12 August, 2011

Mr. Werner Bijkerk
International Organization of Securities Commissions (IOSCO)
Calle Oquendo 12
28006 Madrid
Spain

Subject: Public Comment on Consultation Report: Regulatory Issues Raised by the Impact of Technological Changes on Market Integrity and Efficiency

Thank you for the opportunity to comment on the IOSCO Technical Committee's Consultation Report, "Regulatory Issues Raised by the Impact of Technological Changes on Market Integrity and Efficiency" ("Report"). The Report asks for public comment on the effect of technology, and of certain business models enabled by technology, on securities markets. The Report includes a particularly thoughtful discussion of so-called "high frequency trading" ("HFT") firms and business models. My letter will address those firms and business models.

As the Report notes, the U.S. Securities and Exchange Commission published a concept release on the structure of U.S. equity markets in January, 2010. In April, 2010, I submitted a comment letter on the concept release (http://www.sec.gov/comments/s7-02-10/s70210-107.htm). In that letter I discussed several facets of modern U.S. equity markets that deserved reform. My comment letter described factors responsible for increasing market volatility in recent years and also described ways in which long-term, fundamental investors were being disadvantaged by HFT firms, in some cases aided by the for-profit exchanges themselves.

In the letter I first described how U.S. market centers in the last 10 years had engineered "the replacement of formal and regulated intermediaries with informal and unregulated intermediaries" creating a "HFT-dominated market structure and business model that functionally replaced that of the 20th century exchanges." I noted that "HFT firms invested in ECNs and exchanges, sat on their boards, and held considerable influence over their design and operation," and took particular note of "an exchange that sells real-time data to high frequency trading firms telling those firms exactly where hidden interest rests and in what direction."

Going further, I reported that "A HFT market making firm does not need to register as a market maker on any exchange" and that "market centers pay these firms for supplying liquidity and the liquidity they supply becomes part of the market center business models. " Even registered market makers "have little or none of the regulatory oversight" they used to have, and "do not have any meaningful restriction on moving the market, they have no meaningful capital adequacy standards, no obligation to yield to customer orders, no meaningful obligation to maintain competitive quotes, no dealer position monitoring, nothing." In consequence, "firms are free to trade as aggressively or passively as they like or to disappear from the market altogether," and when markets face liquidity demands their behavior can "increase spreads and price volatility and savage investor confidence." But that, importantly, "They still get valuable privileges if they register as market makers, and they promise in return to merely post any quote, and a penny bid can count as a valid quote."

To the main policy point itself, whether HFT firms are the simple liquidity providers they claim to be, I wrote that despite what they might claim, HFT firms add liquidity only when it suits them, and that they cartwheel from supplying liquidity to demanding it when they spot trading opportunities or need to rebalance inventory, and that these trading practices exacerbated price volatility and directly contributed to the exceptional volatility of recent years.

Subsequent events soon illustrated all of these central points.

The Flash Crash

On May 6, 2010, U.S. markets suffered what's become known as the Flash Crash. Equity markets lost $1 trillion in about 15 minutes, and hundreds of stocks and ETFs traded at penny bids. A significant contributing cause of the Flash Crash was that after partly absorbing a large seller's position, HFT firms aggressively unloaded inventory in a cascading fratricide now sometimes called "hot potato" trading. After dumping inventory, many of these firms also then withdrew altogether from the market. As I later wrote in an invited comment for the FT Trading Room on FT.com, HFT firms "do it because they can. It doesn't happen because of a computer malfunction. It happens because it is designed to happen. It is not a bug, it is a feature."

One week after the Flash Crash, the New Jersey-based brokerage firm and market structure critic Themis Trading published a ground-breaking white paper called "Exchanges and Data Feeds: Data Theft on Wall Street" (http://www.themistrading.com/article_files/0000/0555/5-11-10_Data_Theft_On_Wall_Street_latest.pdf). In their paper, Themis described information leakage in certain exchange proprietary order book data feeds, practices almost entirely unknown to institutional investors and the public. As a consequence of the Themis paper, several market centers -- but not all - in Europe and the U.S. stopped some of their most egregious practices. As for the preceding 10 years, when these feeds leaked trading details to HFT firms with impunity, regulators are unaccountably silent.

Several months after the Flash Crash, Andrei Kirilenko, the Chief Economist of the U.S. Commodity Futures Trading Commission, co-authored one of the most important empirical market microstructure studies yet published entitled "The Flash Crash: The Impact of High Frequency Trading on an Electronic Market" (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1686004).

Kirilenko and his co-authors concluded that "High Frequency Traders exhibit trading patterns inconsistent with the traditional definition of market making," directly contradicting long-standing HFT claims that they are nothing more than simple market makers or liquidity providers. He went on, "Specifically, High Frequency Traders aggressively trade in the direction of price changes," contradicting HFT claims that they dampen volatility. And more, "This activity comprises a large percentage of total trading volume, but does not result in a significant accumulation of inventory. As a result, whether under normal market conditions or during periods of high volatility, High Frequency Traders are not willing to accumulate large positions or absorb large losses." As for signaling behavior, that is, whether all of that HFT trading volume tells other traders anything useful at all, Kirilenko wrote, "Moreover, their contribution to higher trading volumes may be mistaken for liquidity by Fundamental Traders."

And to the heart of the matter, whether HFT firms are healthy for the market, the very policy point HFT firms and their academic and regulatory co-religionists cite as justification for a decade's worth of deregulation in the U.S. - deregulation that enabled and empowered these firms - Kirilenko wrote "when rebalancing their positions, High Frequency Traders may compete for liquidity and amplify price volatility." He then points out that with inventory half-lives of about two minutes, HFT firms rebalance frequently throughout a trading day.

In stable markets, HFT firms do contribute liquidity and improve standard market quality metrics. Several academics have published studies showing a gradual improvement in several of these metrics in recent years. But their statistics are measured over months or years, and the disruptive and destabilizing trading practices of these firms are usually detectable only in episodes of transitory volatility, volatility which averages out in longer term studies. As an example of this kind of statistical artifact, fatalities from a plane crash will not move national mortality statistics a jot, but that is no consolation to the many victims.

HFT firms today rely on four legs to extract rents from the marketplace. As I wrote recently in a letter published in the *Financial Times*, these legs are privileged market access, privileged market data, destabilizing trading practices and deregulation.

Privileged market access

Many exchanges today sell products and services which provide preferential access to their markets. These services include co-location, proprietary data feeds and data interfaces designed specifically for HFT firms, and price tiered high-speed network services.

The effect, as has been widely acknowledged, is to sell more broadly what was once available narrowly, that is, privileged access to a marketplace and to the time, place and information advantages of that access. Advocates of these products -- most especially the for-profit exchanges selling these products and services and the HFT firms who take advantage of them -- claim they do little more than "level the playing field," and that co-location is just a modern variant of an old practice.

GETCO, one of the world's largest high frequency firms, wrote last year that "Historically, professional market participants, market makers in particular, have wanted to be as close as possible to the center of price discovery...Co-location is simply the equivalent of the modem day trading floor... And exchanges grant co-location to anyone willing to pay for it, just as exchanges used to grant trading floor access to those willing to buy a badge or a seat."

In the U.S., market maker access to the "center of price discovery" once came with a variety of regulations, however. Those regulations have been enthusiastically dismantled in the last decade in the U.S. What's left, as has been consistently the case for equity market deregulation in the U.S., is that many of the benefits of privileged access remain while many responsibilities were eliminated.

Co-location minimizes the time it takes for market information to travel from exchange matching engines to HFT firm trading algorithms, and, similarly, minimizes the time it takes for an order to travel from a trading algorithm to the exchange matching engine. In other words, being closer to the center of price discovery gives a firm an information advantage because it gets market information faster than other firms, and it gives a firm a place or situational advantage because it takes less time for an order to arrive at the matching engine. Co-located firms have discrete, identifiable and material advantages over other firms. That these advantages are for sale to anyone does not level the playing field, it merely puts more professionals on high ground while the public sinks.

Because co-location offers material advantages, exchanges must not only offer it on fair and non-discriminatory terms but regulators must define co-located firms as privileged participants, and determine what that privileged status implies for their markets. Co-located firms deploying high frequency strategies of whatever kind should also be subject to certain reporting requirements and trading restraints, as discussed below. Regulators must scrutinize co-location terms and any bundled services for discriminatory practices -- not just among those who buy them, but also for those who do not buy them - and these terms and services should be readily available for public scrutiny.

And regulators should consider whether pure price/time trade priorities any longer serve the public interest, particularly when for-profit exchanges enforce time priority rules and then explicitly charge for speed. Regulators should consider whether different trade priorities or other market microstructure initiatives could better serve the public interest.

To illustrate this point, imagine that an exchange explicitly charged member firms fees based only on how far the firm was located from the exchange, so that a firm five miles away from the exchange paid one rate and a firm 10 yards away from the exchange paid a different rate. Would anyone consider that in the public interest? Now imagine an exchange that explicitly ranked orders in the book based on how far a firm was located from the exchange, so that an order from a firm five miles away from the exchange was ranked below an order from a firm 10 yards away. Is that in the public interest? In practical effect, this is exactly what co-location in the digital era accomplishes.

Privileged market data

Research long ago demonstrated that the complete details of limit order books can be used to predict short term stock price movements. An order book feed gives you much more information than just price and size, such as you get with a consolidated quote. You get order and trade counts and order arrival rates, individual order volumes, and cancellation and replacement activity. You build models to predict whether individual orders contain hidden size. You reverse engineer the precise behavior and outputs of market center matching engines by submitting your own orders, and you vary order types and pore over the details you get back. If you take in order books from several market centers, you compare activity among them and build models around consolidated order book flows. With all of this raw and computed data and the capital to invest in technology, you can predict short term price movements very well, much better and faster than dealers could 10 years ago, even with their exclusive advantages of the time.

HFT firms point out that they do not have control over the order book and they do not have a first look at customer orders. The predictive power of today's analytic models and the vast amount of data available today to feed those models have erased this distinction, however, particularly when combined with the advantages of co-location. A first look at customer orders does not matter if you can use high-speed proprietary order book feeds to predict incoming orders to any profitable degree of accuracy, and if you can act faster than anyone so foolish as to rely on slower national utility data feeds. Last year, JP Morgan wrote quite matter-of-factly, "Some exchanges have products which give clients a faster look at quotes, in exchange for a fee. As a result, some HFTs end up with access to information sooner than institutional or retail investors who rely on more standard venues..." This is a very clear statement of the fact of a two-tiered marketplace, where information advantages are baldly for sale at premium prices.

In the spring and summer of 2010, the specific content of proprietary data feeds became a hot topic, largely because of Themis Trading's bombshell white paper. Their white paper described exactly how hidden trading interests were being leaked -- trading interests widely and reasonably assumed to be confidential. In retrospect, it seems clear that the tension between the public utility functions of an exchange and the private profit-seeking functions of for-profit exchanges created winners and losers in a showdown over who got to see what.

That tension will never be resolved, so regulators must enable investors to protect themselves by opting out of order book feeds. Published aggregated or consolidated quotes at a price should still be published, of course, but investors should be allowed to opt-out of data feeds that expose their individual orders and order activity, including order entry, updates, cancels and fills.

In the U.S., detailed order book information used to be restricted to highly-regulated intermediaries. These intermediaries were prohibited from trading ahead of customers, were restricted from demanding liquidity in most instances, and were held to fiduciary standards. In other markets around the world, this kind of detailed order book information never used to be distributed at all. Now, this data is for sale almost everywhere. As an example of what happens today, a predatory HFT firm can follow the entire lifetime of any order, whether its own or anyone else's, and in many cases predict what kind of entity submitted that order and what its strategy is. JP Morgan wrote that one firm estimated "HFT tracking algorithms can drive execution costs up 1.5 to 3 times, even when institutional investors parcel trades into smaller orders to avoid detection."

Institutional investors try to protect themselves by trading in dark pools -- increasing market fragmentation. Or they try to protect themselves by deploying their own counterforce algorithms. Or regulators can simply enable investors to stop paying rents to HFTs by letting investors opt-out from proprietary order book data feeds. One large institutional investor in the U.S. goes further, asking regulators to implement an opt-in standard, where exchanges must get investor permission before distributing any of their order-level activity (www.sec.gov/comments/s7-02-10/s70210-235.pdf).

Regulators must also insist that the complete details of these order book feeds are publicly disclosed, and that they are not changed without a reasonable chance for the public to weigh in. In the U.S., order

book feed specifications should be embedded in SRO rules or even in National Market System plans, so that the public gets notice of impending changes and has a chance to comment. As it stands now, an exchange could unilaterally decide to broadcast details of even customer reserve (iceberg) orders, and it's not at all clear whether the exchange would be sanctioned. How do we know that? Something like it has already happened.

Destabilizing trading practices and deregulation

Regulators must abandon any notion that market making is passive. This is an anachronism. In the U.S., passive market makers were for the most part driven out of business by 2008. Broadly speaking, while there are three distinct high frequency strategies, namely, market making, short term directional, and arbitrage strategies, any high frequency market making algorithm combines these strategies. For example, market making algorithms include liquidity detection strategies just to function; many also include event-driven strategies and mean reversion strategies (this is likely one reason HFT firms have been repeatedly seen to demand liquidity at price reversals). High frequency market makers active in ETFs combine market making and arbitrage strategies. High frequency market makers in options and futures markets combine market making and cross-asset hedges and arbitrage.

In other words, HFT firms of all stripes are aggressive, profit-seeking traders. As has been the case in the past with older intermediated market structures, if suitable data is made available we can be sure research will show HFT market making firms make money when they demand liquidity and that their liquidity demands are informative about short term price movements. Early research already suggests it. When they demand liquidity, HFT firms behave exactly like informed traders, and are not the passive, uninformed liquidity providers widely assumed about market intermediaries in the past. In stable markets, where buy and sell liquidity demands are in rough equilibrium, classic models apply and market makers do supply liquidity to other informed traders, adjusting spreads and prices to reflect ebbs and flows, and price discovery is driven by other informed participants. In volatile markets, however, market makers behave as if they are themselves, and are in competition with, informed traders, demanding liquidity to balance their inventory and control their risk, and driving price discovery.

Because HFT liquidity supply is frequently offered over time but their liquidity demands are concentrated to when their models alert, market makers cartwheel from seemingly passive, uninformed traders to aggressive, informed traders and drive significant price changes. "Price discovery" in this sense, where intermediary liquidity demands drive prices along with or in competition with other informed traders, should change our ideas of how markets behave. At least in the U.S., most ideas of how market makers operate were crafted in an era when intermediaries were, by rule, prohibited from unrestrained liquidity demands; in other words, these theories presumed regulation *a priori*, whether they knew it or not. In the U.S., that era ended decisively in 2008.

Kirilenko's paper shows how this all works in practice. His is the only trade-by-trade, second-by-second analysis to look at the trading behavior and inventory management of high frequency firms in today's markets. His advantage in this analysis is that, as a regulator, he had much more detailed data than researchers usually see, most importantly data which precisely identified market participants. What he found is deeply troubling, implying structural instability, crashes and liquidity crises large and small, toxic quotes and price discovery in the public markets, and the uncertainty of any order's likely effect on prices. His paper is therefore a decisive empirical justification for reforming how high frequency firms operate in today's markets.

Kirilenko's main finding is that there's a tipping point in volatile markets when, in an instant, high frequency market makers stampede to rebalance their inventories, even cascading positions from firm to firm, while prices collapse. In the Flash Crash itself, markets crashed when high frequency market makers hit internal inventory limits and unloaded onto the next market maker, which then hit limits and unloaded onto the next one, and so on, driving the market down by almost $1 trillion dollars in a few minutes. Though Kirilenko studied e-mini trading in the futures market, the CFTC and SEC staff report on the Flash Crash showed the same behavior at work in the equities markets, doubtless from many of the same firms.

A posted bid or offer has at least four dimensions. Beyond price and size, any resting bid or offer has a lifetime, and the inventory cycle or position resulting from any executed bid or offer has a lifetime. Even if it's at the best price, a bid or offer lasting a fraction of a second hasn't contributed to meaningful price discovery, and market makers use the latest technology to post and cancel thousands of bids and offers per second, even in the same stock. An executed bid or offer where the position is unwound quickly and aggressively isn't meaningful price discovery either; some equities high frequency market makers claim inventory lifetimes of as little as 11 seconds in their stocks. Market maker inventory cycles of a few seconds or minutes, enforced by aggressive trading as time or prices go against the firm, actively destabilize prices, especially so in already volatile markets, just as Kirilenko found in the Flash Crash.

It's important to note that of a quote's four dimensions, in the U.S. only one dimension has materially improved in the last 10 years. Because of decimalization, automation and deregulation, quoted spreads have improved for liquid stocks in stable markets. But quote duration is down, time-in-inventory is down, and for many stocks quote size is flat or down. This is all because, to manage costs, high frequency market maker inventory cycles are engineered down to seconds, and these firms keep their capital commitments low. Frenetic trading isn't a byproduct of their strategies -- it is the strategy.

HFT firms came to dominate with this business model because the business model itself is cheap relative to that of the more highly regulated intermediaries of the past, and because it's cheap HFT firms can post tighter spreads. It's cheap only because the HFT market maker's liquidity is toxic. Its quotes are priced aggressively, showing tight spreads, but only for small quantities with very short lifetimes, with aggressive inventory management behind them to limit the firm's exposure. Firms fine-tune their risk models to make their inventory cycles as short as possible while preserving profits, lowering risk and costs. And since in the U.S. current regulations focus mainly on only one of a quote's four dimensions -- price -- if you're the best price, by rule everyone must do business with you, and you gain market share.

Any firms with longer term inventory models couldn't compete against these lower costs and tighter spreads, so over the last 10 years in the U.S. exchanges deregulated and became hosts for HFT market makers, even cloning themselves to make new hosts. Deregulated, toxic quotes flourished and crowded out more stable quotes. U.S. exchanges today are interchangeable because HFT firms determine market structure through their increasingly toxic quotes. Rather than a diverse ecosystem of market centers, with systemic resilience in that diversity, deregulated U.S. markets are inbreds relying on the same high frequency market maker firms trading the same toxic models.

At minimum, regulators must shed light on all of these business practices. Any firm that co-locates inside an exchange should be presumed to be using HFT strategies and should be required to register with regulators. That firm should be required to report inventory holding periods daily, by stock and in the aggregate, and for all of its activities across all venues in which it trades. Any firm that buys proprietary data feeds from an exchange should have the same reporting requirements.

Regulators should require any firm deploying a market making strategy to register itself as a market maker. In jurisdictions which extend regulatory privileges to registered market makers, privileges which afford any kind of advantage to these firms, market makers should at least be subject to affirmative obligations to maintain genuinely competitive market quotes at all times. Market maker intraday positions should be reported to facilitate review. Most important, market makers must have standards for exactly when and how they can rebalance their inventories, and should be restrained from demanding liquidity if it accelerates the price trend. At minimum, market makers should report to regulators when they supplied liquidity and when they demanded it, by security, across all venues, along with their realized spreads.

Regulators should also carefully evaluate order book trading priorities, particularly the role time priority plays. In the U.S., time priority at a price stands behind many of the problems about which critics are worried. As a simple reform, giving priority to agency orders ahead of principal orders (first ensuring that any firm deploying a market making strategy is treated as principal), and allocating interest to standing principal orders *pro rata,* might bring about much more desirable outcomes than today's never-ending chase for speed. Very bluntly, price/time priority has lost legitimacy. It lost legitimacy because

exchanges learned how to charge for it, and HFT firms simply pay the exchanges for infinitesimal time advantages over other participants.

Regulate them

Equity market reforms and deregulation in the last decade were for good reasons -- monopoly profits were flowing to intermediaries and exchanges, intermediaries were taking advantage of their customers, innovation was being strangled by entrenched interests -- and it was time for reform. As many of us hoped, new participants, technology and business models sprang up. Nobody wants to undo that progress. By analyzing trade and position data from the futures market, Kirilenko showed how a dominant class of these business models can be disruptive, and how these models can be destabilizing enough to create systemic risk as an inherent consequence of their design, just as reformers have said.

A basic function of any market is to produce a quote. Today's HFT quotes are toxic, a hoax on equities markets. No one planned it. It happened as an unanticipated consequence of well-meaning reforms worldwide. There is no competitive solution to this problem within current regulations so long as quote price is a routing table's first regulatory imperative and matching engines rank only by price/time. Competition simply forces market centers to publish more and faster toxic quotes, as market power continues to shift from market centers to HFT firms, and HFT firms use that power to extract rents from investors.

Sincerely,

R.T. Leuchtkafer
New York, NY
USA