Trading Liquidity and Funding Liquidity in Fixed Income Markets: Implications of Market Microstructure Invariance*

Albert S. Kyle
University of Maryland
akyle@rhsmith.umd.edu

Anna A. Obizhaeva
New Economic School
aobizhaeva@nes.ru

First Draft: April 27, 2016

Abstract

This essay applies market microstructure invariance to fixed income markets. An invariance-based illiquidity measure calibrated from stock market data is extrapolated to the markets for Treasury and corporate fixed income securities. By consistently incorporating both leverage neutrality, this illiquidity measure explains both trading liquidity and funding liquidity. Invariance predicts that Treasury markets are about 55 times more liquid than markets for individual corporate bonds and operate about 3,000 times more quickly. Invariance is used to discuss repo haircuts and the flash rally of October 15, 2014.

Keywords: market microstructure, liquidity, bid-ask spread, market impact, transaction costs, order size, invariance, fixed income, banking, systemic risk, repo markets.

*Kyle has worked as a consultant for various companies, exchanges, and government agencies. He is a non-executive director of a U.S.-based asset management company.

The purpose of this paper is to examine trading liquidity and funding liquidity in fixed income markets from the perspective of market microstructure invariance. The principles of market microstructure invariance imply an easily calculated dimensionless illiquidity measure, denoted $1/L$. The quantity $1/L$ is easy to calculate because it depends only on dollar volume, returns volatility, and two "invariant" scaling constants which are the same for all assets. The main idea of this paper is that $1/L$ is a simple, convenient measure of both trading liquidity and funding liquidity. Trading liquidity and funding liquidity are two sides of the same coin, measured in the same way.

The application of microstructure invariance is particularly interesting in the context of fixed income markets because liquidity varies greatly across different fixed income instruments. The market for on-the-run Treasury notes and related futures contracts is one of the most liquid in the world. Corporate bonds, by contrast, are highly illiquid. Trading volume in a typical corporate bond may be about 100,000 times less than the most liquid Treasury market. In its simplest terms, microstructure invariance implies a scaling law with exponents of 1/3; this scaling law relates illiquidity to dollar volume and volatility. The wide variations in volume seen across fixed income markets provides a simple way to examine whether the scaling laws associated with market microstructure invariance give economically reasonable outcomes.

The illiquidity measure $1/L$ is dimensionless. As a measure of trading liquidity, $1/L$ is by definition scaled to measure the average transaction cost of executing a risk-transferring "bet" (meta-order) in a market, expressed as a fraction of the value traded (basis points). Its reciprocal L, which measures liquidity, is proportional to the dollar size of bets which take place in the market. As a measure of funding liquidity, we claim that $1/L$, in addition to being proportional to the cost of liquidating defaulted collateral, is also proportional to the haircut that makes a repo safe and to the standard deviation of returns over intervals of time at which it is practical to mark assets to market. This happens because $1/L$ is proportional to the square root of the returns between the arrival of one bet in the market and the next. Assuming bets are liquidated in "business time" scales proportional to the rates at which bets arrive, this makes $1/L$ proportional to the standard deviation of returns over the horizon that a speculative position would be liquidated in the event of bankruptcy. A haircut proportional to $1/L$ therefore provides similar protection to a liquidator of defaulted collateral both for liquid collateral with a low value of $1/L$ and for illiquid collateral with a high value for $1/L$.

Market microstructure invariance deals with time and leverage in a simple, internally consistent manner which incorporates leverage neutrality in a manner analogous to Modigliani-Miller irrelevance. Since leverage is related to fixed income markets, fixed income markets provide a particularly clean environment for thinking about microstructure invariance. Leverage neutrality is a minimal requirement for any liquidity measure which deals cleanly with funding liquidity or trading liquidity.

The plan of this paper is as follows.

We first compare equity markets with fixed income markets.

We next define and explain the dimensionless liquidity measure $1/L$ and a related measure γ defined as the rate at which bets arrive into the market in calendar time. Two invariant constants are chosen to be consistent with bet size and transition costs estimated by Kyle and Obizhaeva (2016) using a dataset of portfolio transition orders.

To examine whether the principle of market microstructure invariance leads to economically reasonable results, we then extrapolate the liquidity measures for stocks to bonds by keeping the same same invariant constants but substituting dollar volume and returns variance for Treasury or corporate bonds, respectively, for dollar volume and returns volatility for stocks. As a measure of trading liquidity, these extrapolations suggest that the average bet in on-the-run 10-year Treasury notes or futures has a market value of about \$20 million and the market impact cost of executing a bet of \$120 million is about 1 basis point. Individual corporate bonds are about 55 times less liquid than on-the-run Treasuries. The average bet has a market value of about \$342,000, and the market impact cost of executing a \$5 million bet is about 50 basis points.

Funding liquidity is closely related to the speed with which bets are executed in the market because lenders worry about potential losses they might suffer as a result of liquidating defaulted collateral. Market microstructure invariance implies that the speed of a market is proportional to $\sigma^2 \cdot L^2$, where σ measures an asset's returns standard deviation. Our calibration implies that the Treasury market operates about $55^2 \approx 3,000$ times faster than the market for a typical corporate bond. The market for 10-year Treasuries has about 9,000 bets per day, and a typical corporate bond has about 3 bets per day. Using repo haircuts as a measure of funding liquidity, these calculations imply higher repo haircuts for corporate bonds than for Treasuries.

To understand fire sales, it is necessary to understand how speeding up the liquidation of collateral affects temporary price impact. We supplement invariance by using intuition from the smooth trading model of Kyle, Obizhaeva and Wang (2016), which assumes that temporary price impact is proportional to the speed with which sales occur. This fire-sale approach is applied to the flash rally of October 15, 2014, using the flash crash of May 6, 2010, as a "model" for the event. We estimate that an order to purchase \$1.3 billion of 10-year Treasuries over a 12-minute period would temporarily push 10-year Treasury prices up by 1.20%.

The illiquidity measure $1/L$ explains why haircuts for illiquid corporate bonds are higher than haircuts for Treasuries and not that different from haircuts for equities. While corporate bonds may have low returns volatility, their low liquidity implies a long liquidation process in the event of repo default. In the tri-party repo market, haircuts for corporate bonds are not as high as invariance might imply. Consistent with Copeland, Martin and Walker (2010), this suggests that the repo run which disrupted the tri-party repo market occurred because cash lenders relied on the unsecured credit-worthiness of the borrower to justify low haircuts, not the expected cost of liquidating defaulted collateral.

A bank is similar to a tri-party repo which operates in slow motion. Capital requirements are like haircuts, and banking regulators supervise the liquidation of failed banks because the lenders lack the ability to do so. Bank assets like business loans and real estate development loans have very high illiquidity $1/L$, much greater than the illiquidity of corporate bonds; they are typically not traded at all but rather held to maturity. Thus, the liquidation process for failed banks is likely to take years, not days of months.

To summarize, microstructure invariance can be used both to measure how illiquidity varies across different classes of fixed income assets and to explain why trading liquidity and funding liquidity are measured in the same way.

1 Bond Markets

From the perspective of market microstructure, bond markets differ from equity markets in several ways.

Companies tend to issue many different bonds but have one important class of equity. Governments also have many different issues outstanding. As a result, equity liquidity is increased as a result of having one pool of liquidity focussed around one class of common stock while bond liquidity is reduced as a result of having numerous outstanding bond issues with different coupons, maturities, collateral, seniority, and other covenants. In bond markets, recently issued "on-the-run" bonds tend to have more liquidity than older "off-the-run" bonds.

Trading Liquidity. Fixed income markets create trading liquidity by facilitating the exchange of risks. These risks are associated with overall levels of riskfree (Treasury) rates, overall credit market conditions governing spreads between Treasuries securities and investment-grade or non-investment-grade bonds, default risks for specific issuers, and other credit risks related to the collateral, seniority, and covenants in specific bond issues. Trading a specific issue of a corporate bond involves exchanging interest rate risk, overall credit risk, specific company risk, and issue specific risks associated with the specific bond. Markets can un-package all of these risks from the specific bond issue and trade them separately, for example, as T-bond futures, CDX derivatives, specific issue CDS, and the specific bond itself. In this paper, we abstract from these interesting liquidity factors by focussing on only one liquidity factor for each asset class. The costs of exchanging these risks relate to trading liquidity.

The dramatic difference in liquidity between Treasury bonds and corporate bonds has important implications for the evolution and organization of fixed income markets. Traders like liquidity and seek out liquid venues for trading speculative instruments. The huge liquidity of the 10-year Treasury market attracts speculative and non-speculative traders to it. Central banks hold Tresuries for their liquidity. Since the liquidity of Treasuries attracts traders away from other less liquid markets, the tendency for liquidity to pool in 10-year Treasuries not only increases the liquidity of 10-year Treasuries but also reduces the liquidity of other fixed income markets. This tendency of liquidity to pool explains why off-the-run Treasuries have dramatically lower liquidity than on-the run Treasuries and why corporate bonds do not trade much at all.

Funding Liquidity. Fixed income markets create funding liquidity by separating the collateral value of the securities from their risks, essentially creating safe assets out of risky ones. The collateral value can be separated from the assets themselves in several ways: repo markets, securitization, and banks. Safe assets are not necessarily liquid.

Microstructure invariance incorporates the assumption of leverage neutrality. In effect, the illiquidity measure $1/L$ introduced below separates trading liquidity from funding liquidity by making the assumption that adding or subtracting cash from a risky position changes proportionately the trading cost measured in basis points so that the leverage does not affect the dollar cost of exchanging a risk. This makes it possible to add a time dimension to liquidity relating illiquidity to trading volume.

The repo market decomposes a bond into a safe overnight loan (collateralized by the bond) and a risky levered position which bears the overwhelming majority of the bonds' risks.

Securitization of portfolios separates safe collateral from risky tranches by pooling assets over longer horizons with less liquid bonds. Individual components of a securitization may be very illiquid, even when they are structured to be very safe and have AAA credit ratings.

A bank is also a pool of illiquid loans. Individual bank loans—such as a loan to a local real estate developer or an individual consumer—may be almost completely illiquid. Trading bank equity exchanges the pooled risks of the bank's loan portfolio. Bank deposits represent a safe investment backed by the collateral value of the bank's assets.

To summarize, bond markets perform two related functions: facilitating trading liquidity by making it easy to exchange risks and facilitating funding liquidity by moving collateral value among traders and investors. Treasury markets, interest rate swaps, credit derivatives, and markets for specific corporate bonds separate and exchange the various underlying risks. Banks and shadow banks separate the collateral value from the risk. Shadow banking includes money-market funds, securitizations, and repo markets.

2 Liquidity and Microstructure Invariance

Market microstructure invariance implies that both trading liquidity and funding liquidity are related to the asset-specific illiquidity measure $1/L_{jt}$. For asset j at time

t, this measure is a dimensionless quantity defined as

$$\frac{1}{L_{jt}} = \left(\frac{C \cdot \sigma_{jt}^2}{m^2 \cdot P_{jt} \cdot V_{jt}} \right)^{1/3}. \qquad (1)$$

In this definition, the quantities L_{jt}, P_{jt}, V_{jt}, σ_{jt}^2, C, and m^2 all have the following specific meanings. The quantity σ_{jt}^2 is the variance of daily returns; if the asset's daily standard deviation of returns is one percent, then $\sigma_{jt}^2 = 0.01^2 = 10^{-4}$ per day. The quantity $P_{jt} \cdot V_{jt}$ measures trading volume in dollars per day; it is obtained by multiplying the asset price P_{jt} (dollars per share, contract, or notional face value) by units of trading volume (shares, contracts, or notional face value per day). The subscript jt indicates that P_{jt}, V_{jt}, and σ_{jt}^2 are market characteristics of a specific asset j at time t. Markets are illiquid when return volatility σ_{jt} is high and volume $P_{jt} \cdot V_{jt}$ is low.

The constant C is the dollar expected cost of executing a bet. The constant m^2 is a dimensionless scaling parameter. Market microstructure invariance is the empirical hypothesis that these two constants are invariant across assets j and time t.[1] Therefore, neither C nor m^2 has subscripts jt. These two constants are calibrated empirically, as discussed below.

The illiquidity measure $1/L_{jt}$ is defined in a manner that carefully respects consistency of units of measurement. Multiplying the share price P_{jt} (dollars per share) by trading volume V_{jt} (shares per day) removes share units and makes the illiquidity measure immune to splits. The cost of a bet C is measured in dollars and the variance of returns σ_{jt}^2 is measured "per day." Since m^2 is dimensionless, both the numerator and denominator in the definition of $1/L_{jt}$ have the same units of dollars per day. Since the units in the numerator cancel with the units in the denominator, the illiquidity measure $1/L_{jt}$ is dimensionless; it can be interpreted as a fraction. For the purpose of measuring trading liquidity, $1/L_{jt}$ can measures trading costs as a fraction of the value traded. For the purpose of measuring funding liquidity, $1/L_{jt}$ may be proportional to the haircut of a repo transaction, measured as a fraction of the value of the collateral.

Since $1/L_{jt}$ is a dimensionless quantity, it does not depend on the units of volume, currency, and time in terms of which volatility σ_{jt}, price P_{jt}, and volume V_{jt} are measured. If units of shares, dollars, and days are changed to units of \$100 par value, Euros, and years, the same dimensionless value of $1/L_{jt}$ is obtained.

In equation (1), the exponent 1/3 makes the liquidity measure consistent with "leverage neutrality," which captures the same idea as "Modigliani-Miller irrelevance." Consider leveraging up a stock by paying a cash- or riskless-debt-financed dividend equal to half the value of a share. The stock price should should halve and the returns standard deviation should double. If illiquidity $1/L_{jt}$ measures the cost

[1] Since C is measured in dollars, its dollar value will vary over time. Its value is invariant when adjusted for inflation and the productivity of finance professionals.

of trading as a fraction of the value traded or the size of a haircut as a fraction of the value of the asset, then $1/L_{jt}$ should double because the dollar risk of exchanging a given number of shares is constant while the dollar value of a share has halved. An exponent of 1/3 is exactly what is required for $1/L_{jt}$ to behave in this manner.

In effect, this leverage neutrality property of the definition of illiquidity $1/L_{jt}$ separates trading liquidity from funding liquidity in a consistent manner. The definition of $1/L_{jt}$ implies that adding or subtracting cash from a package consisting of the risky asset and the debt used to finance it does not affect the cost of exchanging the underlying risk in the marketplace. The same Modigliani-Miller neutrality in the previous paragraph can be applied to a trader who uses leverage provided by funding liquidity in the following manner. Consider a trader who trades bonds using 10% haircut (equity) capital while borrowing the remaining 90% of the bonds' value in the funding market. Call this 9-to-1 leverage. The trader can measure liquidity using $1/L_{jt}$ based on volume and volatility of the bond. The trader can also measure liquidity based on his own use of leverage. Leveraging a position 9-to-1 is equivalent to reducing the asset price P_{jt} by a factor of 10 because each 100 dollars of asset is attached to 90 dollars of debt. Leveraged dollar volume—the amount of equity traders exchange when buying or selling risky assets—is less than unleveraged volume by a factor of 10. Similarly, leveraged volatility is greater by a factor of 10. The ratio $\sigma_{jt}^2/(P_{jt} \cdot V_{jt})$ inside the exponent of 1/3 in equation ((1)) increases by a factor of 1,000. Taking the cube root changes the factor of 1,000 to a factor of 10, which measures the increase in leveraged illiquidity in a manner that assumes the dollar cost of exchanging a risk is not affected by how the underlying asset is financed. A transaction cost of 20 basis points of the value of the underlying asset becomes a transaction cost of 200 basis points relative to the trader's 10% "equity" or haircut in the trader's position. Again, the exponent must be exactly 1/3 for this concept of leverage neutrality to hold.

A "bet" (meta-order) represents a decision by an institutional investor to buy or sell a specific random quantity. Bets are typically broken up into numerous individual orders and executed over time as numerous trades; therefore the size of a bet may not be proportional to the size of a trade.

Kyle and Obizhaeva (2016) use a dataset of portfolio transition orders to estimate both the size distribution of bets and the transactions costs of executing them. The estimated mean bet size and the estimated transaction costs are approximately consistent with the following specific calibrated values for the invariant constants C and m^2:

$$C = \$2,000, \qquad m^2 = 1/4. \tag{2}$$

Here, the dollar value $C = \$2,000$ is interpreted as the unconditional dollar expected cost of executing a bet of random size, under the identifying assumption that portfolio transition orders have the same random size distribution as bets. Plugging these calibrated constants into equation (1) yields

$$\frac{1}{L_{jt}} = \left(\frac{2,000 \cdot \sigma_{jt}^2}{1/4 \cdot P_{jt} \cdot V_{jt}} \right)^{1/3}. \tag{3}$$

Let $\widetilde{Q}_{jt}$ denote the random number of shares in a bet, with $\widetilde{Q}_{jt} > 0$ denoting buying and $\widetilde{Q}_{jt} < 0$ denoting selling. Assume buys and sells are equally likely, so that $E\{\widetilde{Q}_{jt}\} = 0$. Let $\overline{Q}_{jt} = E\{|\widetilde{Q}_{jt}|\}$ denote the average (unsigned) size of a bet (in shares). The dimensionless quantity m^2 is chosen to scale $1/L_{jt}$ so that the equality

$$P_{jt} \cdot \overline{Q}_{jt} = C \cdot L_{jt} \tag{4}$$

holds. This equation has two interesting interpretations. On the one hand, it implies that the dollar size of a bet $P_{jt} \cdot \overline{Q}_{jt}$ is proportional to L_{jt}, with invariant constant of proportionality C. On the other hand, $1/L_{jt}$ can be interpreted as a transaction cost as follows. Write the equation in the form $1/L_{jt} = C/(P_{jt} \cdot \overline{Q}_{jt})$. Since the numerator C is the average dollar cost of a bet and the denominator $P_{jt} \cdot \overline{Q}_{jt}$ is the average dollar size of a bet, the illiquidity measure $1/L_{jt}$ measures the "dollar-weighted" average transactions cost as a fraction of the dollar value traded. Both of these interpretations of $1/L_{Jt}$ represent strong empirical restrictions implied by invariance.

Kyle and Obizhaeva (2016) find that a hypothetical "benchmark stock" (near the middle of the S&P 500) with $P_{jt} = \$40$ per share, $V_{jt} = 1$ million shares per day and $\sigma_{jt} = 2\%$ per day has an average transactions cost of $1/L_{jt} = 43$ basis points. The definition of $1/L_{jt}$ implies that increasing dollar volume by a factor of 8 (while holding volatility constant) decreases the average transactions cost by a factor of $8^{1/3} = 2$ to about 22 basis points. This empirical restriction is supported in the portfolio transitions data. Furthermore, the values of transactions costs are is similar in magnitude to the transaction costs estimated by Angel, Harris and Spatt (2011) and Angel, Harris and Spatt (2015) using a different dataset.

The distribution of portfolio transition order sizes turns out to be log-normal with log-variance of 2.53.[2] In what follows, we augment the invariance assumptions

[2] Specifically, the distribution of bets inferred from portfolio transition orders for equities resembles a log-normal of the form

$$\ln\left(\frac{|\widetilde{Q}_{jt}|}{V_{jt}}\right) \approx \left(\frac{P_{jt} \cdot V_{jt} \cdot \sigma_{jt}}{(40)(10^6)(0.02)}\right)^{-2/3} \cdot e^{-5.71+\sqrt{2.53}\cdot \widetilde{Z}_{jt}}, \qquad \widetilde{Z}_{jt} \sim N(0,1) \tag{5}$$

The various quantities in equation ((5)) are scaled so that the median bet in a \$40 stock with trading volume of 1 million shares per day and volatility of 2% per day represents a fraction of average daily volume equal to $e^{-5.71} \approx 0.33\%$ of average daily volume. The equation implies $P_{jt} \cdot \overline{Q}_{jt}$ is proportional to L_{jt}, and the distribution of $\widetilde{Q}_{jt}$ is log-normal with log-variance equal to 2.53 for all stocks. The scaling exponent of $-2/3$ and the log-variance 2.53 reflect strong empirical restrictions confirmed in the portfolio transitions data.

by assuming that the shape of the distribution of bets in fixed income markets is the same log-normal as in portfolio transitions, with different means for different assets depending on their liquidity L_{jt}.

Let γ_{jt} denote the number of bets arriving each day. Assume that bet volume and trading volume are the same; this is equivalent to assuming that the other side of each bet trade is taken by a non-bet intermediary, like a specialist or market maker. Since $\overline{Q}_{jt}$ measures average bet size and V_{jt} measures share volume, the ratio $V_{jt}/\overline{Q}_{jt}$ measures the expected number of bets per day, under the assumption that trading volume and bet volume are the same. Equations (1) and (4) imply that the number of bets per day can be written in various ways as

$$\gamma_{jt} = \frac{V_{jt}}{\overline{Q}_{jt}} = \frac{\sigma_{jt}^2 \cdot L_{jt}^2}{m^2} = \left(\frac{P_{jt} \cdot V_{jt} \cdot \sigma_{jt}}{m \cdot C}\right)^{2/3}. \tag{6}$$

The value of γ_{jt} measures the speed of the market, the rate at which "business time" passes. For the benchmark stock, the number of bets is estimated as approximately 85 per day.[3]

We call the quantity $W_{jt} = P_{jt} \cdot V_{jt} \cdot \sigma_{jt}$ in equation (6) "trading activity." Note that trading activity W_{jt} has non-intuitive units of "dollars per unit of time to the 3/2 power." In equation (6), dividing by C removes the dollar units and taking the 2/3 power allows γ_{jt} to be measured per unit of time, consistent with the interpretation that γ_{jt} measures the rate at which bets arrive.

2.1 Trading Liquidity

The illiquidity measure $1/L_{jt}$ is a dimensionless quantity which can be conveniently used to describe different dimensionless aspects of liquidity. As mentioned above, the illiquidity measure $1/L_{jt}$ is scaled to measure trading liquidity as the "dollar-weighted" average transactions cost as a fraction of the dollar value traded.

The transactions cost of executing a specific bet depends on the size of the bet. Let $\widetilde{Z}_{jt} := P_{jt} \cdot \widetilde{Q}_{jt}/(C \cdot L_{jt})$ denote the size of a bet as a multiple of the average dollar bet size $C \cdot L_{jt}$. Let $C^{\$}_{jt}(Z)$ denote the dollar transaction cost of executing a bet that is Z times the size of an average bet, and let $C^{\%}_{jt}(Q_{jt})$ denote the cost of a bet of Q_{jt} shares as a fraction of its value. Invariance implies that, for some function $f(Z) \geq 0$, the dollar transaction cost function has the "invariant" form

$$C^{\$}(Z) = C \cdot |Z| \cdot f(Z), \qquad \text{with} \qquad E\{|\widetilde{Z}| \cdot f(\widetilde{Z})\} = 1, \tag{8}$$

[3]For portfolio transitions involving equities, the constants m and C imply

$$\gamma_{jt} \approx 85 \cdot \left[\frac{P_{jt} \cdot V_{jt} \cdot \sigma_{jt}}{(40)(10^6)(0.02)}\right]^{2/3}. \tag{7}$$

Here, the quantities are scaled so that a \$40 stock with trading volume of 1 million shares per day and volatility of 2% per day has about 85 bets per day, consistent with Kyle and Obizhaeva (2016).

and the percentage transaction cost function equivalently has the form

$$C^{\%}_{jt}(Q_{jt}) = \frac{1}{L_{jt}} \cdot f(Z) \qquad \text{with} \qquad Z := \frac{P_{jt} \cdot Q_{jt}}{C \cdot L_{jt}}. \tag{9}$$

Leveraging a bet up or down does not change Z_{jt} because P_{jt} and L_{jt} change proportionally with changes in leverage. The dollar transaction cost function $C^{\$}(Z)$ does not depend at all on the cash value of the collateral involved in the bet; it depends only on the size of the risk exchanged as a multiple of the risk of an average bet. The percentage transaction cost function $C^{\%}_{jt}(Q_{jt})$ reflects the cash value of the collateral through the liquidity parameter L_{jt}.

If the transactions cost is the sum of a bid-ask spread component and a linear price-impact component, the resulting linear price impact transaction cost model can be written

$$f(Z) = \kappa + \lambda \cdot |Z_{jt}|. \tag{10}$$

The values $\kappa = 0.15$ and $\lambda = 0.0576$, consistent with estimates from portfolio transitions for equities, imply dollar costs given in table 1.[4]

The data in table 1 illustrate the importance of the price impact of large orders. The variance of $\log(\widetilde{Z})$ is 2.53, a very large number. This makes the distribution of unsigned bet size $|\widetilde{Z}|$ have extremely large skewness and gives the symmetric distribution of signed bet size extreme kurtosis (with $\widetilde{Z} > 0$ for buys, $\widetilde{Z} < 0$ for sells). The median bet is only 28% as large as the average bet. The average cost of $2,000 per bet is generated mostly by very high costs for very large bets.

2.2 Funding Liquidity

Some researchers think of trading liquidity as something different from funding liquidity. We will next show that funding liquidity, like trading liquidity, is related to the illiquidity measure $1/L_{jt}$ and the speed of the market. Thus, funding liquidity and trading liquidity are two sides of the same coin.

[4] The function $C^{\%}_{jt}(Q)$ can also be written

$$C^{\%}_{jt}(\widetilde{Q}_{jt}) = \kappa \cdot \sigma_{jt} \cdot \left(\frac{P_{jt} \cdot V_{jt} \cdot \sigma_{jt}}{m \cdot C}\right)^{-1/3} + \lambda \cdot \sigma_{jt} \cdot \left(\frac{P_{jt} \cdot V_{jt} \cdot \sigma_{jt}}{m \cdot C}\right)^{1/3} \cdot \frac{|\widetilde{Q}_{jt}|}{V_{jt}}. \tag{11}$$

For portfolio transitions, the transaction costs with coefficients calibrated to fit a linear price impact model can be written

$$C^{\%}_{jt}(Q) = \frac{\sigma_{jt}}{0.02}\Big(\frac{8.21}{10^4} \cdot \left[\frac{P_{jt} \cdot V_{jt} \cdot \sigma_{jt}}{(40)(10^6)(0.02)}\right]^{-1/3} + \frac{2.50}{10^4} \cdot \left[\frac{P_{jt} \cdot V_{jt} \cdot \sigma_{jt}}{(40)(10^6)(0.02)}\right]^{1/3} \frac{Q}{(0.01)V_{jt}}\Big). \tag{12}$$

This equation is scaled so that a bet of 1% of average volume in a $40 stock with volume of 1 million shares per day and volatility of 2% per day has a bid-ask spread cost of 8.21 basis points and a market impact cost of 2.50 basis points.

Table 1: Transaction Costs with Bid-Ask Spread and Linear Price Impact

Standard Deviation of $\log(\lvert\widetilde{Z}\rvert)$	Scaled Size $\lvert\widetilde{Z}\rvert$	Spread Cost $C \cdot \kappa \cdot \lvert\widetilde{Z}\rvert$ (\$)	Impact Cost $C \cdot \lambda \cdot \widetilde{Z}^2$ (\$)	Total Cost $C \cdot (\kappa \cdot \lvert\widetilde{Z}\rvert + \lambda \cdot \widetilde{Z}^2)$ (\$)
$+0 \cdot \eta =$ median	0.28	85	9	94
$+0.7953 \cdot \eta =$ avg	1.00	300	115	415
$+1 \cdot \eta$	1.38	415	221	636
$+2 \cdot \eta$	6.79	2,039	5,315	7,353
$+3 \cdot \eta$	33.34	10,002	127,957	137,959
$+4 \cdot \eta$	163.59	49,078	3,080,627	3,129,705

This table assumes that scaled bet size $\lvert\widetilde{Z}\rvert$ has a log-normal distribution with $\log(\lvert\widetilde{Z}\rvert) \sim N(-\eta^2/2, \eta^2)$ with $\eta^2 = 2.53$. The transaction cost of executing a bet, $C^{\$}(\widetilde{Z}) = C \cdot \widetilde{Z} \cdot f(\widetilde{Z})$, is assumed to be the sum of a bid-ask spread term and a linear price impact term of the form $f(\widetilde{Z}) = \kappa + \lambda \cdot \lvert\widetilde{Z}\rvert$. The dollar values of the spread and impact terms are calibrated from estimates of portfolio transition orders for equities in Kyle and Obizhaeva (2016). The implied values are approximately $\kappa = 0.15$ and $\lambda = 0.0576$.

Funding liquidity measures the costs of financing asset positions with debt collateralized by the assets themselves. Funding liquidity has two components, an interest rate and a haircut (equity or one-minus-loan-to-value-ratio). Funding transactions are typically structured as repurchase agreements in which the collateral is temporarily sold to lender and bought back at the maturity date of the funding transaction. In repo transactions, the lenders of cash want the loan to be safe. Therefore, the haircut is set to make the probability of default very small, so that the loan is almost risk-free. As a result, haircuts vary across assets with different risks but funding rates do not vary as much. The term of the funding transaction is frequently one day.

Returns Volatility and Marking to Market If the liquidity of the collateral is ignored, a theory of repo haircuts is likely to relate the size of the haircut to the volatility of returns. For example, if the daily returns volatility of the collateral is σ_{jt}, the collateral is marked to market at time intervals ΔT and S returns standard deviation of haircut is needed to make the loan safe, then the haircut is given by

$$\text{Haircut } = \sigma_{jt} \cdot \Delta T^{1/2} \cdot S. \tag{13}$$

Operationally, assets may be marked to market and haircuts recalculated at fixed time intervals of say $\Delta T =$ one day. When ΔT is held fixed, this leads to a theory of haircuts based on returns volatility σ_{jt}, not liquidity $1/L_{jt}$.

We believe that this theory of haircuts fails to take into account how time interacts with funding liquidity. When time is taken into account, the funding haircut should be proportional to $1/L_{jt}$ and not proportional to σ_{jt}, especially for illiquid collateral.

Since borrowers of cash have an incentive to conserve haircut capital and lenders both have an incentive to reduce counter-party credit risk, it is reasonable to conjecture that repos will be marked to market and collateral posted as frequently as is practical. When the collateral consists of very liquid assets, it may be practical to mark to market by posting collateral on an intraday basis. For example, futures contracts post collateral as cash; liquid futures contracts may be marked to market more than once per day in unusually volatile markets.

For illiquid collateral, marking to market on a frequent basis becomes difficult because the value of the asset is difficult to observe. This is especially true at times of market stress, when markets may slow down but posting collateral matters most. As a practical matter, efforts by lenders to collect collateral from cash borrowers whose collateral is falling in value will likely result in disputes about the valuation of the collateral. For thinly traded assets, these valuations disputes are likely to be difficult to resolve, and it will therefore take time for the process of marking to market and posting collateral to take place in an orderly manner. The difficulty of observing accurate market prices is likely related to the rate at which bets are taking place in the market. Implementing collateral agreements which require marking to market thinly traded assets is likely to become difficult over short periods of time when few bets take place. If the borrower of cash defaults, the collateral is liquidated. The speed of the liquidation process is also likely to depend on the speed with which business time passes. A theory of haircuts based on daily volatility σ_{jt} implicitly incorporates the assumption that, in the event of default, the collateral can be liquidated immediately at current market prices with zero transactions costs. In addition to accounting for the risk between dates that the collateral is marked to market, the haircut should also contemplate forced liquidation by containing a component which compensates the lender of cash for costs associated with a forced liquidation. We conjecture that that marking to market, posting collateral, negotiating valuation disputes, and liquidating defaulting collateral are business activities that take place at a rate proportional to the rate γ_{jt} at which bets arrive. When collateral is illiquid, these activities take place very slowly. When business time is properly taken into account, it turn out the haircuts should not be proportional to the daily volatility of the collateral σ_{jt} but rather should be proportional to the illiquidity of the collateral $1/L_{jt}$.

To develop this idea further, consider how to measure volume and volatility in units of business time. Average volume is one bet per unit of business time. Since the average size of a bet is $C \cdot L_{jt}$, we have

$$\text{Dollar Volume per Bet } = C \cdot L_{jt}. \tag{14}$$

Returns variance per unit of business time is $\sigma_{jt}^2/\gamma_{jt}$. As a standard deviation, the returns standard deviation in one unit of business time can be written

$$\text{Volatility per Bet } = \frac{\sigma_{jt}}{\gamma^{1/2}} = \frac{m}{L_{jt}}. \tag{15}$$

These equations show that the constants C and m^2 used in the definition of liquidity can be used to convert L_{jt} into measures of volatility and volume per unit of business time.

This immediately leads to another economic interpretation of the concept of market microstructure invariance. The standard deviation of the change in the dollar mark-to-market value of a bet of random bet of size $\widetilde{Q}_{jt}$ is given by

$$E\left\{\frac{P_{jt} \cdot \widetilde{Q}_{jt} \cdot \sigma_{jt}}{\gamma_{jt}^{1/2}} \Big| \widetilde{Q}_{jt}\right\} = \left(C \cdot L_{jt} \cdot \widetilde{Z}_{jt}\right) \cdot \left(\frac{m}{L_{jt}}\right) \sim m \cdot C \cdot \widetilde{Z}. \tag{16}$$

In the above equation, the notation "$\sim$" means "equal in distribution to." Dropping the subscript jt from $\widetilde{Z}_{jt}$ at the last step incorporates the extra assumption that the probability distribution of $\widetilde{Z}_{jt}$ is invariant across stocks (a log-normal with log-variance 2.53). Of course, the mean of $|\widetilde{Z}_{jt}|$ is one by definition. Thus, the standard deviation of the change in dollar mark-to-market value of an "average bet" is $m \cdot C$; invariance assumes that this number is constant across all assets. The extra assumption implies that the shape of the probability distribution of $\widetilde{Z}_{jt}$ is the same across stocks also. Kyle and Obizhaeva (2016) develop the concept of invariance around the assumption that the size of risks transferred by markets are the same when measured in units of business time.[5]

Since m/L_{jt} measures the returns standard deviation between the arrival of one bet and the next, the standard deviation of returns over the period of time borrowers and lenders mark returns to market, post collateral, and liquidate defaulted collateral can be measured as $m/L_{jt} \cdot B^{1/2}$, where B measures business time as the expected number of bets which take place between the time of making a margin call and finishing liquidation of defaulted collateral. Note that B does not have a subscript jt because business-time units are the same for all assets.

What matters for funding liquidity is returns volatility per unit of business time, measured by $1/L_{jt}$, not returns volatility per unit of calendar time, measured by σ_{jt}. This analysis leads to the intuition that trading liquidity and funding liquidity are fundamentally the measured in the same way because both are measured by the illiquidity measure $1/L_{jt}$.

Fire Sales In a situation in which defaulted repos are being liquidated, the cash lender who is liquidating the collateral may be more interested in being made whole than in minimizing the transactions costs of liquidation. To reduce the risk of adverse price fluctuations over the liquidation horizon, the cash lender may have an incentive to speed up collateral sales even if this increases transactions costs. The result is a "fire sale."

To further understand fire sales, it is necessary to understand how transactions costs are affected by shortening the liquidation horizon. The invariance-implied average transactions cost of $1/L_{jt}$ makes the implicit assumption that the seller liquidates a large position gradually enough to achieve reasonable or normal transactions costs.

While the theory of market microstructure invariance does not (yet) have a well-developed approach to modeling fire sales, we can take some first steps towards developing such a theory here by examining a simplistic example. The transaction cost for executing a bet is likely to depend on the time period over which it is executed. The transaction cost function $f(Z)$ above implicitly assumes that the bet is executed at a "normal" speed designed to optimize the typical trade-off between the urgency of the bet and the costs associated with increased speed of execution. In addition to Z, consider adding another parameter to the function $f(Z)$ to measure the speed of liquidation. Let H_{jt} denote the horizon of execution of a bet, such as a forced collateral liquidation, in days. Rather than using calendar time, it is more convenient to measure this time horizon in "natural" units of business time measuring the number

[5] If all returns volatility results from the price impact of bets, then equation (16) requires $m^2 := E\{\widetilde{Z}^2\}$. Use of the letter m reminds us that m^2 is the second "moment" of $\widetilde{Z}$. This refinement involves understanding empirical relationships related to the fraction of volume which results from bets, the fraction of volatility which results from bets, and the level of dealer profits on intermediation trades. These issues may require re-calibrating the constants m^2 and C. They take us beyond the scope of this paper.

of bets expected to arrive during the execution horizon H_{jt}. Since γ_{jt} bets arrive per day, this horizon is given by $B_{jt} := H_{jt} \cdot \gamma_{jt}$

A generalized formula for the transactions cost function, including a time horizon, can be written

$$C^{\$}(Z, B) = C \cdot Z \cdot f(Z, B), \quad \text{where} \quad Z = \frac{P_{jt} \cdot Q_{jt}}{C \cdot L_{jt}}, \quad B = H_{jt} \cdot \gamma_{jt} = \frac{H_{jt} \cdot \sigma_{jt}^2 \cdot L_{jt}^2}{m^2}. \tag{17}$$

Now consider the shape of the function $f(Z, B)$. The function $f(Z, B)$ should have the properties that transaction costs are greater if the bet size Z is larger and the execution horizon B is shorter. Since linearity makes things simple, consider a transactions cost formula in which things are as linear as possible.

The smooth trading model of Kyle, Obizhaeva and Wang (2016) implies an intuitive theory of how speed affects transactions costs, and it works particularly well for fire sales. We will use this theory here. In the smooth trading model, temporary price impact is proportional to the rate at which sales occur. This implies the functional form

$$f(Z, B) = \kappa + \lambda \cdot \overline{B} \cdot \frac{Z}{B}, \tag{18}$$

which is bilinear in scaled bet size Z and scaled speed $1/B$. Doubling the speed of asset sales $1/B$ doubles the temporary market impact cost. The quantity $\overline{B}$ is the "normal" rate at which bets take place; if $B = \overline{B}$, then equation ((18)) is the same a equation ((10)). Since market impact dominates transaction costs for large transactions, we can think of the transaction cost $f(Z, B)$ as resulting mostly from market impact costs so that $\kappa \approx 0$.

PK::

In a smooth trading equilibrium with trading occurring under "normal" (equilibrium) market conditions, traders with private information choose the degree of temporary price impact so that over the long run the temporary price impact approximately becomes permanent price impact as their private information gradually dissipates into the market. This implies a choice of liquidation horizon B. The smooth trading model has an equilibrium in which each trader choose a trading speed such that his actual inventory converges toward a target inventory at a constant rate. We can think of this constant rate as determining a "natural" value for B. For example, if typical bets are executed at a speed which represents one percent of volume over the horizon, then $B = 100$ for a typical bet.

The trader's internal model correctly recognizes that trades have temporary price impact averaging $1/L_{jt}$ for trades of normal size executed at a normal speed (assuming this covers almost all bets). In principle, the trader could measure his price impact costs by comparing pre-trade prices with execution prices (implementation shortfall). An economist studying market efficiency and price impact would discover empirically that most of the price impact appeared to be permanent. In a fire sale situation, a trader chooses to sell at a speed much more rapid than normal. An economist would observe that temporary price impact is reversed after the trade takes place.

The above analysis suggests that natural repo haircuts are the sum of two components:

$$\text{Haircut } = \frac{1}{L_{jt}} \cdot \left(S \cdot m \cdot B^{1/2} + \lambda \cdot f(Z, B)\right). \tag{19}$$

The first component represents the returns standard deviation during the interval over which prices are marked to market, collateral posted, and collateral liquidated in default. The second component represents the cost of liquidating defaulted collateral. This is the haircut model consistent with market microstructure invariance. It implies that the haircut is proportional to illiquidity $1/L$; it also implies a direct connecting between trading and funding liquidity implied by the cost of executing a bet.

3 Trading and Funding Liquidity in U.S. Treasury Market

The values $C = \$2,000$ and $m^2 = 1/4$ are chosen to generate values for $1/L_{jt}$ and γ_{jt} consistent with estimates of bet size and transactions costs for individual stocks based on equity market portfolio transitions. In this section, we use these parameters to calculate implied values for $1/L_{UST}$ and γ_{UST} for the market for on-the-run U.S Treasuries. The purpose is to examine whether microstructure invariance generates scaling laws which apply to securities other than equities and to securities with dramatically different liquidity characteristics.

3.1 Measuring Size and Liquidity of Fixed Income Markets

The size or liquidity of a market is properly measured by the sizes of the risks it exchanges. To focus discussion, consider a simplistic one-factor Treasury bond market in which the only risk of concern is a parallel movement in the yield curve. The return volatility of a bond is the product of the volatility of yields and the bond's duration. Thus, if the volatility of yields is $\sigma_Y = 5$ basis points per day per day, the return volatility of a bond with a 10-year duration is $\sigma_{10} = 50$ basis points per day, the volatility of a bond with a 5-year duration is $\sigma_5 = 25$ basis points per day, and the volatility of a bond with a 2-year duration is $\sigma_2 = 10$ basis points per day.

To simplify discussion, suppose the Treasury market consists of volume in two-year, five-year, and ten-year bonds. Suppose the bonds are are perfect substitutes from the perspective of interest rate risk with durations of approximately two, five, and ten years respectively. Suppose that arbitragers make sure that the bonds are constantly priced with the correct arbitrage relationship.

How should the aggregate size of the market be measured? Clearly, since durations differ by factors of 2 and 5, it is intuitively appropriate to begin with a weighted average of dollar volume where the weights differ by factors proportional to the durations of the bonds. Since returns volatilities are proportional to durations, volatilities provide a good weighting scheme. Let $P_{10} \cdot V_{10}$, $P_5 \cdot V_5$, and $P_2 \cdot V_2$ denote dollar volume per day in on-the-run Treasury notes with maturities of 10, 5, and 2 years, respectively. Here, P_{10} measures the dollar price of one dollar of face value and V_{10} measures the total face value or notional value in dollars per day. [6] A reasonable aggregate measure of the sizes of the risks transferred by the entire market, consistent with our previous definition of trading activity, is given by

$$\text{Treasury Trading Activity} = P_{10} \cdot V_{10} \cdot \sigma_{10} + P_5 \cdot V_5 \cdot \sigma_2 + P_2 \cdot V_2 \cdot \sigma_2 \tag{20}$$

Since, for example, $\sigma_{10}/\sigma_2 = 10/2 = 5$, this measure of aggregate risk transfer has the desirable property that it is unchanged if \$100 billion of two-year note volume is converted into economically equivalent \$20 billion of ten-year note volume.

Assume that bond markets price the exchange of risk consistently in the sense that the dollar market impact cost of exchanging the same economic risk is the same dollar amount. Suppose that buying \$600 million of ten-year notes has a market impact cost 5/256 of one percent of the par value. Then buying \$3 billion of two-year notes will have a price impact cost of 1/256 of one percent of the par value. In both cases, the dollar price impact cost is $\$500 \cdot 10^6 \cdot 0.01 \cdot 5/64 = \$390,625$.

Measured as a fraction of the market value of the two bonds, the market impact cost of buying or selling the ten-year note is five times greater than the two-year note. The fractional cost of 5/256 of one percent of par for the ten year note is $5/256 \cdot 1\% = 0.0195\% = 1.95$ basis points. The fractional cost of 1/265 of one percent of par for the two-year note is $1/256 \cdot 1\% = 0.0039\% = 0.39$ basis points.

Investors participate in fixed income markets for two distinct motives. On the one hand, some investors trade fixed income securities to exchange risks associated with changes in interest rates or credit spreads. Such investors are interested in minimizing the market impact costs of exchanging a risk of a given size. On the other hand, many investors hold fixed income assets for their liquidity. Such investors prefer to hold very safe assets for short time periods. They therefore have an incentive to choose assets whose trading costs are low as a fraction of the value of the asset.

Volume. Applying microstructure invariance to the entire U.S. Treasury bond market at once requires approximations for volume and returns volatility.

Trading volume is split between the futures market and the cash market and also split across different maturities, with on-the-run maturities of 2, 5, and 10 years dominating trading volume. High frequency traders keep the cash and futures market tightly arbitraged so that the entire market functions as one integrated market.

[6]For stocks, P_{jt} measures the price in dollars per share and V_{jt} measures the number of shares traded per day. For bonds, one "share" is one dollar of face value or one dollar of notional value. Either way, the product $P_{jt} \cdot V_{jt}$ is measured in dollars per day.

According to charts in the Joint Staff Report (2015), summarized in our table 2, daily dollar volume in the cash and futures markets for 10-year, 5-year, and 2-year maturities combined sums to approximately \$232 billion in the years leading up to the flash rally. To convert the daily volume numbers ten-year-equivalent duration-weighted amounts, we weight five-year volume by 0.50 and weight two-year volume by 0.20. These simplistic approximations result in ten-year-equivalent volume of approximately \$168 billion.

Table 2: Daily Treasury Bond Trading Volume

Cash or Futures	Maturity (Years)	Daily Volume (\$ billion)	Weight	Weighted Volume (\$ billion)
Futures	10	80	1.00	80
Futures	5	40	0.50	20
Futures	2	16	0.20	8
Cash	10	40	1.00	40
Cash	5	40	0.50	20
Cash	2	20	0.20	4
Sum		232		168

Volatility. Charts in the Joint Staff Report (2015) imply a median daily absolute yield change of about 5 basis points and a median daily high-low range of about 8 basis points. This implies a median daily standard deviation of returns of 5 or 6 basis points. To convert basis points into approximate percentage returns, multiply a duration ten years by 5 basis points to obtain an approximate daily standard deviation of 50 basis points per day, implying $\sigma_{UST} = 0.0050 = 50$ basis points or 32/64 of one percent of par.

Liquidity and Bet Size. Now plug $P_{UST} \cdot V_{UST} = \$168$ billion and $\sigma_{UST} = 0.0050$ into equations ((1)) to obtain $L_{UST} = 9435.388 \approx 10^4$, or

$$\frac{1}{L_{UST}} \approx 10^{-4} = 1 \text{ basis point.} \tag{21}$$

The "magic number" $1/L_{UST} \approx 10^{-4}$ generates many implied predictions about the Treasury bond market.

The average size of a bet, expressed as ten-year equivalents, is predicted to be $C \cdot L_{UST} \approx \$2,000 \cdot 10^4 = \$20$ million. The value weighted average transaction cost is $1/L_{UST} \approx 10^{-4} = 1$ basis point.

Assume that the shape of the distribution of Treasury bet sizes is the same as the shape of the distribution of portfolio transitions. The log-normal distribution of unscaled bet size has a log-variance of 2.53. This makes the distribution of bet size very skewed. The vast majority of bets are smaller than average and have smaller transactions costs than the value-weighted average. The implied sizes, transactions costs, and probabilities of bets of different sizes are presented in table 3.

Table 3: Implied Probability Distribution of U.S. Treasury 10-Year Bet Sizes

Standard Deviation	Scaled Size	\$ Million Size	Probability Larger	T-Cost (bp)
$+0 \cdot \eta =$ median	0.28	6	0.50	0.17
$+0.7953 \cdot \eta =$ average	1.00	20	0.21	0.21
$+1 \cdot \eta$	1.38	28	0.16	0.23
$+2 \cdot \eta$	6.79	136	0.023	0.54
$+3 \cdot \eta$	33.34	667	0.0013	2.07
$+4 \cdot \eta$	163.59	3271	0.000032	9.57
$+4.6113 \cdot \eta$	432.56	8651	0.0000020	25.04

This table measures bets in Treasuries in units of standard deviation of $\log(\widetilde{Z})$ in column 1, in scaled size Z in column 2, and dollar size in column 3. The scaled sizes and probability in column 4 is based on a log-normal distribution with log-variance $\eta^2 = 2.53$. The transaction cost in column 5 assumes the same model of linear price impact as table 1.

According to table 3, the median bet has a size of \$6 million and the average bet a size of \$20 million. It takes a bet of \$8.6 billion to move prices 25 basis points if the bet is executed at a "normal" speed. The implied transaction cost of 0.17 basis point for the median bet consists mostly of bid-ask spread costs.

This extrapolation from data for portfolio transitions in stocks assumes that institutional details of the Treasury market do not affect scaled transaction costs. One relevant institutional detail may be tick size. The most liquid venue for trading 10-year Treasury risk is probably the futures market for 10-year Treasury notes. The tick size in this market is 1/64 of one percent of par. Thus, one tick represents 1.5625 basis point of return, and one-half tick represents 0.78125 basis points of return. The implied transaction cost of 0.17 basis points is approximately 1/5 the size of a half-tick. Thus, the implied transaction cost for small bets implies that traders use smart-enough order execution systems so that trades are timed to save approximately 80 percent of the cost that would be incurred from hitting bids or lifting offers randomly. Since the tick size in both the cash and futures market is large (relative to $1/L_{UST}$), the bid-ask spread is typically one tick, and there are large quantities available at both the bid and offer.

Now consider a rather large "two-standard-deviation" bet of size $\log(\widetilde{Z}) = 2 \cdot \eta$.

According to table 3, such a bet is 6.79 times larger than an average bet and has a dollar size of \$136 million. The implied execution cost is 0.54 basis points, less than one half-tick of 0.78 basis points.

Tables in the Joint Staff Report (2015) suggest that when trading returned to somewhat normal patterns in the afternoon after the flash rally of October 15, 2014, three-tick-level market depth in the futures market was about \$1 billion and three-level market depth in the cash market was about \$400 million. Dividing by three ticks and dividing again by two to convert to half ticks implies one-half-tick depth of about \$200 million. This is consistent with the interpretation that simultaneously hitting bids or offers in the cash and futures market may be a reasonable trading strategy for executing relatively large bets. It takes only a modest amount of timing ability to reduce costs from 0.78 basis points to 0.54 basis points. Indeed, while the large tick size may increase the cost of small orders, it may decrease the cost of larger orders.

It will require further study to verify whether the liquidity estimates implied by invariance are reasonable. These simple calculations suggest that they are.

Number of Bets. The number of bets per day is given by

$$\gamma_{UST} = \frac{P_{UST} \cdot V_{UST}}{C \cdot L_{UST}} \approx 8,900 \text{ bets per day.} \tag{22}$$

This is approximately 100 times greater than the 85 bets per day for a typical stock.

Interesting market properties may be inferred from its speed. Multiplying the number of bets per day by the cost of a bet C may generate a good approximation to the quantity of real resources devoted to a particular market.

Large bets are typically not executed immediately. Instead, they are broken into small pieces and executed gradually over time. It is likely that the time scale of execution of bets is proportional to the rate at which bets arrive. If a large bet of \$136 million is executed as a fraction of 1% of bet volume over its execution horizon, then it will be executed over a horizon during which 679 bets arrive. For a typical stock with 85 bets per day, this horizon is approximately 8 days. For the U.S. Treasury bond market, the horizon is approximately 100 time shorter and equal to approximately 1/12 of a day, a time frame of less than one hour of active volume. While hitting bids and lifting offers may be a reasonable way to execute a bet of \$135 million, timing over approximately one hour may reduce transaction costs for a passively executed bet.

It is reasonable to conjecture that market resiliency, defined as the speed or half-life with which the market recovers from a the price impact of an informationless bet, is proportional to the rate at which bets arrive. Suppose that it take approximately a year for the price of a typical stock to recover from the price impact of an informationless bet. Since the U.S. Treasury bond market operates about 100 times faster than the market for the typical stock, it may take only two or three days for the U.S. Treasury bond market to recover from the price impact of an informationless bet executed in a slow and deliberate manner to avoid "transitory" impact associated with execution of the bet itself.

The Flash Crash and the Flash Rally The "flash rally" of October 15, 2014, focussed attention on the functioning of the U.S. Treasury bond market from the perspective of market microstructure. During the "flash rally" of October 15, 2014, the entire Treasury market was placed under a great deal of stress for about 12 minutes.

In the beginning of the day, bond prices exhibited significant volatility unprecedented in the recent history of the Treasury market. Over the 25 minutes after the U.S. retail sales report for the month of September at 8:30 a.m. ET, prices started to drift upward and the 10-year yield declined by about 11 basis points. Between 9:33 and 9:45 a.m. yields dropped sharply further by about 16 basis points and then rebounded. Taking into account duration of the 10-year U.S. Treasuries, this change in yields corresponded to the 120-basis-point decline and rebound in prices compressed into a 12-minute time period. The Joint Staff Report (2015) says there were no obvious explanations for this event.

Market microstructure invariance can be used to calibrate may have happened during the Treasury market flash rally of October 15, 2014. The observed price patterns are consistent with the execution of a large order over a short period of time. Our approach is to use the stock market "flash crash" of May 6, 2010, as as a "model" flash event. The scaling laws associated with microstructure invariance can then be used to extrapolate from the U.S. stock futures markets to the U.S. fixed income market.

On May 6, 2010, a single large trader sold 75,000 contracts (over $4 billion) of S&P E-mini future over a period of 20 minutes, using an algorithm which targeted participating in 9% of trading volume over the period of execution. During the 20-minute period, prices collapsed by about 5%, then recovered. As prices collapsed, arbitragers bought futures contracts and sold ETFs and individual stocks. During the previous twelve months, individual traders sold equal or larger single-day quantities only twice. On one of these occasions, the same large traders sold 75,000 contracts over a period of five hours.

Using equation (5), it can be shown that 75,000 contracts is a 4.30-standard deviation event in the entire U.S. stock market consisting of both markets for individual U.S. stocks and the U.S. stock futures. The overall dollar volume of the entire U.S. stock market is about $300 billion per day and volatility is about one percent per day. Similar calculations are done by Kyle and Obizhaeva (2013). A 4.30-standard-deviation order should normally be executed over a period of about five hours.

Using (7), it can be shown that there are about 20,000 bets executed per day in the U.S. stock market. The U.S. Treasury market runs more than twice slower than the U.S. stock market, since there are only about 8,900 bets executed per day. Plugging the dollar daily volume of about $168 billion per day and volatility of 50 basis points per day for the U.S. Treasury market into equation (5) implies that an equivalent 4.30-standard deviation order in this market would be about \$5.6 billion. Taking into account the difference in speed between two markets, it would take about 12 hours to executed this order if it is executed at a natural pace.

Next suppose that the "natural" execution horizon increases linearly with the order size. Assume also that the price impact increases linearly with the speed by how much execution is speeded up relative to the natural speed.

The flash-rally occurred over the period of 12 minutes. How big had to be a large bet executed over 12-minute period to change price by about 120 basis points? Plugging daily volume and daily volatility as well as GDP deflator of 0.85 to adjust for inflation into an exponential version of the transaction cost formula (12), we obtain an order of about \$1.3 billion:

$$1 - \exp\Big[-\frac{5.00}{10^4} \cdot \Big(\frac{168 \cdot 0.85 \cdot 10^9}{40 \cdot 10^6}\Big)^{1/3} \cdot \Big(\frac{0.0050}{0.02}\Big)^{4/3} \cdot \frac{1.3 \cdot 10^9}{0.01 \cdot 168 \cdot 10^9}\Big] \approx 0.0120.$$

If the order were to be executed over a one-hour period, for example, from 9 a.m. to 10 a.m., then its size might have been about \$3 billion. [7]

Funding Liquidity in the Treasury Market. Because of its great liquidity, funding costs for U.S. Treasury collateral are very low relative to other assets. Our analysis above suggested that funding liquidity is proportional to $1/L$, which equals approximately 1 basis point for 10-year Treasuries. Although our main point about funding liquidity is that illiquidity $1/L$ is more important than daily returns volatility σ in setting repo haircuts, borrowing against Treasury collateral may be an exception to this principle. Unless positions are tens of billions of dollars, minimizing haircuts for Treasury collateral would probably result in marking to market many times per day and liquidating default collateral over a period of hours or minutes, not days. As a practical matter, it is probably convenient to mark to market, post collateral, and liquidate collateral based on institutionally convenient decision intervals of exactly one day.

The standard deviation of returns over one day is about 50 basis points. A 4-standard-deviation haircut of 2% is therefore likely to be reasonable for all but the largest collateral sizes. Suppose, for example, that the collateral for an overnight repo is \$500 million of 10-year Treasury securities. According to table 3, the transaction

[7] Similar calculations for the Flash crash suggest a price impact of about 8 percent

$$1 - \exp\Big[-\frac{5.00}{10^4} \cdot \Big(\frac{(300) \cdot 0.90 \cdot 10^9}{40 \cdot 10^6}\Big)^{1/3} \cdot \Big(\frac{0.01}{0.02}\Big)^{4/3} \cdot \frac{75,000 \cdot 50 \cdot 1,164}{0.01 \cdot (300) \cdot 10^9} \cdot 15\Big].$$

Here we assumed that the execution was speeded up by a factor of 15 (20 minutes instead of 5 hours). The prices have changed by only about 5 percent, potentially due to the effect of triggered circuit breakers that arrested the price decline.

cost of liquidating the position is only about 2 basis points. Since the liquidation can take place over a time horizon of minutes, the liquidation costs and risks of liquidation are small relation to the 200 basis point haircut.

Liquidation costs and liquidation horizons become relevant when the collateral size is large. Consider the last line of table 3, which considers a bet large enough to move prices temporarily by 25 basis points. The bet size required to do so is 433 times large than the average bet, representing a collateral position of approximately \$8.6 billion. Given about 8,900 bets per day, liquidating such collateral in one day would require participating in around 4.87% of average daily volume for an entire day. A lower participation rate might result in lower liquidation costs, but spreading the liquidation over several days would increase the probability that underlying market movements make the haircut inadequate.

4 Trading and Funding Liquidity in Corporate Bond Market

4.1 Corporate Bond Markets

Corporate bond markets are much less liquid than U.S. Treasury markets and operate at a much slower speed. To get a sense of the difference, make the back-of-the-envelope assumption that institutional bet volume in a typical bond is $P_{CB} \cdot V_{CB} = \$1$ million per day and the idiosyncratic standard deviation of daily returns is $\sigma_{CB} = 50$ basis points. This standard deviation includes firm-specific and issue-specific risks but is meant to ignore interest rate risk. The interest rate, while significant, can be hedged using government bonds.

Using the invariance assumptions $C = \$2,000$ and $m^2 = 0.25$, we calculate $L \approx 171$. This implies that average transactions costs are given by

$$\frac{1}{L_{CB}} \approx 0.0058 = 58 \approx 55 \text{ basis points.} \tag{23}$$

These assumptions imply an average transaction cost of $1/L_{CB} \approx 55$ basis points and bet size of $C \cdot L_{CB} = \$342,000$, which implies about $P_{CB} \cdot V_{CB}/(C \cdot L_{CB}) \approx 3$ bets per day. A median bet of about \$100,000 has an implied cost of about 10 basis points, a trade of \$1 million has an implied cost of about 20 basis points, and a trade of \$10 million has an implied cost of about 100 basis points.

The results of this back-of-the envelope calculation are meant to be approximately consistent with the detailed study of corporate bond transactions costs by Harris (2015), who reports about 2.5 trades per day and an average effective half-spread of 50.7 basis points on trades of size greater than \$1 million. While our implied costs for large trades are similar or perhaps slightly smaller in magnitude, our implied costs of 10 basis points for \$100,000 trades are much smaller than the 85 basis point cost for retail trades (less than \$100,000) reported by Harris (2015). This is consistent with the interpretation that the antiquated dealer structure of the corporate bond market has the effect of increasing costs for small customers by a factor of five or ten.

Our measure of liquidity $L_{UST} \approx 1$ basis point is about 55 times higher than the liquidity measure for a single issue of a corporate bond $L_{CB} \approx 55$ basis points. Thus, the average implied trading cost is 55 times higher in the corporate bond market. Our assumption that Treasury bonds and corporate bonds (idiosyncratic risk) have the same volatility $\sigma_{UST} = \sigma_{CB} = 50$ basis points per day implies that bets in the Treasury market are on average 55 times larger than bets in the corporate bond market. The speed of the Treasury market is faster by a factor of 55-squared, or approximately 3,000 (approximately consistent with a comparison of 3 bets per day for a corporate bond with 8,900 bets per day for 10-year Treasuries).

Compare an equally-weighted portfolio of 55 different corporate bonds with a portfolio of 10-year Treasuries of the same dollar size. Suppose that the Treasury portfolio is so gigantic that it would take one full day to liquidate. Then a diversified portfolio of corporate bonds of the same size would take 3,000 days to liquidate, and the total liquidation cost would be 55 times greater than than Treasury portfolio; as a practical matter, the gigantic portfolio of corporate bonds would probably be held to maturity since 3,000 trading days is more than ten years.

Funding Liquidity. Measures of funding liquidity should take account of time when the collateral is illiquid. Suppose, for example, that the collateral is the hypothetical portfolio of 55 corporate bonds mentioned above. The transaction costs associated with liquidation are likely to be significant, and the haircut level in the example above should be modified to take this into account. The transactions cost associated with liquidation is proportional to $1/L_{CB}$, some multiple of about 55 basis points. Furthermore, the horizon over which the collateral is liquidated is likely to be weeks or months, depending on the size of the positions. During these weeks and months, the value of the collateral may decline further.

Let us apply this to the portfolio of 55 corporate bonds. The standard deviation of returns of 50 basis points per day on each of the 55 bonds reflects the credit risk only, not the interest rate risk on the portfolio. There is an implicit assumption that the interest rate risk has been hedged. In the context of a repo transaction, the owner of the corporate bond portfolio may have funded the transaction by selling short an equal quantity of Treasury bonds to hedge the interest rate risk. Thus, if the the repo transaction becomes distressed and must be liquidated, the lender is likely to the sell the corporate portfolio and buy back the Treasuries. Since the Treasury portfolio was large, its assumed transaction cost of liquidation was about 2 times as large as the average transactions cost of $1/L_{UST} = 1$ basis point. Invariance implies that the transaction cost of liquidating each of the 55 bonds is about twice the average transaction cost of 55 basis points, or about 110 basis points. Furthermore, the liquidation horizon is likely to be very long. Invariance implies that the sales may take place over a period of several weeks, say 25 business days. This implies that the standard deviation of returns for each bond over the liquidation horizon is about $\sigma_{CB} \cdot 25^{1/2} = 0.0250 = 250$ basis points. With a 3-standard-deviation cushion, the repo lender needs an extra haircut of $110 + 3 \cdot 250 = 860$ basis points to make the repo a safe lending proposition. Compared with the case of a Treasury portfolio, the 860 basis points of haircut needed to cover liquidation risk is much larger than the 150 basis point haircut needed to cover market movements (under the assumption that the returns on the corporate bonds will be perfectly correlated in a liquidation situation).

Although the stated maturity of the assumed overnight repo is one day, the effective maturity is likely to be much longer for illiquid collateral when stress situations are contemplated when the repo is entered into.

Now let us apply this fire-sale intuition to the problem of liquidating a large corporate bond portfolio. If the corporate bond portfolio is sold at normal speed over 25 days, the execution cost is 110 basis points and the standard deviation of returns over the entire horizon is 250 basis points. Now suppose that the cash lender speeds up liquidation of the collateral by a factor of 4. Then temporary price impact increases by a factor of 4, magnifying the cost to approximately $4 \cdot 110 = 440$ basis points. The standard deviation of returns over the investment horizon decreases from 250 basis points to $250 \cdot 4^{1/2} = 125$ basis points. If the lender must return extra haircut collateral to the borrower after a successful liquidation, it is easy to see why a lender with and 860 basis point haircut will speed up liquidation. With a normal rate of liquidation, as assumed above, the lender suffers losses proportional to the amount by which a random variable with volatility of 250 basis points exceeds three standard deviations. With the speeded up liquidation, the lender suffers losses proportional to the amount by which a random variable with volatility of 125 basis points exceeds $(860 - 440)/125 = 3.36$ standard deviations. The lender is better off both because the number of standard deviations of coverage is greater (3.36 versus 3.00) and because the standard deviation itself is smaller (125 basis points versus 250 basis points). As the repo becomes more poorly collateralized, these incentives magnify.

In this fire sale situation, we are likely to observe a transitory price impact of approximately 440 basis points, followed by a reversal of approximately the same amount. This fire sale discount and reversal are likely to unfold and reverse over a period of time equal to 1/4 the normal horizon of 25 days, or approximately 6 days.

Of course, the borrower benefits from a slower rate of selling because the value of the implicit call option increases. The borrower benefits not only from the increased volatility associated with slower execution but also from the more favorable striking price associated with lower execution costs passed along to him in the event the seller is made whole.

Given the divergence of interests between the buyer and the seller when a repo defaults, the borrower has an incentive to keep the haircut on the repo well-funded. When the value of the collateral falls, the asset owner can keep the repo well-funded by either by adding more collateral to restore the haircut based on marking to market or by selling off a portion of the asset so that the value of the existing haircut as a fraction of the value of the remaining position increases. De-leveraging by selling off a portion of the collateral may also result in fire-sale prices if the de-leveraging transactions are done at a faster speed than the normal speed for such assets. When a distressed borrower owns different collateral with different degrees of liquidity, this logic suggests that the asset owner will liquidate collateral with faster markets at a faster speed. Invariance implies that faster markets are not exactly the same thing as more liquid markets when measured by L. Recall from the discussion above that the cost of liquidating a typical portfolio at a typical speed is measured by $1/L \sim (\sigma^2/(P \cdot V))^{1/3}$ while the horizon of liquidation for a typical portfolio at a typical speed is proportional—not by $1/L$—but rather to $1/(\sigma^2 \cdot L^2) \sim (P \cdot V \cdot \sigma)^{2/3}$. Speed and liquidity are both monotonically increasing trading volume $P \cdot V$. Thus, given two assets with the same volatility σ, a distressed seller will likely liquidate the assets with higher liquidity $1/L$ first or at a faster pace. If two assets have the same trading volume $P \cdot V$ but different volatilities σ, a distressed seller is likely to liquidate the higher volatility assets first or at a faster pace.

For example, in the 10-year Treasury market, where we estimate 8,900 bets take place per day, observing accurate market prices is likely to be easy because there are numerous arms-length prices and the path of prices is reasonably continuous. In an illiquid asset with 3 bets per day, it is possible that the buyers and the sellers might be the borrowers or lenders of the repo collateral, which implies that the prices may not be arms length. Indeed, the borrower of cash my be adding to his position to prop up prices, as occurred during the JPMorgan episode involving the London Whale. Meanwhile, the lender of cash—anticipating taking possible ownership of the collateral if the borrower defaults or making purchases if the borrower exits with a fire sale—may be selling the collateral in anticipation of reacquiring it later. The borrower will, of course, call this front-running and complain.

These perverse incentives explain why an orderly approach to handling defaulting repo transactions is for the cash lender to take possession of the collateral by purchasing the collateral from the borrower in exchange for relieving the borrower of further obligations to repay the transaction. Such orderly exchanges will most likely when the defaulting borrower has little other net worth, as might be the case of a distressed hedge fund. Of course, such exchanges work best when the mark-to-market haircut value is enough to cover expected liquidation costs. If so, then the lender can take over the collateral and either liquidate it gradually or hold it to maturity.

This thinking is consistent to what happened to Long Term Capital Management (LTCM) in 1998 and also to what happened to the failing BSAM hedge funds in summer 2007. In the case of LTCM, the creditors took over the vast majority of ownership in the underlying positions in fall 1998, held the positions for a long time, and did not receive a cash bailout from the government. In the case of the BSAM hedge funds, Bear Stearns bailed out the hedge funds bearing its initials to protect its reputation even though the original repo counterparties were banks different from Bear Stearns. Over the following months, the market lost confidence in Bear Stearns, Bear Stearns lost its ability to fund its positions in the repo market, JPMorgan acquired Bear Stearns in a distressed sale in March 2008, and the U.S. government participated in a controversial bailout by using taxpayer dollars to buy off $30 billion of assets JPMorgan refused to accept at the market prices Bear Stearns had assigned to them. By taking over the $30 billion in assets and not liquidating them in a fire sale, the Fed avoided adding immediate price pressure to already distressed assets and perhaps postponed the collapse of the financial system until the fall of 2008.

5 Short-Selling, Bilateral- and Tri-Party Repos, Derivatives

Copeland, Martin and Walker (2010) provide a detailed analysis of the tri-party repo market. Here we discuss some related issues.

There are direct connections between the trading and funding markets. Traders placing speculative long bets use the repo market to borrow cash. Short sellers placing speculative short bets use the bilateral repo market to borrow securities to be sold short. There is an important asymmetry between the markets for borrowing cash and borrowing securities.

In the market for borrowing cash, the lenders are interested in being protected by an adequate haircut. In addition, some lenders may be particularly interested in their ability to unwind the repo transaction and be repaid in a timely manner. Lenders interested in being able to unwind their loan quickly are likely to demand liquid Treasury securities as collateral. As a result, the repo rate for Treasury securities is likely to be lower than the repo rate for less liquid securities. Indeed, if forced liquidation horizons are something like 20–25 business days for illiquid non-Treasury collateral, the rate on repos with illiquid collateral is likely to be similar to the rate on 30-day commercial paper of similar credit-worthiness. Commercial paper, while often safe, is not liquid and in most cases held to maturity.

To attract money-market investors seeking high interest rates, money market funds do have an incentive to accept less liquid collateral in exchange for a higher interest rate. When collateral is illiquid, valuing the collateral is difficult. Large and important market participants, such as money market funds, may not have the sophistication to value illiquid collateral. The tri-party repo market structure directly addresses this issue by placing a third party, a custody bank, between the borrower and lender of cash. The third party does its job well when it keeps collateral safe and values collateral accurately for lenders of cash. During the financial crisis, the tri-party repo market ceased to function effectively. There was a run on the market. To some extent, this run resulted from investors shifting funds from money market funds holding safe but less liquid assets to money market funds investing in government securities. To some extent, this run also resulted from money market funds demanding higher haircuts than borrowers were willing to pay. Copeland, Martin and Walker (2010) suggest the money market funds took account of both the unsecured creditworthiness of the borrower and the haircut. When the creditworthiness of borrowers deteriorated (due to ratings downgrades), the haircuts did not adjust due to structural rigidities in the market; as a result, the lenders engaged in a shadow-bank run. The analysis in this paper suggests that the higher haircuts should have resulted from a slowing down of business time related to increase volatility and reduced dollar volume "when the music stopped" and the financial crisis began.

An additional perspective on the bilateral and tri-party repo markets is obtained by examining the market from the perspective of short sellers who want to borrow specific collateral so that they can sell it short. For liquid markets like Treasuries, the market for negotiating repo transactions is itself likely to be liquid. Under normal circumstances, a short-seller should be able to locate numerous lenders of collateral willing to engage in a bilateral repo at a rate reflecting a small rental-rate premium for specific collateral. Under abnormal conditions, the repo market for specific Treasury securities may reflect a high rental rate or "specialness" due to owners of collateral being unwilling to lend it due for strategic reasons.

For illiquid securities, such as specific corporate bonds or specific tranches of securitizations, the number of investors holding institutional-size positions is likely to be small and the bilateral repo market is likely to be illiquid. Owners of illiquid collateral may also have a strategic interest in keeping the market price of the collateral high. For example, the owners may face capital requirements based on marking the value of fixed income investments to market. A higher market price makes it look like a trader with a long position is making larger trading profits, and it conserves haircut capital. For these strategic reasons, the owner may prefer not to lend the securities to a borrower of securities who might sell the securities short. Instead, the owner has an incentive to finance the securities with a tri-party repo in which the third party promises to keep the collateral "off the street" in custody account, where short sellers cannot get their hands on it. Thus, tri-party repos are a funding mechanism which can be used to prop up the value of securities by keeping the collateral out of the hands of short sellers. Incentives to do this are strongest when long positions are most concentrated.

Now let us examine tri-party repos from the perspective of microstructure invariance. Fixed income securities are often chopped into many small issues. A single corporation often issues dozens of different securities, even though it only has one type of equity outstanding. Similarly, a given pool of mortgage assets is often sliced into 20 or more tranches which trade as separate securities. Such slicing and dicing of corporate and mortgage securities reduces the trading volume of each security. Furthermore, fixed income securities are often purchased and held to maturity. To some extent, this occurs precisely because the securities are illiquid. The combined result is that trading volume is not only low because issue size is low but also is low because the turnover rate of the securities is low.

Now invariance predicts that as trading volume falls, bet size also falls, but falls at a rate proportional only to the 1/3 power of trading volume (holding volatility constant). Consider what happens when the issue size is reduced by a factor of 8. This reduces the size of each bet by a factor of 2 and doubles transactions costs when measured in basis points. Thus, the size of the smaller amount outstanding, expressed as a number of average-sized bets, decreases by a factor of 4. Business time slows down by a factor of 4. The less competitive market for borrowing collateral discourages short sales.

The result is low trading volume, low market liquidity, and low funding liquidity.

Derivatives To deal with these issues, the finance industry has created credit default swaps and "pay-as-you-go" swaps. Credit default swaps allow different specific issues to be delivered against one generic issuers. They essentially allow a short seller substitute one specific issue for another. Pay-as-you-go swaps are structured to create a bet on an assets cash flows without the short seller having to borrow the underlying collateral which is shorted. In effect, the pay-as-you go swap is like a repo whose term is the maturity of the asset, except that the exchange of physical collateral is replaced with a collateral agreement in which notional collateral is exchanged. While credit default swaps and pay-as-you-go swaps solve the problem of borrowing and lending collateral, they do not solve the problem of determining an accurate market price for illiquid collateral.[8] They may, however, help clarify where valuations are likely to be most inaccurate.

Consider an asset which is being held at an inflated value by an owner who does not want the market to recognize that the value of the asset has fallen. Suppose that the asset is a tranche of a securitization, for which the pay-as-you-go derivative is priced at 80 percent of its par value. The owner may nevertheless bid a price of 99 in the physical market and finance the bond in a tri-party repo arrangement. If the intermediary values the asset at 99, smart lenders of cash will demand an extra $99 - 80 = 19$ percentage points of par in additional haircut to reflect its "true" value in the derivatives markets.

6 Banking

A commercial bank is like the borrower of cash in a tri-repo transaction, except that many parts of the business move more slowly. Copeland, Martin and Walker (2010) compare tri-party repos with banking. There are both similarities and differences between banks and tri-party repos.

[8] Credit default swaps, unlike pay-as-you-go swaps, also do not have an effective mechanism for determining the cash flows which an investor receives in bankruptcy.

The lenders of cash in banking are frequently retail depositors. Retail depositors are less sophisticated than the money-market-fund asset managers and other lenders in the tri-party repo market. Both retail depositors and asset managers need help monitoring the creditworthiness of the banks and tri-party repo arrangement, respectively. The clearing banks which administer tri-party repos help the money market funds by obtaining independent, third-party valuations of assets, holding the assets, and supervising the creditworthiness of the daily unwind of positions. Bank regulators help retail depositors by providing deposit insurance and supervising capital and liquidity requirements. Capital requirements for a bank play a role similar to haircuts in a tri-party repo arrangement.

For a bank, capital requirements are like haircuts. Capital requirements should be based both on risk σ and liquidity L of assets, but liquidity L is much more important. What does microstructure invariance imply about bank regulation?

Consider the distinction between the banking book and the trading book. The trading book holds assets which are liquid enough to have a meaningful market price. In principle, dollar volume $P \cdot V$ and volatility σ can be estimated, from which $1/L$ can be calculated.

We can think of illiquidity as beginning with returns standard deviation σ, then adjusting it for time by dividing by the standard deviation per bet $m \cdot \gamma^{1/2}$. Various concepts in bank regulation also require time horizons. Invariance suggests that these time horizons should related to the speed of the market γ.

Consider the concept of value-at-risk. For our purposes, value-at-risk is a first cousin of dollar standard deviation over a particular horizon; this gives both value-at-risk and standard deviation dimensions of dollars per unit of square-root-of-time. Thus, like illiquidity $1/L$, value at risk requires a time horizon. Invariance suggests that different types of assets require different horizons, and these horizons should be inversely proportional to $\gamma^{1/2}$. If so, then value-at-risk becomes a concept similar to illiquidity $1/L$.

Consider assets in the banking book and not the trading book. Many of these assets are so illiquid that they almost never trade. In the even of bank failure, it is less costly to hold them to maturity and try to collect in bankruptcy than to sell a portfolio of illiquid bad loans in the market. Stress tests are a mechanism for determining an appropriate level of bank capital. If the bank's assets are going to be so illiquid that they cannot be sold at all, then an appropriate stress test should freeze the bank's portfolio, then simulate defaults over a long horizon. Since market prices are impossible to determine, it is appropriate that stress tests have the flavor of an accounting exercise which simulate the dynamics of accounting write-offs over some horizon. Since the underlying assumption is that the asset cannot be sold, the appropriate horizon for a stress test should not be two or three years, but should instead be much longer, say ten years.

It is appropriate for regulators to try to force a bank to raise equity when the bank becomes undercapitalized. When a bank is in financial distress, it is highly leveraged.

This makes the market value of its assets low, which lowers dollar trading volume. The volatility of the bank's assets is high, both because the bank has high leverage (due to low capital) and because distressed fixed income assets have higher volatility than non-distressed assets. The resulting low dollar volume and high returns volatility will be associated with high illiquidity $1/L$. Assuming the costs of issuing equity are related to $1/L$, invariance therefore suggests that issuing equity will be very costly for a weak bank.

In practice, the apparent illiquidity may be greater than actual illiquidity. Market participants may expect the bank to be given forbearance, allowing it to delay forced equity issuance. When equity is actually issued, the price will plunge not only because of forced sales in an illiquid market, but also because the market digests the news shock of not receiving more forbearance.

7 Conclusion

This paper has applied the illiquidity measure $1/L$ to the study of both trading liquidity and funding liquidity.

When liquidity estimates for individual stocks, estimated from portfolio transition orders, are extrapolated from stocks to bonds, the results substantiate the use of $1/L$. Estimates for market impact costs for large bets in highly liquid 10-year Treasuries imply are somewhat less than the depth instantaneously available in the limit order book, suggesting that timing of modest timing trades over microstructure horizons is important for controlling transaction costs optimally. Estimates for larger trades in corporate bonds are consistent with Harris (2015).

In both the Treasury market and the corporate market, the organization of the market inflates costs for small traders. In the Treasury market, these costs are inflated by large tick size. In the corporate bond market, these costs are inflated by an antiquated dealer structure which makes it impossible for small trades to have limit order protection at all.

Invariance is also relevant for thinking about funding liquidity. When the speed at which trading activity takes place is taken into account, funding liquidity is proportional to the same illiquidity measure $1/L$ which determines trading liquidity. The illiquidity of corporate bond market implies that they function very slowly and therefore funding liquidity is very low. Bank portfolios are even more illiquid that portfolios of corporate bonds. The defining characteristic of commercial banking is its slowness.

References

Angel, James J, Lawrence E Harris, and Chester S Spatt. 2011. "Equity Trading in the 21st Century." *The Quarterly Journal of Finance*, 1(01): 1–53.

Angel, James J, Lawrence E Harris, and Chester S Spatt. 2015. "Equity Trading in the 21st Century: An Update." *The Quarterly Journal of Finance*, 5(01): 1550002.

Copeland, Adam M, Antoine Martin, and Michael Walker. 2010. "The Tri-Party Repo Market before the 2010 Reforms." *Federal Reserve Bank of New York Staff Reports*, 477: available at
`http://www.econstor.eu/bitstream/10419/60888/1/641201133.pdf`.

Harris, Lawrence. 2015. "Transaction Costs, Trade Throughs, and Riskless Principal Trading in Corporate Bond Markets." Working Paper, available at
`\\http://dx.doi.org/10.2139/ssrn.2661801`.

Joint Staff Report. 2015. "Joint Staff Report: The U.S. Treasury Market on October 15, 2014." U.S. Department of the Treasury, Board of Governors of the Federal Reserve System, Federal Reserve Bank of New York, U.S. Securities and Exchange Commission, U.S. Commodity Futures Trading Commission. Available at
`https://www.treasury.gov/press-center/press-releases/Documents/Joint_Staff_Report_Treasury_10-15-2015.pdf`.

Kyle, Albert S., and Anna A. Obizhaeva. 2013. "Large Bets and Stock Market Crashes." Working Paper, University of Maryland.

Kyle, Albert S., and Anna A. Obizhaeva. 2016. "Market Microstructure Invariance: Empirical Hypotheses." *Econometrica*, forthcoming, available at
`https://www.econometricsociety.org/publications/econometrica/journal--materials/forthcoming--papers`.

Kyle, Albert S., Anna A. Obizhaeva, and Yajun Wang. 2016. "Smooth Trading with Overconfidence and Market Power." Working Paper, University of Maryland, available at `http://dx.doi.org/10.2139/ssrn.2423207`.