Intermediation in Over-the-Counter Markets

Nicola Branzoli*

E-mail: nbranzol@ssc.wisc.edu

September 29, 2012

Abstract

This paper studies the effect of intermediaries’ market power on asset prices and volumes traded in over-the-counter markets. I develop a search-and-bargaining model of trading and show that the market power of dealers reduces volumes traded and increases price volatility. I estimate the model using a new dataset containing all trades executed in a secondary market of municipal bonds whose liquidity plummeted in 2008, with significant consequences for investors wealth and issuers financing costs. The model provides a good fit of the data. I simulate the model to study the outcome of alternative market structures. I find that proportional fees significantly increase volumes traded with small effects on the level of prices. The introduction of a centralized market can have an ambiguous effect that depends on which side of the market is allowed to post orders. Volumes double and prices increase on average by 18% under buyers posting, but they fall 60% and 12% respectively under sellers posting.

JEL:

Key Words:

*Ph.D. Candidate - University of Wisconsin Madison. William H. Sewell Social Science Building 1180 Observatory Drive Madison, WI 53706-1393
1 Introduction

Intermediaries have a pivotal role in financial markets. They find buyers and sellers, hold inventories and influence prices and volumes. Their behavior affects issuers financing costs, investors wealth and, in extreme cases, can have significant consequences for the entire economy. Dealers improve the liquidity of assets acting as middleman, buying, selling and being compensated for matching the two sides of the market. However they may also exploit arbitrage opportunities generated by information asymmetries. In this case their behavior can increase price volatility and reduce volumes traded. The recent financial crisis has fueled the debate over their regulation. Ultimately, which role they play depends on their market power and the structure of the market in which they operate.

In this paper I analyze the effect of the market power of dealers in over-the-counter markets. I develop a model of asset trading to study how the market power of intermediaries affects asset prices and volumes. In the model, investors receive liquidity shocks that change their valuation of an asset over time. Once the investor valuation changes, the investor can adjust the amount of the asset owned by searching for a dealer willing to trade and bargaining over the price and the quantity traded. Dealers can trade with investors and in a competitive inter-dealer market. The time needed to find a dealer willing to trade generates the market power of dealers. This market power imposes a cost of trading that varies with investors idiosyncratic valuation of the asset at the time of trade. On one hand, the higher is the market power of dealers, the closer is the transaction price to the valuation of the investor involved in the trade. On the other hand, the lower is the market power of dealers, the closer is the transaction price to the price in the inter-dealer market. Thus higher market power of dealers is associated with higher price volatility. Moreover, if an investor expects that her valuation will have large changes in the future, she also expects to pay large costs of trading to the intermediary in terms of discounted price. Thus she chooses to buy an amount of the asset according to her expected valuation rather than her current valuation, to reduce these costs\(^1\). Thus higher market power of dealers is also associated with lower volumes traded.

\(^1\)The general intuition is that how the value of each trade is divided between agents involved in a meeting \textit{ex-post} determines the action taken \textit{ex-ante}: the amount bought. This interpretation is similar to Mortensen [1982] who study search intensity and not volumes. See Section 3 for a discussion.
I use this model to analyze trading behavior in the secondary market of Auction Rate Municipal bonds. These bonds are debt contracts issued by public administrations and tax-exempt institutions to finance a wide range of social services: financing utility companies, non-profit hospitals, student loans and the maintenance of public roads are only few examples. In 2008 these bonds provided roughly $60 billion to the public sector in the US.

Trading in the Auction Rate Municipal bond market has plummeted after the unravelling of the Great Recession, as shown in Figure 1, with significant consequences for investors’ wealth and issuers’ financing costs. Before the recent financial crisis, these bonds were traded in weekly auctions that allowed investors to trade directly with each other. Since 2008 these bonds are traded in a secondary market organized exclusively through financial intermediaries. The model is able to rationalize the increase in price dispersion and the fall in volumes traded observed after the failure of the centralized auction market. I estimate the structural model using a new dataset containing all trades executed in this market between 2009 and 2011. My aim is to measure the market power of dealers, its effect on volumes traded and how different policy interventions could restore the demand in this market.

In a preliminary analysis I test the main implications of the model by studying the relationship between the characteristics of bonds, prices and volumes. The results suggest the existence of two markets: a retail market for investors and an inter-dealer market involving trades between dealers. The characteristics of the bonds are not significant predictors of prices in the retail market, indicating that the main determinant of the prices in these trades are investors’ idiosyncratic valuations of the bond. On the contrary, in the inter-dealer market prices are significantly related to observable characteristics of the bonds, showing that dealers set prices according to the real value of these assets. I also construct several measures of the liquidity of these bonds and show that they positively affect the distribution of trade sizes, i.e. investors buy larger amounts of more liquid bonds. Thus the evidence from this market is consistent with the implications of the model.

I then estimate the structural model. The results show that the model is able to match the main

\footnote{ McConnell and Soretto [2010] and Han and Li [2010] provide a detailed analysis of the events that led to the failure of the centralized auction market.}
moments characterizing the distribution of prices, frequency of trading and the size of trades. The estimated bargaining power of dealers is 0.61, which suggests that dealers extract more than half of the surplus generated by each trade. Thus dealers exercise significant market power, which implies potentially large losses in terms of volumes traded.

I use the estimated model to simulate the distribution of volumes and prices under four alternative market structures. I consider the case in which investors pay intermediation fees, either fixed or proportional to the amount traded. The results indicate that volumes could increase by 60%-70% depending on the fee structure. The large quantitative effect of these market reforms hinders on the significant bargaining power that dealers exercise in this market, thus the effects on the total volumes traded are large. When investors pay a predetermined fee, so that the cost of trading is independent of traders’ valuations at the moment of trade, investors expect lower losses from intermediation. Thus investors buy and sell more and the volumes traded increase.

Furthermore I consider an environment in which buyers post prices and quantities they want to trade, and one in which sellers post prices and quantities. In the former case, Bertrand competition among buyers drives the price they post to their reservation value so sellers obtain all the surplus from a trade. In the latter, Bertrand competition among sellers drive the price they post to their reservation value so buyers obtain all the surplus. Volumes traded would roughly double under buyers’ posting, while they would fall by 61% under sellers’ posting. The intuition is the following. When buyers are posting, thus revealing their preferences first, sellers extract all the gains from trade. Therefore all investors tend to buy more of the asset when they can, to avoid being buyers again in case their valuations of the asset will increase. The demand of the bond increases and so does its price. However investors with valuation higher than the average valuation have also a higher probability of being sellers in the future, thus they expect lower costs of adjusting their portfolio in the future. Thus higher types take on more of the increase in the demand than low types, and the distribution of asset holdings becomes more dispersed, increasing the volumes traded. A specular argument gives the intuition for why volumes traded would fall under posting by sellers.

This paper makes two main contributions. First, I measure the distortions induced by the market
power of dealers, and in particular the effect on the liquidity of these bonds. Second, I contribute to the recent literature that develops and estimates structural models of investments in decentralized markets. In this paper, I estimate a search-and-bargaining model with endogenous volumes traded built on Lagos and Rocheteau [2009] using publicly available trade data. The structural estimates allow me to discuss the effects of several policy intervention that could increase volumes traded in a municipal bond market that was severely hit by the recent financial crisis.

The results show that there are potentially large gains by reforming this market. Possible ways to avoid the inefficiency associated with the market power of dealers are restricting intermediaries ex-ante to fixed or proportional fees, as it is already done in other sectors, or introducing an electronic exchange for bonds. I also show that market centralization is not necessarily preferable to intermediation, since a market where sellers post terms of trades implies a significant reduction in volumes traded with respect to the market with intermediation.

Relation with existing literature - This paper is related to three strands of literature analyzing: i) trading in the secondary market of municipal bonds ii) models of asset trading in over-the-counter markets and iii) the effect of market microstructure on asset liquidity. Harris and Piwowar [2006] and Green et al. [2007] analyze prices in the secondary market for municipal bonds. I confirm their results on the large costs of trading in this market and the significant market power exercised by dealers. I also extend their analysis by studying the effect of the liquidity of bonds on trade sizes and by estimating a structural model of investment with search frictions.

This paper is also related to the search-and-bargaining approach of trading in over-the-counter markets, initiated by Duffie et al. [2005]. The structural model presented in Section 3 is built on the model of asset trading with unrestricted asset holdings and search frictions developed by Lagos and Rocheteau [2009]. I modify the model to adapt it to the market of Auction Rate Municipal bonds, derive new

---

3 I collected the dataset used in this paper writing a web-scraper that downloaded trade information from a public website. Section 2 discusses how the data was collected.

4 An electronic exchange for corporate bonds has already been introduced by the London Stock Exchange in 2010 (see the official announcement), and it has been recently discussed for the US corporate bond market.

5 An alternative approach, not pursued in this paper, models over-the-counter markets as networks. In particular Gofman [2011] finds that bargaining in network markets may lead to an inefficient allocation of assets. My paper finds a similar result in a decentralized, search-and-bargaining market.
results about the effect of dealers’ market power and show the identification the structural parameters. Thus this research is also related to the recent branch of literature estimating structural model of asset trading with search frictions, like Feldhütter [2012] and Gavazza [2012]. These papers estimate search-and-bargaining models in over-the-counter markets with unit demand, i.e. each investor either sell or buy the asset without adjusting the quantity traded. My paper represents the first attempt to structurally estimate a search model that allows agents to bargain over both the price and the quantity they want to trade.

Finally, this paper addresses the question of how volumes and prices react to changes in the structure of a market. There is a large theoretical literature that studies how market structure affects prices and volumes but I am not aware of any attempt to quantify it. The closest theoretical paper is Madhavan [1992] who analyzes price formation under two trading mechanisms: a quote-driven system where dealers post prices before order submission and an order-driven system where traders submit orders before prices are determined. To my knowledge, this paper is the first attempt to use a structural model to produce a quantitative answer to the effect of the two market structures on investors’ behavior.

The rest of the paper is organized in 6 Sections. Section 2 shows the results of the reduced form analysis supporting the assumptions of the model. Sections 3 and 4 present the model and its identification. Sections 5 presents the result and the simulations under alternative market structures. Section 6 concludes.

2 Overview of the market and empirical evidence

In the first part of this Section I briefly describe the history of the Auction Rate Municipal Bond market and the sequence of events that led to the massive withdraw of liquidity during the recent financial crisis. McConnell and Soretto [2010] and Han and Li [2010] provide a detailed analysis of these events.

In the second part of the Section I analyze the two main variables that characterize trading in asset markets: prices and volumes.
I analyze bonds that are traded exclusively through dealers in the period from 2009 to 2011. These bonds constitute the majority of this market. First I study prices at which investors trade with dealers and prices at which dealers trade among themselves. The results show that there are two markets in which these bonds are traded: a retail market where dealers trade with investors, and an inter-dealer market where dealers trade among themselves. I show that the characteristics of bonds are not significant predictors of prices at which investors trade in the retail market. More than 40% of the variation of prices remains unexplained even after controlling for bonds’ fixed effect. On the contrary, the characteristics of bonds are significant explanatory variables of price variation for inter-dealers trades. Moreover, for these trades an hedonic regression of prices on measures of risk and returns of the bonds explains more 70% of the price variation.

Second, I analyze the effect of various measures of the liquidity of these bonds on the distribution of trade sizes. I show that bonds that are more liquid have also larger trades. These facts are consistent with the model of asset trading with search frictions and endogenous asset holdings described in Section 3.

2.1 Brief history of the market

Auction Rate Municipal bonds are relatively safe assets with low returns, mostly below 2%. Issuers are non-profit organizations like municipalities, utilities, and student loans facilities. The first Auction Rate Municipal Bonds was issued by the Warrick County in Indiana in 1985 to fund the local gas and electric company, however Auction Rate Municipal bonds became a major financing instrument for the public sector after 2000. Outstanding debt issued in Auction Rate Municipal Bonds increased from $16 billion in the first quarter of 2002 to $30 billion by the end of the fourth quarter of 2003. At the end of the 2007 the market was worth more than $60 billion.

Their secondary market is organized in weekly or monthly auctions and in an over-the-counter market between these auctions. The auctions provide a centralized market to facilitate trading. However these bonds have the following rule: the auction fails in case of excess supply. After an auction failure, buyers

---

6In the rest of the paper I will refer to the auctions as the centralized market and to the market organized around dealers’ intermediation as the over-the-counter market.
and sellers cannot trade in the centralized market but they have to wait for the next auction, hoping that it will succeed, or contact a registered dealer in the over-the-counter market to find a trader.

In February 2008, the four largest investment banks organizing the centralized market\(^7\) declined to buy any excess supply in the auctions. From that day onward, trade volumes have fallen dramatically and new issuance has been reduced to zero. Since 2008, 95\% of Auction Rate Municipal bonds are traded only through dealers because their auctions always fail\(^8\).

The Municipal Rule Making Board (the government agency responsible for regulating the market for municipal bonds, MSRB henceforth) has recently released all individual trade data of Auction Rate Municipal bonds\(^9\). Trades must be executed through registered dealers\(^10\) and dealers are required to report trades to the MSRB.

### 2.2 Description of the data

There are 777 bonds in the dataset. I observe all trades executed in the secondary market of Auction Rate Municipal bonds from the 31st of January 2009 to the 31st of December 2011, for a total of 9,841 trades\(^11\). Each observation is an executed trade and all trades are executed through dealers. The data collected indicate whether the trade involved an investor buying from a dealer, selling to a dealer or was executed between two dealers. I have also collected several characteristics of the assets, such as the identity of the issuer, ratings, maturity, purpose of the bond, tax status and so on. Tables 1 and 2 describe the information collected about all bonds.

---

7 Citigroup, UBS, Morgan Stanley and Merrill Lynch.
8 McConnell and Soretto [2010] and Han and Li [2010] analyze the events that led to the collapse of the market and the determinants of auction failure.
9 The data were obtained directly from the MSRB website that distributes transaction information. The dataset was assembled through a web-spider that automatically collects all data. Figure 2 provides an example of the web-page containing information about bonds and transactions.
10 There are roughly 1000 dealers registered to trade in this market.
11 The total number of Auction Rate Municipal bonds in my dataset is 1722. From these bonds I remove 40 bonds for which auctions succeed every week with at least one buy order executed by an investor different from the issuer of the bond. I have also removed 880 bonds because they were recalled by the issuer during the period analyzed, and 22 bonds with maturity year before 2020. In March 2008, class action lawsuits were filed against several of the large banks alleging that they have misinformed investors in this market about the liquidity of these bonds. Some of these banks have settled the lawsuits offering to buy at par auction rate securities they had sold to their retail clients. These trades are not part of market-driven transactions, and so are not considered in this analysis. In the Appendix I explain how I have identified these trades.
These bonds are subject to the same trading rules\textsuperscript{12} and are relatively more homogeneous than bonds traded in the corporate bond market in terms of observable characteristics. First, there are a total of 144 issuers who are all tax-exempt institutions. The issuers in this market are mostly municipalities, utilities companies and hospitals, who are backed by public administrations. Second, most of these bonds are insured by a third party who guarantees the bond in case of issuer’s default. Third, all bonds have a coupon rate which is defined as a fixed difference above market interest rates (mostly the LIBOR rate). In the period between 2009 and 2011 these rates have been very low and stable, thus the coupon rates paid by Auction Rate Municipal bonds have also been low and stable for the entire period. For example, 75\% of the bonds pay a coupon rate between 0.5\% and 1\% per year, as shown in Figure 3.

Nonetheless there is large price heterogeneity charged by intermediaries to investors. There are two prices at which dealers trade: the bid price, i.e. the price at which the intermediary buys the bond, and the ask price, i.e. the price at which he sells it. If the intermediary had no cost of intermediation, this difference would represent his profits. However dealers have to find some investors to trade the bond. This intermediation activity is particularly important, since there is no centralized market where investors can post their orders to trade and wait to be contacted by new investors\textsuperscript{13}.

Table 3 reports summary statistics of the trades. A key characteristic of the data is the large variation of both trades’ size and price\textsuperscript{14}. The trade amount can be as small as $5,000 and as big as $152,750,000. Prices are also significantly heterogeneous, going from 40\% of the par-value traded to a 100\%. The median price for investors is just below 90\%. Figure 4 shows the probability and cumulative distribution functions of bid and ask prices. The graphs show significant mass on both tails of the support. In fact the 25th and 75th percentile are 81 and 100 respectively for the ask prices and 83 and 100 for the bid price\textsuperscript{15}. Prices in trades between dealers are much less dispersed. For example, the 25th percentile is 90\%, ten percentage points higher then the 25th percentile of trades with investors.

\textsuperscript{12}See the Rule G-17 issued by the MSRB that regulates all dealers operating in the Auction Rate Municipal bond market.

\textsuperscript{13}For example Bloomberg, a standard source of asset prices, does not provide quotes for these bonds, therefore investors and dealers not only do not know whether there are investors interested in trading the bond, but also what would be a market price for it.

\textsuperscript{14}There are two prices at which dealers trade: the bid price, i.e. the price at which the intermediary buys the bond, and the ask price, i.e. the price at which he sells it. I report summary statistics for buy and sell orders separately.

\textsuperscript{15}The truncation at a 100 is due to the characteristics of these bonds, that have weekly auctions in which traders can buy and sell the at par, i.e. at a price of 100. Selecting the bonds with only auction failures results in a distribution of prices that is truncated at 100.
2.3 Bonds’ prices: investors’ trades and the inter-dealer market

This dispersion in the distribution of prices can be determined by investors’ idiosyncratic valuations or by preferences over assets’ characteristics. The lack of information about the identity of investors and dealers prevents me from controlling for traders’ idiosyncratic characteristics. However I can use the trade size as a proxy for investor’s type.

In the corporate bonds’ market, Goldstein et al. [2007] and Bessembinder et al. [2009] show that trades bigger than $100,000 are predominantly institutional trades such as banks or mutual funds. To investigate this issue I first show the scatter plot of order size and price in Figure 5 pooling all bonds. These graphs show the large variation in prices even conditional on the size of the trade. For example, for orders between $100,000 and $70,000,000 the range of prices observed (maximum price minus minimum price) is the same, roughly 50%.

Price variation in this market is so large that conditioning on one bond’s characteristic at a time helps understanding price dispersion in this market. The coupon rate is a likely determinant of the value of these bonds for the investors\textsuperscript{16} given the relative low risk associated with issuers’ default in this market\textsuperscript{17}. Most of these bonds pay a fixed difference over market returns (mostly the LIBOR rate) thus bonds paying higher coupon rates also provide a higher stream of expected payments for the entire duration of the debt. Figure 6 shows the evidence about trade prices conditioning on the return of the bond. The left graph shows the scatter plot of trades (par-value traded and price) for bonds which were paying, at the moment of the trade, less the 0.5%, while the right graph shows trades for the bonds that have an average return less then 0.3% over the entire sample\textsuperscript{18}. We should expect these trades to be at relatively similar prices. However, conditional on the size of the trade, the range of prices is still in the order of 25 percentage points. Similar results hold for all the other observable characteristics: conditioning on bonds issued by the same organization, with the same maturity or rating and so on\textsuperscript{19}.

To quantify the effect of all characteristics on prices paid by investors in the secondary market of Auc-

\textsuperscript{16}The timing of these payments is different from bond to bond and it depends on the frequency of the auctions, which is mostly weekly. All coupon rates are expressed on an yearly level. Investors selling the bond between payment dates receive from the dealer the price of the trade plus the interests accrued until the day they trade.

\textsuperscript{17}In the period under study, there have been no default on payments from issuers of Auction Rate Municipal bonds.

\textsuperscript{18}I find similar evidence using bonds paying high coupon rates.

\textsuperscript{19}Figures 6 and 7 show the plots for the most traded bond and shows that, even conditioning on a single asset, there can be more then 30 percentage points difference in the prices of trades in the same bond and of the same amount.
tion Rate Municipal bonds, I estimate the following hedonic regression:

\[ p_{i,b,t} = \beta_0 + \beta_1 q_{i,b,t} + \beta_2 \text{Coupon Rate}_{b,t} + \beta_3 \text{Rating Category}_{b,t} + \beta_4 \text{Days to Maturity}_{b,t} + \beta X_t X_t + \epsilon_{i,t} \]  

(1)

where \( p_{i,b,t} \) is the price of the \( i \)-th transaction of bond \( b \) at date \( t \) and \( q_{i,b,t} \) is a dummy variable for the category of the quantity traded in such transaction\(^{20}\). The other variables included are the coupon rate at the moment of trade, dummies for the rating category of the bond at the time of the trade and the number of days to the maturity of the bond. Finally \( X_t \) contains several other characteristics of bonds and market indices that might be important for investors’ valuation of the bond\(^{21}\). I have estimated two specifications of the Eq.1: an OLS model (with and without bonds’ fixed effect) and a logit model (with and without bonds’ fixed effect) to account for the fact that prices are between 0 and 100.

The results are reported in Table 4. None of the bonds’ characteristics are significant in explaining prices paid by buyers. The only significant variables for sellers are the tax status and the insurance status, although both are significant only in one of the specifications considered. Finally, note that the \( R^2 \) of the model is around 0.4 for both buyers and sellers, indicating that more then half of the variation of prices is not explained by bonds’ observable characteristics. The fit of the model increases to 0.6 if I include bonds’ fixed effect. Thus controlling for unobservable assets characteristics still leaves unexplained a large fraction of the variation of prices charged by dealers in the secondary market.

I perform the same analysis described above for trades in the inter-dealers market. To identify inter-dealers trades that are initiated by an investor buying or selling the asset I follow the methodology proposed by Green et al. [2007] and consider the remaining trades between dealers as trades in the

\(^{20}\)To define trade categories, I follow two references: Goldstein et al. [2007] and Bessembinder et al. [2009] who find that trades bigger than $100,000 are predominantly institutional trades; Feldhütter [2012] uses 6 classes to fit the data in the corporate bond market ($0–10,000, $10,000–50,000, $50,000–100,000, $100,000–500,000, $500,000–1,000,000, and more than $1,000,000).

\(^{21}\)More precisely, \( X_t \) contains issuer’s fixed effect, the number of program dealers, the frequency of reset of the coupon rate, and aggregate variables that control for the risk-free rate (the interest paid by 3-month US Treasury bills and their effective yield in the secondary market) and bonds’ market activity. To control for market activity I include several bonds’ indices constructed by Bank of America, Marrill Linch and Moody. These indices are publicly available from the Federal Bank of Saint Louis. I have included 19 indices that are related to market activity in fixed income securities.
inter-dealer market. A detailed description of their methodology is given in the Appendix B.

Table 5 reports the results of estimating the model in Eq.1 using inter-dealers trades. All bonds’ characteristics are significant predictors of prices, with sign consistent with economic theory. The coupon rate increases the expected stream of payments received by bonds’ holders, thus it has a positive effect on the price paid by dealers. It is significant at 1% level in the model without bonds’ dummies, while it becomes insignificant when I include a dummy for each bond. This loss of significance is likely a consequence of the constant mark-up of the coupon rate over the LIBOR rate. All other characteristics, maturity, insurance status and tax status, are significant in the specification with bonds’ fixed effects. The fit of the regression is significantly higher than the one using trades with investors. These results indicate that prices in the inter-dealer market reflect the value of each bond and I use this evidence to support the assumption in the structural model that dealers can trade in an inter-dealer market where bonds are priced competitively.

### 2.4 Trade size and liquidity

Defining the liquidity of an asset is a complex task. Intuitively, a liquid asset is an asset that can be traded easily. However finding a single measure of liquidity of an asset has been proven to be a hard task\(^\text{22}\). Thus in this Section I use several proxies of liquidity to understand its effect on the distribution of trade sizes.

The first two variables used measure of number of trades by buyers and dollar volumes bought by retail investors in each bond. I construct the total dollar amount bought by investors for each bond and each month and use its lagged value as a proxy for the liquidity of the bond at the time of the trade. To construct the variable related to the number of trades, I consider all trades in the over-the-counter market since the first failed auction in 2008 for each bond. I calculate the total number of buy orders executed by investors at or above the average price of the bond in each month\(^\text{23}\). A third measure for the illiquidity of the bonds is based on the amount of the bond owned by dealers.

---

\(^{22}\) See the review of the literature in Amihud et al. [2005].

\(^{23}\) An asset is illiquid if a seller accepts large discounts to sell it. Thus I condition this measure on relatively high prices to exclude orders executed at significant discounts, which would measure the illiquidity of a bond rather than its liquidity. Conditioning on other price levels such as the number of buy orders executed at par give similar results.
I can only track the aggregate inventories of the dealers since I do not have information about dealers identity. Moreover, some bonds have been issued before the start of my data, so I can track dealers inventories only imperfectly. Thus I decide to start constructing the total amount of the bond held by dealers since the first auction failed. For each day, I construct the total amount bought and sold by investors, and consider the difference as the change in dealers’ inventories. The total sum of these changes since the first auction failed give the aggregate amount of the bond held by dealers.

Finally I consider a measure of liquidity based on Kyle [1985] and implemented first by Brennan and Subrahmanyam [1996]. Kyle [1985] analyzes price determination in a market with informed and uninformed traders. Uninformed trader submits a normally distributed quantity $u$ and the informed trader submits an optimal demand $x$ knowing that her order reveals her information about the true value of the asset $v$. The market maker sets the price depending on the order flow earning zero profits\(^{24}\). Kyle shows that there exists a linear equilibrium in which the market maker sets the price equal to:

$$p_t = E(v|u_t + x_t) = p_{t-1} + \lambda_t(u_t + x_t)$$

where $\lambda_t$ measures the price impact per unit of order. Based on this idea, Brennan and Subrahmanyam [1996] propose to estimate $\lambda$ by regressing trade-by-trade price change $\Delta p_t$ on the signed quantity traded\(^{25}\). The estimate of the effect of the quantity traded on changes in price of the asset measures the price impact. I slightly modify their approach. Instead of using the entire quantity traded, I consider only the amount bought or sold by a dealer, thus considering the quantity that 'enters' the inter-dealer market\(^{26}\). Thus $\lambda$ is a measure of illiquidity of the bond since it measures the price impact of the quantity traded in the inter-dealer market. I estimate one $\lambda$ for each bond.

\(^{24}\)Thus the market maker works as a competitive market.

\(^{25}\)The expression 'signed quantity' means that buy orders are considered positive quantities while sell orders are considered negative quantities. The regression also includes $D_t - D_{t-1}$ where $D$ is equal to 1 for buy orders and $-1$ for sell orders.

\(^{26}\)If the buy and the sell transaction size match exactly, the trade is excluded from the computation.
Using these four measures of liquidity, I estimate the following model:

\[
q_{i,b,t} = \beta_0 + \sum_{i=1}^{4} \beta_{m_i} m_{i,b,t} + \beta_2 \text{Coupon Rate}_{b,t} + \beta_3 \text{Rating Category}_{b,t}^I + \\
+ \beta_4 \text{Days to Maturity}_{b,t} + \beta_{X,t} X_t + \epsilon_{i,t}
\]

(2)

where the \(m_{i,b,t}\)'s correspond to the four measures of bond liquidity, the other variables are self-explanatory and \(X_t\) contains several characteristics of bonds and market indices described in Footnote 21.

Table 6 summarizes the results. Measures of liquidity have a positive effect on the average trade size while measures of illiquidity have a negative effect on the average trade size. The effect of the number of buy orders executed since the first auction failed have a positive effect on trades, although it is weakly significant. The total dollar volume bought in the previous month is instead strongly significant. The measures of illiquidity have opposite effects: the higher is the amount of inventories of a bond, the harder should be to sell it, and so the average trade size is smaller. The Kyle’s measure of illiquidity have also a negative effect. When all measures are included, dollar volumes becomes insignificant. Finally the effect all measures of liquidity decrease in absolute values.

To conclude, this Section documents the effect of bonds liquidity on the sizes of trades executed by retail investors. It shows that investors tend to trade larger amounts for more liquid bonds. The next Section describes a structural model consistent with this evidence.

3 Theory

In this Section I describe the model, which is built on Lagos and Rocheteau [2009]. The model characterizes the optimal behavior of two types of agents, investors and dealers, operating in a secondary market for an asset. There are two differences between investors and dealers: first investors can trade only with dealers while dealers can trade in an inter-dealer market. This assumption captures the institutional restriction that requires investors to trade only through dealers’ intermediation. Sec-
ond, investors have an intrinsic valuation of holding the asset that changes over time, causing them to become buyers and sellers at random dates. Dealers instead derive utility from a bid-ask spread they charge to trade it. This assumption captures the intermediation activity of dealers\textsuperscript{27}. The rest of this Section presents the model, while all the details are reported in Appendix A.

### 3.1 Investors’ preferences and the inter-dealer market

Time is continuous and unbounded. There is a unit mass of agents in the economy with heterogeneous valuations for the asset. An investor holding an amount \( a \) of the asset at instant \( t \) receives an instantaneous utility \( \epsilon_t u(a) = \epsilon_t^{a 1 - \sigma} \) where \( \epsilon_t \) is the agent’s current idiosyncratic valuation for the asset. The assumption that an investor derives utility directly from the amount of the asset owned is standard in the literature of search in asset markets\textsuperscript{28} and can be interpreted in different ways. In the case of municipal bonds, \( \epsilon_t u(a) \) can stand for the utility that the investor derives from the asset for several reasons: the coupon rate paid by the bond, tax-advantages, a savings plan, or simply she might enjoy the participation to local projects such as student financing programs or a non profit hospital\textsuperscript{29}.

Investors’ valuation \( \epsilon \) changes according to a Poisson process with arrival rate \( \alpha \). In the rest of the paper I shall refer to a shock that changes investor’s utility as a liquidity shock. An investor hit by a liquidity shock draws a new type \( \epsilon \) from a distribution with c.d.f and p.d.f \( \Pi(\epsilon) \) and \( \pi(\epsilon) \) respectively\textsuperscript{30}. When an investor changes her type she wish to change the amount of the asset she owns.

Market decentralization is modelled through a random arrival rate of trading possibilities at the same rate as the liquidity shocks, \( \alpha \). This assumption captures the idea that trading depends ultimately on the arrival rate of other investors and that their arrival rate depends on the frequency of their liquidity shocks\textsuperscript{31}. The arrival rate of trading possibilities can be interpreted as the time taken by a dealer to

\textsuperscript{27}The existence of intermediaries is assumed, i.e. I do not investigate why they operate in this market or how their number is determined in equilibrium. The interested reader should look at Wong and Wright [2011] and Babus [2012] for the analysis of how and how much intermediation arises endogenously in decentralized and network markets.

\textsuperscript{28}See for example Duffie et al. [2005].

\textsuperscript{29}Another possible interpretation is that the investor is able to borrow at an interest rate \( r \) and use the asset as capital that produces the consumption good that gives her utility \( \epsilon_t u(a) \). Note also that investors’ heterogeneity enters in a log-linear way into the utility function. This assumption is made mainly for tractability: it implies that higher types have a higher marginal valuation for any quantity of the asset they own, keeping the coefficient of risk aversion constant.

\textsuperscript{30}Thus investors’ type follows a compound Poisson process with jump size distribution \( \pi(\epsilon) \).

\textsuperscript{31}Note that the probability that an investors receives a liquidity shock and find a dealer willing to trade in the same moment is zero, making the time elapsed between a liquidity shock and a trade uncertain from the point of view of an
execute an order once he is contacted by an investor interested in buying or selling the asset.

Let \( V(\epsilon_t, a) \) be the discounted utility attainable by an investor of type \( \epsilon \) at time \( t \) with asset holdings \( a \).

The value function satisfies:

\[
V(\epsilon_t, a) = \mathbb{E} \left\{ \int_t^{T_\alpha} e^{-r(s-t)} \epsilon_s u(a) ds + e^{-r(T_\alpha-t)} [V(\epsilon_{T_\alpha}, a') - p(a' - a)] \right\}
\]

where \( T_\alpha \) is the next (random) date in which the investor can adjust her asset holdings, \( a' \) is her new asset position and \( p \) is the price at which she trades the amount \( a' - a \). If the amount traded is positive, the investor is a buyer and pays the price \( p(a' - a) \), if it is negative she is a seller and receives a payment.

The first term of Eq.3 captures the utility of an investor with valuation \( \epsilon \) keeping her asset holdings \( a \) between time \( t \) and the next time she meets a dealer, which happens after a random period of length \( T_\alpha - t \). The second term captures the utility she gets from the trade with the dealer at a price \( p \), and the continuation value with the new asset position. The price and the amount traded is the outcome of a bargaining game described below. The expectation operator is taken over the distribution of types the investor can draw between time \( t \) and \( T_\alpha \) and the distribution of \( T_\alpha \).

The evidence provided in Section 2 shows that dealers are highly interconnected in the municipal bond market and that they can trade shares of municipal bonds in a relatively competitive environment where prices represent the value of bonds. Thus I model the inter-dealer market as a competitive market where dealers can buy and sell the asset at price \( \xi \).

A centralized asset market for dealers simplifies the bargaining problem between investors and dealers described below for the following reason. Since dealers can trade in the inter-dealer market at a competitive price, the terms of trade between a dealer and a seller will not depend directly on the dealer’s asset holdings involved in the trade, but on the price \( \xi \) in the inter-dealer market. The price in this market depends on investors valuations of the bond and the how frequently investors can trade it. See investor. A more detailed model where investors can trade continuously with dealers, who can accumulate inventories to equate supply and demand, would significantly complicate the analysis with bargaining and unrestricted asset holdings. The analysis of dealers’ inventory problem is left for future developments of the analysis of the data.

\[32\] Which is exponentially distributed expected value \( 1/\alpha \).
below for a detailed derivation.

3.2 Meeting between investors and dealers

Once an investor meets a dealer willing to trade, the parties start an alternating offer game over the quantity traded and the price. The parties choose the quantity traded and the price in order to maximize the joint surplus of the trade. The solution of the game can be characterized by the generalized Nash-bargain problem given by:

$$\max_{p,a'} [V(\epsilon_t, a') - V(\epsilon_t, a) - p(a' - a)]^{1-\eta} [(p - \xi)(a' - a) - c]^{\eta}$$

(4)

where $\eta$ is an exogenous parameter that measures the bargaining power of the dealer and $c$ is the dealer’s fixed cost of executing the trade. The maximization is subject to the participation constraints:

$$V(\epsilon_t, a') - V(\epsilon_t, a) - p(a' - a) \geq 0$$

(5a)

$$(p - \xi)(a' - a) - c \geq 0$$

(5b)

The solution of the bargaining problem is:

$$a' = \arg\max [V(a', \epsilon_t) - \xi a']$$

(6a)

$$p(a' - a) = (1 - \eta)[\xi(a' - a) + c] + \eta(V(\epsilon_t, a') - V(\epsilon_t, a))$$

(6b)

Eq. 6b shows the main source of identification of dealers market power. Each bond has a price in the inter-dealer market. If dealers have no bargaining power, the price paid by the investors is equal to this price plus the cost of executing the trade. If, instead, dealers have full bargaining power, the prices observed for trades with investors does not depend on the price in the inter-dealer market (as long as the there are gains from trade). Thus, the lower is the bargaining power of dealers, the higher is the correlation between the bid and ask prices and the price at which dealers trade in the inter-dealer
market. In Section 4 I discuss the details of the implementation.

3.3 Equilibrium

Substituting the solution of the bargaining problem (Eqs.6a and 6b) into the value function of investors (Eq.3), I obtain the following expression:

\[
V(\epsilon_t, a) = \mathbb{E} \left\{ \int_t^{T_\alpha} e^{-r(s-t)} \epsilon_s u(a) ds + e^{-r(T_\alpha-t)} [(1 - \eta) \max_{a'} [V(\epsilon_{T_\alpha}, a') - \xi(a' - a)] + \eta V(\epsilon_{T_\alpha}, a)] \right\}
\]

Eq. 7 highlights the main inefficiency associated with dealers’ market power. It shows that the problem of an investor is the same as the one in an environment in which she meets a dealer with probability \(\alpha\) and can adjust her asset holdings at a price \(\xi\) with probability \(1 - \eta\), thus obtaining the entire surplus of her trade, and with probability \(\eta\) the investor does not adjust her portfolio and waits for another arrival of a possibility to trade.

The inefficiency due to the bargaining in decentralized markets was first noticed by Mortensen [1982]. Mortensen analyzes a decentralized economy where individuals of two types choose how much to search for a mate of the opposite type to produce a consumption good. When they meet, the agents split evenly the surplus of the match. How the surplus of the match is divided affects the search intensity of unmatched agents, which affects the aggregate number of matches. Mortensen shows that bargaining over the match surplus causes an inefficient level of search intensity and thus an inefficient level of matches. The intuition is that how the value of each meeting is divided between agents involved in a meeting ex-post determines the action taken ex-ante, the search intensity. Thus, for example, an individual with no bargaining power will spend less resources in searching then an individual that can get all the surplus from a meeting with a trading counterpart. The same intuition carries over to this environment. The fact that at the moment of future trades, investors will give part of the gains from trades to the dealers cause them to choose an inefficient level of investment to incur in lower costs of trades. Thus, all things equal, investors trade less in today’s trades to hold an amount of the asset
optimal for their expected valuation\textsuperscript{33}.

Given the above interpretation, Eq.7 can be rewritten recursively as:

\[ V(\epsilon_t, a) = \bar{u}(\epsilon_t, a) + e^{-r(T_\kappa - t)}E \left\{ \max_{a'} [V(\epsilon_{T_\kappa}, a') - \xi a'] \right\} \quad (8) \]

where:

\[ \bar{u}(\epsilon_t, a) = \left[ \frac{(r + \kappa)\epsilon_t + \alpha \bar{\epsilon}}{r + \kappa + \alpha} \right] \frac{a^{1-\sigma}}{1 - \sigma} \]

and \( \kappa = \alpha(1 - \eta) \) and \( \bar{\epsilon} \) is the mean value of investors types \( \epsilon \)’s.

I can plug in the solution into the bargaining problem in Eq.6a to obtain:

\[ a(\epsilon) = \arg \max_a [\bar{u}(\epsilon, a) - r\xi a] \quad (9) \]

Eq.9 allows to derive the optimal asset holdings for an investor of type \( \epsilon \):

\[ \frac{\partial \bar{u}(\epsilon, a)}{\partial a} = \bar{\epsilon} a^{-\sigma} = r\xi \quad (10) \]

Where \( \bar{\epsilon} = \left[ \frac{(r + \kappa)\epsilon + \alpha \bar{\epsilon}}{r + \kappa + \alpha} \right] \).

I can now summarize the effect of the bargaining power on volumes traded in the next proposition:

**Proposition 1** Total volumes traded decrease with \( \eta \)

The intuition for this proposition is based on the optimal asset holdings of the investors. An increase in \( \eta \), the bargaining power of dealers, increases the cost of trade for the investors. As future adjustments in their portfolio become more costly, investors choose to invest an amount determined by their expected valuation, to avoid these costs. Thus the current valuation of the asset becomes less important in determining the amount they invest. Hence every time they get a liquidity shock and trade, they trade smaller amounts.

\textsuperscript{33}Mortensen [1982] also discusses how correct this inefficiency. The simulations discussed below are partly related to his proposals.
Equation 10 can be combined with the equilibrium condition that the sum of asset holdings by all types should equate the total supply of the asset, $A$:

$$\int a(\epsilon)\pi(\epsilon)d\epsilon = A$$  \hspace{1cm} (11)$$

This gives also an expression for the price of the asset in the inter-dealer market:

$$r\xi = \left[\frac{E(\bar{\epsilon}^{1/\gamma})}{A}\right]^\gamma$$  \hspace{1cm} (12)$$

Dealers’ bargaining power also increases the variability of prices. In Appendix A I prove this formally under some conditions on the variance of the intermediation costs $c$ and using the delta method to approximate the variance of the change in investors’ valuation of the asset.

The intuition is that prices are a convex combination of dealers’ transaction costs and the change in the valuation of investors, where the weights depend on the bargaining power. The higher is the bargaining power of the dealer, the closer is the price to the idiosyncratic change in the investor valuation rather than the price in the inter-dealer market. Thus if the variance of intermediation costs is small enough relative to the distribution of investors valuations, the variability of prices also increases with the bargaining power of dealers.

4 Econometric specification and identification

The method used to estimate the model described in the previous section combines steps from the Simulated Method of Moments as studied by McFadden [1989] and Pakes and Pollard [1989] and Indirect Inference as in Gourieroux et al. [1993] and Gourieroux et al. [2002].

I analyze trading data from 2009 to 2011 and divide the sample into “periods” of 3 months. Thus I observe 12 ”periods” of trading, which a consider random samples of draws from the same distribution of liquidity shocks and consumers valuations.

In the next Section I discuss the identification of the model, describing the mapping between the parameters and the moments used to estimate them. In Section 4.2 contains the discussion about the
implementation.

4.1 Identifying moments

Under the assumption of a Poisson arrival rate of liquidity shocks, the distribution of the number of trades is informative about the value of \( \alpha \). In the implementation I include both the mean number of trades and their variance\(^{34}\). The estimation of the bargaining power is based on the co-movement between the prices in the inter-dealer market and prices paid by investors. Consider the \( i - th \) transaction in the bond \( b \). The following equilibrium restriction provides the intuition:

\[
p_{i,b} = (1 - \eta)[\xi_b + c_{i,b}] + \eta \left[ \frac{V(\epsilon_{i,b}, a') - V(\epsilon_{i,b}, a)}{a' - a} \right]
\]  

Consider the two extreme cases of \( \eta = 1 \) and \( \eta = 0 \):

\[
\eta = 1 : \quad p_{i,b} = V(\epsilon_{i,b}, a') - V(\epsilon_{i,b}, a)
\]

\[
\eta = 0 : \quad p_{i,b} = \xi_b + c_{i,b}
\]

On one hand, if dealers have full bargaining power, prices paid for retail trades depend on investors’ idiosyncratic valuations. On the other hand, if dealers have no bargaining power, the price at which investors buy and sell the bond is equal to the price at which dealers trade in the inter-dealer market, plus a random disturbance which represent the cost of executing the trade for the dealer. I assume that the cost \( c_{i,b} \) of executing the trades are i.i.d. across trades and bonds\(^{35}\). The identification is based on the fact that inter-dealer trades are directly observable in the data as those trades between dealers that are not associated with a buyer or a seller\(^{36}\). Thus \( \xi_b \) is observable. The main issue is what measure

\(^{34}\)I exclude trades in the inter-dealer market from the computation of the spells since they do not represent the arrival of neither buyers nor sellers.

\(^{35}\)Conversations with professionals trading fixed income securities confirm that the costs of executing the trades are independent of the bond traded but are not independent of the type of transactions. Unit costs are generally lower for larger trades. Harris and Piwowar [2006] find that transaction costs decrease with trade size and do not depend significantly on trade frequency. To address this issue I will exclude larger trades from the estimation. See Section 4.2 for details.

\(^{36}\)Green et al. [2007] propose three different methodologies to identify trades in the inter-dealer market. The results
of correlation to use. One possibility is the classic Pearson’s correlation. However in the Monte Carlo simulations I find that this measure of correlation is not a very informative about $\eta$, especially when I have a small number of trades in a market. The reason is following. Unless $\eta$ is exactly 0, the price in the interdealer market is a convex combination of $\xi$ and the idiosyncratic valuation of the asset of the investor. If variations in $\xi$ are significantly larger then the variations in the idiosyncratic valuations, the Pearson’s correlation is high no matter what is the level of $\eta$. Instead, if variations in $\xi$ are significantly smaller then the variations in the idiosyncratic valuations, the Pearson’s correlation is low independently of $\eta$. Thus a simple measure of correlation would be weekly informative for intermediate levels of $\eta$. A better measure is the coefficient of the regression of prices in the retail market on the price in the inter-dealer market. The idea is that if changes in the inter-dealer price across bonds correlate with the prices paid in the retail market of investors, then dealers have low bargaining power. If, instead, changes in $\xi_b$ across bonds are not correlated with prices paid in the retail market of investors, then dealers have high bargaining power, i.e. dealers are buying at prices that reflect investors valuations at the moment of the trades rather then the valuation of the asset.

The estimation of the investors’ intertemporal elasticity involves more details on the model. Eq.10 implies that:

$$\log(a' - a) = \frac{1}{\sigma} \log(\xi) + \log(\tilde{\epsilon}^{1/\sigma} - \tilde{\epsilon}^{1/\sigma})$$

Thus a measure of correlation between the log of trade sizes and the price in the inter-dealer market can be used as informative moment about $\sigma$. I use the coefficient of the regression of the log of trade sizes on the price in the centralized market instead of the simple Pearson’s correlation for the same reason discussed in the above paragraph. The intuition for why this moment is informative about the level of risk aversion is the following. Buying the asset is like entering a lottery for the investor: the return is uncertain because of the random time she might spend with asset holdings that are not optimal. The price of the lottery in terms of the numeraire good is $\xi$ (plus a random component determined by her present type and the type she was when she last adjusted her asset holdings). By exploiting the

---

reported in Section 5 are obtained using the augmented round-trip transaction rule. Its details are reported in Appendix B. I plan to study the robustness of the results to this assumption in the next version of the paper.
effect of the price in the centralized asset on the distribution of asset holdings I thus identify \( \sigma \). For given distribution \( \pi(\epsilon) \), investors are risk-averse if variations in the price in the centralized market have small effects on trade sizes: it takes a large change in cost of the lottery to induce traders to invest more. On the contrary, investors are not risk-averse if small changes in \( \xi \) have large effects on the distribution of trade sizes.

Investors’ heterogeneity is assumed to come from an exponential distribution with parameter \( \lambda_b \) specific to each bond. The optimal amount invested by investors of type \( \epsilon \) can be expressed as a function of the observable price in the centralized market and investors’ valuation only:

\[
a(\tilde{\epsilon}') - a(\tilde{\epsilon}) = \frac{\tilde{\epsilon}'^{1/\sigma} - \tilde{\epsilon}'^{1/\sigma}}{(r\xi)^{1/\sigma}}
\]  

(17)

The more dispersed is the distribution of trade sizes for a given bond, the more heterogeneous are investors’ preferences. To get the intuition, suppose that \( \sigma \) is close to 1. Assuming that \( \epsilon \) is exponential with parameter \( \lambda \) implies that the difference between draws from a distribution \( \pi(\epsilon) \sim \text{Exp}(\lambda) \) is Laplace \( (0, \frac{1}{\lambda}) \)\(^{37} \). The maximum likelihood estimator for the parameter \( \lambda \) of a Laplace \( (0, \frac{1}{\lambda}) \) distribution is \( \frac{1}{N} \sum_{i=1}^{N} |x_i| \). Thus for values of \( \sigma \) close to 1, the mean of the (absolute value of the) trade size will be informative about the distribution of investors’ heterogeneity \(^{38} \).

Finally, the discount factor is traditionally hard to estimate, so I set it equal to the average risk-free rate in the economy during the period.

\(^{37}\)The Laplace distribution is similar to the distribution of trades, as shown in the fit of the model. Note also that the exponential distribution is known for describing the time between events in a Poisson process. The interpretation used here is different and it is based on the second theorem of the extreme value theory, known as Pickands–Balkema–de Haan theorem, see Balkema and de Haan [1974] and Pickands [1975]. The theorem shows that, under mild conditions, the distribution of draws from the right tail of a distribution \( F \) converges to the generalized Pareto distribution. The exponential distribution belongs to the generalized Pareto family. The theorem complements the first theorem of the extreme value theory which shows that the maximum of a sequence of random draws is distributed according to the generalized extreme value distribution. Thus the exponential distribution can also approximate the probability distribution of a set of the largest draws of a random variable.

\(^{38}\)The choice of the exponential distribution is convenient because it requires estimating one parameter for each bond. However it implies that both the mean and the variance of investors’ heterogeneity increase with \( \lambda_b \). A more flexible specification could consider \( \epsilon \) as the sum of two components. The first is specific of each bond and affects the mean of the valuation of all investors in the same bond in the same way. The second affect each investor independently of the bond. Thus the 'type' of investor \( i \) holding bond \( b \) can be parametrized as:

\[
\epsilon_{i,b} = \theta_b + \nu_i
\]  

(18)

The specification assumed in the main text fits the data well. I plan to estimate the more flexible specification in future developments of the project.
4.2 Estimator

The reduced form results presented in Section 2 give mixed evidence on the relation between observable bonds’ characteristics and the distribution of prices. On one hand, the coupon rate, the time to maturity of the bond, the insurance and tax status do not significantly affect them. On the other hand, some dummies capturing issuers’ fixed effect are significant covariates of prices. I discuss the estimation of the model considering each bond as a separate market and investigate empirically whether alternative definitions of markets, i.e. all ARS bonds and bonds with the same issuer, give different results.

I consider the arrival rate of investors and the distribution of investors’ valuation as specific to the market\textsuperscript{39}. I divide the parameters in two sets: one that I consider market-specific and one that is not. In Table 7 I summarize the main parameters of the model, whether they are market-specific and the moments used to estimate them.

The simulated method of moments involves matching a set of observed and simulated moments. The SMM procedure selects the structural parameter vector that minimizes a metric of the distance between these two sets of moments.

The algorithm searches over the values of $\eta$, $\alpha_m$, $\lambda_m$ and $\sigma$ to minimize the distance between the observed moments and the simulated moments. The moments used are:

1. the coefficient of the regression of the prices in the retail investors’ market on the prices in the inter-dealer ($\beta_{\xi,p}$);

2. the mean and the variance of the number of trades per market per period ($\mu_N, \sigma_N$);

3. the mean and the variance of the trade sizes for each bond divided by the total outstanding amount of the bond (from Eq.17, $\mu_q, \sigma_q$);

4. the coefficient of the regression of the mean of the logarithm of the trade sizes on the prices in the centralized market (from Eq.16, $\beta_{\xi,q}$).

\textsuperscript{39}The specification with heterogeneous arrival rates of liquidity shocks can be interpreted as a model in which investors with different liquidity needs sort themselves into seemingly similar bonds as in Vayanos and Wang [2007]. I have investigated whether observable bonds’ characteristics can explain the number of trades through a Poisson model for the number of trades on bonds’ observable characteristics. The results are similar to the those about prices in the retail market.
Formally, the vector of moments for market \( m = 1, \ldots, M \) is given by \( \psi^m \equiv (\beta_{\xi,p}, \beta_{\xi,q}, \mu^{m}_q, \mu^{m}_N, \sigma^{m}_N, \sigma^{m}_q) \) and defined over \( \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^3 \). The vector of moments \( \Psi \) used is obtained by stacking the vectors \( \psi^m \) from each market. The vector of parameters is given by \( \Theta \equiv (\eta, \sigma, \alpha_m, \lambda_m) \). Note the dimension of \( \Psi \) is \( 2 + 4M \) while the vector of parameters is of dimension \( 2 + 2M \). Thus the number of parameters is lower then the number of moments to be matched.

The parameters are found minimizing the criterion function:

\[
\Lambda = (\Psi(\Theta) - \hat{\Psi})' W (\Psi(\Theta) - \hat{\Psi})
\]  

where \( W \) is the weighting matrix. In this version of the paper \( W \) is the identity matrix\(^{40}\).

The algorithm is implemented in the following way:

1. initialize the values of all parameters;

2. fix two sets of random draws from an exponential distribution, one with a number of draws equal to the number of markets times the number of periods considered, and one with an arbitrarily large number of draws from the distribution of investor heterogeneity (fixing the draws is used to minimize the sampling error);

3. minimize over \( \alpha_m \), the distance between the mean and the variance of the number of trades of the observed trades per bond and period and the one generated by the model. Given the value of \( \alpha_m \), compute the number of trades for market \( m \) in each period;

4. taking as given the value of \( \sigma \) and \( \eta \), repeat for each market the following steps:

   (a) for each trade take two values from the distribution of investors’ valuations and compute their values given \( \lambda_m \). These two values represent the valuation of the investor at the moment of trade and the last time she was able to adjust her asset holdings;

   (b) use Eq.17 to compute the size of each trade\(^{41}\);

---

\(^{40}\)This approach should be improved by using a weighting matrix for each moment.

\(^{41}\)The value of \( E(\tilde{e}^{1/\sigma}) \) in Eq.17 is approximated numerically.
(c) compute the mean and the variance of the trade size and take the distance from the observed moments. Repeat until convergence for market $m^{42}$;

5. minimize over $\sigma$ the distance between the coefficient of the regression of the logarithm of trade sizes (in absolute values) and the price of bonds in the centralized market and the same parameter estimated using simulated data;

6. minimize over $\eta$ the distance between the coefficients of the regression of the prices in the retail market on prices in the inter-dealer market using the observed and the simulated data. Go back to step 4 and repeat until convergence of all moments.

I have investigated numerically the robustness to different initialization values$^{43}$.

5 Results

Table 8 reports the estimates from the simulated method of moments for three specifications of the model. The three specifications differ in what I consider a market: all Auction Rate Securities bonds, all bonds from the same issuer and the each bond.

The bargaining power of dealers and the intertemporal elasticity of substitution of investors are the same for all markets, therefore the table reports the point estimates of these parameters$^{44}$. Investors’ preferences and the arrival rate of trading possibilities, instead, are different across markets, to account for assets’ heterogeneous characteristics$^{45}$. Therefore the table reports the mean and the variance of the distribution of these parameters.

The results are very similar across specifications, suggesting that auction rate municipal bonds can be

---

$^{42}$Convergence is defined up to an error. In this version I have set that the sum of the distance between the observed and the simulated moments should not be higher then 0.01.

$^{43}$As starting values, I considered: $\eta \in \{0, 0.5, 1\}$, $\sigma \in \{0.5, 1, ..., 2.5\}$. These give me 15 combinations of starting values for the non-market specific parameters. For the market specific parameters, I have computed the minimum and the maximum of the number of trades and the size of the trades, considering both buy and sell orders. Then I have initialized the $\alpha_m$’s and the $\lambda_m$’s with these values. These give me 4 combinations of starting values for the market specific parameters. The total number of starting point investigated is thus 60.

$^{44}$I plan to add standard deviation of the estimates in the next version of the paper.

$^{45}$I could have parametrized the dependence of these parameters on bonds characteristics. The estimates would have been useful to investigate how bonds’ characteristics affect liquidity and investors’ valuation. I plan to investigate this issue in the future.
treated as assets traded in one single market where trades possibilities arrive at the same rate for all bonds and the heterogeneity in the number of trades is essentially due to randomness. It also suggests that the results are not sensitive to the choice of the period considered.

The point estimate of the parameter $\eta$ is ranges from 0.58 to 0.61, implying a significant bargaining power on the dealers’ side. To interpret this number consider its meaning in the bargaining game underlying it. The alternating offer game à la Rubinstein described in Section 3.2 assumes that the bargaining power of investors depends, among other things, on the time necessary to contact another dealer during the bargaining stage. Thus in a market where investors have two wait a significant amount time before finding another dealer to trade with, dealers should extract a significant fraction of the surplus generated by a trade. The high estimate of the dealers’ bargaining power captures this characteristic of the auction rate municipal bond market. Similar results are found by Green et al. [2007], who analyze the determinants of the bid-ask spreads in the municipal bond market in general (i.e. not only for auction rate securities). They also find that dealers exercise substantial market power in this market.

The estimates of $\sigma$ are between 1.01 and 1.11, which implies an intertemporal elasticity of substitution between 0.99 and 0.90. There is a wide set of estimates of the intertemporal elasticity of substitution in the literature and these estimates fall in this range. Differences are mainly driven by the access of consumers to financial markets. My results are slightly above the values found in the literature for individuals with access to financial markets. For example Vissing-Jorgensen [2002] estimates an intertemporal elasticity of substitution between 0.8 and 1 for bondholders using data from the U.S. Consumer Expenditure Survey.

Finally there is no direct interpretation the bond-specific estimates. Thus in the next section I asses whether these estimates can fit the observed distribution of volumes and prices.

5.1 Goodness of fit

I asses the goodness of fit of the estimates by comparing the distribution of some relevant variables generated by the model with the one observed in the data. The variable chosen are the ones also ana-

---

46 For CRRA utility function, the intertemporal elasticity of substitution is given by $1/\sigma$.
47 See the discussion in Guvenen [2006].
lyzed in Section 2: trades sizes, prices and the number of trades per bond. I report the results for the model estimated considering the bonds with the same issuer as a single market. The other specifications give similar results. Figures 8-10 report these distributions.

The graphs shows that the model estimated on each bond is able to replicate the main features of the empirical distributions. In particular, the exponential distribution for consumer heterogeneity seems to be a good approximation, given the estimate of \( \sigma \). Indeed this assumption implies a Frechet-type distribution generated by the model for the trade sizes, which is a good first approximation of the data. The simulated distribution of prices captures also the seemingly bimodal distribution of observed prices, with one mode around 85 and one at 100. The data, however, are spread on the entire support, while the model predicts that prices should be more concentrated on the tails of the support, especially on the lower part of the distribution. This is likely a consequence of the assumption of a constant bargaining power for all trades. Green et al. [2007] shows that the bargaining power depends also on the characteristics of trades. For example dealers exercise an higher bargaining power in trades involving more then one dealer. My model does not account for this heterogeneity and thus partly misses some of the determinant of prices.

Finally the distribution of the number of trades generated by the model captures the main characteristics of data, namely the size of the support and the declining trend. However it seems to over-estimate the probability of low number of trades.

Table 9 shows that the 25th, the mean and the 75th quantile of the observed and simulated distributions are also very similar.

To formally assess the fit of the model, I run a series of Kolmogorov-Smirnov test for these variables. More precisely, for each market with more then 10 trades, I test whether the observed and the simulated distribution of prices and volumes come from the same underlying distribution. For the number of trades, instead, I compare the observed and simulated distribution for all bonds. The last column of Table 9 reports the results, showing the mean of Kolmogorov-Smirnov test statistics for trade size and prices.

All quantiles in the distributions generated by the model match closely those observed in the data. However the model overestimate the 25th quantile of the price distribution, as suggested also by Figure
The error is 1.25 percentage points of the par value traded, which is not negligible. However, the rest of the statistics fits the data exactly at the second decimal digit. Moreover the Kolomogorov-Smirnov test do not reject the null hypothesis of the same distribution for observed and generated trades for prices and trade sizes for all bonds. Next I discuss the prediction of the model on how the possible market reforms could increase volumes trades, and their effect on prices.

5.2 Counterfactual Simulations

In this Section I compare the distribution of trade sizes and prices under three alternative market structures that avoids bargaining.

5.2.1 Trading under predetermined fees

The first scenario is a market in which investors trade at the equilibrium prices in the inter-dealer market and pay a fee to the brokers to place their orders. Normally, brokerage fees can be based on either a percentage of the transaction, a flat fee or a combination of the two.

*Proportional fees* - The effect of a fee proportional to the transaction is straightforward: it increases the price at which buyers buy the bond and lowers the price at which sellers sell it. However this fee structure may still be desirable for investors, since it makes the prices independent from their valuation at the moment of the trade.

The first two columns of Table 10 reports the results for fees of 1 to 3% of the trade amount. The results shows a significant increase in the total volumes traded, which is between 60 and 67% of the observed values, while the effect on prices would be small. The driver of the magnitude of these results is the large value of the bargaining power for dealers. As shown in the previous section, the bargaining power of dealers exacerbate trading frictions because it allows dealers to capture more than half of the gain from trades every time they can trade. A proportional fee fixed ex-ante, instead, increases the gain from trade kept by investors, therefore induce them to invest more in the bonds.

*Fixed fees* - The third column of Table 10 shows the results under the assumption of a fixed fee. I

---

48 Brokers’ profits from trading in fixed income securities comes from bid-ask spreads both in the municipal and corporate bond market, so there is no typical fee for these securities. Thus I have taken these fees as typical range fro trading fees in the stock market.
have used an indicative value of $5,000 since there is no typical fee for fixed income securities. The Table shows that volumes would increase only by 4%.

The effect of a fixed fee have two conflicting effects on the amount of volumes traded: on one side it prevents some investors to adjust their asset holdings, since only those that expect a gain from trade higher then the fee will decide to trade. This reduces the number of trades and thus the overall liquidity of the bond. On the other hand, a fee fixed ex-ante guarantees higher gains from trade to investors with large changes in their valuations, increasing the amount they buy and sell each time. The results from the simulation shows that the two conflicting effects almost compensate each other.

5.3 Buyers and sellers posting terms of trades

The second set of simulations looks at the market equilibrium after introducing an electronic exchange for bonds where buyers and sellers can post quantities and prices. Computing the equilibrium of a game in which strategic investors post terms of trades can be very complicated, as shown for example by Goettler et al. [2005]. Thus, to simplify the analysis, I make two assumptions. First I consider two separate cases: one in which only buyers post their terms of trades and one in which only sellers post their terms of trades. Second, I assume that an investor, either the buyer or the seller, post the optimal amount she wants to trade and the minimum price at which she wants to trade such amount. This would be the equilibrium prices if investors would compete à la Bertrand with other investors of the same type (for example if there is more then one buyer of a type $\epsilon$). Thus, sellers capture all trading surplus when buyers post prices and quantities, and buyers capture all trading surplus when sellers post terms of trades. These assumptions make the outcome comparable with the previous simulations, in which investors always adjust their quantities optimally, while keeping the model tractable.

The fourth and fifth columns of Table 10 reports the results. When buyers post terms of trades, total volumes more than double while they fall by 63% when are sellers those who post. To interpret the result note that investors are better off being sellers under prices posting by buyers, so that they can capture the surplus from their trades. Thus the total demand of the asset increases and the prices of the bonds increase on average by 18%. This is because buyers loose all the surplus from trading. However investors with valuation higher then the average valuation have also a higher probability of being sellers
and a larger amount of the asset to sell in the future. However they are better off then investors with valuation below average, who are more likely to be buyers in the future. Thus higher types take on more of the increase in the demand than low types, and the distribution of asset holdings spread out, increasing also the volumes traded. A graphical representation is given in Figure 11. Under sellers posting we have exactly the opposite effect and the total volumes traded decrease.

6 Conclusions

In this paper I study the effect of dealers’ bargaining power on prices and volumes in over-the-counter markets. I estimate a search-and-bargaining model build on Lagos and Rocheteau [2009] to quantify dealers’ market power and the effect of dealers’ intermediation on volumes traded in the Auction Rate Municipal bond market. The results show that dealers extract more then half of investors’ gains from trade and that this market power reduces volumes traded. I simulate the model under alternative market structures to assess possible solutions to this problem. I find that introducing predetermined fees proportional to the amount traded can increase volumes traded by 60%, while fixed fees have small effects. I also find that an electronic exchange for bonds can have potentially large impact on volumes traded, but only if we allow buyers to post their orders. Under this structure, volumes could double and prices would increase on average by 18%. However if we allow sellers to post their orders, volumes could fall as much as 60% and prices would decrease by 12%. These results can be useful in guiding policy interventions aiming at restoring the demand this market, which still suffers from the turmoil caused by the Great Recession.

This paper represents the first attempt to estimate a structural model of trading in over-the-counter markets with endogenous trade volumes. I show that the model can fit the data reasonably well. However the identification relies on two key assumptions that it will be interesting to investigate in the future. First, I model the inter-dealer market as competitive market. Thus I ignore a potentially important determinant of prices: dealers’ inventories. Second, I consider the measure of investors in this market

49This assumption is supported by the large number of inter-dealers trades that are not related to the investors’ buying or selling and to the fact that for these trades bonds’ characteristics affect prices in an economic meaningful way. On the contrary I show that bonds’ characteristics are not significantly related to the prices paid by and to investors.
as exogenous. It’s reasonable to expect that, as volumes traded increase, more traders would turn their attention to this market. I leave the analysis of these issues for future research.

References


7 Tables & Figures

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>Date (month, day, year, hour and minute) of the trade; All trades must be reported within 15 minutes of the execution</td>
</tr>
<tr>
<td>Par amount</td>
<td>Par value traded, i.e. the nominal vale of the bond traded at maturity. The par-value of a bond is the value of the bond at maturity, i.e. bond holdings of par value 100,000$ entitle the owner to a payment of 100,000$ from the issuer at the date of maturity of the bond. Therefore if a trade reports a buy order by an investor of 100,000$ at price of 97, it means that an individual bought shares of the bonds with total par value of 100,000$ and paid at the moment of trade 97,000$.</td>
</tr>
<tr>
<td>Price</td>
<td>Percentage of the par amount traded paid by the buyer</td>
</tr>
<tr>
<td>Trade Type</td>
<td>Identifier of a transaction between a dealer and a seller, a dealer and a buyer, or between two dealers</td>
</tr>
<tr>
<td>Bond</td>
<td>The bond traded</td>
</tr>
</tbody>
</table>

Table 1 – Transaction Data: this table describes the data collected about transactions executed in the over-the-counter market of Auction Rate Municipal bonds.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cusip code</td>
<td>A 9-character alphanumeric code that uniquely identifies a bond</td>
</tr>
<tr>
<td>Issuer description</td>
<td>The identity of the issuer, identified as the first 6 digits of the Cusip code. For most bonds, the name also reports the geographic location of the municipality or public institution issuing the bond</td>
</tr>
<tr>
<td>Total Par Amount</td>
<td>The total amount the bond issued</td>
</tr>
<tr>
<td>Issuance and Maturity</td>
<td>The date of issuance and maturity of the bond</td>
</tr>
<tr>
<td>Program Dealers</td>
<td>The identity of the dealers who handled the initial placement in the secondary market and that usually facilitate trades in the secondary market</td>
</tr>
<tr>
<td>Reset period</td>
<td>Frequency of reset of the coupon rate</td>
</tr>
<tr>
<td>Rating</td>
<td>Ratings by Fitch and S&amp;P</td>
</tr>
</tbody>
</table>

Table 2 – Bond Data: this table describes the data collected about bonds’ characteristics.

<table>
<thead>
<tr>
<th>Type of order</th>
<th>N.Obs.</th>
<th>Traded Amount in million of $</th>
<th>Traded Price % of the par-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>25th prct.</td>
<td>Median</td>
</tr>
<tr>
<td>Investor buy</td>
<td>2649</td>
<td>0.15</td>
<td>0.75</td>
</tr>
<tr>
<td>Investor sold</td>
<td>2435</td>
<td>0.10</td>
<td>0.5</td>
</tr>
<tr>
<td>Inter-dealer trades</td>
<td>4757</td>
<td>0.005</td>
<td>0.125</td>
</tr>
</tbody>
</table>

Table 3 – Summary statistics of trades
<table>
<thead>
<tr>
<th>Variable</th>
<th>Buyers’ trades</th>
<th>Sellers’ trades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coupon Rate</td>
<td>−0.19*</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.24)</td>
</tr>
<tr>
<td>Days to maturity</td>
<td>−0.25</td>
<td>−27.82</td>
</tr>
<tr>
<td></td>
<td>(3.19)</td>
<td>(38.79)</td>
</tr>
<tr>
<td>Ratings’ categories †</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Tax Status</td>
<td>−0.21</td>
<td>−1.05</td>
</tr>
<tr>
<td></td>
<td>(1.28)</td>
<td>(2.06)</td>
</tr>
<tr>
<td>Insurance Status</td>
<td>2.30</td>
<td>−2.77</td>
</tr>
<tr>
<td></td>
<td>(2.17)</td>
<td>(6.56)</td>
</tr>
<tr>
<td>Cusip Fixed Effect</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>N. Obs.</td>
<td>2606</td>
<td>2606</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.43</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Table 4 – Reduced-form estimates of the hedonic regression for investors: The dependent variable is the price of the transactions executed with retail investors. The model also contains issuer’s fixed effect, the number of program dealers, the frequency of reset of the coupon rate, the interest paid by 3-month US Treasury bills and their effective yield in the secondary market and 19 bonds’ indices constructed by Bank of America, Merrill Lynch and Moody publicly available from the Federal Bank of Saint Louis. OLS estimates. *: p-value < .1  **: p-value < .05  ***: p-value < .01. †: I use 3 ratings agencies, and divide ratings into 3 categories for each rating (9 categories). This row shows the number of significant categories at 5% level.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Dealers’ trades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coupon Rate</td>
<td>0.18***</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
</tr>
<tr>
<td>Days to maturity</td>
<td>2.80*</td>
</tr>
<tr>
<td></td>
<td>(1.50)</td>
</tr>
<tr>
<td>Ratings’ categories †</td>
<td>8</td>
</tr>
<tr>
<td>Tax Status</td>
<td>2.57***</td>
</tr>
<tr>
<td></td>
<td>(0.49)</td>
</tr>
<tr>
<td>Insurance Status</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>(0.48)</td>
</tr>
<tr>
<td>Cusip Fixed Effect</td>
<td>0</td>
</tr>
<tr>
<td>N. Obs.</td>
<td>3783</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Table 5 – Reduced-form estimates of the hedonic regression for dealers: The dependent variable is the price of the transactions executed in the inter-dealer market. The exogenous variables not shown are the bonds characteristics described in Table 4. OLS estimates. *: p-value < .1  **: p-value < .05  ***: p-value < .01. Standard errors clustered at bond’s level. †: I use 3 ratings agencies, and divide ratings into 3 categories for each rating (9 categories). This row shows the number of significant categories at 5% level.
<table>
<thead>
<tr>
<th>Measure of bond liquidity</th>
<th>Dependent Variable: trade size (in million)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Number of trades by buyers</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
</tr>
<tr>
<td>Amount bought by investors in the previous month</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Aggregate dealers’ inventories</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Kyle’s λ</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Issuer Fixed Effect</td>
<td>1</td>
</tr>
<tr>
<td>N. Obs.</td>
<td>2467</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Table 6 – Estimates of the effect of liquidity measures on trades’ size hedonic regression for dealers: The dependent variable is the size of the trade. The number of trades by buyers counts all trades in the over-the-counter market for each bond since the first failed auction in 2008. I consider the total number of buy orders executed by investors at or above the average price of the bond in each month. See Section 2.4 for details on the construction of the Kyle’s λ. The exogenous variables not shown are the bonds characteristics described in Table 4. OLS estimates. * : p-value < .1 ** : p-value < .05 *** : p-value < .01. Standard errors clustered at bond’s level.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Market-specific</th>
<th>Moments</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>Arrival rate trading possibilities</td>
<td>Yes</td>
<td>Number and variance of the number of trades per period</td>
</tr>
<tr>
<td>( \eta )</td>
<td>Bargaining power of dealers</td>
<td>No</td>
<td>Correlation between demeaned inter-dealer prices and demeaned bid and ask prices</td>
</tr>
<tr>
<td>( \pi(\epsilon) )</td>
<td>Distribution of investors’ heterogeneity</td>
<td>Yes</td>
<td>Exponential distribution, estimated using the distribution of trade sizes</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>Inter-temporal elasticity of substitution</td>
<td>No</td>
<td>Mean of the inter-dealers’ market price</td>
</tr>
<tr>
<td>( r )</td>
<td>Discount factor</td>
<td>No</td>
<td>Calibrated</td>
</tr>
</tbody>
</table>

Table 7 – Parameters of the structural model: This table summarizes the parameters of the model, their description and the moments used to identify them.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Market defined by:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All bonds</td>
<td>Issuer’s code</td>
</tr>
<tr>
<td>( \eta )</td>
<td></td>
<td>0.58</td>
</tr>
<tr>
<td>( \sigma )</td>
<td></td>
<td>1.01</td>
</tr>
<tr>
<td>( \alpha_m ) (variance across markets)</td>
<td></td>
<td>0.0003</td>
</tr>
<tr>
<td>( \lambda_m ) (variance across markets)</td>
<td></td>
<td>0.059</td>
</tr>
</tbody>
</table>

Table 8 – Simulated Method of Moments estimates: this table reports the estimates of three specifications of the model. The specifications differ in the definition of a market; recall the arrival rate of liquidity shocks and the distribution of investors valuations are specific to each market. Column 1 consider all bonds as a single market, column 2 consider all bonds from the same issuer as a single market, column 3 consider each bond as a single market. The number in parenthesis show the variance across markets. I still have to obtain the standard errors of the estimates. All models use as a period of trading 3 months.
<table>
<thead>
<tr>
<th>Distribution</th>
<th>Observed data</th>
<th>Simulated data</th>
<th>Kolmogorov-Smirnov test results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25th q.</td>
<td>Mean</td>
<td>75th q.</td>
</tr>
<tr>
<td>Trade size</td>
<td>0.003</td>
<td>0.063</td>
<td>0.049</td>
</tr>
<tr>
<td>Prices</td>
<td>81.75</td>
<td>89</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 9 – Fit of the model: This table presents the results about the fit of the model. The results refer to the model using all bonds from the same issuer as a market. It compares the observed and the simulated distribution of the trade sizes, prices and the number of trades. The top panel shows the mean, the 25th and the 75th quantiles of the distributions of trade sizes and prices for all bonds. The column on the Kolmogorov-Smirnov tests shows the mean of the K-S test statistic computed for each bond with more than 10 trades per period. The number in parenthesis shows the percentage of bonds for which the test accept $H_0$ (equality of the distributions) at a 5% confidence level. The bottom panel shows the mean, the 25th and the 75th quantiles of the distributions of the number of trades and the results on the Kolmogorov-Smirnov test comparing the observed and the simulated distributions. The *** indicates that the test accepts the equality of the two distribution at 1% confidence level.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Proportional fee 1%</th>
<th>Proportional fee 3%</th>
<th>Fixed Fee</th>
<th>Buyers Posting</th>
<th>Sellers posting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volumes</td>
<td>67%</td>
<td>60%</td>
<td>4%</td>
<td>108%</td>
<td>-63%</td>
</tr>
<tr>
<td>Prices</td>
<td>0.9%</td>
<td>0.9%</td>
<td>-0.2%</td>
<td>18%</td>
<td>-12%</td>
</tr>
</tbody>
</table>

Table 10 – Counterfactual simulations: This table shows the percentage change of volumes traded and price of the bonds across different market structures. The values reported are averages across all bonds. The first two columns consider the introduction of a fee proportional to the amount traded: 1% of the value traded for column 1, 3% of the value traded for column 2. The changes is computed with respect to the observed distribution of volumes and prices. The third column shows the results for a fixed fee of $5000 per trade. The last two columns report the results for buyers and sellers posting terms of trades. See the main text for a description of how these values are computed.
**Figure 1** – Total volumes traded (both inside the auctions and through dealers) in the market for Auction Rate Municipal Bonds, in billion of dollars.

**Figure 2** – Example of the web-page from where the data was scraped. The red circles highlight the different parts of the page that provide the variables listed in Tables 1-2.
Figure 3 – Distribution of bonds’ coupon rates: for each bond I have computed the average coupon rate within a month (Coupon rates are adjusted to the market interest rates at the date of an auction when the auction fails). This figure plots the 25-th, the median and the 75-th percentile of the distribution of coupon rates for each month.

Figure 4 – Distribution of bid and ask prices: each trade is one observation
Figure 5 – Scatter plot of buy orders and sell orders: each dot is a trade, on the x-axis there is the amount traded (in million of dollars) and on the y-axis the unit price (in percentage points), ask prices on the left, bid prices on the right graph.

Figure 6 – Scatter plot of buy orders and sell orders for bonds with similar coupon rate. Left graph: trades for which the coupon rate of the bond was less than 0.3% yearly at the time of the trade. Right graph: trades for which the mean coupon rate of the bond was less then 0.5% yearly on the entire sample. Both graphs covers roughly 25% of all trades.
Figure 7 – Scatter plot of buy orders and sell orders for the most traded bonds. The left graphs shows trade for the most traded bond (140 trades). The right graph shows the second most traded (116 trades).

Figure 8 – Distribution of the size of trades, data on the left and simulated on the right. The trade size is divided by the total outstanding bond.
**Figure 9** – Distribution of prices, data on the left and simulated on the right.

**Figure 10** – Distribution of the number of trades per bond, data on the left and simulated on the right.
Figure 11 – This graphs shows the optimal asset holdings as a function of investors’
types. The estimates comes from a bond issued by the New Jersey Higher Education
Student Assistance Authority. It considers two market structures: intermediation and
buyers posting terms of trades. Investors are better off being sellers under prices posting
by buyers, so that they can capture most of the surplus from the trades. Thus the demand
of the asset increases for all types and the price of the bond increases (not shown). In-
vestors with valuation higher then the average valuation have also a higher probability
of being sellers in the future, thus they expect lower costs of trading. Thus higher types
take on more of the increase in the demand than low types, and the distribution of asset
holdings becomes more dispersed, increasing the volumes traded.

A Proofs

A.1 Solution of the bargaining problem

Here I derive the solution of the bargaining problem given in Eq.6. Taking the logarithm of the
objective function in Eq.4 I obtain:

\[
\max_{p,a'} (1 - \eta) \log \{ [V(\epsilon_t, a') - V(\epsilon_t, a) - p(a' - a)] \} + \eta \log \{ [(p - \xi)(a' - a) - c] \}
\]

Setting the first order condition with respect to \( p \) equal to 0 implies:

\[
p(a' - a) = (1 - \eta)[\xi(a' - a) + c] + \eta[V(\epsilon_t, a') - V(\epsilon_t, a)]
\]
which is Eq. 6b in the text. Eq. 6a is obtained substituting the above equation into the objective function.

A.2 Recursive formulation of the investor’s problem and the price in the inter-dealer market

I start by deriving the recursive representation of the value function of the investor, given in Eq. 8. Let \( \kappa = \alpha (1 - \eta) \). The value function of the investor can be written as:

\[
V(\epsilon_t, a) = \mathbb{E} \left\{ \int_t^{T^*} e^{-r(s-t)} \epsilon_s u(a) ds + e^{-r(T^*-t)} \left[ \max_{a'} \left[ V(\epsilon_{T^*}, a') - \xi (a' - a) \right] \right] \right\}
\]

Consider the recursive formulation of \( u(\epsilon_t, a) = \int_t^{T^*} e^{-r(s-t)} \epsilon_s u(a) ds \):

\[
(r + \kappa + \alpha) \bar{u}(\epsilon_t, a) = \epsilon_t u(a) + \alpha \mathbb{E}[\bar{u}(\epsilon_s, a)]
\]

(20)

where the expectation is take over the distribution of types. Multiply by \( \pi(\epsilon) \) and integrate over \( \epsilon \) to obtain:

\[
(r + \kappa + \alpha) \int \pi(\epsilon) \bar{u}(\epsilon_t, a) = \int \pi(\epsilon) \epsilon_t u(a) + \alpha \mathbb{E}[\bar{u}(\epsilon_s, a)] \Rightarrow \mathbb{E}[\bar{u}(\epsilon_s, a)] = \frac{\mathbb{E}(\epsilon) u(a)}{r + \kappa}
\]

Plugging this expression in 20 I obtain the expression for \( \bar{u}(\epsilon_t, a) \). Plug-in the recursive formulation of the investor’s problem into the optimization problem over the amount held:

\[
a = \arg \max \left[ \bar{u}(\epsilon_t, a) + \mathbb{E} \left\{ e^{-r(T^*-t)} \left[ \max_{a'} \left[ V(\epsilon_{T^*}, a') - \xi (a' - a) \right] \right] - \xi a \right\} \right]
\]

\[
a = \arg \max \left[ \bar{u}(\epsilon_t, a) + \mathbb{E} \left\{ e^{-r(T^*-t)} \xi a - \xi a \right\} \right]
\]

The price of the asset is given by:
\[ \xi = \int_0^\infty e^{-(r+\kappa)s} \xi(t + s)ds \]

thus:

\[ a = \arg \max [\pi(\epsilon_t, a) - r\xi a] \]

### A.3 Proof of Proposition 1

This proof shows that an increase in \( \eta \) has two effects: on one hand types lower then the mean type hold more of the asset in equilibrium, with the increase in asset holdings proportional to the type. On the other hand, types higher then the mean type hold less of the asset, with the decrease in asset holdings proportional to the type. Thus the distribution of asset holdings become less dispersed. Thus each investor trades less to adjust her asset holdings every time she can trade, and volumes traded decrease. The exact opposite effect happens if \( \eta \) decreases. I omit the algebra whenever it does not have an economic interpretation, the details are available upon request.

The analysis is divided in two cases. I initially keep the asset price in the centralized market constant. This case is presented to give the main intuition. Then I prove the proposition taking into account the effect of a change in \( \eta \) on the price in centralized market, showing that the main intuition carries through.

Recall that a type \( \tilde{\epsilon} \) demands in equilibrium:

\[ a(\tilde{\epsilon}) = \left( \frac{\tilde{\epsilon}}{r\xi} \right)^{1/\sigma} \tag{21} \]

Thus assuming that the price in the centralized market \( \xi \) remains constant, I have:

\[ \frac{\partial a(\tilde{\epsilon})}{\partial \eta} = \left( \frac{1}{r\xi} \right)^{1/\sigma} \frac{1}{\sigma \tilde{\epsilon}^{1/\sigma - 1}} \frac{\partial \tilde{\epsilon}}{\partial \eta} \]

The sign of this expression depends on \( \partial \tilde{\epsilon} / \partial \eta \). After some algebra, I find that:
\[
\frac{\partial \bar{\epsilon}}{\partial \eta} = \frac{\alpha (\bar{\epsilon} - \epsilon)}{(r + \alpha (1 - \eta) + \alpha)^2}
\]

Thus the types lower than the mean type hold in equilibrium more of the asset, with the increase in asset holdings proportional to the difference with the mean type.

I now consider the more general case which takes into account the change in the centralized market price.

Substituting the equilibrium value of the market price, Eq.21 becomes:

\[
a(\bar{\epsilon}) = \frac{\bar{\epsilon}^{1/\sigma}}{\mathbb{E}(\bar{\epsilon}^{1/\sigma})} A
\]

Taking the derivative of the above expression with respect to \(\eta\) I obtain that:

\[
\begin{align*}
\epsilon < \frac{1 - \Lambda \alpha}{1 + \frac{\Lambda}{1}[r + \alpha (1 - \eta)]} \bar{\epsilon} & \quad \Leftrightarrow \quad \frac{\partial a(\bar{\epsilon})}{\partial \eta} > 0 \quad (22) \\
\epsilon > \frac{1 - \Lambda \alpha}{1 + \frac{\Lambda}{1}[r + \alpha (1 - \eta)]} \bar{\epsilon} & \quad \Leftrightarrow \quad \frac{\partial a(\bar{\epsilon})}{\partial \eta} < 0 \quad (23)
\end{align*}
\]

where:

\[
\begin{align*}
\Lambda = \int \pi(\epsilon)[(r + \alpha (1 - \eta))\epsilon + \alpha \bar{\epsilon}]^{1/\sigma} d\epsilon \quad ; \quad \Upsilon = \int \pi(\epsilon)[(r + \alpha (1 - \eta))\epsilon + \alpha \bar{\epsilon}]^{1/\sigma - 1}(\bar{\epsilon} - \epsilon) d\epsilon
\end{align*}
\]

Note that the distribution of types has been integrated out in both \(\Lambda\) and \(\Upsilon\), so they do not depend on \(\epsilon\).

The last steps of tedious algebra show that:

\[
0 < \frac{1 - \frac{\Lambda}{1} \alpha}{1 + \frac{\Lambda}{1}[r + \alpha (1 - \eta)]} < 1
\]
Thus again the types lower than the mean type hold in equilibrium more of the asset. In the more general case in which I consider the effect on the centralized market price, the threshold value is a scaled value of the mean type.

A.4 On the relationship between the variance of prices and dealers’ bargaining power

**Proposition 2** If the variance of intermediation costs is small enough, the variance of prices increases with $\eta$.

**Proof:** Using the solution to the bargaining problem and the recursive formulation of the investor’s problem, the price of a trade of an investor changing her asset holdings from $a$ to $a'$ can be written as:

$$p(a' - a) = (1 - \eta) [\xi + c] (a' - a) + \eta \left[ \frac{(r + \kappa)\epsilon_t + \alpha \omega}{r + \kappa + \alpha} a'^{1-\sigma} - a^{1-\sigma} \right]$$

Thus, assuming that the intermediary’s cost of executing trade - $c$ - is independent of the asset holdings of the investor, the variance of prices is:

$$\text{Var}(p) = (1 - \eta)\text{Var}(c) + \eta \text{Var} \left[ \frac{\epsilon(a'^{1-\sigma} - a^{1-\sigma})}{(a' - a)(1 - \sigma)} \right]$$

where $\epsilon = \frac{(r+\kappa)\epsilon_t + \alpha \omega}{(r+\kappa+\alpha)}$. Combine this expression with the optimal asset holdings of each type and the price in the centralized market to obtain:

$$\text{Var}(p) = (1 - \eta)\text{Var}(c) + \eta \text{Var}(\Psi)$$

where:

$$\Psi = \frac{\epsilon'}{(1 - \sigma)} \left( \frac{\mathbb{E}(\epsilon^{1/\sigma})}{A} \right)^{\frac{\sigma}{\sigma - (1-\sigma)/\sigma}} \frac{\epsilon^{(1-\sigma)/\sigma} - \epsilon^{(1-\sigma)/\sigma}}{\epsilon^{1/\sigma} - \epsilon^{1/\sigma}}$$

Thus the variance of prices is a convex combination of the variance of intermediation costs $c$ and a random variable $\Psi$ that depends, among other things, on the distribution of investor types. As $\eta$
approaches one, the variance of prices converges to the variance of this expression. Sufficient, but not necessary, conditions to have the variance of prices increasing with the $\eta$ are: the variance of $c$ is small relative to the variance of $\Psi$ and the variance of $\Psi$ is increasing in $\eta$.

For notational convenience, let me write $\Psi = \Psi_\eta(\tilde{\epsilon}, \epsilon)$. For a given $\epsilon$, I look at what happens to this variance as $\eta$ increases. To approximate the variance of $\Psi_\eta(\tilde{\epsilon})$, I use the delta method, which states that:

$$Var(f(X)) \approx f'(\mathbb{E}(X))^2 Var(X)$$

where $f$ is a twice differentiable function and the mean and variance of $X$ are finite.

Under this approximation note that:

$$\frac{\partial Var(\Psi_\eta(\tilde{\epsilon}))}{\partial \eta} \approx 2\Psi_\eta'(\tilde{\epsilon}) \frac{\partial \Psi_\eta'(\tilde{\epsilon})}{\partial \eta}$$

The derivative appearing in the above expression is given by:

$$\Psi_\eta'(\tilde{\epsilon}) = \left[ \frac{\mathbb{E}(\tilde{\epsilon}^{1/\sigma})}{A(1-\sigma)^{1/\sigma}} \right] \sigma \{ [r + \alpha(1-\eta)]\Gamma + [(r + \alpha(1-\eta))\epsilon' + \alpha \tilde{\epsilon}]\Phi \}$$

where:

$$\Gamma = \frac{[(r + \alpha(1-\eta))\epsilon' + \alpha \tilde{\epsilon}]^{(1-\sigma)/\sigma} - [(r + \alpha(1-\eta))\epsilon + \alpha \tilde{\epsilon}]^{(1-\sigma)/\sigma}}{[(r + \alpha(1-\eta))\epsilon' + \alpha \tilde{\epsilon}]^{1/\sigma} - [(r + \alpha(1-\eta))\epsilon + \alpha \tilde{\epsilon}]^{1/\sigma}}$$

$$\Phi = \frac{(r + \alpha(1-\eta))(r + \alpha(1-\eta))\epsilon' + \alpha \tilde{\epsilon}]^{(1-\sigma)/\sigma} \Delta}{\sigma \{(r + \alpha(1-\eta))\epsilon' + \alpha \tilde{\epsilon}]^{1/\sigma} - [(r + \alpha(1-\eta))\epsilon + \alpha \tilde{\epsilon}]^{1/\sigma}\}^2}$$

$$\Delta = [(r + \alpha(1-\eta))\epsilon + \alpha \tilde{\epsilon}]^{(1-\sigma)/\sigma} - \sigma [(r + \alpha(1-\eta))\epsilon' + \alpha \tilde{\epsilon}]^{(1-\sigma)/\sigma}$$

$$- (1-\sigma) \frac{[(r + \alpha(1-\eta))\epsilon + \alpha \tilde{\epsilon}]^{1/\sigma}}{[(r + \alpha(1-\eta))\epsilon' + \alpha \tilde{\epsilon}]^{1/\sigma}}$$

Taking the derivative of the above expression with respect to $\eta$ is can be seen that:

$$\text{sign} \left( \frac{\partial \Psi_\eta'(\tilde{\epsilon})}{\partial \eta} \right) = \text{sign} (\Psi_\eta'(\tilde{\epsilon}))$$
Thus $\frac{\partial \text{Var}(\Psi_{\eta}(\hat{\epsilon}))}{\partial \eta} > 0$.

B Green et al. [2007] methodology to match buy and sell orders

Green et al. [2007] use transactions from the secondary market of all municipal bonds from the 1st of May 2000 to January 10th 2004. In this paper, I analyze transactions from the Auction Rate Municipal bond market between 2009 and 2011. The two dataset are very similar in terms of information available and frequency of trading. Thus I use their methodology to construct the dataset about transactions in the retail market of investors and the inter-dealer trades associated with them. I consider all inter-dealer trades not associated with retail transactions as transactions in the inter-dealer market for the purpose of section 2.

They associates trades using three methods, called "immediate matches", "round-trip transactions" and the "first-in-first-out" (FIFO) rule. In this paper I consider all their three rules, plus what I call the "augmented round-trip transactions", described below.

Immediate matches associate trades in the same bond, for the same par amount, on the same day, with no intervening trades in that bond. For such trades, dealers do not take inventories at the end of transaction day. I consider any inter-dealer trade in the same bond, for the same par amount, on the same day, with no intervening trades between the retail trades in that bond as part of these transactions, and so they are not considered as part of the inter-dealer market.

Round-trip transactions drop the requirements that there should not be any intervening trades in that bond between the buy and sell order. Thus a buy and a sell order (and any inter-dealer trade between the two) are considered part of the same deal if they are in the same bond, for the same par amount and on the same day.

The augmented round-trip transaction rule relaxes the requirement that the amount of trades must be the same. Thus only the sum of the par amounts traded must be equal to the largest trade on the other side of the market in the same date and in the same bond. Thus under this rule I consider a sell order of $200,000 and two buy orders of $100,000 as part of the same deal. In this case all transactions between dealers summing to one of these par amounts are considered transactions for the retail market.
If more than one combination of inter-dealer trades in the same day matches one of these transactions, I include the combination with the smallest number of inter-dealer trades.

Finally, the FIFO sample of retail trades is created in the following way. First all trades associated with the round-trip transaction rules are removed. Then the remaining buy trades from investors are associated to sell orders if they come after a sell order in the retail market and they have a par amount equal or smaller of the originating trade. Thus consider for example a sell order of $100,000 not previously associated using the round-trip rule or exhausted by other buy orders using the FIFO rule. The FIFO rule associate to this trade the first buy order of $70,000 two days later and consider that there are still $30,000 of outstanding par amount of this trade. The first buy order less or equal to $30,000 is associated with this trade. All inter-dealer transactions of par amount equal or smaller then the par amounts associated with these trades in the days between these trades are considered part of the retail market.