Manipulation of Illiquid Asset Indexes

K. Jeremy Ko, Igor Kozhanov, and Sean Wilkoff*†

February 24, 2016

Abstract

We develop a model of manipulation of indexes, whose components are illiquid in that they require fair valuation based on comparable instruments. Such an index may be susceptible to manipulation since distorting the prices of only a few assets could potentially shift its value. Our model provides a measure of the manipulability of an index and identifies which assets are most likely to be manipulated. We apply our model to analyze the manipulability of national municipal bond indexes subject to various bond size thresholds.

Keywords: **Manipulation, Index, Derivatives, Exchange-traded Funds, Fixed Income, Municipal Bonds, Liquidity, Regulation**

JEL Classification: G18, G23, G28

---

*We thank seminar participants at the Financial Conduct Authority for their comments. All remaining errors are our own. K. Jeremy Ko and Igor Kozhanov are at the Division of Economic and Risk Analysis at the Securities and Exchange Commission, and Sean Wilkoff is at Cornerstone Research. Please address correspondence to: 100 F Street NE, Washington, DC 20549, `kok@sec.gov`, 202-551-7895.

1 Introduction

The ability to profit from derivatives positions by manipulating underlying assets has been a longstanding concern of those who oversee and participate in financial markets. Numerous academic studies have articulated concerns regarding manipulation of underlying markets for both physical and cash-settled derivatives.[1] A number of financial institutions have been accused over time of engaging in this type of manipulation. For example, the United Kingdom's Financial Conduct Authority fined Barclay's in 2014 for placing manipulative orders during the daily gold price fixing in order to avoid paying out on a derivative position.[2] In the US, the CFTC imposed sanctions in 2001 on the energy trading firm, Avista, for manipulating the NYMEX electricity futures contract. The CFTC stated in its order: "To the extent that Avista Energy's Traders could distort the price for futures contracts with an order at the Close on Options Expiration Day that was smaller than the positions created by Avista Energy's OTC derivatives contracts, Avista Energy's Traders thought that they might be able to profit via an artificially created increase in the value of its OTC derivatives contracts."[3] Other manipulations of this type have involved firms misreporting transactions and prices to distort derivative settlement values in commodities, interbank lending, and other markets.[4]

The aforementioned examples all involve manipulation of a single reference asset or instrument. Manipulation may also be feasible for derivatives which reference an index of assets when either economic or structural issues make the index vulnerable to manipulation. For example, Dutt and Harris (2005) develop a model of manipulation which they apply to narrow-based index derivatives. Manipulation of a concentrated or narrow-based index may be feasible since the cost of manipulating a small number of assets may be sufficiently low. Our paper examines another setting of such vulnerability, in which the index consists of illiquid assets. We use the term illiquid to refer to assets with two features. First, these assets are costly to trade in terms of market impact, bid-ask, and other transaction costs. Second,

[1] See Kyle (1984), Kumar and Seppi (1992), Pirrong (2001), and Dutt and Harris (2005).

[2] See: https://www.fca.org.uk/news/barclays-fined-26m-for-failings-surrounding-the-london-gold-fixing.

[3] See: http://www.cftc.gov/files/enf/01orders/enfavista-speaking.pdf.

[4] Regarding the well-publicized LIBOR manipulation, see: http://stream.wsj.com/story/the-libor-investigation/SS-2-32262/. Regarding a recent alleged manipulation of the oil market, see: http://www.bloomberg.com/news/articles/2013-05-24/bp-shell-statoil-face-u-s-lawsuit-after-eu-oil-price-probe.

they are traded infrequently and, as a result, generally require fair valuation based on similar assets.[5] A trader can, in principle, manipulate an index of such assets by distorting prices in just a few select components of the index. The prices of other assets in the index should shift as a result of this distortion since their prices are linked through fair valuation.[6]

In this paper, we develop a model of such manipulation. We assume that the manipulator holds a derivative position which can be liquidated costlessly. This position can represent a cash-settled or redeemable derivative or liquid exchange-traded product such as an ETF. Prices are distorted either indirectly through trade (as in the aforementioned Barclay's and Avista manipulations) or directly by misreporting quotes and transactions (as in the LIBOR manipulation). Our model can address two critical questions. First, to what degree is a particular portfolio of illiquid assets manipulable? Second, which assets within such a portfolio are most likely to be manipulated? Addressing the first question will give market supervisors and others information about which derivatives and investment vehicles are prone to manipulation. Addressing the second question can help market supervisors focus their efforts to detect manipulation.

In this paper, we apply our model to assessing the manipulability of fixed-income indexes. Our motivation derives from the growing class of indexed investment funds referencing fixed-income and other illiquid assets.[7] Although these funds can be mutual funds, a growing number are exchange-traded funds (ETFs) that list their shares on national securities exchanges.[8] Exchange proposals to list new exchange-traded products are subject to Sec. 6(b)(5) of the Securities Exchange Act of 1934, which requires, among other things, that the exchange's rules be designed to prevent manipulative acts and practices.

Some of these ETFs reference municipal bond indexes with bond supply thresholds that vary significantly across funds in terms of issuance amount or par amount outstanding. One

[5] This notion of illiquidity does not necessarily conform to the definition provided by the Investment Company Act of 1940 and Commission guidelines. The 40-Act definition of an illiquid asset is one which cannot be disposed of in the ordinary course of business within 7 days for its carrying value. Long-standing Commission guidelines have required 40-Act funds (e.g., mutual funds and ETFs) to hold no more than 15% of their net assets in illiquid securities and other illiquid assets. Our definition admits a much larger class of securities. In this paper, we examine indexes, which can potentially underlie 40-Act funds yet are composed almost entirely of illiquid assets according to our definition.

[6] In this paper, we also consider the case whereby a manipulator distorts the index by trading in related securities outside of the index as we discuss later.

[7] See: http://www.bloomberg.com/news/articles/2012-07-10/fixed-income-etfs-to-own-2t-in-10-years.

[8] See SR-NYSEArca-2015-25 (http://www.sec.gov/rules/sro/nysearca/2015/34-74730.pdf) and SR-NYSEArca-2015-28 (http://www.sec.gov/rules/sro/nysearca/2015/34-74701.pdf).

purpose these thresholds serve is as a proxy for liquidity. Morgan Stanley and BlackRock also recently proposed the creation of "single issuer" fixed-income trusts, which would hold bonds issued by a single corporate or municipal entity.[9] These funds are currently prohibited by concentration limits within US tax code and have been proposed to address bond illiquidity exacerbated by CUSIP proliferation. However, the extent to which these highly concentrated funds would be prone to manipulation is still unknown.

In this paper, we study how varying bond supply thresholds affects the manipulability of a particular hypothetical bond index (based on the S&P National AMT-Free Municipal Bond Index). A lower threshold will admit more components to the index. All else equal, such a change should make this index less susceptible to manipulation, as having to distort more components should increase the cost to the manipulator. However, a lower threshold should also decrease the average liquidity of the index components, thus making the value fo these components more sensitive to manipulative activity. Such a change should increase susceptibility to manipulation, all else equal. Therefore, the net effect on manipulability of lower supply thresholds is a priori unclear.

In this paper, we generate a number of novel findings. First, market impact cost does, indeed, appear to decrease with bond supply. However, municipal bond maturity has a much stronger relationship with market impact cost. This measure of liquidity is relevant for a manipulator because it determines both the cost of a manipulation and the benefit in terms of the magnitude of the price distortion. This result is consistent with the findings of the European Securities and Market Authority (ESMA) in a 2014 consultation paper (ESMA (2014)). The ESMA study documents a relationship between bond issuance size and liquidity as measured by trading volume for sovereign and corporate bonds. Market impact cost also decreases as credit rating increases. However, municipal bond maturity appears to have the strongest relationship with market impact cost than these other two variables in our sample. Second, we document a slight decrease in manipulability as the number of index components increase in response to decreasing bond supply thresholds. This decrease in manipulability stems from a decrease in index concentration, which more than offsets any changes in market impact costs of the index components. We should add the caveat that our results hold for

[9] See: http://www.reuters.com/article/2015/05/15/us-morgan-stanley-bonds-liquidity-idUSKBN0O025A20150515 and http://www.blackrock.com/corporate/en-us/literature/whitepaper/viewpoint-bond-etfs-benefits-challenges-opportunities-july-2015.pdf.

the particular index we examine and may not hold generally across all conceivable municipal bond indexes.

Our paper is related to a number of others which examine manipulation in financial markets. First, our model is based on that of Dutt and Harris (2005), which applies to indexes underlying cash-settled derivatives. We focus on indexes of illiquid assets that are routinely fair-valued so that the index value moves when a manipulator trades in a small number of comparable assets. There are also related papers which develop models of manipulation of single reference assets for cash-settled derivatives. The paper of Kumar and Seppi (1992) was the first to do so in a one-period model of trade in the underlying market. Horst and Naujokat (2011) and Gallmeyer and Seppi (2000) both develop dynamic models of manipulation to benefit cash-settled derivative positions. In their models, multiple manipulators trade to distort the price of a single reference instrument over multiple periods – the former in continuous time and the latter in discrete time. In contrast, we focus on one-shot manipulation that occurs prior to the realization of gains. Our focus is on the manner in which a manipulator distorts specific instruments to move an index and not on the dynamics of the manipulation.

A related empirical study is that of Cornell (1997), which examines the possible manipulation of the Bond Buyer Index (BBI) of municipal bonds underlying cash-settled futures contracts listed on the Chicago Board of Trade. There are also a number of papers which employ data analysis or econometric methods to detect manipulation.[10] Our proposed method for detection differs in that it is grounded in a theory of profit-maximization by a manipulator rather than deriving from an ad-hoc empirical model.

The remainder of this paper proceeds as follows. Section 2 desribes the general form of our model, which only assumes that the manipulator exerts a linear impact on prices. Section 3 describes a "first-principles" version of the model based on Kyle (1985) in which the market forms rational expectations from informed trades. We then apply our model to our hypothetical municipal bond indexes in Section 4. Section 5 concludes.

[10]See Cao et al. (2014), Cao et al. (2013), Comerton-Forde and Putnins (2011), Ogut et al. (2009), Aitken et al. (2009), Diaz et al. (2011).

2 General Model

The model described in this section is based on that of Dutt and Harris (2005) with a few additional features including the fact that the assets in the index require fair valuation as a result of illiquidity. We assume that the manipulator in our model holds a fixed endowment of liquid index derivative securities, which could represent cash-settled futures or options, ETFs with liquid secondary market trading, open-end mutual funds, etc. This position has a notional value of κ dollars.

The manipulator trades in the illiquid underlying assets (or related securities) to manipulate the index. However, our model can generalize to the case where the manipulator is a market maker who misstates quotes as we discuss at the end of this section. We assume that the manipulator trades in a segmented dealer market with delayed public revelation of trades after execution as with a number fixed-income markets.[11] Trades are assumed to have a linear impact on prices as in the Kyle (1985) model. Namely, trading θ_j dollars of asset j causes a return (i.e., the fractional change in price) on this asset of magnitude:

$$r_j = \gamma_j \cdot \theta_j \tag{1}$$

Econometric models of market impact typically use a non-linear impact function.[12] We employ a linear model for simplicity although our results generally apply for a non-linear impact function as well.

There are N assets, in which the manipulator could trade to distort the index. However, the manipulator only trades in a subset, $n \leq N$, of these assets with different dealers.[13] We make this assumption to capture discrete costs of trading each asset. The subset of assets, in which the manipulator trades, is denoted $\Im_n = \{i_1, i_2, \ldots, i_n\}$. In vector notation, the

[11]For example, both the Municipal Securities Rulemaking Board and FINRA post trades and prices in the municipal and corporate bond markets, respectively, to the public up to 15 minutes after each transaction.

[12]See Almgren (2003) and the references therein.

[13]We assume that the manipulator does not execute multiple trades in the same asset with different dealers. This assumption can be justified by adding the presence of similar substitutable assets with identical market impact cost. Trading in such assets can yield greater benefit to the manipulator if these assets are not perfectly correlated. Namely, the index will react less to two trades in the same asset than to two trades in imperfectly correlated assets. This property of inferences from multiple variables will be a feature of the first-principles model outlined in the next section.

manipulator makes the following dollar trades in these assets:

$$\vec{\Theta}_{\Im_n} = \begin{bmatrix} \theta_{i_1} \\ \theta_{i_2} \\ \vdots \\ \theta_{i_n} \end{bmatrix} \tag{2}$$

Trading in a single asset is assumed to influence not only the price of that asset, but also other correlated assets in the index after revelation of the trade. This assumption is meant to capture fair valuation for illiquid assets based on similar assets, such as "matrix pricing" for bonds. We again assume a linear price impact so that, after revelation, the manipulator's trade causes a return (or fractional change in value) on the index of the following magnitude:

$$r_0 = \vec{\Phi}^T_{\Im_n} \vec{\Theta}_{\Im_n} \tag{3}$$

where $\vec{\Phi}_{\Im_n}$ is the n-vector linear price impact coefficient for this trade on the index. We account for the dollar amount of the manipulator's trades in terms of the "true" (or pre-trade) value of the assets. Therefore, the net cost of this trade to the manipulator is given by:

$$C\left(\vec{\Theta}_{\Im_n}\right) = \sum_{m=1}^{n} r_{i_m} \theta_{i_m} = \sum_{m=1}^{n} \gamma_{i_m} \theta^2_{i_m} \tag{4}$$

since r_{i_m} is the amount by which the manipulator overpays for each dollar of asset i_m. We can re-write this cost in vector notation as:

$$C\left(\vec{\Theta}_{\Im_n}\right) = \vec{\Theta}^T_{\Im_n} \Gamma_{\Im_n} \vec{\Theta}_{\Im_n} \tag{5}$$

where $\Gamma_{\Im_n}$ is a diagonal matrix such that $(\Gamma_{\Im_n})_{m,m} = \gamma_{i_m}$. We assume that the value of the manipulator's derivative position moves in proportion to the change in the index. In other words, information migrates from the underlying market to the derivative market as a result of index arbitrage. Therefore, the profit of the manipulator is given by:

$$\Pi_{\Im_n}\left(\vec{\Theta}_{\Im_n}\right) = \kappa r_0 - C\left(\vec{\Theta}_{\Im_n}\right) = \kappa \vec{\Phi}^T_{\Im_n} \vec{\Theta}_{\Im_n} - \vec{\Theta}^T_{\Im_n} \Gamma_{\Im_n} \vec{\Theta}_{\Im_n} \tag{6}$$

The first-order condition for this profit yields the manipulator's optimal trade quantities:

$$\vec{\Theta}^*_{\Im_n} = \frac{\kappa}{2}\Gamma^{-1}_{\Im_n}\vec{\Phi}_{\Im_n} \tag{7}$$

Therefore, the manipulator's maximum profit for the subset of assets $\Im_n$ is given by:

$$\Pi_{\Im_n}\left(\vec{\Theta}^*_{\Im_n}\right) = \frac{\kappa^2}{4}\vec{\Phi}^T_{\Im_n}\Gamma^{-1}_{\Im_n}\vec{\Phi}_{\Im_n} \tag{8}$$

The manipulator will further maximize this profit over the subset of assets traded, so that this optimized profit is given by:

$$\Pi^* = \max_{\Im_n}\frac{\kappa^2}{4}\vec{\Phi}^T_{\Im_n}\Gamma^{-1}_{\Im_n}\vec{\Phi}_{\Im_n} = \frac{\kappa^2}{4}M(n) \tag{9}$$

For a given n, the term $M(n) = \max_{\Im_n}\vec{\Phi}^T_{\Im_n}\Gamma^{-1}_{\Im_n}\vec{\Phi}_{\Im_n}$ provides a measure of the manipulability of this index. In addition, the optimal subset of assets and traded quantities derived from this model can help market supervisors ascertain which assets are most likely to manipulated. As mentioned earlier, this information can be used in determining the likelihood that certain market actions were intended to manipulate prices for the purposes of investigating and prosecuting offenses.

One can directly estimate the market impact coefficients for trade in an asset on the transacted price for that asset, $\gamma_.$, using transaction and pricing data for the underlying markets. We estimate these coefficients for the aforementioned municipal bond indexes in an event-study framework in Section 4. Estimation of $\vec{\Phi}_{\Im_n}$, in contrast, is not as straightforward. Generally, one will observe trading activity in different sets of assets each trading period in real-world transactions data. It is not a priori clear what the associated data imply for the market impact on the index from a trade in the specific set of assets, $\Im_n$. For this reason, we require a model to determine what information should be inferred when the manipulator trades in $\Im_n$. We consequently develop, in the next section, a first-principles model of informed trade in multiple assets, in which the market forms rational expectations based on this trade.

As mentioned previously, our model can generalize to the case where the manipulator is a market maker who misstates quotes. In this case, the manipulator distorts prices directly and not indirectly through trading. We begin our argument by observing that the manipulator's cost function can be rewritten in terms of price distortions rather than traded quantities. Namely, we invert equation (1) and re-write in vector notation as: $\vec{\Theta}_{\Im_n} = \Gamma_{\Im_n}^{-1}\vec{r}_{\Im_n}$ where $\vec{r}_{\Im_n}$ is a vector capturing price distortions for these assets in terms of the returns caused by trading. Substituting into the manipulator's cost from equation (5) yields:

$$C\left(\vec{r}_{\Im_n}\right) = \vec{r}_{\Im_n}\Gamma_{\Im_n}^{-1}\vec{r}_{\Im_n} \tag{10}$$

Therefore, the manipulator's profit in equation (6) can also be rewritten as follows:

$$\Pi_{\Im_n}\left(\vec{r}_{\Im_n}\right) = \kappa r_0 - C\left(\vec{r}_{\Im_n}\right) = \kappa r_0 - \vec{r}_{\Im_n}\Gamma_{\Im_n}^{-1}\vec{r}_{\Im_n} \tag{11}$$

The manipulator's cost in this objective function is quadratic in price distortions and proportional to the liquidity of the distorted assets (as captured by $\Gamma_{\Im_n}^{-1}$). In addition, the manipulator's revenue is proportional to the price impact on the index from the manipulation. We argue that a manipulator who distorts prices directly will face an objective function that is roughly equivalent to this one. It is reasonable to surmise that such a manipulation will involve costs that are convex in the the magnitude of the distortion and proportional to the liquidity of the distorted assets. For example, misreporting the price of a more liquid asset will be more likely to arouse suspicions of manipulation.[14] In addition, manipulating prices directly can lead the market to make similar inferences as trades that lead to those price distortions, as long as prices impound information efficiently. Therefore, the price impact on the index should be similar whether conducted through trading or prices. We conclude that our model can generalize to the case of a market maker manipulating prices directly.

3 First-Principles Model

The following model is a one-period model of trade in a competitive dealer market similar to Kyle (1985). Our variation of his model features trading in multiple assets through segmented dealers. In addition, information about all transactions is publicly revealed prior

[14]Misstating the price of a more liquid asset may also attract informed trade attempting to exploit this distorted price if based on marketable quotes.

to assets paying off. We assume that trading by a manipulator in this market is sufficiently rare so that it has no material effect on the equilibrium for prices. Under normal market conditions, a number of informed traders arrive randomly to the market independently of one another. For simplicity, all agents in this model are assumed to be risk-neutral, and the risk-free rate is assumed to be zero.

Dealers offer competitive prices in N risky assets. At time 1, asset j provides a terminal payoff of $\Re_j$, which is normally distributed with mean of 1. At time 0, each informed trader transacts with a different dealer after observing information about one of the risky assets. Specifically, a trader observes a signal of $s_j = \Re_j + \varepsilon_j$ where ε_j is normally distributed with mean zero. This trader then transacts with a dealer by soliciting a quote for their desired dollar quantity, θ_j. In addition to their own trade, the informed trader must also intermediate a trade of v_j in dollar terms, which is distributed normally with mean zero. This additional demand is equivalent to noise trade in the model of Kyle (1985). This demand could represent external trades passed through to dealers such as creation and redemption requests made to asset managers.

Following the analysis of Kyle (1985), the linear equilibrium for prices and trade in this market is as follows:

$$\begin{aligned} &\Delta p_j = r_j = \gamma_j \cdot \theta_j \\ &\theta_j = \tfrac{1}{2\gamma_j}\left(E[\Re_j|s_j] - 1\right) + v_j \end{aligned} \tag{12}$$

where $\gamma_j = \frac{1}{2}\left\{\frac{var(E[\Re_j|s_j]}{var(v_j)}\right\}^{\frac{1}{2}}$. We can equate price changes from the Kyle (1985) model to returns above since the price of each risky asset is marked to its unconditional mean of 1 prior to trade.[15]

We denote the terminal value of the index by $\Re_0$, which is a normally distributed average of the individual asset payoffs. After public revelation of trades, the value of the index should change by the following amount dependent upon the conditional expectation on $\Re_0$ based on the observed trade of $\vec{\Theta}^T_{\Im_n}$:

$$r_0 = E\left[\Re_0|\vec{\Theta}^T_{\Im_n}\right] - 1 = \vec{\Phi}^T_{\Im_n}\vec{\Theta}_{\Im_n} \tag{13}$$

[15] For the same reason, we can equate the dollar value of trade (marked at its pre-trade value as in the previous section) to share demand from the Kyle (1985) model.

where $\vec{\Phi}_{\Im_n} = \Sigma^{-1}_{\Im_n,\Im_n} \vec{\Sigma}_{\{0\},\Im_n}$.

$\vec{\Sigma}_{\{0\},\Im_n}$ is the vector covariance between the index value and trade in these assets with elements given by:

$$\begin{aligned} \left(\vec{\Sigma}_{\{0\},\Im_n}\right)_m &= cov(\Re_0, \theta_{i_m}) \\ &= \frac{\delta_{i_m}}{2\gamma_{i_m}} cov(\Re_0, \Re_{i_m}) \end{aligned} \tag{14}$$

where $\delta_j = \frac{var(\Re_j)}{var(\Re_j)+var(\varepsilon_j)}$ is a parameter ranging between 0 and 1, which measures the informativeness of signal $s_j = \Re_j + \varepsilon_j$.

$\Sigma_{\Im_n,\Im_n}$ is the covariance matrix of trade in these assets with elements given by:

$$\begin{aligned} \left(\vec{\Sigma}_{\Im_n,\Im_n}\right)_{l,m} &= cov(\theta_{i_l}, \theta_{i_m}) \\ &= \frac{\delta_{i_l}\delta_{i_m}}{4\gamma_{i_l}\gamma_{i_m}} cov(\Re_{i_l}, \Re_{i_m}) + \mathbf{1}(l = m) var(v_{i_m}) \end{aligned} \tag{15}$$

Therefore, the index market impact coefficient vector, $\vec{\Phi}_{\Im_n}$, can be estimated using the covariance of returns, trades, or some combination thereof.

4 Application to Municipal Bond Indexes

4.1 Index Information and Data

In this section, we apply our model to assessing the manipulability of municipal bond indexes as discussed earlier. Specifically, we analyze the effects of bond supply thresholds on the manipulability of a specific hypothetical national municipal bond index based on the S&P National AMT-Free Municipal Bond Index. This index admits all outstanding non-callable investment grade U.S. municipal bonds rated by the big three credit rating agencies (i.e., S&P, Moody's, and Fitch) with deal size greater than \$100 million. We consider four possible supply thresholds at issuance for this index: 1) \$100 million (bond or CUSIP-level), 2) \$100 million (deal-level), 3) \$75 million (deal-level), and 4) \$20 million (deal-level).

The first threshold roughly corresponds to the generic standard for fixed-income ETF indexes on the NYSE Arca exchange, the primary listing venue for ETFs. Conforming to this standard allows the exchange to list the ETF without external regulatory approval. The precise threshold for this standard is \$100 million par amount outstanding per bond (or

CUSIP). However, we lack precise daily data on amount outstanding. Therefore, we use bond size at issuance as a proxy for this threshold. The second threshold obviously corresponds to the existing threshold for the S&P National Index. There are municipal bond indexes with listed ETFs which correspond to the last two thresholds.

We divide our index into "matrix categories" (i.e., groupings of similar bonds) based on these four thresholds as well as two other characteristics. Specifically, we further subdivide bonds by high and low credit ratings (AA or higher and below AA) and high and low maturity (>15 years and ≤15 years). Therefore, there are $16 = 4 \times 2 \times 2$ matrix categories in our analysis. We assume that bonds are homogeneous within each category with respect to their market impact and their trade and return distribution.

We use daily fixed-income price and transaction data for the 2014 calendar year from three sources. First, trade size and price data are obtained for the bonds in our sample from the Municipal Securities Rulemaking Board (MSRB). In keeping with the model of Kyle (1985), we eliminate inter-dealer trades and use customer transactions only. We also winsorize this data, eliminating extreme trade prices and sizes. Specifically, any trade with a price greater than or equal to $148 or less than $70 (less than 0.01% of the sample each) is dropped.[16] In addition, any trade below $5000 par (0.05% of our sample) and above $1 million par (2.5% of our sample) is eliminated. Small and large trade sizes are omitted from our sample for two reasons. First, the market impact function is likely closer to linear than with the inclusion of extreme trade sizes. Second, there are potential timing issues with large trades. Namely, our discussions with a bond transaction data provider (MSRB) indicate that dealers may become aware of large trades prior to their execution. Therefore, large trades may be reported too late relative to their associated price movement. Our second data source is Thomson Reuters DataScope, which provides ratings data from the big three credit rating agencies and daily valuations from Thomson Reuters for the bonds in our sample. We finally obtain the following daily macro interest-rate series from the St.Louis Federal Reserve Bank: the 3 month and 10 year Constant Maturity Treasury Bond yields and the Aaa and Baa Moody's Bond yields. Based on this data, we control all municipal bond returns for daily changes in three macro interest-rate factors: 1) the change in the 3-month Treasury yield, 2) the change in the 10-year Treasury yield, and 3) the change in the Baa – Aaa credit

[16]We assume such prices represent errors in the data.

spread. Our goals is to isolate the municipal bond specific component of price movements, as information embedded in municipal bond trades is most likely related to these components.

4.2 Parameter Estimation

We estimate the market impact coefficients for individual bonds using the following panel regression:

$$r_{j,t} = \gamma_{n(j)} \cdot \theta_{j,t} + \epsilon_{j,t} \tag{16}$$

where $r_{j,t}$ is the return on bond j at the time of transaction relative to the most recent transaction or valuation price, $\theta_{j,t}$ is the dollar size of the trade (marked at the most recent previous price), and $\gamma_{n(j)}$ is the market impact coefficient for bond j's matrix category, $n(j)$.

Our estimates for these market impact coefficients are shown in Figure 1 and Table 1. Other than for low credit rating/low maturity bonds, market impact costs increase as bond supply thresholds decrease. Therefore, municipal bond supply appears to be a reasonable proxy for liquidity in the context of manipulation. As mentioned previously, this result is consistent with the findings of a 2014 ESMA consultation paper (ESMA (2014)). The ESMA analysis documents a relationship between bond issuance size and liquidity as measured by trading volume for sovereign and corporate bonds. In addition, market impact cost is higher for low credit rating and high maturity bonds. Maturity serves as the best proxy for liquidity in our sample, as this relationship is particularly strong.

We next turn our attention toward estimation of the index impact vector, $\vec{\Phi}_{\Im_n} = \Sigma^{-1}_{\Im_n,\Im_n} \vec{\Sigma}_{\{0\},\Im_n}$. Estimates of $\vec{\Sigma}_{\{0\},\Im_n}$ and $\Sigma_{\Im_n,\Im_n}$ based on trade covariances in equations (14) and (15) are shown in Table 2. One can see from these estimates that only 5 of 16 of the elements of $\vec{\Sigma}_{\{0\},\Im_n}$ are significantly different from zero. We conjecture that there is low power in the non-zero alternative since other factors which drive returns (including concurrent trades in other assets) create noise in these covariance estimates. As a result, we use an estimate for $\vec{\Sigma}_{\{0\},\Im_n}$ based on the return covariance in equation (14). One only needs to estimate $\delta.$ for each matrix category to employ this approach.

We estimate $\delta.$ (and $var(v.)$) using a GMM procedure which equates the estimates of the elements of $\vec{\Sigma}_{\Im_n,\Im_n}$ in equation (15). These estimates for $\delta.$ are shown in Table 1. Our GMM procedure yields estimates for $\delta.$, which all range within the prescribed bounds of 0 to 1. One can see that there is little variation in these estimates across matrix categories. Within the Kyle (1985) model, one would interpret this finding to mean that variability in liquidity across assets is driven by the volume of noise trading and not by the accuracy of information.

4.3 Manipulability Analysis

We now discuss our main findings regarding the effect of bond supply thresholds on the manipulability of our hypothetical municipal bond index. Our analysis assumes that the manipulator holds a \$10 million notional position in a municipal bond index derivative, which can be traded or redeemed costlessly. In addition, the manipulator distorts prices by executing eight trades in our 16 possible matrix pricing categories. We believe that a manipulator in the municipal bond market could reasonably execute eight trades within a 15-minute time span before these trades are revealed to the public. We have performed this analysis for both five and ten simultaneous trades and produced similar findings. One implicit assumption in our analysis is that the manipulator can distort an index with a given supply threshold by trading in any bond within our universe, whether or not this bond is a component of the index.

Table 3 shows results for our four bond supply thresholds in terms of M-factor, manipulator profit, and index distortion in percent. We see that manipulability decreases slightly as bond supply thresholds decrease. In addition, the manipulator executes trades in two matrix categories. Specifically, one trade is in a bond with high credit quality, low maturity, and at least \$100 million CUSIP-level issuance size; seven trades are in bonds with high credit quality, low maturity, and between \$100 million CUSIP-level issuance size and \$100 million deal-level issuance size. The manipulator trades in this manner irrespective of the index threshold. This finding is noteworthy because the manipulator trades in these categories whether or not they are included in the index. The reason is related to the high correlation between size-weighted returns of these matrix categories as shown in Table 4. Therefore, the manipulator selects matrix categories based on attributes other than their degree of correlation with (or inclusion in) the index.

In particular, the two optimal matrix categories offer the two lowest market impact costs for the manipulator. Hence, we find minimizing market impact costs to be a primary determinant of the manipulator's trading behavior. One can see from equation (14) that the manipulator's impact on the index is inversely related to this market impact cost (as represented by $\gamma.$). Therefore, decreasing market impact cost, all else equal, increases the manipulator's benefit while decreasing the cost of the manipulation.[17]

The decrease in manipulability of the index across lower thresholds is caused by a decrease in index concentration. It is clearly not caused by a shift in the manipulator's trading strategy to more liquid instruments as this strategy stays fixed across thresholds. However, the decrease in M-factor is slight and does not fall in proportion to the number of index components. The reason is again related to the high correlation between matrix categories. As a result, the "effective" concentration of the index does not decrease dramatically as more components are admitted.

5 Conclusion

In this paper, we developed a novel model of manipulation of indexes of illiquid assets. Manipulation becomes feasible with such an index in our framework since shifting prices of a few assets should move the index as a result of fair valuation. Our model offers a quantitative measure of manipulability of a particular index and a prediction about which particular assets would be targeted for manipulation. We applied our model to a specific hypothetical national municipal bond index with varying bond supply thresholds. Our model indicated that lower thresholds can decrease susceptibility to manipulation by decreasing the concentration of the index. We should add the caveat that our findings may not hold for other municipal bond indexes with greater concentration in terms of geographical region, maturity, rating, bond type, and purpose. We should also note that our empirical analysis was based on an ad hoc model of asset returns with coarse matrix categories. This empirical implementation could be improved using the actual matrix model employed by pricing vendors in this market. One could also extend the model to multiple periods as in Horst and Naujokat (2011) and Gallmeyer and Seppi (2000) to understand the dynamic strategy of a manipulator holding a

[17]This relationship only holds if the accuracy of information (as represented by $\delta.$) is unrelated to market impact cost (as represented by $\gamma.$) as we find in our estimates.

derivative whose "delta" could change over time.

Our framework is relevant to a number of current regulatory issues. For example, exchanges are seeking to expand the range of fixed-income index derivatives available for trading. An understanding of the potential manipulation of indexes becomes increasingly important as regulation expands its oversight to new exchange-listed derivatives and direct supervision of index providers.[18]

[18]The Financial Conduct Authority in the United Kingdom has begun to regulate eight benchmarks used to price financial instruments and contracts. The European Commission has also recommended the regulation of all benchmarks used for pricing in the EU. A principal stated goal of these new regulations is to deter manipulation.

Table 1: Market Impact ($\gamma_{\cdot}$) and Information Accuracy ($\delta_{\cdot}$) Estimates

This table shows estimates for market impact coefficients ($\gamma_{\cdot}$) and information accuracy ($\delta_{\cdot}$) parameters for 16 matrix categories using data from Jan-Dec 2014. The numbers after "Cat" (i.e., matrix category) represent the following. In the left (tens) digit: 1 means at least \$100 million CUSIP-level size; 2 means at least \$100 million deal-level but less than \$100 million CUSIP-level size; 3 means at least \$75 million but less than \$100 million deal-level size; 4 means at least \$20 million but less than \$75 million deal-level size. In the ones digit: 1 means high maturity (>15 years) and high rating (AA or above); 2 means high maturity and low rating (AA– or below); 3 means low maturity (≤15 years) and high rating; 4 means low maturity and low rating.

	Cat 11	Cat 12	Cat 13	Cat 14	Cat 21	Cat 22	Cat 23	Cat 24
Market Impact ($\times 10^{-7}$)	14.5	17.7	4.49	7.67	14.3	20.1	5.40	7.27
Information Accuracy	0.148	0.145	0.134	0.132	0.154	0.152	0.136	0.139

	Cat 31	Cat 32	Cat 33	Cat 34	Cat 41	Cat 42	Cat 43	Cat 44
Market Impact ($\times 10^{-7}$)	15.7	19.9	5.70	7.80	20.9	22.7	7.63	9.38
Information Accuracy	0.139	0.152	0.131	0.139	0.156	0.143	0.140	0.140

Table 2: Trade Covariances

This table shows sample covariance ($\times 10^5$) of daily average dollar transaction size for each of 16 matrix categories using data from Jan-Dec 2014. The last row is the sample covariance of daily average dollar transaction size with the return on our national municipal bond index with threshold of $20 million per offering at issuance. The numbers after "Cat" (i.e., matrix category) represent the following. In the left (tens) digit: 1 means at least $100 million CUSIP-level size; 2 means at least $100 million deal-level but less than $100 million CUSIP-level size; 3 means at least $75 million but less than $100 million deal-level size; 4 means at least $20 million but less than $75 million deal-level size. In the ones digit: 1 means high maturity (>15 years) and high rating (AA or above); 2 means high maturity and low rating (AA– or below); 3 means low maturity (≤15 years) and high rating; 4 means low maturity and low rating. For each row, p-values are shown in parenthesis below each covariance; ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

	Cat 11	Cat 12	Cat 13	Cat 14	Cat 21	Cat 22	Cat 23	Cat 24
Cat 11	1,428.8***							
	(0.000)							
Cat 12	153.9**	927.4***						
	(0.035)	(0.000)						
Cat 13	118.9	370.6**	7,320.9***					
	(0.563)	(0.025)	(0.000)					
Cat 14	41.5	157.9	-41.6	5,222.2***				
	(0.811)	(0.258)	(0.916)	(0.000)				
Cat 21	294.0***	255.5***	199.3	217.9	897.1***			
	(0.000)	(0.000)	(0.220)	(0.112)	(0.000)			
Cat 22	139.0***	147.1***	98.8	3.5	269.1***	491.8***		
	(0.009)	(0.001)	(0.412)	(0.973)	(0.000)	(0.000)		
Cat 23	147.3**	145.5***	543.8***	322.5***	259.6***	177.7***	590.4***	
	(0.011)	(0.002)	(0.000)	(0.004)	(0.000)	(0.000)	(0.000)	
Cat 24	194.1***	158.6***	360.6***	398.2***	191.0***	210.1***	359.4***	602.1***
	(0.001)	(0.001)	(0.006)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Cat 31	4.6	-87.3	349.7	177.4	178.8**	186.8***	212.4***	188.7***
	(0.967)	(0.333)	(0.167)	(0.407)	(0.043)	(0.004)	(0.003)	(0.009)
Cat 32	197.6**	141.8*	6.8	95.3	261.9***	226.8***	250.6***	158.4***
	(0.037)	(0.063)	(0.975)	(0.600)	(0.000)	(0.000)	(0.000)	(0.010)
Cat 33	173.6*	38.9	302.6	358.1**	237.5***	127.2**	328.9***	276.5***
	(0.060)	(0.602)	(0.148)	(0.042)	(0.001)	(0.019)	(0.000)	(0.000)
Cat 34	233.8**	132.3	126.5	282.0	159.0*	56.9	313.8***	379.8***
	(0.033)	(0.135)	(0.611)	(0.179)	(0.067)	(0.378)	(0.000)	(0.000)
Cat 41	257.8***	155.9***	143.7	146.1	260.4***	119.9***	199.4***	161.7***
	(0.000)	(0.001)	(0.294)	(0.207)	(0.000)	(0.001)	(0.000)	(0.000)
Cat 42	104.3**	125.6***	403.3***	158.2*	60.7	80.3***	111.3***	100.1***
	(0.030)	(0.001)	(0.000)	(0.086)	(0.112)	(0.004)	(0.000)	(0.001)
Cat 43	131.9**	125.4***	380.1***	308.1***	195.9***	114.7***	319.5***	279.6***
	(0.018)	(0.005)	(0.002)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)
Cat 44	96.0*	112.2***	248.8**	94.2	129.4***	146.3***	279.8***	226.9***
	(0.065)	(0.007)	(0.034)	(0.344)	(0.002)	(0.000)	(0.000)	(0.000)
Index Returns	-46.13	29.63	-139.45	-115.43	-112.60**	-68.52*	-98.64**	-102.29***
	(0.443)	(0.541)	(0.306)	(0.316)	(0.018)	(0.052)	(0.010)	(0.008)

	Cat 31	Cat 32	Cat 33	Cat 34	Cat 41	Cat 42	Cat 43	Cat 44
Cat 31	2,176.2***							
	(0.000)							
Cat 32	81.3	1,569.9***						
	(0.489)	(0.000)						
Cat 33	270.9**	102.7	1,487.1***					
	(0.017)	(0.290)	(0.000)					
Cat 34	166.5	186.4	250.4**	2,097.0***				
	(0.220)	(0.105)	(0.025)	(0.000)				
Cat 41	163.9**	79.2	122.6**	196.5***	636.2***			
	(0.028)	(0.212)	(0.047)	(0.007)	(0.000)			
Cat 42	86.0	47.6	42.2	93.5	134.4***	404.0***		
	(0.148)	(0.346)	(0.391)	(0.109)	(0.000)	(0.000)		
Cat 43	149.1**	151.7***	341.7***	273.4***	200.9***	89.0***	542.5***	
	(0.030)	(0.009)	(0.000)	(0.000)	(0.000)	(0.003)	(0.000)	
Cat 44	183.0***	171.4***	196.7***	175.1***	119.7***	72.4***	233.8***	471.0***
	(0.004)	(0.002)	(0.000)	(0.005)	(0.000)	(0.009)	(0.000)	(0.000)
Index Returns	-69.55	-102.84	-50.96	-26.59	-63.64	-62.56**	-126.89***	-34.23
	(0.349)	(0.102)	(0.407)	(0.715)	(0.112)	(0.050)	(0.001)	(0.322)

Table 3: Manipulability of National Municipal Bond Indexes

This table shows, for various supply thresholds: the M-factor, the manipulator's profit, total size of trading activity, and amount in percent by which the index would be distorted. We assume that the manipulator: 1) holds a \$10 million position in a municipal bond index derivative which can be liquidated costlessly; 2) the manipulator distorts prices by executing eight trades in 16 possible matrix pricing categories.

Size threshold	\$100 M CUSIP	\$100 M deal	\$75 M deal	\$20 M deal
M factor ($\times 10^{-9}$)	5.89	5.33	5.28	5.13
Manipulator profit	\$147,279	\$133,201	\$132,071	\$128,367
Index distortion (percent)	2.95%	2.66%	2.64%	2.57%
Number of components	1,348	68,587	86,841	203,319
Trade size	\$149,523	\$142,189	\$141,584	\$139,584

Table 4: Correlation between Matrix Category Returns

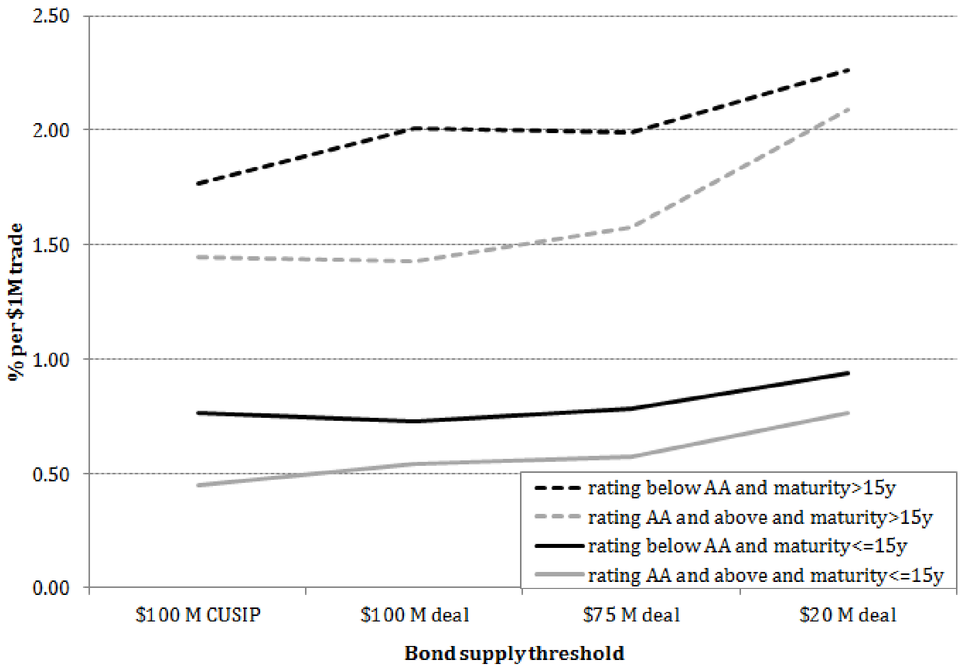
This table shows the correlation between size-at-issuance weighted returns for each matrix category using data from Jan-Dec 2014. The numbers after "Cat" (i.e., matrix category) represent the following. In the left (tens) digit: 1 means at least \$100 million CUSIP-level size; 2 means at least \$100 million deal-level but less than \$100 million CUSIP-level size; 3 means at least \$75 million but less than \$100 million deal-level size; 4 means at least \$20 million but less than \$75 million deal-level size. In the ones digit: 1 means high maturity (>15 years) and high rating (AA or above); 2 means high maturity and low rating (AA– or below); 3 means low maturity (≤15 years) and high rating; 4 means low maturity and low rating. All entries are significant at the 1% level.

	Cat 11	Cat 12	Cat 13	Cat 14	Cat 21	Cat 22	Cat 23	Cat 24
Cat 11	1							
Cat 12	0.956	1						
Cat 13	0.804	0.777	1					
Cat 14	0.777	0.791	0.918	1				
Cat 21	0.961	0.938	0.858	0.828	1			
Cat 22	0.957	0.970	0.835	0.832	0.980	1		
Cat 23	0.828	0.795	0.974	0.939	0.887	0.861	1	
Cat 24	0.811	0.810	0.956	0.963	0.872	0.871	0.982	1
Cat 31	0.956	0.935	0.840	0.806	0.981	0.971	0.865	0.849
Cat 32	0.953	0.967	0.831	0.828	0.972	0.991	0.857	0.864
Cat 33	0.824	0.790	0.971	0.935	0.884	0.857	0.998	0.980
Cat 34	0.813	0.804	0.956	0.956	0.874	0.869	0.984	0.995
Cat 41	0.961	0.942	0.851	0.824	0.996	0.982	0.881	0.867
Cat 42	0.954	0.967	0.828	0.831	0.974	0.991	0.858	0.865
Cat 43	0.820	0.788	0.971	0.936	0.882	0.855	0.998	0.981
Cat 44	0.803	0.795	0.954	0.960	0.865	0.859	0.983	0.995

	Cat 31	Cat 32	Cat 33	Cat 34	Cat 41	Cat 42	Cat 43	Cat 44
Cat 31	1							
Cat 32	0.963	1						
Cat 33	0.863	0.853	1					
Cat 34	0.852	0.867	0.982	1				
Cat 41	0.983	0.974	0.880	0.869	1			
Cat 42	0.963	0.989	0.854	0.867	0.976	1		
Cat 43	0.860	0.852	0.999	0.983	0.877	0.852	1	
Cat 44	0.841	0.857	0.982	0.996	0.860	0.861	0.983	1

Figure 1: Market Impact Cost for Matrix Categories

This figure shows estimates for market impact coefficients ($\gamma_{\cdot\cdot}$) as percentage of bond price change per \$1,000,000 trade for 16 matrix categories using data from Jan-Dec 2014. Each line represents a different rating/maturity category (in Table 1 notation, Cat ·2, Cat ·1, Cat ·4, and Cat ·3, from top to bottom), and each category on the x-axis represents a bond supply (issuance size) category. The x-axis categories are ordered in the order of decreasing bond issuance size (in Table 1 notation, Cat 1·, Cat 2·, Cat 3·, and Cat 4·, from left to right).



References

Aitken, M. J., Harris, F. H. and Ji, S. (2009). Trade-based manipulation and market efficiency: A cross-market comparison, *Working paper*.

Almgren, R. F. (2003). Optimal execution with nonlinear impact functions and trading-enhanced risk, *Applied Mathematical Finance* **10**: 1–18.

Cao, Y., Li, Y., Coleman, S., Belatreche, A. and McGinnity, T. (2013). A hidden Markov model with abnormal states for detecting stock price manipulation, *2013 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, IEEE, pp. 3014–3019.

Cao, Y., Li, Y., Coleman, S., Belatreche, A. and McGinnity, T. (2014). Detecting price manipulation in the financial market, *Working paper*.

Comerton-Forde, C. and Putnins, T. J. (2011). Measuring closing price manipulation, *Journal of Financial Intermediation* **20**(8): 135–158.

Cornell, B. (1997). Cash settlement when the underlying securities are thinly traded: A case study, *The Journal of Futures Markets* **17**(8): 855–871.

Diaz, D., Theodoulidis, B. and Sampaio, P. (2011). Analysis of stock market manipulations using knowledge discovery techniques applied to intraday trade prices, *Expert Systems with Applications* **38**(10): 12757–12771.

Dutt, H. R. and Harris, L. E. (2005). Position limits for cash-settled derivative contracts, *The Journal of Futures Markets* **25**(10): 945–965.

ESMA (2014). Consultation paper on MiFID II/MiFIR, *ESMA 2014/1570*. Available at http://www.esma.europa.eu/system/files/2014-1570_cp_mifid_ii.pdf.

Gallmeyer, M. and Seppi, D. (2000). Derivative security induced price manipulation, *Working paper*.

Horst, U. and Naujokat, F. (2011). On derivatives with illiquid underlying and market manipulation, *Quantitative Finance* **11**(7): 1051–1066.

Kumar, P. and Seppi, D. J. (1992). Futures manipulation with "cash settlement", *The Journal of Finance* **47**(4): 1485–1502.

Kyle, A. S. (1984). A theory of futures market manipulations, *in* R. Anderson (ed.), *Industrial Organization of Futures Markets*, Lexington Books, pp. 141–173.

Kyle, A. S. (1985). Continuous auctions and insider trading, *Econometrica* **53**(6): 1315–1335.

Ogut, H., Doganay, M. M. and Aktas, R. (2009). Detecting stock-price manipulation in an emerging market: The case of Turkey, *Expert Systems with Applications* **36**(9): 11944–11949.

Pirrong, C. (2001). Manipulation of cash-settled futures contracts, *Journal of Business* **74**(2): 221–244.