DGCU2, CARC, GRIM, INTR, AGRO, LONG, SEMI, GARO, POLL, RIGO, GBAN, SAMI, METR, PATY, CGPA2; **Clúster 3:** ALUA, BBAR, TEGNO4, TRAN, YPFD.

Analizando los momentos de la distribución al interior de cada clúster y comparativamente, se aprecia que el clúster 3, compuesto por 5 acciones, exhibe el spread bid-ask más pequeño relativo al resto del mercado (-0.00262). Además, presenta una pronunciada simetría positiva (1.90), indicativo que la mayor frecuencia de transacciones ocurre a spreads más bajos que la media, así como también una distribución

[4]El paquete sadi (Stata) contiene la función ari que permite verificar la correlación entre la composición de los clústeres para cada iteración.

pronunciadamente leptocúrtica (curtosis 2.14) comparativamente, lo que revela que la mayor parte de las operaciones se concentran a un spread similar y pequeño. Puede argumentarse entonces que, la mayor certidumbre acerca del precio al que se operará la firma – reforzado por el menor desvío estándar comparativo- se traduce en bajos costos de entrada y salida, lo que caracteriza a inversiones con un perfil de riesgo principalmente averso, propio de inversores pequeños que operan con un reducido presupuesto y buscan acciones con baja volatilidad y el mínimo costo de entrada y salida.

De forma opuesta, el clúster 1 está compuesto por las 8 firmas que comparten el perfil de inversión más riesgoso de toda la muestra. Además de exhibir el spread relativo promedio más alto de los 3 conjuntos, el mayor desvío estándar (1.52) se da conjuntamente con una asimetría negativa (-0.54), lo que pone en evidencia que la mayor frecuencia de las operaciones intra-clúster se ejecutan a un valor más alto que el spread promedio, y simultáneamente con una distribución platocúrtica (curtosis -0.30) –marcada por el desvío-, que pone de manifiesto la menor certeza sobre el precio que se pueda conseguir en el mercado para dichas acciones. Este clúster, en efecto, aparece comparativamente con los mayores costos de entrada y salida, aceptable para un perfil de inversión más amante del riesgo. En este sentido, es posible que concentre en su mayor parte grandes inversores que buscan oportunidades de largo plazo (pudiendo soportar la mayor volatilidad de corto plazo) y presenten entonces menor necesidad de liquidez.

Por su parte, el clúster 2, que aglutina la mayor parte de las compañías, se ubica a la mitad de los extremos de la segmentación en cuanto al spread relativo y desvío estándar (0.30), aunque si en términos comparativos se habla, se encuentran valores notoriamente más cercanos en cuanto a curtosis al grupo del conjunto menos líquido (-0.19), aunque contrariamente con una ligera simetría más positiva si se la observa en contraposición al clúster 1 (-0.12). En efecto, este conjunto de acciones, si bien revela una cierta heterogeneidad en su interior, presenta en comparación al clúster 3 un perfil de la inversión más riesgoso y menos líquido.

4.2.- Mejor modelo de exceso de retornos por clúster y determinación del índice de liquidez

A continuación, se presentan los resultados obtenidos para el mejor modelo de exceso de retornos, junto con un análisis de descomposición de varianza para evaluar la contribución

de cada medida de liquidez a la variabilidad de los retornos, siguiendo la medida de Cavallo y Valenzuela (2007). En este sentido, la contribución de cada covariable se define como:

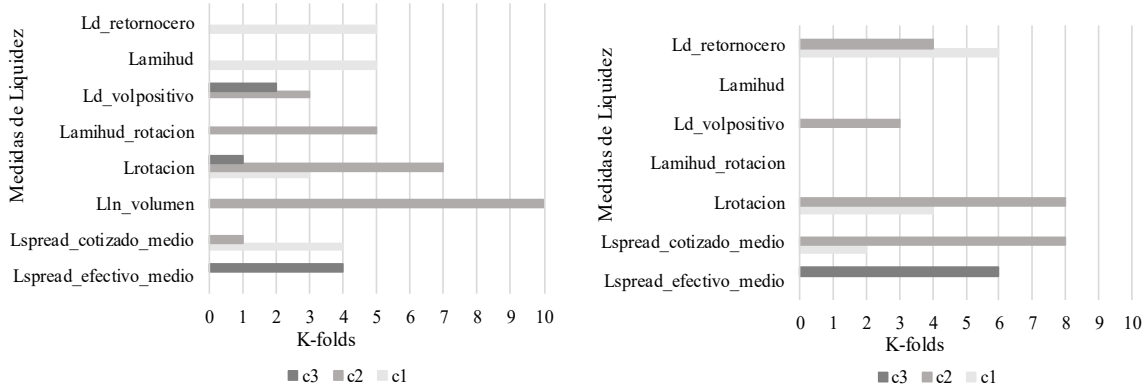
$$\beta_{z_k}\rho_{y,z_k} = \beta_{z_k} * \left(\frac{\sqrt{var_{z_k}}}{\sqrt{var_y}}\right) * \rho_{y,z_k}$$

, donde Z_k es cada variable independiente $Z = [x_1, \dots\dots, x_n]$, β_{z_k} representa a cada coeficiente de regresión múltiple estandarizado de Z_k y ρ_{y,z_k} el coeficiente de correlación de Pearson simple entre la variable dependiente y Z_k. Asimismo, se analiza la participación promedio de cada medida para los tres clústeres, tomando en consideración las k=1…10 veces que las covariables fueron predictores óptimos.

Tabla 2.- Mejores predictores del exceso de retornos para la ecuación general y su representatividad en las K=1...10 estimaciones, según clúster

	Clúster 1			Clúster 2			Clúster 3			
Variables	**Coef.**	**K-grup.**	**Contrib.**	**Coef.**	**K-grup.**	**Contrib.**	**Coef.**	**K-grup.**	**Contrib.**	**Part. prom.**
spread_efectivo_medio	5.5539	5	6.02%	4.1425	5	2.40%			0.00%	33.00%
ln_volumen	0.0005	5	0.08%	0.0021	10	0.17%	0.0008	5	0.06%	67.00%
rotacion	0.0001	9	8.26%	0	6	0.94%	0.0001	5	7.14%	67.00%
amihud_rotacion	0.0003	3	0.77%			0.00%	-0.0516	10	4.06%	65.00%
amihud	0.135	10	18.30%	0.0144	9	11.45%	3.5339	10	56.41%	97.00%
tamaño_negocio	-0.0018	2		0.0003	6					
d_retornocero			0.00%			0.00%	-0.4142	5	1.31%	17.00%
apalancamiento							0.0014	1		
indice_global	0.0048	1		0.0078	1		0.0235	2		
FX				-0.0004	1					
Lretornos	-0.093	5					-0.0406	3		
Lspread_cotizado_medio	-0.6713	4	0.38%	0.2273	1	-0.04%			0.00%	17.00%
Lspread_efectivo_medio			0.00%			0.00%	2.7519	4	0.84%	13.00%
Lrotacion	0	3	0.81%			0.16%	0	1	-0.08%	37.00%
Ld_volpositivo			0.00%	-0.0292	3	0.12%	-0.2669	2	-0.03%	17.00%
Lamihud	0.0176	5	-0.29%			0.00%			0.00%	17.00%
Ld_retornocero	0.0849	5	0.38%			0.00%			0.00%	17.00%
Lmv_bv	-0.002	5		-0.0006	4					
Ltamaño_negocio							-0.0537	7		
LROA							-0.0176	1		
Lindice_global				-0.0665	9		-0.0189	3		
const	0.0177	10		3.1512	9		0.0171	10		
Índice de (I) Liquidez	0.1501			0.0778			-0.1685			
Observaciones	170			857			110			
Rmse_out	0.085			0.123			0.105			

Gráfico 2.- Representatividad de las variables de liquidez en los *K-fold*



En términos generales, los estimadores contemporáneos más representativos resultaron ser la medida de Amihud, con casi 97% de peso para los tres clústeres en los diferentes sets de entrenamiento, seguido del volumen (logarítmico), la rotación y la medida de Amihud corregida – con más del 65% de participación en los K=1...30 – considerando los tres segmentos-, y en menor medida del spread efectivo medio y los días con retorno cero. La capacidad predictiva de los indicadores de liquidez con un rezago, medido en términos de K folds, se presenta mucho más homogéneamente distribuida entre los diferentes indicadores de liquidez, siendo los efectos del volumen y la rotación lo más relevantes. Más específicamente, de los resultados obtenidos se desprende que:

(i) el spread efectivo medio aproximado a partir de datos diarios no contribuye a explicar la variabilidad del exceso de retornos como sí lo hacen las medidas cuya frecuencia es de naturaleza diaria y no intra-diaria, tales como la medida de Amihud (2002), empleada en diversos estudios. No obstante, el spread efectivo medio va creciendo en importancia a medida que el clúster se vuelve más ilíquido, lo que es consistente con la clusterización delineada.

(ii) la medida de Amihud resulta no solamente el mejor predictor en términos de K-folds (en su forma contemporánea) sino también por la contribución a la variabilidad del exceso de retornos, al interior de cada clúster.

(iii) las medidas de liquidez son más relevantes en su forma contemporánea, tanto en términos de participación en los mejores modelos como de contribución. Esto se encuentra asociado a la frecuencia de las observaciones y la naturaleza de las variables de liquidez.

(iv) Para definir el sentido del índice construido, se regresaron los retornos contra el índice de liquidez por clúster. En efecto, si el incremento del índice de liquidez implica

mayores retornos se define como un indicador compuesto principalmente explicado por variables de liquidez en sentido negativo. Como resultado, el índice de liquidez presentó una co-variabilidad positiva con los retornos para los tres clústeres (el clúster 1 exhibió un coeficiente positivo del índice de +0.9365; el clúster 2 de +0.2741 y el clúster 3, +0.0067).

4.3.- Mejor modelo de determinantes de liquidez por clúster

Tomando como variables dependientes dentro del sistema de ecuaciones a los retornos y el índice de liquidez, a continuación, se exhiben los regresores que resultaron estadísticamente significativos para explicar la liquidez en el largo plazo (ello debido a la estabilidad paramétrica otorgada por la técnica *K-fold validation*), según cada clúster. Asimismo, se presentan los resultados de la prueba de Cumby-Huizinga.

Tabla 3.- Mejor modelo de determinantes de liquidez por clúster, efectos totales

	Clúster 1		**Clúster 2**		**Clúster 3[5]**	
Variables	**Coeficiente**	**pvalue**	**Coeficiente**	**pvalue**	**Coeficiente**	**pvalue**
mercado_liquidez	0.0004	0.588	0.0006	0.05**		
industria_liquidez			0.0035	0.009***		
retornos_desviod	1.4314	0.004***	1.4118	0***		
retornos	0.3554	0***	0.0976	0***	2.5051	0***
pagodiv_dum					0.0355	0.031**
PE_norm					0.0005	0***
ROA	0.0914	0.001***				
Beta			0.0020	0.027***	-0.0922	0.01***
tasa_interes_pp			0.0157	0***		
tasa_dividendos			-0.0420	0.009***		
MV_BV	0.0028	0.008***	-0.0004	0***		
apalancamiento	0.0237	0.001***	-0.0030	0.1690		
Lliquidez			0.0772	0.001***	-0.0109	0.012**
prop_activosliq			0.0294	0.06*		
const	-0.0007	0.858	-0.0008	0.5430	-0.0083	0.3880
Observaciones	191		1368		117	
P value <0.01 *** <0.05** <0.1 *						
R_sqr	0.578		0.294		0.565	
Rmse	0.055		0.043		0.104	

[5] Debido a que el clúster 3 es el que concentra gran parte de la liquidez tanto a nivel de mercado como industrial, al incluir dichas variables se observaba un sobre ajuste "artificial" del modelo por la presencia de una alta correlación, por lo que fuera removidas a fin de evitar que capturar efectos de otros potenciales predictores y distorsionar los resultados.

Tabla 4.- Prueba de Cumby-Huizinga para autocorrelación de residuos

	Clúster 1		Clúster 2		Clúster 3	
Rezagos	chi2	p-val	chi2	p-val	chi2	p-val
1-1	0.183	0.67	0.859	0.35	0.004	0.95
1-2	4.633	0.10	1.616	0.45	1.261	0.53
1-3	4.634	0.20	7.305	0.06	3.032	0.39
1-4	5.091	0.28	7.646	0.11	4.435	0.35

Test robusto a la heterocedasticidad. Ho: q=0 (errores serialmente incorrelacionados). Ha: c.s. presente en un rango específico

De acuerdo con los resultados expuestos, la estimación del modelo de determinantes de la liquidez resulta fundamentalmente robusta a la presencia de autocorrelación, brindando así estimadores insesgados. Más específicamente, la implementación de la prueba de Cumby-Huizinga revela que, para los tres clústeres, no hay evidencia de autocorrelación evaluando los errores del modelo para un amplio rango de rezagos, i.e., de 1-4, lo que convalida la insesgadez de los estimadores de 3SLS[6]. Analizando la relación entre la liquidez y los diferentes determinantes hallados como estadísticamente significativos, los efectos exhiben matices muy diversos e interesantes en relación con estudios previos de la literatura especializada. Asimismo, de la clasificación delineada en la sección 2.3, se demuestra que todas las clases de determinantes estuvieron presentes en alguno de los clústeres. En términos de bondad del ajuste, el modelo de determinantes de la liquidez permite explicar, en promedio, el 58% de la variabilidad del índice de liquidez en el clúster 1, el 30% en el clúster 2 y el 57% en el clúster 3.

Desde un foco transversal, se verifica la relación negativa entre los retornos y la liquidez al interior de cada segmento (i.e., un aumento de 100 puntos porcentuales de los retornos incrementa, en promedio, el índice de iliquidez en 0.3554 en el clúster 1, 0.0976 en el clúster 2 y 2.5051 en el clúster 3). Asimismo, y de acuerdo con Stoll (2000), la volatilidad diaria de los retornos es un regresor estadísticamente significativo y con un efecto negativo sobre la liquidez para los grupos más ilíquidos, en donde se espera un mayor desvío de los retornos diarios (cada 100 pp. que sube el desvío estándar el índice de iliquidez crece el índice, en promedio, en 1.4314 y 1.4118 unidades para el clúster 1 y 2,

[6] Únicamente para el clúster 2, del rango de 1-3 rezagos, se observa la presencia de autocorrelación significativa débilmente al 95%, lo cual no se considera sea evidencia robusta de insesgadez.

respectivamente). Por otra parte, la liquidez en el clúster con más acciones y heterogeneidad sectorial presenta una co-variabilidad significativa y positiva con el índice de (i) liquidez del mercado y la industria. Asimismo, el componente autorregresivo de la variable endógena se hizo presente en los clústeres 2 y 3.

En cuanto a las variables macroeconómicas, fundamentales y sistémicas, un conjunto variado de determinantes resultó estadísticamente significativos para explicar la variabilidad del índice de liquidez al interior de cada clúster. En referencia al clúster de acciones más ilíquidas, asociadas a perfiles de la inversión más amantes del riesgo, puede apreciarse que tanto el ROA, como el ratio valor mercado a valor libro y el apalancamiento han resultados estadísticamente significativos, exhibiendo una relación positiva con el índice de (i)liquidez. Más específicamente, un incremento de 100 puntos porcentuales en estos indicadores disminuye la liquidez, en promedio, 0.0914, 0.028 y 0.0237 unidades (o incrementa el índice), respectivamente. En primera instancia, se pone en evidencia que el apalancamiento, contrariamente a incrementar la liquidez por una menor asimetría de información al reducir el costo de agencia entre gerentes y accionistas como lo sostiene Goyenko et al. (2009) – en efecto, hay mejores decisiones corporativas y de gestión frente a un mayor riesgo de default-, es percibido como una señal de riesgo al aumentar las posibilidades de quiebra y por tanto de disminución del valor del patrimonio de la firma. Por su parte, la relación negativa entre el ROA y la liquidez es consistente con lo hallado para mercados emergentes por Al Salamat (2016). Analizando conjuntamente los efectos de estos determinantes, sumado a la evidencia de un impacto negativo del valor mercado a valor libro sobre la liquidez (como medida también de rentabilidad), sugiere que las firmas más pequeñas que se encuentran dentro de este clúster puedan estar adoptando una estrategia de financiar rentabilidad mediante deuda, lo que deriva en un mayor riesgo idiosincrático que se traduce en una menor liquidez de mercado.

Por su parte, para el clúster 2, un conjunto más amplio de variables resultó significativas para explicar la liquidez. De acuerdo con Dicle (2008), un incremento de 1 punto porcentual de la tasa de interés de referencia de corto plazo impacta empeorando la liquidez, en promedio, en 0.0157 unidades del índice, ya que eleva el costo de inventario (o el riesgo de financiamiento de las operaciones), así como una mayor co-variabilidad de las acciones con los movimientos del mercado reflejado por el coeficiente beta se

convierte en un factor de riesgo al no permitir la diversificación al interior de los portafolios, incrementando aún más el costo de inventario. En relación con la política de dividendos de las firmas, los resultados obtenidos revelan que las acciones que pagan una mayor tasa de dividendos satisfacen mejor la liquidez de los accionistas y por lo tanto capturan su atención. Este aspecto es coincidente con lo planteado por lo descripto por Griffin (2010), quien señala que esto ocurre en mayor medida en papeles con menor liquidez relativa. En efecto, el coeficiente de estimación indica que un aumento de la tasa de dividendos de 100 puntos porcentuales disminuye el índice de (i)liquidez, en promedio, 0.0420 unidades. Contrariamente a los evidenciado en el clúster 1, el ratio de valor mercado a valor libro permite mejorar la liquidez, algo que convalida la hipótesis que una mejora en este indicador de rentabilidad implica un mejor retorno potencial para el accionista. Cabe considerar, por último, que la proporción de activos líquidos en relación con los activos totales de la hoja de balance exhibe una relación negativa con la liquidez (una variación de 100 pp. de este porcentaje eleva, en promedio, 0.0294 unidades al índice). De este modo, la evidencia indica que el efecto "incertidumbre de inversión" desarrollado por Myers y Rajan (1998) prevalece por sobre el efecto "incertidumbre de valuación". Analizando este aspecto en el marco del mercado local, es esperable que en países con cambios frecuentes en el régimen macroeconómico adoptado y en el marco regulatorio, un mayor flujo de fondos líquido sea un indicativo o bien de una ralentización futura de la inversión o de una mayor exposición a la redistribución no adecuada u expropiación de dicho capital, y que sumado a la baja calidad de gestión y gobernanza corporativa de firmas especialmente más riesgosas, pequeñas y por tanto con menor liquidez relativa, por consiguiente la certidumbre del valor razonable de la firma se vea opacada por un conjunto mucho más amplio de información financiera.

Finalmente, y en atención al conjunto de regresores que resultaron estadísticamente significativos para explicar la variabilidad del índice de liquidez al interior del clúster 3, que captura un perfil de inversión marcadamente averso al riesgo, los hallazgos relevan coincidencias con algunas de las teorías propuestas por la literatura especializada. Asimismo, permiten abordar conclusiones sumamente interesantes desde la óptica del comportamiento de ciertas variables frente a la liquidez para distintos perfiles de la inversión. En primera instancia, se verifica que un cociente Precio /Beneficio normalizado (o PE) más elevado impacta negativamente en la liquidez al otorgar menos oportunidades de retornos futuros. Contrariamente a las formulaciones delineadas para el clúster 2, las

acciones que pagan dividendos exhiben, en promedio, una menor liquidez (o una diferencia positiva en el índice de 0.0355 unidades con respecto a aquellas que no realizan pago de dividendos), lo que convalida la hipótesis que el pago de dividendos no permite aprovechar al máximo el menor costo del *equity* que presentan las firmas más líquidas para invertir en proyectos que potencian el resultado futuro de la compañía. Análogamente, los resultados obtenidos arrojan una relación positiva entre el coeficiente beta y la liquidez (un incremento de 1 unidad en el coeficiente sistémico reduce, en promedio, el índice de liquidez compuesto en 0.0922). De este modo, a la evidencia aportada por Chan et al. (2013) puede añadirse que son los inversores con mayor aversión al riesgo aquellos que prefieren acciones con fuerte co-variabilidad con los movimientos del mercado, o, en otros términos, que les permitan predecir su comportamiento futuro dado el limitante al acceso a la información idiosincrática que poseen.

5.- CONCLUSIONES

A lo largo de años de investigaciones en el campo de la microestructura de los mercados, la evidencia presentada tanto para países desarrollados como para mercado emergentes ha dejado sin lugar a la duda que la liquidez constituye un factor central para comprender el proceso de formación de precios al interior de los mercados accionarios. En efecto, es bien sabido que los activos que operan con una menor profundidad, amplitud y flexibilidad requieren de retornos diferenciales para compensar el mayor riesgo asociado, principalmente, al costo de procesamiento, inventario y selección adversa producto de la información asimétrica entre los participantes. En especial, abordar el fenómeno de la liquidez en mercados concentrados como el argentino y con diversas y persistentes fricciones que se traducen en volatilidad y riesgo que desalientan la participación y su crecimiento, llama a prestar una especial atención. La presente investigación, basada en la implementación de una nueva metodología que combina la segmentación de las acciones según perfiles de riesgo de la inversión, el uso de las técnicas de *machine learning* más adecuadas para la selección óptima de variables y de econometría avanzada para la inferencia causal e interpretabilidad de resultados financieros, permite abordar conclusiones que resultan de interés para la identificación de aquellos factores que inciden en la liquidez, desde una visión de política diferenciada. A partir de esta nueva perspectiva y de forma concluyente, los resultados alcanzados al interior de cada clúster de acciones, que permiten evidenciar efectos diversos y consistentes con la literatura especializada

para los distintos perfiles de la inversión, prestan evidencia de la importancia de abordar nuevas metodologías que den lugar a hallar las verdaderas razones entre variables.

6.- REFERENCIAS BIBLIOGRÁFICAS

Al Salamat, W. A. (2016). Macroeconomics, Firm-Specific Factors and Stock Liquidity: An Empirical Evidence from Jordan. *International Journal of Financial Research*, *7*(5), 110-123.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets*, *5*(1), 31-56.

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of financial Economics*, *17*(2), 223-249.

Amihud, Y., Mendelson, H., & Lauterbach, B. (1997). Market microstructure and securities values: Evidence from the Tel Aviv Stock Exchange. *Journal of Financial Economics*, *45*(3), 365-390.

Amihud, Y., Mendelson, H., & Pedersen, L. H. (2005). Illiquidity and Stock Returns: Liquidity and asset pricing. *Foundation and Trends in Finance*, *1*, 269-364.

Arellano, M. (1990). Testing for autocorrelation in dynamic random effects models. *The review of economic studies*, *57*(1), 127-134.

Armitage, S., Brzeszczyński, J., & Serdyuk, A. (2014). Liquidity measures and cost of trading in an illiquid market. *Journal of Emerging Market Finance*, *13*(2), 155-196.

Baltagi, B. H. (1995). Testing for Fixed Effects in Logit and Probit Models Using an Artificial Regression. *Econometric Theory*, *11*(5), 1179-1179.

Banerjee, S., Gatchev, V. and Spindt, P. (2007) 'Stock market liquidity and firm dividend policy', Journal of Financial and Quantitative Analysis, Vol. 42, No. 2, pp.369–384.

Batten, J. A., & Vo, X. V. (2014). Liquidity and return relationships in an emerging market. *Emerging Markets Finance and Trade*, *50*(1), 5-21.

Bekaert, G., Harvey, C. R., & Lundblad, C. (2007). Liquidity and expected returns: Lessons from emerging markets. *The review of financial studies*, *20*(6), 1783-1831.

Biais, B., Glosten, L., & Spatt, C. (2005). Market microstructure: A survey of microfoundations, empirical results, and policy implications. *Journal of Financial Markets*, *8*(2), 217-264.

Brennan, M. J., & Subrahmanyam, A. (1996). Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of financial economics*, *41*(3), 441-464.

Brennan, M. J., Chordia, T., & Subrahmanyam, A. (1998). Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics*, *49*(3), 345-373.

Brennan, M. J., Chordia, T., Subrahmanyam, A., & Tong, Q. (2012). Sell-order liquidity and the cross-section of expected stock returns. *Journal of Financial Economics*, *105*(3), 523-541.

Butler, A. W., Grullon, G., & Weston, J. P. (2005). Stock market liquidity and the cost of issuing equity. *Journal of Financial and Quantitative Analysis*, 331-348.

Cavallo, E. A., & Valenzuela, P. (2007). The Determimants of Corporate Spreads in Emerging Markets: An Option-Adjusted Spreads Analysis. *Available at SSRN 963982*.

Chakravarty, S., Harris, S. P., & Wood, R. (2001). Decimal trading and market impact. *Available at SSRN 266877*.

Chan, K., Hameed, A., & Kang, W. (2013). Stock price synchronicity and liquidity. *Journal of Financial Markets*, *16*(3), 414-438.

Charoenwong, C., Ding, D. K., & Siraprapasiri, V. (2011). Adverse selection and corporate governance. *International Review of Economics & Finance*, *20*(3), 406-420.

Choi, W. G., & Cook, D. (2006). Stock market liquidity and the macroeconomy: Evidence from Japan. In *Monetary Policy with Very Low Inflation in the Pacific Rim* (pp. 309-340). University of Chicago Press.

Chordia, T., Sarkar, A., & Subrahmanyam, A. (2005). The joint dynamics of liquidity, returns, and volatility across small and large firms. *Finance*.

Chordia, T., Subrahmanyam, A., & Anshuman, V. R. (2001). Trading activity and expected stock returns. *Journal of financial Economics*, *59*(1), 3-32.

Chung, K. H., & Van Ness, R. A. (2001). Order handling rules, tick size, and the intraday pattern of bid–ask spreads for Nasdaq stocks. *Journal of Financial Markets*, *4*(2), 143-161.

Claessens, S., Dasgupta, S., & Glen, J. (1998). The cross-section of stock returns: evidence from the emerging markets, Emerging Markets Quarterly 2 (Winter): 4–13.

Coughenour, J. F., & Saad, M. M. (2004). Common market makers and commonality in liquidity. Journal of Financial economics, 73(1), 37-69.

Crowe, C. W., Johnson, S., Ostry, J. D., & Zettelmeyer, J. (2010). *Macrofinancial linkages: trends, crises, and policies*. International Monetary Fund.

Cumby, R. E., & Huizinga, J. (1992). Investigating the correlation of unobserved expectations: Expected returns in equity and foreign exchange markets and other examples. Journal of Monetary Economics, 30(2), 217-253.

Datar, V. T., Naik, N. Y., & Radcliffe, R. (1998). Liquidity and stock returns: An alternative test. *Journal of Financial Markets*, *1*(2), 203-219.

De Cesari, A., Espenlaub, S., & Khurshed, A. (2011). Stock repurchases and treasury share sales: Do they stabilize price and enhance liquidity?. *Journal of Corporate Finance*, *17*(5), 1558-1579.

Dicle, M. F. (2008). Commonality of liquidity around the world: Evaluation of possible reasons.

Easley, D., & O'Hara, M. (2003). Microstructure and asset pricing. *Handbook of the Economics of Finance*, *1*, 1021-1051.

Easley, D., Hvidkjaer, S., & O'hara, M. (2002). Is information risk a determinant of asset returns?. *The journal of finance*, *57*(5), 2185-2221.

Eleswarapu, V. R., & Reinganum, M. R. (1993). The seasonal behavior of the liquidity premium in asset pricing. *Journal of Financial Economics*, *34*(3), 373-386.

Eleswarapu, V.R., 1997. Cost of transacting and expected returns in the NASDAQ market. Journal of Finance 52, 2113–2127

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, *33*(1), 3-56.

Froot, K. A., O'connell, P. G., & Seasholes, M. S. (2001). The portfolio flows of international investors. *Journal of financial Economics*, *59*(2), 151-193.

Gluzmann, P., & Panigo, D. (2015). Global search regression: A new automatic model-selection technique for cross-section, time-series, and panel-data regressions. *Stata Journal*, *15*(2), 325-349.

Gopalan, R., Kadan, O., & Pevzner, M. (2012). Asset liquidity and stock liquidity. *Journal of Financial and Quantitative Analysis*, 333-364.

Goyenko, R. Y., Holden, C. W., & Trzcinka, C. A. (2009). Do liquidity measures measure liquidity?. *Journal of financial Economics*, *92*(2), 153-181.

Grandes, M., Panigo, D., & Pasquini, R. (2006). The cost of equity beyond CAPM: Evidence from Latin American stocks (1986-2004). *Center for Financial Stability, Working Paper No18*.

Griffin, C. H. (2010). Liquidity and dividend policy: international evidence. International Business Research, 3(3), 3.

Hasbrouck, J. (2007). *Empirical market microstructure: The institutions, economics, and econometrics of securities trading*. Oxford University Press.

Hillert, A., Maug, E., & Obernberger, S. (2016). Stock repurchases and liquidity. *Journal of financial economics*, *119*(1), 186-209.

Huang, A. Y., & Chang, C. L. (2015). Dynamics of Stock Liquidity.

Jacoby, G., Fowler, D. J., & Gottesman, A. A. (2000).The capital asset pricing model and the liquidity effect: A theoretical approach. *Journal of Financial Markets*, *3*(1), 69-81.

Jain, P. K. (2003). Institutional design and liquidity at stock exchanges around the world. *Available at SSRN 869253*.

Jun, S. G., Marathe, A., & Shawky, H. A. (2003). Liquidity and stock returns in emerging equity markets. *Emerging Markets Review*, *4*(1), 1-24.

Jung, Y., & Hu, J. (2015). AK-fold averaging cross-validation procedure. *Journal of nonparametric statistics*, *27*(2), 167-179.

Kearney, C. (2012). Emerging markets research: Trends, issues and future directions. *Emerging Markets Review*, *13*(2), 159-183.

Ko, K., & Lim, T. (2006). Short selling and stock prices with regime switching in the absence of market makers: The case of Japan. *Japan and the World Economy*, *18*(4), 528-544.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, 1315-1335.

Lam, K. S., Tam, L. H., & Dong, L. (2019). Liquidity and Stock Returns: Evidence from the Chinese Stock Market. *中国会计与财务研究*, *21*(4).

Lee, J. H., Lin, S. Y., Lee, W. C., & Tsao, C. Y. (2006). Common factors in liquidity: Evidence from Taiwan's OTC stock market. *International Review of Financial Analysis*, *15*(4-5), 306-327.

Lee, K. H., Sapriza, H., & Wu, Y. (2016). Sovereign debt ratings and stock liquidity around the world. *Journal of Banking & Finance*, *73*, 99-112.

Lesmond, D. A. (2005). Liquidity of emerging markets. Journal of Financial Economics, 77(2), 411-452.

Lesmond, D. A., Ogden, J. P., & Trzcinka, C. A. (1999). A new estimate of transaction costs. *The review of financial studies*, *12*(5), 1113-1141.

Liem, P. F., & Basana, S. R. (2012). Price Earnings Ratio and Stock Return Analysis (Evidence from Liquidity 45 Stocks Listed in Indonesia Stock Exchange). Jurnal Manajemen dan Kewirausahaan, 14(1), 7-12.

Lipson, M. L., & Mortal, S. (2009). Liquidity and capital structure. *Journal of financial markets*, *12*(4), 611-644.

Lischewski, J., & Voronkova, S. (2012). Size, value and liquidity. do they really matter on an emerging stock market?. *Emerging Markets Review*, *13*(1), 8-25.

Liu, F. (2010). *Statistical analysis on market microstructure models* (Doctoral dissertation).

Madhavan, A. (2000). Market microstructure: A survey. *Journal of financial markets*, *3*(3), 205-258.

Makles, A. (2012). Stata tip 110: How to get the optimal k-means cluster solution. The Stata Journal, 12(2), 347-351.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The journal of finance*, *42*(3), 483-510.

Myers, S. C., & Rajan, R. G. (1998). The paradox of liquidity. *The Quarterly Journal of Economics*, *113*(3), 733-771.

Ni, K., Bresson, X., Chan, T., & Esedoglu, S. (2009). Local histogram based segmentation using the Wasserstein distance. *International journal of computer vision*, *84*(1), 97-111.

Ni, K., Bresson, X., Chan, T., &Esedoglu, S. (2009). Local histogram based segmentation using the Wasserstein distance. International journal of computer vision, 84(1), 97-111.

Pástor, Ľ., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political economy*, *111*(3), 642-685.

Rehman, I. U., Mahdzan, N. S., & Zainudin, R. (2016). Is the relationship between macroeconomy and stock market liquidity mutually reinforcing? Evidence from an emerging market. *International Journal of Monetary Economics and Finance*, *9*(3), 294-316.

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance*, *39*(4), 1127-1139.

Rouwenhorst, K. G. (1999). Local return factors and turnover in emerging stock markets. *The journal of finance*, *54*(4), 1439-1464.

Stoll, H., 2000. Friction. Journal of Finance 56, 1–31

Tang, K., & Wang, C. (2011). Corporate governance and firm liquidity: evidence from the Chinese stock market. *Emerging Markets Finance and Trade*, *47*(sup1), 47-60.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267-288.