Market Making Under the Proposed Volcker Rule

Darrell Duffie*
Stanford University

January 16, 2012

Abstract

This submission discusses implications for the quality and safety of financial markets of proposed rules implementing the market-making provisions of section 13 of the Bank Holding Company Act, commonly known as the "Volcker Rule." The proposed rules[1] have been described by the Office of the Comptroller of the Currency, the Board of Governors of the Federal Reserve System, the Federal Deposit Insurance Corporation, and the Securities and Exchange Commission. The Agencies' proposed implementation of the Volcker Rule would reduce the quality and capacity of market making services that banks provide to U.S. investors. Investors and issuers of securities would find it more costly to borrow, raise capital, invest, hedge risks, and obtain liquidity for their existing positions. Eventually, non-bank providers of market-making services would fill some or all of the lost market making capacity, but with an unpredictable and potentially adverse impact on the safety and soundness of the financial system. These near-term and longer-run impacts should be considered carefully in the Agencies' cost-benefit analysis of their final proposed rule. Regulatory capital and liquidity requirements for market making are a more cost effective method of treating the associated systemic risks.

*Dean Witter Distinguished Professor of Finance, Graduate School of Business, Stanford University. This submission is also a report requested from the author by SIFMA. Rather than compensating the author, SIFMA will make a charitable contribution of $50,000 to the The Michael J. Fox Foundation for Parkinson's Research. For other potential conflicts of interest, see www.stanford.edu/∼duffie/ I am pleased to acknowledge comments from Viral Acharya, Yakov Amihud, Markus Brunnermeier, Vincent de Martel, Peter DeMarzo, Peter Fisher, Michael Fleming, Andrew Lo, Gene Ludwig, Jeff Meli, Andrew Metrick, Lasse Heje Pedersen, Jacques Rolfo, Gabriel Rosenberg, Jeremy Stein, John Taylor, and Haoxiang Zhu. The opinions expressed here are entirely my own, and do not necessarily reflect the views of anyone else.

[1]See PROHIBITIONS AND RESTRICTIONS ON PROPRIETARY TRADING AND CERTAIN INTERESTS IN, AND RELATIONSHIPS WITH, HEDGE FUNDS AND PRIVATE EQUITY FUNDS, authored by Office of the Comptroller of the Currency, Treasury (OCC); Board of Governors of the Federal Reserve System (Board); Federal Deposit Insurance Corporation (FDIC); and Securities and Exchange Commission (SEC). Reference: BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM 12 CFR Part 248, Docket No. R-1432, RIN: 7100 AD 82. From this proposal document, I focus primarily on Questions 80, 81, 82, 83, 84, 87, 89, 92, 93, 96, and 97 posed by the Agencies.

1 Executive Summary

In a section of the Dodd-Frank Act commonly known as "the Volcker Rule," Congress banned proprietary trading by banks and their affiliates, but exempted proprietary trading that is related to market making, among other exemptions. Proprietary trading is the purchase and sale of financial instruments with the intent to profit from the difference between the purchase price and the sale price. Market making is proprietary trading that is designed to provide "immediacy" to investors. For example, an investor anxious to sell an asset relies on a market maker's standing ability to buy the asset for itself, immediately. Likewise, a investor who wishes to buy an asset often calls on a market maker to sell the asset out of its inventory. Market makers handle the majority of trading in government, municipal, and corporate bonds; over-the-counter derivatives; currencies; commodities; mortgage-related securities; currencies; and large blocks of equities. (The Volcker Rule exempts currencies, United States treasuries, federal agency bonds, as well as certain types of state and municipal bonds.) Most market making, both in the U.S. and abroad, is conducted by bank-affiliated broker-dealers.

Several federal agencies are now writing the specific rules by which they will implement the Volcker Rule, which comes into force in July, 2012. In particular, these agencies are charged with designing rules that implement the exemption for market making. I believe the restrictions on market making by banks in their proposed rules would have two major unintended consequences:

1. Over the years during which the financial industry adjusts to the Volcker Rule, investors would experience higher market execution costs and delays. Prices would be more volatile in the face of supply and demand shocks. This loss of market liquidity would also entail a loss of price discovery and higher costs of financing for homeowners, municipalities, and businesses.

2. The financial industry would eventually adjust through a significant migration of market making to the outside of the regulated bank sector. This would have unpredictable and potentially important adverse consequences for financial stability.

I will elaborate on these consequences and suggest an alternative approach, of using capital and liquidity requirements to conservatively buffer market-making risks. Market making risks, and other risks taken by a bank, are unsafe whenever they are large relative to the capital and liquidity of the bank.

2 Summary

This report discusses implications for the quality and safety of financial markets of proposed rules for market making by banks under section 13 of the Bank Holding Company Act, the "Volcker Rule." These rules have been proposed by the Office of the Comptroller of the Currency, the Board of Governors of the Federal Reserve System, the Federal Deposit Insurance Corporation, and the Securities and Exchange Commission (the "Agencies"). The Agencies' proposed implementation of the Volcker Rule would reduce the quality and capacity of market making services that banks provide to investors. Investors and issuers of securities would find it more costly to borrow, raise capital, invest, hedge risks, and obtain liquidity for their existing positions. Eventually, non-bank providers of market-making services would fill some of the resulting void in market making capacity, but with an unpredictable impact on the safety and soundness of financial markets. I believe these near-term and long-run impacts should be considered carefully in the Agencies' cost-benefit analysis and final rule making.

Perhaps in light of these potential adverse consequences, Congress exempted proprietary trading related to market making and certain other client-oriented services from its proprietary trading restrictions on banks. The Agencies state that they have therefore "endeavored to develop a proposed rule that does not unduly constrain banking entities in their efforts to safely provide such services." In my opinion, the proposed implementing rules would not succeed in this respect. I suggest instead rigorous capital and liquidity requirements for market makers, combined with effective supervisory monitoring, with the objective of ensuring that banks have abundant capital and liquidity to cover their market-making risks.

The Agencies' proposed implementation of the Volcker Rule seems to be written from the viewpoint that a trade involving significant risk of gain or loss, or taken with the objective of profiting from expected changes in market prices, is not consistent with bona fide market making. This is not the case. Market making is inherently a form of proprietary trading. A market maker acquires a position from a client at one price and then lays off the position over time at an uncertain average price. The goal is to "buy low, sell high." In order to accomplish this goal on average over many trades, with an acceptable level of risk for the expected profit, a market maker relies on its expectation of the future path of market prices. Future prices are uncertain because of unforeseen changes in economic fundamentals and market conditions. The length of time over which a position must be held is subject to the unpredictable timing and direction of client demands for immediacy. These risks vary significantly across time because of changes in market volatility and significant variation in the sizes of positions that market making clients may wish to acquire or liquidate. A market maker is also sometimes exposed to investors that are better informed than itself.

The greater the extent to which the proposed rule is successful at reducing market making risk, the more it will reduce the effective amount of market making services provided to clients. This would not benefit our financial system, relative to the alternative of capital requirements that force a market maker to safely absorb its own losses.

In order to provide significant immediacy to its customers, a market maker requires substantial discretion and incentives regarding the pricing, sizing, and timing of trades. It must also have wide latitude and incentives for initiating trades, rather than merely reacting to customer requests for quotes, in order to properly risk manage its positions or to prepare for anticipated customer demand or supply. Likewise, in order to efficiently provide liquidity to its clients, a market maker relies heavily on the option to buy and sell from other market makers.

While the Agencies accurately describe the relevance of these forms of market-making discretion and make some allowance for them, the criteria and metrics that are proposed would nevertheless substantially discourage the use of market making discretion. Banks would frequently find that meeting a client's demands for immediacy would be unattractively risky relative to the expected profit. In particular, a bank that continues to offer substantial market making capacity to its clients would face a risk of regulatory sanction (and the attendant stigma) due to significant and unpredictable time variation in the proposed metrics for risk and for profit associated with changes in market prices. Likewise, the norms that are likely to arise from the proposed regulatory metrics would discourage discretion by individual market making traders in the face of career concerns. A trader's incentives and discretion would also be dampened by the proposed approach to compensation.

Consequently, some banks may wish to exit the market making business. Alternatively, under the proposed rule, a bank could significantly reduce the amount of capital that it devotes to market making, merely offering this service within modest risk limits in order to cream-skim the easiest market-making opportunities. Having modest risk limits is inconsistent with the ability to provide substantial immediacy to clients.

The resulting increase in investors' execution costs and loss of market liquidity would also cause issuers of securities to be harmed by lower prices. The fact that the Volcker Rule exempts U.S. government securities is a recognition by Congress that it would harm the U.S. government as an issuer if it were to apply the Rule to its own debt issues. The Bank of Japan and Japanese Financial Services Agency have written[2] to the Agencies about their concern "that the proposed Restrictions would have an adverse impact on Japanese Government Bonds (JGBs) trading. They would raise the operational and transactional costs of trading

[2]See the letter of Masamichi Kono, Vice Commissioner for International Affairs Financial Services Agency, Government of Japan, and Kenzo Yamamoto, Executive Director Bank of Japan, dated December 28, 2011. The Canadian government has written to the Agencies with a related concern about the impact of the proposed restrictions on the liquidity of non-U.S. government bonds. See the letter of Julie Dickson, Superintendant, Office of the Superintendant of Financial Institutions, Government of Canada, December 28, 2011.

in JGBs and could lead to the exit from Tokyo of Japanese subsidiaries of US banks. Some of the Japanese banks might be forced to cease or dramatically reduce their US operations. Those reactions could further adversely affect liquidity and pricing of the JGBs. We could also see the same picture in sovereign bond markets worldwide at this critical juncture. We would appreciate your expanding the range of exempted securities substantially, to include JGBs." The Agencies' proposed restrictions would likewise adversely affect U.S. corporations and home buyers who, like the United States and foreign governments, benefit from liquid capital markets through lower interest expense. If investors anticipate a secondary market with higher execution costs and delays due to a lack of market making capacity, along with higher price volatility, then they will demand higher bond yields on new issues. The markets for U.S. corporate bonds and non-agency mortgage-related securities are particularly important examples of markets that would be harmed by the proposed rule. Corporations would likewise face a higher cost of capital due to lower liquidity in the secondary market for their common shares.

Although treasury, agency, and some types of municipal debt securities are exempted, the proposed rule would reduce the liquidity of markets for interest rate swaps and other derivatives used to hedge these securities. Thus, the rule could somewhat elevate government borrowing costs.

The proposed rule would also hamper efficient price discovery, lowering the quality of information about economic fundamentals that is revealed by markets. For example, during the financial crisis of 2007-2009, the reduced market making capacity of major dealer banks caused by their insufficient capital levels resulted in dramatic downward distortions in corporate bond prices.

In the long term, the proposed disincentives for market making by U.S. banks would probably lead to a significant migration of market making and investment activities. Some of these activities could move outside of the United States. Within the U.S., the proposed rule could spur the emergence of large non-bank broker dealers. For example, the proposed rule may lead some current banks whose business models depend heavily on market making to give up their banking charters. Given the difficulty of competing when subject to the proposed market making rules, other large banks could choose to spin off their market making businesses.

Some of the lost market-making capacity might be filled by existing non-bank firms such as hedge funds or insurance companies. Insurance firms might not, under the proposed rule, be significantly constrained in their effective market-making activities. Insurance firms fall under a system of regulatory transparency, capital, and liquidity requirements which is not designed to treat market making risk. Hedge funds have extremely limited regulatory oversight. Some market making could be replaced by a new form of brokerage conducted

by large asset-management firms. For example, an investor who wishes to enter or exit a position could notify the associated trading desk of a large asset-management firm. By a prior contractual arrangement with the clients of the asset-management firm, that trading desk could have been given the discretion to temporarily adjust the clients' portfolios within specified asset-allocation bands so as to accommodate the desired trade.

These outcomes seem inconsistent with congressional intent, and have unpredictable and potentially adverse consequences for the safety and soundness of our financial system. Leading up to the financial crisis of 2007-2009, the United States was unique in having several of the world's largest broker-dealers outside of its regulated banking sector. The failure of some of these and near failure of others dramatically exacerbated that crisis. By spurring a somewhat unpredictable transition to non-bank dealers, the proposed rule could reduce financial stability. This concern is reduced somewhat by the prospect that large non-bank dealers will be designated as systemically important by the Financial Stability Oversight Council. Access to the liquidity support of the central bank, however, is more cumbersome to arrange for non-banks, especially given the Dodd-Frank prohibition of emergency liquidity provision by the Federal Reserve to individual non-banks. Further, Basel III liquidity and capital requirements do not apply to non-bank broker dealers.

Thus, it is premature at best to assume that non-bank market makers will have regulatory supervision, access to liquidity, and capital and liquidity requirements that are as effective as those for regulated banks. The failure or sudden loss of capacity of a large broker dealer is at least as adverse for the economy as the failure of a similarly large financial institution devoted to conventional lending and deposit taking. I believe the costs and benefits of the potential migration of market making services to non-banks should be carefully considered by the Agencies before their rules are finalized.

The proposed rule would directly discourage the discretion of market makers to efficiently absorb significant risks from their clients through the provision of immediacy. As a consequence, the rule would also reduce the allocation of capital to market making businesses. These direct and indirect effects would increase trading costs for investors, reduce the resiliency of markets, reduce the quality of information revealed through security prices, and increase the interest expense and capital-raising costs of corporations, individuals, and others. These outcomes would lead to somewhat lower expected economic growth. The migration of a significant amount of market making outside of the regulated banking sector was not intended by Congress, would be likely under the proposed rule, and has potential adverse consequences for systemic risk.

This report is not a comprehensive analysis of the proposed rule. Rather, my objective is to focus on some key principles. I do not propose alternative metrics for detecting "risky market making." Although some forms of trading that clearly serve no market making intent

can be proscribed, an attempt to separate "legitimate and acceptable" market making from "speculative and risky" market making is not productive, in my opinion. The objective should be to ensure that market makers clearly have abundant capital and liquidity to cover the risks they take.

The next section of this report describes how and why market makers provide immediacy, and illustrates the adverse price distortions that can be caused by a limited supply of immediacy. In the following section, I discuss the impact of the proposed rules on the ability or incentives of market makers to provide immediacy, and the likely negative consequences. Finally, after a concluding section, I raise and respond to some questions that may be raised by this report.

3 The Provision of Immediacy by Market Makers

As opposed to a broker, who merely matches buyers and sellers, a market maker itself buys and sells assets, placing its own capital at risk. The service that it provides is "immediacy," the ability to immediately absorb a client's demand or supply of an asset into its own inventory. At any given point in time, the set of other investors who would in principle be prepared to bid competitively for the client's trade is not generally known or directly accessible to the client. The client could conduct an auction or a search for another suitable counterparty, but this takes time. Even if interested counterparties could be quickly identified, they would not necessarily have the infrastructure or balance-sheet capacity required to quickly take the client's trade. The client is therefore often willing to offer a price concession to a market maker in order to trade immediately rather than suffer a delay that exposes the client to price risk. If the client wishes to liquidate a position for cash, it may also have an opportunity cost for delayed access to the cash.[3]

If the asset is traded on an exchange, the client could obtain some degree of immediacy from the exchange limit-order book, but with an adverse price impact that is increasing in the client's trade amount. A market maker can often handle large "block" trades with lower price impact than an exchange. The vast majority of transactions in over-the-counter (OTC) markets are with a market maker. The OTC market covers essentially all trade in bonds (corporate, municipal, U.S. government, and foreign sovereign bonds), loans, mortgage related securities, currencies, and commodities, and about 60% of the outstanding notional amount of derivatives.

When a market maker serves a client's demand for immediacy its inventory often moves away from a desired target level. If the inventory is abnormally high or low, the market

[3]For a supporting theoretical model, see Duffie, Gârleanu, and Pedersen (2005). A client may also seek immediacy from a market maker in order to avoid a broader release of information about its positions or trading intentions, which could harm its average execution price.

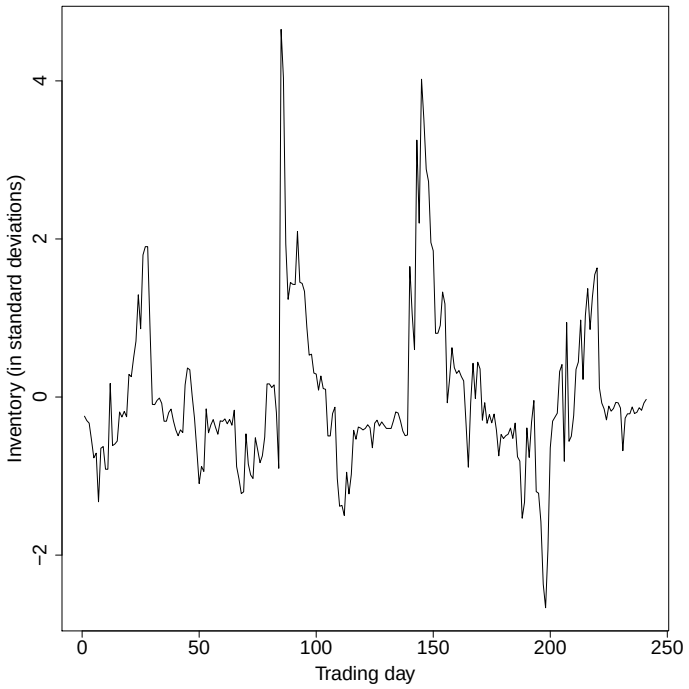

Figure 1: A plot of the inventory of the U.S.-dollar position of a block market making desk of a major broker-dealer for a single equity, Apple Inc., including effective positions implied by derivatives (on a "delta-equivalent" basis) and other effective exposures. The inventory levels are shown after scaling by the sample standard deviation of the dollar inventory levels for the sample period, a contiguous period of 2010-2011. Source: SIFMA-member data.

maker typically shifts its bid and ask quotes with the goal of moving its inventory back toward its target over time. The market maker may wish to accelerate the reduction of an inventory imbalance, lowering its risk, by requesting trades from others, including other market makers. Inventory risk management includes hedging with related financial instruments. In the meantime, the market maker continues to absorb supply and demand shocks from its clients. The general objective is to buy low and sell high, balancing the risk of loss against expected profit.

Demands for immediacy by customers can vary from moderate to extremely large, as illustrated in Figures 1 and 2, which were prepared by a major broker-dealer at the request of the author for the purpose of this report, based on the actual daily U.S.-dollar inventory[4] of common shares of Apple Incorporated held by that broker-dealer during a contiguous period of 2010-2011. Figure 1 shows the daily inventory[5] in units of sample standard deviations. Figure 2 is a frequency plot of unexpected shocks to inventory, showing the number of

[4]Derivatives are included on a "delta-equivalent" basis.

[5]The inventories shown include the effect of derivatives (on a "delta-equivalent" basis) and other effective exposures.

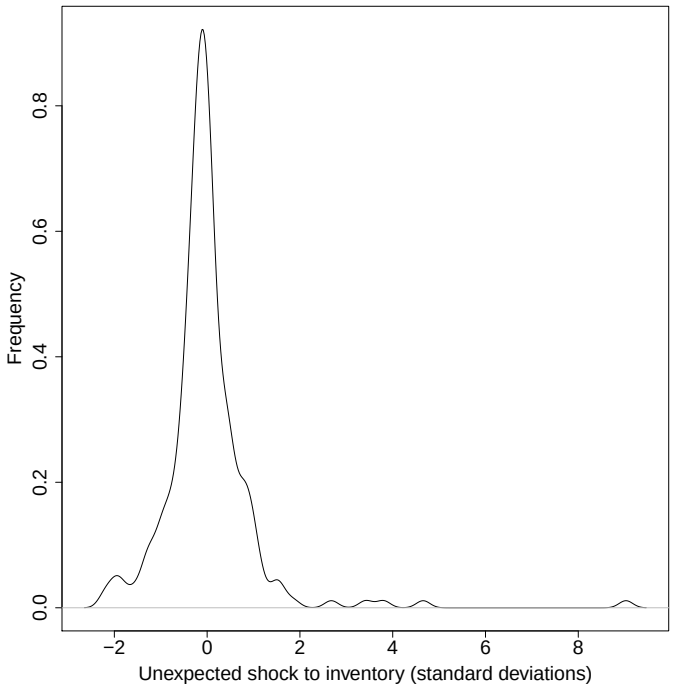

Figure 2: A frequency plot of unexpected shocks to the U.S.-dollar position of a block market making desk in the common shares of Apple Inc., including effective positions implied by derivatives and other effective exposures, based on the data shown in Figure 1. The shocks are scaled by their sample standard deviation. Source: SIFMA-member data.

standard deviations by which the inventory changed unexpectedly from one day to the next. These "shocks" are estimated using a simple statistical model,[6] which indicates that the market maker's inventory of this security is expected to revert approximately 20% of the way toward normal each day.[7] This implies a roughly 3-day "expected half-life" of inventory imbalances. Across other individual equities handled by the same market maker, the same statistical analysis shows that the expected half life of inventory imbalances is greatest for those equities with the highest bid-ask spreads and the lowest trading volume, as one would expect for a provider of immediacy.

Most market making done by large banks involves substantial granularity in both trade frequency and trade size. Particularly in fixed-income markets, trades are widely and unpredictably spaced in time, and sometimes are effectively "by appointment." For example, research by Goldstein, Hotchkiss, and Sirri (2007), Bao, Pan, and Wang (2011), and Chen,

[6]The autoregressive model $X_{t+1} = a + bX_t + Z_t$ was fit to the time series of inventory X_t on each trading day t during the sample period. The "persistence parameter" b is estimated at 0.80, with a standard error of 0.04. Figure 2 is a density plot of estimates of the "inventory surprise" Z_t, using kernel smoothing with a band width of 0.146. The Appendix provides a "QQ" plot of the quantiles of these shocks, more clearly indicating the "fat tails."

[7]Evidence of the targeting of inventory by market makers is abundant, beginning with the work of Amihud and Mendelson (1980).

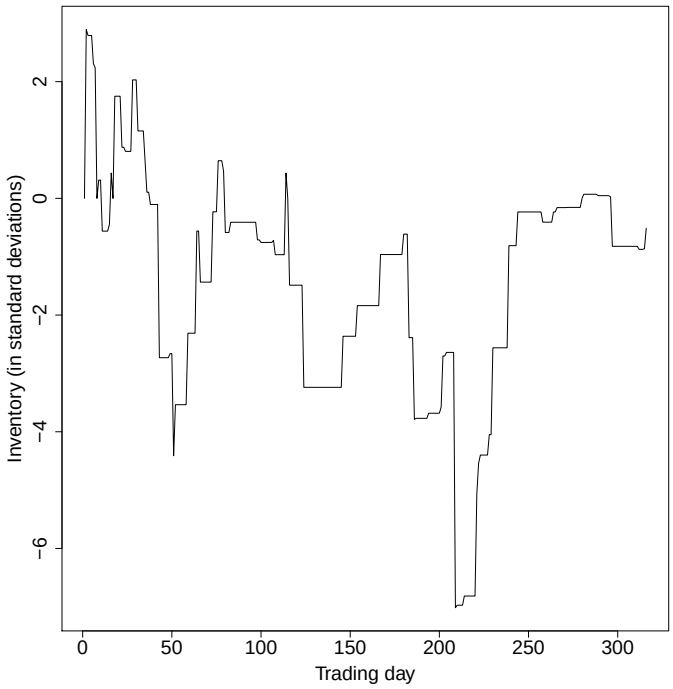

Figure 3: A plot of the inventory of the U.S.-dollar position of a market making desk of a major broker-dealer for a single investment-grade corporate bond. The inventory levels are shown after scaling by an estimate of the sample standard deviation of the dollar inventory levels for the sample period, a contiguous period of 2010-2011. Source: SIFMA-member data.

Fleming, Jackson, Li, and Sarkar (2011) shows that trades in individual U.S. corporate bonds or individual corporate credit default swaps typically occur a few times per day at most, in total across the entire market.[8]

Figure 3 shows the market making position in a particular investment-grade corporate bond for the broker-dealer that provided the data for Figure 1. During the illustrated time period, the market maker facilitated significant client sales that caused the market maker's inventory to become negative (that is, the market maker was "short"). As illustrated, the market maker targeted reductions in the resulting inventory imbalances between these client-sale events, subject to the constraints of illiquidity and continuing to provide immediacy. Because demands for immediacy in individual corporate bonds are sparsely spaced in time, as illustrated by the "step-like" inventory path shown in Figure 3, and because of the rel-

[8]For the sample of BBB-rated corporate bonds studied by Goldstein, Hotchkiss, and Sirri (2007), the fraction of days on which a given bond was traded was 26.9%, on average across bonds. The sample of more actively traded bonds studied by Bao, Pan, and Wang (2011) were traded on average 174 times per month, in total across all market makers. For the credit default swap study of Chen, Fleming, Jackson, Li, and Sarkar (2011), "The 48 actively traded corporate reference entities traded an average of 10 times daily, with the top reference entity trading an average of 22 times per day. Less actively traded reference entities traded on average 4 times daily and infrequently traded reference entities traded on average less than once per day. The actively traded sovereign reference entities traded on average 30 times daily; less actively traded sovereigns traded on average 15 times per day and infrequently traded sovereign contracts traded an average of 2 times daily."

ative illiquidity of the corporate bond market in other respects, the expected half life of inventory imbalances in a corporate bond is typically much longer than those for equities. For the illustrated corporate bond, the expected half life of inventory shocks is estimated at approximately two weeks, which is typical of the cross section of investment-grade corporate bonds handled by this broker-dealer.[9]

In general, a market maker's target inventory level and preferred rate of reversion of inventory levels toward the target vary with the asset type, current market conditions, and the level of capital that the market maker currently allocates to the associated trading desk. Whenever the market maker has limited capacity to warehouse risk on its balance sheet, its target inventory level is low, and it avoids requests for immediacy from clients that would move its inventory far from the target inventory level. The lower is the market maker's tolerance for risk, the less capacity it has to absorb supply and demand imbalances from the market, and the more it may demand immediacy for itself from other investors. Given the size and volatility of modern financial markets, market liquidity relies on the presence of highly capitalized market makers.

In compensation for bearing the risk that it will suffer a loss on its inventory due to unforeseen changes in fundamental or market conditions, or due to trades with a particularly well informed client, a market maker requires an expected return. Absent this compensation, it would be irrational for the market maker to supply immediacy to the client. The greater the inventory risk relative to the capital or risk limits allocated to the market making desk, the greater is the required expected return, other things equal.[10] A market maker's bids and offers apply to trade sizes up to a moderate and conventional "round-lot" amount, which varies by asset type. For clients who wish to trade a larger amount, a price and quantity negotiation is likely to result in a trade for an amount less than that desired by the client, or a larger price concession to the market maker for taking additional risk, or no trade. Even moderate-sized trades may require a larger-than-normal expected return to the market maker if they threaten to increase an imbalance in inventory that is already close to the market maker's risk limit for the asset type or broader asset class.

Because an astute market making trader is aware of changes in market conditions, he or she can often anticipate periods of time over which an imbalance in the demand for immediacy on one side of the market is likely to present an opportunity to profit by allowing inventory to diverge significantly from normal. The imbalance is later reduced over time through trades at prices that are expected to result in a net profit. This positioning of

[9]The estimated persistence coefficient of the autoregressive (AR1) model applied to weekly inventory data for the illustrated corporate bond is 0.73. The median of the weekly inventory persistence coefficients across all investment-grade corporate bonds in the firm's sample is 0.75. When estimated on a daily basis, the sample median of the estimated persistence coefficients is 0.938, which corresponds to roughly the same effective half life in weeks (because 0.938^5 is approximately 0.73).

[10]For supporting empirical evidence on the determination of federal fund loan rates, see Chapter 2 of Duffie (2012), based on research conducted for Ashcraft and Duffie (2007).

inventory to profit from expected changes in market prices is an essential aspect of market making that improves market liquidity and benefits market participants, as supported by considerable theoretical and empirical research.[11] If market makers were to refrain from absorbing supply and demand imbalances into their inventory in anticipation of likely price improvements, the price impacts suffered by those seeking immediacy would be deeper, and the corresponding distortions in prices would be larger and more persistent. Brunnermeier and Pedersen (2009) consider the adverse consequences on market liquidity of tightening a market maker's inventory risk limit. As Comerton-Forde, Hendershott, Jones, Moulton, and Seasholes (2010) explain and support with empirical evidence, "market makers face short-run limits on the amount of risk they can bear. As their inventory positions grow larger (in either direction, long or short), market makers become increasingly hesitant to take on more inventory, and quote accordingly. Similarly, losses from trading reduce market makers equity capital. If leverage ratios remain relatively constant, as suggested by the evidence in Adrian and Shin (2007), market makers' position limits decrease proportionately, which should similarly reduce market makers' willingness to provide liquidity."

Some of the supply and demand shocks absorbed by market makers are idiosyncratic, tied to investor-specific trading motives. Other supply or demand shocks are more episodic, related to market-wide events. As a motivating example, Figure 4 illustrates the average price impact of deletions of equities from the S&P 500 stock index, and the associated average price reversal over time. These deletions occur when the list of firms comprising the S&P500 index is adjusted. The underlying data, provided to me by Professor Jeremy Graveline, cover the period from December 1990 through July 2002, and include 61 such deletions. At these events, index-tracking investors are effectively forced to immediately sell large blocks of the deleted equities. Suppliers of liquidity including market makers were therefore offered substantial price concessions for absorbing the supply shocks into their own inventories of the equity. They hoped to subsequently profit by laying off their positions over time at higher prices.[12] While the illustrated average path of recovery in prices after

[11] Grossman and Miller (1988) provide a seminal model. Subsequent theoretical foundations have been provided by Weill (2007), Gromb and Vayanos (2002), He and Krishnamurthy (2009), Gromb and Vayanos (2010), Lagos, Rocheteau, and Weill (2009), Rinne and Suominen (2009), Brunnermeier and Pedersen (2009), and Duffie (2010a). Nagel (2009), Lou (2009), Rinne and Suominen (2010), and Bao, Pan, and Wang (2011) offer supporting evidence of return reversals due to price pressure. A wealth of empirical evidence of price surges and return reversals caused by specialist inventory imbalances has been provided by Andrade, Chang, and Seasholes (2005), Comerton-Forde, Hendershott, Jones, Moulton, and Seasholes (2010), Hendershott and Seasholes (2007), and Hendershott and Menkveld (2009).

[12] As reported by Chen, Noronha, and Singhal (2004) for a similar data set, deleted stocks suffered a loss of approximately 8% on the deletion announcement date and an additional loss of 6% between the announcement date and the effective deletion date. Quoting from Chen, Noronha, and Singhal (2004), who cite several studies that further support this remarkable price impact and reversal, "The negative effect of deletions disappears completely 60 days after the effective date. The cumulative abnormal return from announcement to 60 days after the effective date is not significantly negative, and always economically small." Related studies of price impacts and recoveries associated with index recompositions, including both debt and equity indices, include those of Shleifer (1986), Harris and Gurel (1986), Madhavan (2001), Greenwood (2005), Mitchell, Pulvino, and Stafford (2002), Wurgler and Zhuravskaya (2002), Kaul, Mehrotra, and Morck (2000), Chen, Lookman, Schürhoff, and Seppi (2009), and Feldhütter (2009). Petajisto (2009) provides a model in which the pressure is borne by intermediaries, and applies his model to explain the empirical evidence on index deletions.

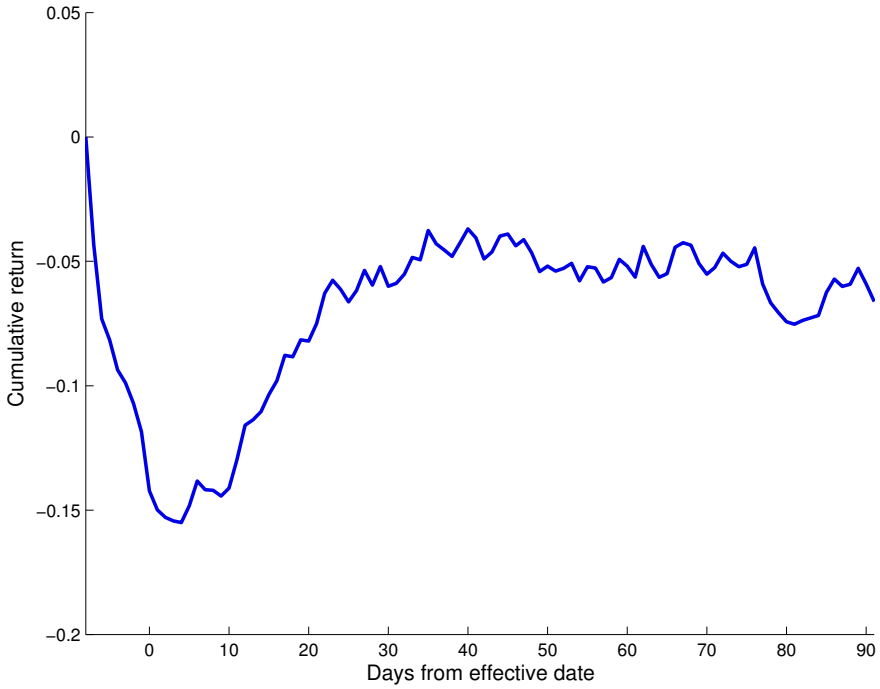

Figure 4: Average cumulative returns for deleted S&P 500 stocks, 1990-2002. The average number of days between the announcement and effective deletion dates is 7.56. The passage of time from announcement to deletion for each equity is re-scaled to 8 days before averaging the cumulative returns during this period across the equities. The original data provided by Jeremy Graveline were augmented by Haoxiang Zhu. Source: Duffie (2010a).

deletions represents a significant enticement to providers of immediacy on average, there was nevertheless substantial uncertainty regarding the profitability of supplying liquidity at any particular deletion event. Were market makers to stand back from the opportunity to offer immediacy to investors anxious to unload large quantities of the affected equity, the initial price impact of the supply shock would be greater and the time period over which the price distortion is expected to persist would be greater.[13]

As investors learn over time about trading opportunities presented by a specific type of supply shock such as an index recomposition, asset-management practices adjust and tend to reduce the cost of large demands for immediacy. The role of liquidity provision by market makers in the face of the particular type of supply or demand shock then declines. New forms of demand and supply shocks emerge, however, from changes in the institutional structure of markets and the macroeconomy, for which market makers are once again at the front line of liquidity provision. This is especially true in bond and OTC derivatives markets, where essentially all demands for immediacy are served by market makers.

As motivated by the last example, once a market maker has absorbed part of a large supply

[13]Duffie (2010a) provides a model of the impact on the expected price impact of a supply shock and the subsequent time pattern of price distortions associated, including the effect of reducing the risk tolerance or quantity of providers of immediacy.

shock into its inventory, it begins to lay off its position to other investors over time at higher anticipated prices. (The case of a demand shock is symmetric.) Immediacy-seeking investors will trade at the market maker's ask price. For these trades, the market maker hopes to profit from both the bid-ask spread and also from the expected recovery in price from the time at which the market maker first expanded its inventory. The price is expected to increase during this period because of the diminishing overhang of inventories held by suppliers of immediacy. The market maker may at the same time seek immediacy from other investors, including other market makers, in order to reduce its inventory in a prudently rapid manner. When it seeks immediacy from others to lower its excess inventory, the market maker expects to profit from any price recovery since the original supply shock, *less* the effective spread that it pays to its counterparties. A more passive approach of waiting to reduce its inventory over time exclusively through trades initiated by clients would expose the market maker to the additional risk associated with a more prolonged exposure to unexpected changes in price.

The incentive of a market maker to provide immediacy is increasing in the expected profit associated with *both* anticipated changes in market prices and from the net effect of bid-ask spreads (received net of paid).

As another illustration, Figure 5, from Kulak (2008), shows the average pattern of equity prices around the time of seasoned equity offerings. In this case, anticipation of the announced supply shock causes the price to decline, on average, as the issuance date approaches. During this period, market makers and other providers of liquidity generally wish to reduce their inventory below a normal target level in order to "make space" on their balance sheets for the anticipated new supply. Once suppliers of immediacy have absorbed the supply shock at a relatively deep average price concession, they lay off their inventory over time to other investors at an expected profit. The longer they are willing to hold inventory, the greater the expected profit, accompanied of course by an extended exposure to loss associated with unexpected fundamental news.[14] Market makers and underwriters are among the most important providers of liquidity.

Figure 6, provided to the author by Professor Honjun Yan, shows the impact of U.S. Treasury note auctions on the associated treasury yields. Note yields go up as the date of the anticipated new supply of treasuries approaches, and then recover in subsequent days. Fleming and Rosenberg (2007) show that Treasury dealers adjust their positions to absorb these issuance supply shocks. They describe how "dealers seem to be compensated for the risks associated with these inventory changes via price appreciation the subsequent week." The figure shows that the auction supply temporarily raises not only the yields of the security issued, but also those of the previously issued ("off the run") treasuries of the same maturity

[14]That secondary offerings are made at substantial price concessions has been documented by Mikkelson and Partch (1985). At least as early as the work of Scholes (1972), researchers have focused on the presence of temporary price impacts at secondary equity issuances. Additional empirical evidence is offered by Loughran and Ritter (1995), Chaisurote (2008), and Gao and Ritter (2010).

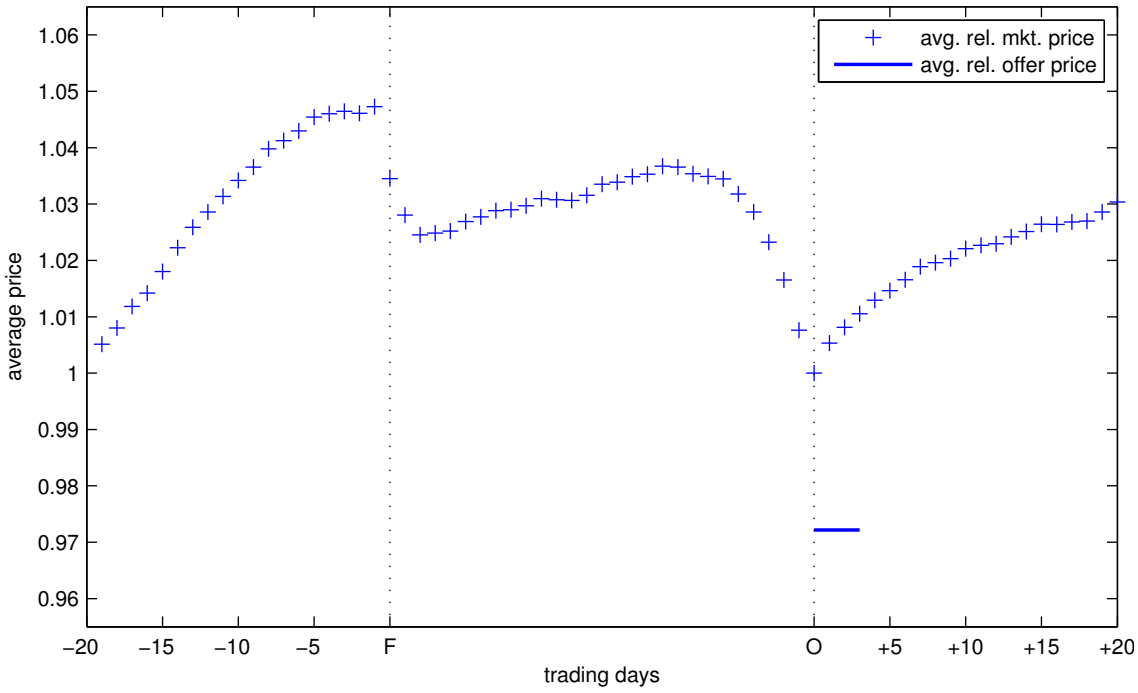

Figure 5: Average price dynamics around seasoned equity offerings. The figure, kindly supplied to the author by Jan Peter Kulak, covers 3850 U.S. industrial firms that undertook a firm-commitment public seasoned offering in the United States between 1986 and 2007. The plotted line shows the average across issuances of the ratio of secondary market price of the equity to the closing price of the equity on the offering date. Because offerings differ in the number of trading days between the filing announcement and the offering date, the times between filing and offering date are rescaled interpolated to the average across the sample of the number of trading days between the filing and the issuance date. Source: Kulak (2008), published in Duffie (2010a).

class, because their returns are highly correlated with those of the issued note. Although treasury securities are exempted from the proposed rule, the same principles apply to other markets, to an even greater degree given the high liquidity of treasury markets relative to other security markets.

For example, Figure 7, from Newman and Rierson (2003), shows the expected pattern of yield impacts around the time of a large corporate bond issuance. In this example, the illustrated impact is for corporate bonds of firms *other than the issuer*, that are in the same industry as the issuer, the European telecom industry. When a company in this sector scheduled a significant issuance of bonds during the period 1999-2001, the entire related market for European telecom bonds suffered from higher bond yields. The figure shows the estimated path of yield impacts on European telecom bonds, not including those of the issuer, Deutsche Telekom, associated with a particular 16-billion-Euro issuance. As for the case of treasury note issuances, yields increased as the issuance date approached, and then recovered toward normal. The degree to which the yields of corporate bonds are adversely affected

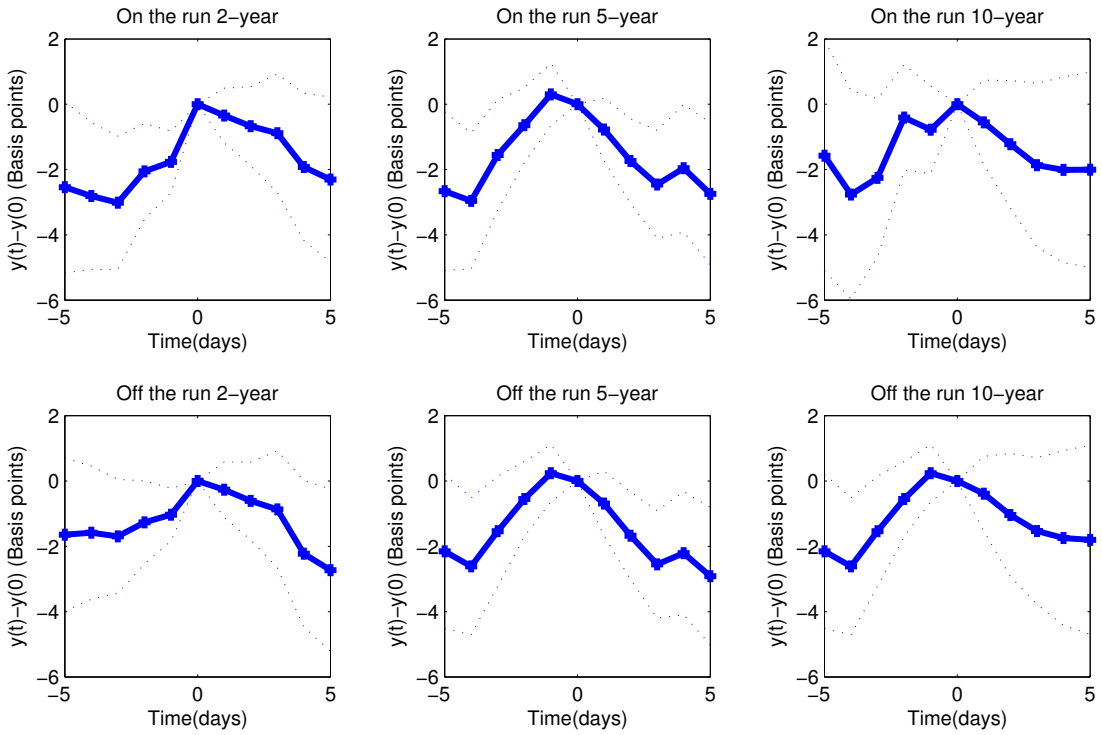

Figure 6: Yield elevation at the issuance of U.S. Treasuries, with 95% "confidence" bands. The figure, kindly provided to the author by Professor Honjun Yan, covers U.S. Treasury issuances from January 1980 to March 2008. Yields are based on averages of bid and ask prices obtained from CRSP. Auction dates are from the U.S. Treasury Department. The sample includes 332 2-year note auctions, 210 5-year note auctions, and 132 10-year note auctions. For each maturity, the differences between the yield on the issuance date and the yield on dates within 5 days of the issuance date are averaged across issuances, for both on-the-run and off-the-run notes. Source: Honjun Yan, published in Duffie (2010a).

by issuance shocks is greater than that for treasuries because the liquidity of the corporate bond market is lower by comparison, and because corporate bonds are riskier than treasuries, exposing suppliers of immediacy to greater inventory risk. If market makers were to lower their risk limits, or have inflexible risk limits in the face of market-wide supply shocks, the yield impacts of these and other supply shocks would be deeper and more persistent.[15]

Figure 8 illustrates the concept that, particularly in an over-the-counter market, the provision of immediacy is facilitated by a network of market makers and inter-dealer brokers. A market maker is able to provide immediacy more efficiently (at lower cost to clients and at lower risk to itself), through the opportunity to lay off positions with other market makers,

[15]Chen, Lookman, Schürhoff, and Seppi (2009) document the impact on the yields of corporate bonds in the automotive sector caused by the downgrade of General Motors in 2005. Because some institutional investors in corporate bonds are required to hold only investment-grade bonds, the prospect of a downgrade caused forced sales. Chen, Lookman, Schürhoff, and Seppi (2009) are able to demonstrate the impact of this supply shock, above and beyond the implications of the information related to the downgrade.

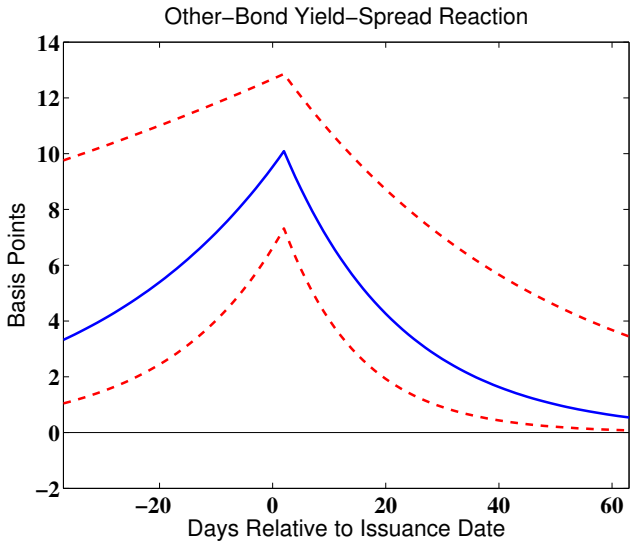

Figure 7: Capital immobility in the telecom debt market. The estimated impacts on the yields of European Telekom issuers, not including Deutsche Telekom, associated with a particular 16-billion-Euro issuance by Deutsche Telekom, using an econometric method explained by Newman and Rierson (2003). Source: Newman and Rierson (2003), published in Duffie (2010a).

who may be better aware of ultimate investors who are interested in trading in the opposite direction. This intermediation of immediacy occurs through direct dealer-to-dealer trades, or indirectly through inter-dealer brokers. Because of search and contracting frictions as well as the benefit of confidentiality in reducing price impacts for large trades, it is often inefficient for client investors to negotiate simultaneously and directly with a large number of market makers. It is even more costly for ultimate investors to conduct large trades, or trades in illiquid products, directly with ultimate investors. Instead, investors may request quotes from one or a subset of market makers.[16] These contacts can lead to a trade with a particular market maker, who may then wish to rebalance its inventory relatively quickly through the inter-dealer network. This is often more efficient for the market maker than requesting immediacy from another ultimate investor, or waiting for an ultimate investor who might wish to trade in the opposite direction. In effect, the inter-dealer network acts as a broader mechanism for transmitting supply and demand shocks from ultimate investors to ultimate investors.[17] Bech and Garratt (2003) provide strong evidence of the inter-dealer network effect in re-distributing supply and demand shocks in the federal funds market.

[16] Large institutional investors can initiate "requests for quotes" or "dealer runs," sometimes through swap execution facilities (SEFs). The cost of a sequential search, one market maker at a time, is analyzed by Zhu (2012).

[17] Concerns over the transparency and competitiveness of OTC markets remain, and have been partially addressed by recent requirements for price transparency in corporate bond markets, and by the Dodd-Frank requirements for transactions disclosure and the use of swap execution facilities in the standardized OTC derivatives market.

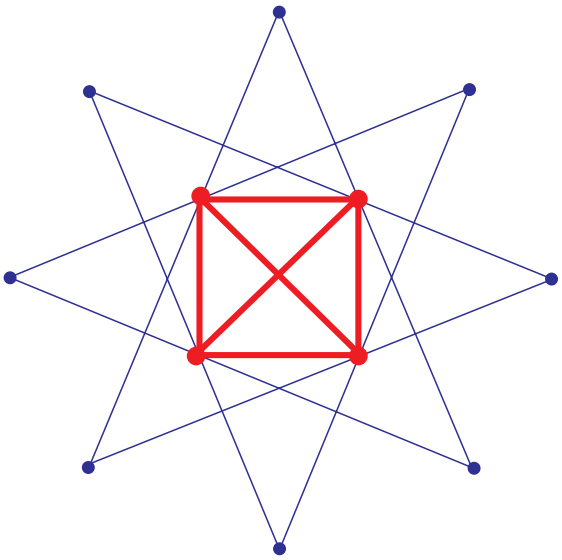

Figure 8: A schematic of an over-the-counter market with a "core" inter-dealer market in which market makers and inter-dealer brokers act as a network that collectively provides immediacy to ultimate investors.

4 Impact of the Proposed Rule on Investors and Issuers

The proposed rule would discourage the provision of immediacy by market makers, particularly through the threat of sanctions for significant increases in market making risk or for significant profits caused by price changes (as opposed to profits associated with a bid-ask spread revenues).

At page 94, the Agencies write that "Market making and related activities seek to generate profitability primarily by generating fees, commissions, spreads and other forms of customer revenue that are relatively, though not completely, insensitive to market fluctuations and generally result in a high level of revenue relative to risk over an appropriate time frame." This statement does not accurately characterize market making. The Agencies' explanation of their proposed rules clearly indicates the intention to use the various proposed risk and profit metrics to restrict market making activities to those consistent with this definition. For example, at page 94, immediately before this characterization of market making, one reads: "The Agencies expect that these realized-risk and revenue-relative-to-realized-risk measurements would provide information useful in assessing whether trading activities are producing revenues that are consistent, in terms of the degree of risk that is being assumed, with typical market making related activities." At page 92, the Agencies suggest they will use the proposed risk metrics to "to determine whether these activities involve prohibited proprietary trading because the trading activity either is inconsistent with permitted market making-related activities or presents a material exposure to high-risk assets or high-risk

trading strategies." At page 93: "Significant, abrupt or inconsistent changes to key risk management measures, such as VaR, that are inconsistent with prior experience, the experience of similarly situated trading units and managements stated expectations for such measures may indicate impermissible proprietary trading."

Were this approach to be reflected in the Agencies' final rule, the intent of Congress to exempt market making by banks would be thwarted and U.S. financial market liquidity would suffer, with the adverse consequences outlined in Section 2 of this report.

Under the proposed implementing rules, market makers would retain the ability and incentive to absorb only moderately sized demands for immediacy. It is precisely through their ability to service heightened demands for immediacy, however, that market makers mitigate the most significant associated price distortions and execution costs to investors. The ability of market makers to buffer unexpectedly large supply and demand imbalances depends on significant and flexible market making capacity and on the incentive to profit from expected price changes. Were the proposed rule to be implemented, market makers who absorb large demand and supply shocks into their inventories would experience a "deterioration" in the proposed metrics for their market-making risk, and the associated threat of regulatory sanction. They would also be less inclined to absorb the associated risks given the likely sanctions for significant profits from price changes. Further, under the proposed rules for trader compensation, market making traders would have significantly lower incentives to accept trades involving significant increases in risk or profit.

Under the proposed rule, imbalances in the demand or supply of immediacy would therefore cause larger and more persistent distortions in market prices. Price discovery would suffer. Home owners, businesses, and some municipalities would face higher borrowing costs. Firms would face higher costs for raising new capital. These increased costs would occur directly in the form of higher price impacts at the point of financing, and indirectly from the lower appetite of investors to own securities that would trade in thinner and more volatile secondary markets.

In addition to the research that I have already cited, there is significant empirical evidence that a limited risk-taking capacity of marker makers leads to price distortions.[18] As a relatively extreme but illustrative example, Mitchell and Pulvino (2009) describe a dramatic distortion in corporate bond yields that arose during the financial crisis due to an insufficient risk-taking capacity of market makers. As shown in Figure 9, corporate bond yields were elevated well above those implied by credit default swap (CDS) rates.[19] The difference

[18]For example, Meli (2004) found evidence that changes in dealer capital are strongly related to changes in swap spreads (the difference between swap rates and treasury rates). Etula (2009) describes how variation over time in broker-dealer assets is significantly correlated with crude oil returns. Further evidence on the relationship between dealer risk-bearing capacity and distortions in risk premia is provided by Adrian, Etula, and Shin (2009) and Adrian, Moench, and Shin (2011).

[19]In a frictionless market, the CDS rate is, within a small tolerance for technical contract differences, equal to the yield spread on a par bond of the maturity of the CDS of the same issuer, that is, the bond yield less the associated risk-free yield. If, for example, the basis for a particular corporate bond becomes negative, as illustrated in Figure 9, one could short a risk-free

between the CDS-implied bond yield and the actual bond yield is known as the "basis." The exceptional CDS basis violations that appeared during the financial crisis across broad portfolios of investment-grade and high-yield bonds were due to the extremely low levels of capital of dealer banks.[20] Investment-grade corporations issuing bonds in late 2008 and early 2009 had to pay roughly 2% higher interest rates due to this market inefficiency. For lower rated firms, as illustrated, the distortion in borrowing rates would have been far greater, for any that actually attempted to issue bonds during this period. As large dealers regained some balance-sheet capacity, the CDS basis went back toward normal, as illustrated.

Musto, Nini, and Schwarz (2011) show that even U.S. treasury prices were severely distorted at the height of the financial crisis through a loss of market liquidity. Particularly around December 2008, portfolios of treasuries promising equivalent cash flows were often trading at substantial price differences. The cornerstone of treasury market liquidity is the market making desks of primary dealers. Although U.S. treasuries are exempted from the Volcker Rule, many important classes of securities, that already trade in less liquid markets than those for U.S. treasuries, will be affected. As mentioned in Section 2, foreign governments have asked that their bonds also be exempted.

The incentive and discretion to supply immediacy by taking extra risk in light of extra expected profit is also important at the level of an individual trader on a market-making desk. The proposed rule world lead the compensation of market-making traders to be more like that of flow-based brokerage agents. Coupled with the reputational risk of exceeding likely regulatory norms for "low-risk market making" that would arise from the proposed metrics, a market making trader would often avoid taking the discretion needed to meet a customer's demand for immediacy. Under the proposed rule, a trader would frequently fail to offer two-sided markets for significant quantities at efficient prices. For example, the proposed rule would encourage a trader faced with the extra risk of taking a large position to quote prices for only a limited fraction of the customer's desired amount. When the efficient approach to a trade enquiry with extra risk is a widening of the bid-ask spread, especially when facing a well informed client, the proposed metrics would discourage the trader from taking the position at all, or encourage the trader to take the position at a small expected profit relative to the risk of loss, out of fear of drawing attention to himself or herself over trades that adversely affect the regulatory metrics of the proposed rule. Indeed, one of the

bond, invest the proceeds in the corporate bond, and buy default protection on the corporate bond with a credit default swap. Putting aside some technical issues and ignoring counterparty risk, the net income of this strategy per year, at no net initial investment, is the principal debt position multiplied by the absolute magnitude of the basis. If the basis becomes negative, the opposite trade is likewise highly profitable, although holding a short position in corporate bonds is somewhat cumbersome and can involve extra costs or risks. Institutional details can cause the basis to diverge somewhat from zero. See Duffie (1999). The CDS basis can also be elevated by counterparty risk, although this effect is tiny by comparison with the basis shown in Figure9.

[20]Exploiting the CDS basis "arbitrage" calls for a substantial amount of balance-sheet capacity at dealer banks, both to make markets in the underlying bond (which calls for finding or holding the underlying bonds) and to handle two CDS counterparty positions, one with the arbitrageur and one with a counterparty taking the opposite position. Exacerbating the capital shortage of dealers, the amount of capital necessary to hold corporate bonds increased because of an increase in the "haircut" applied to finance corporate bonds in the repo markets, as explained by Mitchell and Pulvino (2009).

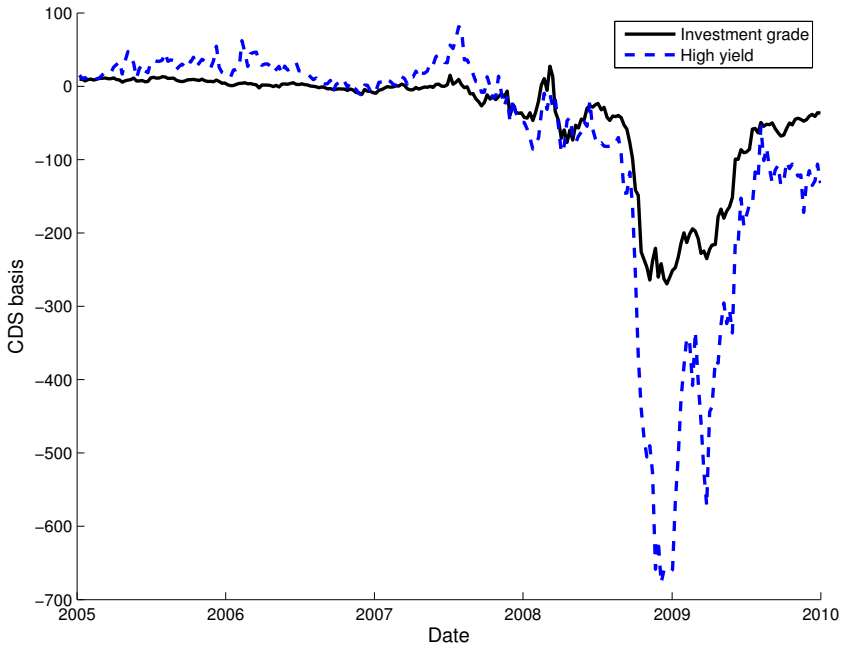

Figure 9: Average basis of U.S. corporate bond portfolios. The CDS basis for a given bond is the difference between the yield spread of a bond that is implied by the associated credit default swap (CDS) rate and the actual bond yield spread. The CDS basis is near zero in frictionless markets. As shown, the average CDS basis across portfolios of U.S. investment-grade bonds and high-yield bonds widened dramatically during the financial crisis and then narrowed as the crisis subsided. The underlying data, kindly provided to the author by Mark Mitchell and Todd Pulvino, cover an average of 484 investment-grades issuers per week and 208 high-yield issuers per week. Source: Mitchell and Pulvino (2010)., published in Duffie (2010a).

proposed metrics seems to suggest that trades should not be unduly profitable, relative to what they would be at historically normal bid-ask spreads. In the event that a trade turns out to be "overly profitable" because of an unexpectedly favorable price change, would a trader then have an incentive to incur an offsetting loss in order to avoid scrutiny? Similarly, in the face of a likely market-wide imbalance of supply or demand, a market making trader should have the discretion and incentive to significantly reposition his or her firm's inventory in order to absorb some of the supply imbalances. The proposed rule, including its compensation norms, would reduce the trader's discretion and incentive to do so, exacerbating the adverse consequences that I have described.

A trader's incentives for undue risk taking can be held in check by vesting incentive-based compensation over a substantial period of time. Pending compensation can thus be forfeited if a trader's negligence causes substantial losses or if his or her employer fails. The pool of pending compensation is thus effectively contributing to the capital of the firm, consistent with a recommendation of the Squam Lake Group.[21]

[21] See *The Squam Lake Report: Fixing the Financial System*, Princeton University Press, 2010. I am one of 15 authors.

5 Concluding Remarks

Section 13 of BHC Act (Section 619 of the Dodd-Frank Act) exempts market making from its proprietary trading restrictions on banks "to the extent that any such activities permitted by this subparagraph [including "market making related activities"] are designed not to exceed the reasonably expected near term demands of clients, customers, or counterparties." From the viewpoint of impact on market participants, including ultimate investors and those seeking to raise capital and finance themselves, I believe the Agencies' interpretation of this language is overly narrow and would cause undue costs to the economy. The Agencies did not provide a cost-benefit analysis that suggests otherwise. The potential for systemic risk and costs to the Deposit Insurance Fund associated with market making by banks can be treated more effectively through regulatory capital and liquidity requirements. In any case, if implemented, the proposed rule could inadvertently *increase* systemic risk because of a migration of market making activities to outside of the regulated banking sector, as I have outlined in Section 2.

Capital and liquidity requirements are a more direct and effective means of handling the legislated exemption for market making. The proposed restrictions on market making instead attempt to identify and eliminate specific patterns of trading. This attempt to disentangle those trades that have market making intent from those that do not is likely to be effective only in reducing the capacity of market making services provided by banks. Capital and liquidity requirements directly consider the soundness of a financial institution and its potential for causing systemic risk and costs to the Deposit Insurance Fund. In the case of market making, capital requirements treat risk on a portfolio-wide basis, an appropriate approach.

Leading up to the financial crisis of 2007-2009, the regulatory capital and liquidity requirements of financial institutions were clearly insufficient. These requirements should continue to be strengthened as deemed appropriate by regulators to robustly protect the Deposit Insurance Fund and the soundness of the financial system. An alternative to heightened capital and liquidity requirements could be some form of "ring-fencing" requirement that allows separately capitalized bankruptcy-remote market-making affiliates, an approach under adoption in the United Kingdom. This approach is significantly less efficient from the perspective of risk diversification, although generally consistent with the primary legislative motive of insulating banks from proprietary trading risks. In any case, whether market making is conducted by banks or others, market makers should be required to meet robust capital and liquidity requirements. A crucial point is that the market making and other risks taken by a financial institution are unsafe precisely when they are large relative to the institution's capital and liquidity buffers.

6 Additional Questions and Answers

I offer some questions that may be raised by my report, and responses.

1. *If significant market making activities are permitted, wouldn't banks be in a position to conduct proprietary trading that has no market making intent?* The legislated exemption for market making creates an unfortunate moral hazard that cannot be cured by the Agencies' rule writing. Some forms of proprietary trading that are clearly unrelated to market making can be identified and proscribed. Even with the Agencies' proposed restrictions, however, there will remain an incentive and ability to disguise as "exempted market making" certain forms of speculative trading that do not serve an ultimate objective of providing market making services to clients. As the Agencies recognize, effective market making involves some trades that are similar or identical to trades that would be conducted without market-making intent.[22] Intent is difficult to measure and therefore to regulate. The proposed rule attempts to do so with the use of criteria that "are intended to ensure that the banking entity is engaged in bona fide market making." I expect that this intent would be not achieved. Instead, an application of the proposed criteria would lead to *less* market making.

2. *Hasn't the financial crisis shown us that derivatives trading by large banks is an important source of systemic risk?* The Dodd-Frank Act addresses systemic risk in the market for OTC derivatives by heightened requirements for collateral, a requirement for the central clearing of standardized products, requirements for post-trade price transparency, and the requirement to trade standardized derivatives in swap execution facilities. Strong collateral standards and effective clearing will lower counterparty risk. All of these requirements are likely to reduce the degree of concentration of market making among a small set of systemically important banks. The Basel III accord substantially increases the capital and liquidity requirements associated with OTC derivatives. These measures therefore significantly alter the cost-benefit tradeoffs to be considered when implementing the Volcker Rule. In any case, further improvements in the cost-benefit tradeoff associated with market making risk are more efficiently achieved through further improvements in capital and liquidity requirements, wherever deemed appropriate by regulators, than by the proposed rule.

3. *Are the Basel III regulatory capital and liquidity requirements associated with market making sufficient?* This is a subject for more study. The Basel Committee on Banking

[22] At page 53, the Agencies write: "In particular, it may be difficult to determine whether principal risk has been retained because (i) the retention of such risk is necessary to provide intermediation and liquidity services for a relevant financial instrument or (ii) the position is part of a speculative trading strategy designed to realize profits from price movements in retained principal risk."

Supervision (2011) is currently conducting a "fundamental review" of capital requirements for the trading books of regulated banks. Their results are to be released in 2012. It makes sense for the Agencies' to adopt a conservative approach from the viewpoint of safety and soundness of the financial system, and to "harmonize" capital and liquidity requirements across regulatory jurisdictions so as to avoid a significant incentive for market making to migrate or to "morph" unsafely.

4. *Don't higher capital requirements lower the incentives of banks to provide banking services?* Higher capital requirements are costly to current shareholders because they lower the value of the limited-liability option held by equity owners. This leads to a rational reluctance by banks to raise capital even in some cases for which additional capital would significantly reduce distress costs, a problem known as "debt overhang." (See Chapter 4 of Duffie (2010b).) Relatively few banking activities that are profitable at low capital levels would would cease to be profitable at higher capital levels, at least across the range of capital requirements that are likely to be considered. I have not seen any reliable evidence or a conceptual foundation for the contrary view. The reduced return on equity of a banking activity implied by higher equity levels does not itself change the set of profitable banking activities, a point explained in detail by Admati, DeMarzo, Hellwig, and Pfleiderer (2011).[23] There is an exception, to the extent that a bank is "too big to fail." In this case, a higher capital requirement also reduces the effective government subsidy to the bank associated with lower debt financing rates charged to the bank by creditors who consider the likelihood of government support in lowering their expected default losses. A reduction of this effective subsidy through higher capital requirements would reduce the set of profitable investments by a bank, including some of those associated with lending and market making. I have not considered the impact of higher capital requirements through the potential loss of this subsidy. Leading up to the financial crisis of 2007-2009, it seems apparent that regulatory capital and liquidity requirements were not effective, and that many of the largest U.S. financial institutions were not well supervised. This could be viewed as an argument against the effectiveness of capital and liquidity requirements, and therefore in favor of reducing market making risk by other means, such as the proposed implementation of the Volcker Rule. In my view, the failure of capital and liquidity requirements to be effective in the financial crisis of 2007-2009 can be corrected. The Basel III requirements are an example of that.

5. *Isn't it true that the losses incurred by banks through market making have been responsible for past banking crises?* No. Most banking crises are caused by losses that banks incur through loan defaults, as explained by Reinhart and Rogoff (2009). Losses due

[23] Bolton and Sanama (2010) describes why "contingent capital" may be a relatively cost effective approach to meeting capital requirements.

to borrower defaults on conventional banking activities, such as loans to sovereigns, mortgages, and loans to commercial real estate projects, tend to be far greater in magnitude than losses on market making. This was certainly true in the financial crisis of 2007-2009. That crisis was nevertheless exacerbated by the proprietary trading losses of some large broker dealers, particularly Bear Stearns, Lehman Bothers, Merrill Lynch, and the broker-dealer affiliates of Citibank and some foreign banks.[24] Although I have not seen a systematic study of the available data, most of the largest trading losses seem to have been associated with forms of proprietary trading that are not market making or otherwise exempted by section 13 of the BHC Act. (The case of Bear Stearns may be an exception.) According to the United States Government Accountability Office (2011), trading losses during the last financial crisis were relatively small for the largest bank holding companies, including market making and all other proprietary trading-related gains or losses. Figure 10 shows total industry securities trading gains and losses from 2007 to 2011, breaking out those for the largest dealers.[25] I am not aware of reliable data bearing on the market-making component of these total trading gains and losses. Market making risks make relatively high demands on a bank's liquidity, in proportion to assets, because of contractual margin and collateral requirements, the potential adverse effects of fire sales and other market dislocations, and the need for a market maker to continue to offer clients immediacy, including through trades that drain cash from the market maker. A market maker that refuses to provide significant liquidity to clients risks signaling its financial weakness, which would likely exacerbate its own liquidity position by creating an incentive for creditors, counterparties, and clients to further withdraw effective financing. (See Duffie (2010b), Chapters 2 and 3.)

6. *Don't the proposed risk metrics provide useful additional information to the Agencies for supervising the market making risks of regulated banks?* Yes. Some of the proposed metrics, such as the "Risk and Position Limits" metric, VaR, Stress VaR, or Risk Factor Sensitivities, would provide useful supervisory information, especially if they are measured effectively and for a carefully considered menu of asset classes. The United States Government Accountability Office (2011) points out that the largest six bank holding companies had proprietary trading losses that frequently exceeded their VaR estimates, more frequently than consistent with an effective risk measure. The design and supervision of these risk measures should be revisited, given that they are used for

[24] The significant losses of the Royal Bank of Scotland in credit trading are reviewed in Section 4.1 of the report on the failure of RBS of the Financial Services Authority (2011).

[25] As of the fourth quarter of 2008, the "Major Firms" are BANC OF AMERICA SECURITIES LLC, BARCLAYS CAPITAL INC., CITIGROUP GLOBAL MARKETS INC., CREDIT SUISSE SECURITIES (USA) LLC, DEUTSCHE BANK SECURITIES INC., GOLDMAN, SACHS & CO., J.P. MORGAN SECURITIES INC., MERRILL LYNCH, PIERCE, FENNER & SMITH INCORPORATED, MORGAN STANLEY & CO. INCORPORATED, UBS FINANCIAL SERVICES INC., UBS SECURITIES LLC, and WACHOVIA SECURITIES, LLC. Since 2009, SIFMA does not report the individual names of the "top 10" firms.

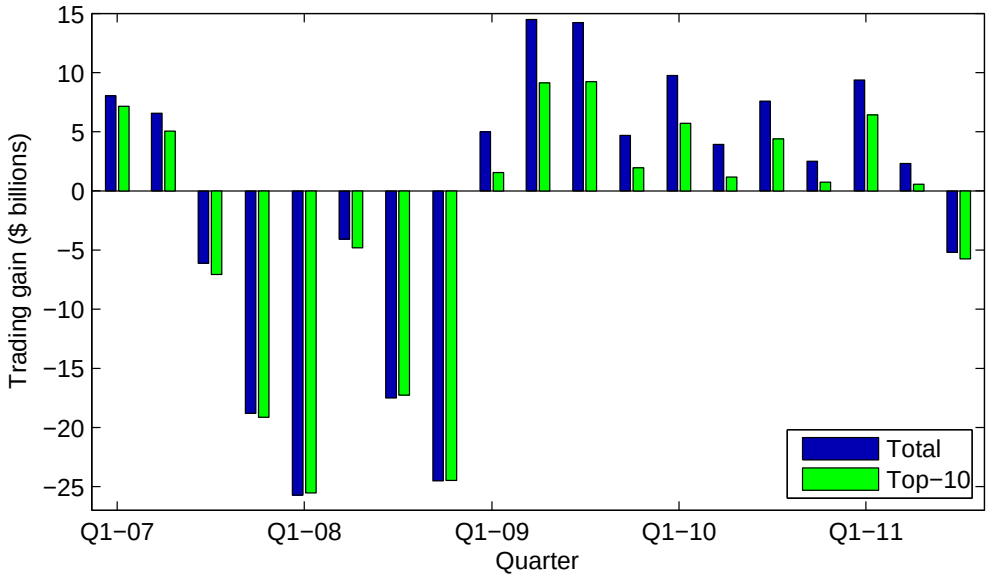

Figure 10: Quarterly trading gains and losses of US broker dealers, 2007-2011, in total, and for the largest dealers. FINRA defines these data as "realized and unrealized gains and losses on securities held for sale in the ordinary course of business (net of dividends and interest earned on such securities but not reduced by floor costs or taxes)." The data are from the SECs Financial and Operational Combined Uniform Single (FOCUS) Report regulatory filings, and cover the U.S. domestic operations of broker-dealer units doing a public business. Before 2009, the data shown here for the "Top 10" are instead reported by SIFMA for "major firms," which are sometimes 12 or 13 in number. Since 2009, SIFMA provides data for the "Top 10" without reporting the individual firm names comprising these top 10 firms. Data source: SIFMA DataBank.

supervisory purposes and also for determining capital requirements. I also suggest the use of counterparty risk exposure measures, not only to the risk of counterparty default but also to potential gains and losses to major counterparties for each of a specified list of systemically important scenarios. These measures should cover both exposure to changes in market value and also exposure to cash flows. The collection and use by regulators of these and other risk measures for supervisory purposes, if done broadly across bank and non-bank financial firms, could improve the ability of regulators to detect and mitigate risks to individual institutions and to the financial system as a whole. The collection and use of these and similar metrics is already authorized under existing broad supervisory mandates of the Agencies, including those applicable to banks, registered broker dealers, and non-bank financial firms that will be designated by the Financial Stability Oversight Council as systemically important.

7. *Wouldn't it be prudent to lower the risk to the economy associated with bank failures by forcing banks to stop making markets?* Congress concluded otherwise by exempting market making from the Volcker Rule. I believe that Congress got this right. Although

separating market making from traditional banking would make banks less complex and thus simpler for regulators to supervise, systemic risk could nevertheless rise. Large broker dealers would be outside of the regime of Basel III capital and liquidity requirements, with a different supervisory regime and with reduced access to lender-of-last-resort liquidity from the central bank. As demonstrated during the financial crisis of 2007-2009 and in the current Eurozone crisis, access to central bank liquidity can be crucial in mitigating the damage caused by a financial crisis. If there is an argument in favor of separation of market makers from conventional regulated banks, it would be more easily based instead on the view that systemically crucial market-making services offered by banks could suddenly be impaired when a bank suffers large losses on its conventional lending. (The current situation in the Eurozone includes this risk.) This argument is in my view trumped by the potential systemic risk posed by the migration of market making outside of the regulated banking environment.

A Technical Annex

Figure 11, based on the same market making inventory data for a single equity shown in Figure 1, illustrates the fact that unexpected shocks to inventory are "fat tailed," meaning that there are the inventory sometimes increases or drops dramatically.

References

Admati, A., P. M. DeMarzo, M. F. Hellwig, and P. Pfleiderer (2011). Fallacies, Irrelevant Facts, and Myths in the Discussion of Capital Regulation: Why Bank Equity is Not Expensive. Working Paper, Graduate School of Business, Stanford University.

Adrian, T., E. Etula, and H. S. Shin (2009). Risk Appetite and Exchange Rates. Working paper, Federal Reserve Bank of New York.

Adrian, T., E. Moench, and H. S. Shin (2011). Macro Risk Premium and Intermediary Balance Sheet Quantities. *IMF Economic Review 58*, 179–207.

Amihud, Y. and H. Mendelson (1980). Dealership Market: Market Making with Inventory. *Journal of Financial Economics 8*, 31–53.

Andrade, G., C. Chang, and M. Seasholes (2005). Predictable Reversals, Cross-Stock Effects, and the Limits of Arbitrage. Working Paper, Haas School of Business, University of California.

Ashcraft, A. and D. Duffie (2007). Systemic Dynamics in the Federal Funds Market. *American Economic Review, Papers and Proceedings 97*, 221–225.

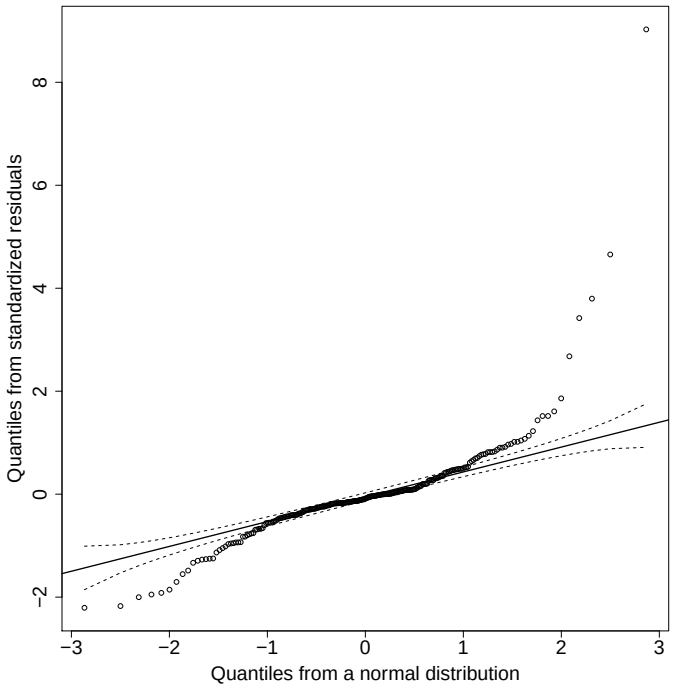

Figure 11: A "QQ" plot of quantiles of the distribution of unexpected shocks to the U.S.-dollar position in a Apple Inc. common shares of a block market making desk of a major broker-dealer, relative to the quantiles of the standard normal distribution. The quantile outcomes outside of the dashed-line "error" bands indicate a statistically significantly divergence from normality. The raw data include effective positions implied by derivatives and other effective exposures, and are those for the plot shown in Figure 1. The shocks are scaled by their sample standard deviation. Source: SIFMA-member data.

Bao, J., J. Pan, and J. Wang (2011). The Illiquidity of Corporate Bonds. *Journal of Finance 66*, 911–946.

Basel Committee on Banking Supervision (2011). Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems. Bank for International Settlements.

Bech, M. and R. Garratt (2003). The Intraday Liquidity Management Game. *Journal of Economic Theory 109*, 198–219.

Bolton, P. and F. Sanama (2010). Contingent Capital and Long Term Investors. Working Paper, Columbia University.

Brunnermeier, M. and L. H. Pedersen (2009). Market Liquidity and Funding Liquidity. *Review of Financial Studies 22*, 2201–2238.

Chaisurote, S. (2008). Liquidity Provision by International Institutional Investors: Evidence from Brazilian Privatization. Working paper, Stanford University.

Chen, H., G. Noronha, and V. Singhal (2004). The Price Response to S&P 500 Index Additions and Deletions: Evidence of Asymmetry and a New Explanation. *Journal of Finance 59*, 1901–1929.

Chen, K., M. Fleming, J. Jackson, A. Li, and A. S. Sarkar (2011). An Analysis of CDS Transactions: Implications for Public Reporting. Staff Report Number 517, Federal Reserve Bank of New York.

Chen, Z., A. A. Lookman, N. Schürhoff, and D. J. Seppi (2009). Why Ratings Matter: Evidence From Lehman's Index Rating Rule Change. Working paper, CMU.

Comerton-Forde, C., T. Hendershott, C. Jones, P. Moulton, and M. Seasholes (2010). Time Variation in Liquidity: The Role of Market Maker Inventories and Revenues. *Journal of Finance 65*, 295–331.

Duffie, D. (1999). Credit Swap Valuation. *Financial Analysts Journal,*, January–February, pp. 73–87.

Duffie, D. (2010a). Asset Price Dynamics with Slow-Moving Capital. *Journal of Finance 65*, 1238–1268.

Duffie, D. (2010b). *How Big Banks Fail – And What to Do About It*. Princeton University Press.

Duffie, D. (2012). *Dark Markets: Asset Pricing and Information Transmission in Over-the-Counter Markets*. Princeton University Press.

Duffie, D., N. Gârleanu, and L. H. Pedersen (2005). Over-the-Counter Markets. *Econometrica 73*, 1815–1847.

Etula, E. (2009). Broker-Dealer Risk Appetite and Commodity Returns. Working Paper, Federal Reserve Bank of New York.

Feldhütter, P. (2009). The Same Bond at Different Prices: Identifying Search Frictions and Selling Pressures. Working Paper, University of Copenhagen.

Financial Services Authority (2011). The Failure of the Royal Bank of Scotland. Financial Services Authority Board Report, United Kingdom.

Fleming, M. J. and J. V. Rosenberg (2007). How do Treasury Dealers Manage their Positions? Federal Reserve Bank of New York.

Gao, X. and J. R. Ritter (2010). The Marketing of Seasoned Equity Offerings. *Journal of Financial Economics 97*, 33–52.

Goldstein, M., E. Hotchkiss, and E. Sirri (2007). Transparency and Liquidity: A Controlled Experiment on Corporate Bonds. *Review of Financial Studies 86*, 643–682.

Greenwood, R. (2005). Short and Long Term Demand Curves for Stocks: Theory and Evidence. *Journal of Financial Economics 75*, 607–650.

Gromb, D. and D. Vayanos (2002). Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs. *Journal of Financial Economics 66*, 361–407.

Gromb, D. and D. Vayanos (2010). Limits of Arbitrage: The State of the Theory. Working paper, London School of Economics.

Grossman, S. and M. Miller (1988). Liquidity and Market Structure. *Journal of Finance 43*, 617–633.

Harris, L. and E. Gurel (1986). Price and Volume Effects Associated with Changes in the S&P 500: New Evidence of Price Pressures. *Journal of Finance 41*, 815–830.

He, Z. and A. Krishnamurthy (2009). A Model of Capital and Crises. Working Paper, Northwestern University, forthcoming, *Review of Economic Studies*.

Hendershott, T. and A. Menkveld (2009). Price Pressures. Working Paper, Haas School of Business.

Hendershott, T. and M. S. Seasholes (2007). Market Maker Inventories and Stock Prices. *American Economic Review 97*, 210–214.

Kaul, A., V. Mehrotra, and R. Morck (2000). Demand Curves for Stocks do Slope Down: New Evidence from an Index Weights Adjustment. *Journal of Finance 55*, 893–912.

Kulak, J.-P. (2008). Unpublished Notes on Secondary Equity Offerings. University of Lausanne.

Lagos, R., G. Rocheteau, and P. Weill (2009). Crashes and recoveries in illiquid markets. NBER working paper.

Lou, D. (2009). A Flow-Based Explanation for Return Predictability. Working Paper, Yale School of Managment.

Loughran, T. and J. R. Ritter (1995). The New Issues Puzzle. *Journal of Finance 50*, 23–51.

Madhavan, A. (2001). The Russell Reconstitution Effect. *SSRN eLibrary*.

Meli, J. (2004). Do Capital Constraints on Market Makers Matter? Evidence From the U.S. Treasury Market. Working paper, Graduate School of Business. University of Chicago.

Mikkelson, W. H. and M. M. Partch (1985). Stock Price Effects and Costs of Secondary Distributions. *Journal of Financial Economics 14*, 165–194.

Mitchell, M. and T. Pulvino (2009). Arbitrage Crashes and the Speed of Capital. Working Paper, AQR Capital Management.

Mitchell, M., T. Pulvino, and E. Stafford (2002). Limited Arbitrage in Equity Markets. *Journal of Finance 57*, 551–584.

Musto, D., G. Nini, and K. Schwarz (2011). Notes on Bonds: Liquidity at all Costs in the Great Recession. Working paper, University of Pennsylvania.

Nagel, S. (2009). Evaporating Liquidity. Working paper, Stanford University.

Newman, Y. and M. Rierson (2003). Illiquidity Spillovers: Theory and Evidence from European Telecom Bond Issuance. Working Paper, Graduate School of Business, Stanford University.

Petajisto, A. (2009). Why Do Demand Curves for Stocks Slope Down? *Journal of Financial and Quantitative Analysis 44*, 1013–1044.

Reinhart, C. and K. Rogoff (2009). *This Time Is Different: Eight Centuries of Financial Folly*. Princeton University Press.

Rinne, K. and M. Suominen (2009). A Structural Model of Short-Term Reversals. Working Paper, Helsinki School of Economics.

Rinne, K. and M. Suominen (2010). Short-Term Reversals, Returns to Liquidity Provision and the Cost of Immediacy. Working paper, Helsinki School of Economics.

Scholes, M. S. (1972). The Market for Securities: Substitution versus price pressure and the effects of information on share price. *Journal of Business 45*, 179–211.

Shleifer, A. (1986). Do Demand Curves for Stocks Slope Down? *Journal of Finance 41*, 579–590.

United States Government Accountability Office (2011). Proprietary Trading: Regulators Will Need More Comprehensive Information to Fully Monitor Compliance with New Restrictions When Implemented. Report to Congressional Committees, July, 2011.

Weill, P.-O. (2007). Leaning Against the Wind. *Review of Economic Studies 74*, 1329–1354.

Wurgler, J. and E. Zhuravskaya (2002). Does Arbitrage Flatten Demand Curves for Stocks? *Journal of Business 75*, 583–608.

Zhu, H. (2012). Finding a Good Price in Over-the-Counter Markets. *Review of Financial Studies* forthcoming.