Working Paper Series

Private Information in Over-the-Counter Markets

WP 16-16

Zachary Bethune
University of Virginia

Bruno Sultanum
Federal Reserve Bank of Richmond

Nicholas Trachter
Federal Reserve Bank of Richmond

FEDERAL RESERVE BANK OF RICHMOND®
Richmond ▪ Baltimore ▪ Charlotte

Private Information in Over-the-Counter Markets*

Zachary Bethune
University of Virginia

Bruno Sultanum
Federal Reserve Bank of Richmond

Nicholas Trachter
Federal Reserve Bank of Richmond

Working Paper No. 16-16

Abstract

We study trading in over-the-counter (OTC) markets where agents have heterogeneous and private valuations for assets. We develop a quantitative model in which assets are issued through a primary market and then traded in a secondary OTC market. Then we use data on the US municipal bond market to calibrate the model. We find that the effects of private information are large, reducing asset supply by 20%, trade volume by 80%, and aggregate welfare by 8%. Using the model, we identify two channels through which the information friction harms the economy. First, the distribution of the existing stock of assets is inefficient because some of the efficient trades, which should occur, do not. Second, the total stock of assets is inefficiently low because resale value and liquidity go down due to the information friction. We investigate how much a simple tax/subsidy scheme that spurs issuance of new assets can help mitigate the cost associated with private information and find that it lowers the welfare cost from 8% to approximately 1%.

JEL Classification: D53, D82, G14

Keywords: Decentralized markets, bilateral trade, asset issuance, liquidity, asymmetric information.

*Contact: Zachary Bethune: zab2t@virginia.edu, Bruno Sultanum: bruno@sultanum.com, Nicholas Trachter: nicholas.trachter@rich.frb.org. This paper previously circulated under the title "Decentralized Trade with Private Values." We thank Jim Albrecht, Julien Hugonnier, Ben Lester, Susan Vroman, Pierre-Olivier Weill, Randall Wright, and seminar participants at Georgetown, UC Santa Barbara, UC Santa Cruz, the Summer Workshop on Money, Banking, and Asset Markets at FRB-Chicago, SED Toulouse, Wisconsin School of Business Money, Banking, and Asset Markets Conference, the Cornell/Penn State Macroeconomics Conference, and the UCSB LAEF conference on "What we know about information in financial markets" for helpful suggestions and feedback. Any opinions or views expressed in this article are not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

1 Introduction

A key feature of over-the-counter (OTC) markets is that trade is bilateral. It is well-known that efficiency in bilateral trade is not easy to obtain when both parties posses private information. What is less known is how and to what extent this friction, at the bilateral level, impacts aggregate trading dynamics and market efficiency. Is private information quantitatively important in terms of impacting welfare and market outcomes, such as trade volume and yields? What can a policymaker do in order to mitigate the implied inefficiencies? To address these questions, we build a novel model of OTC trade with private information and use data on the US municipal bond market to quantify the importance of private information in OTC markets. We find that private information creates a large distortion in financial markets, reducing asset supply, trade volume, and liquidity. However, the introduction of a simple policy that subsidizes asset issuance can regain a significant portion of the welfare losses induced by the information frictions, despite private information continuing to affect asset trade after issuance.

To be consistent with many features of OTC markets, we develop a model in which trade occurs in both a primary market, where issuers sell assets with a known and common dividend to investors, and in a secondary market, where investors trade assets amongst themselves. Both markets are decentralized in the sense that agents trade in bilateral meetings that are subject to search frictions. As in Hugonnier et al. (2014), we allow for a general form of heterogeneity in the trading motives of individuals. Heterogeneity in trading motives reflects differences in opportunity costs for buying and selling the asset. In the primary market, issuers are heterogenous with respect to the cost of creating and issuing a new asset. In the secondary market, investors are heterogenous in the utility flow of holding the asset. These forms of heterogeneity are meant to capture a wide range of differences that lead to gains in trade, such as different tax and regulatory advantages or funding and liquidity costs. We assume, as in the static model analyzed in Myerson (1981), that trading motives are private information. Therefore, as famously described in Myerson and Satterthwaite (1983), the lack of common knowledge about the gains from trade implies that there is no trading arrangement that guarantees ex-post efficient outcomes in a meeting between a seller and buyer of an asset.[1]

Private information regarding trading motives affects the market outcome in two ways: it creates a distortion in the efficiency of individual trades and reduces total asset supply. Under complete information, agents engage in trade whenever the potential seller of the asset values it less than the potential buyer. However, trade under private information requires the buyer's valuation to be larger than the seller's valuation *plus* a term that captures the informational rents possessed by the buyer. Hence, when the gains from trade in a meeting do not exceed the informational rent, no trade occurs. This mechanism was first studied by Myerson (1981) in a static setup, and we incorporate it into our dynamic general equilibrium setting. Moreover, private information decreases asset issuance as the value to investors of holding an asset in the secondary market falls. This decrease in valuation happens because (i) the resale value of the asset falls because a seller must give up informational rents; and (ii) the value of waiting and buying the asset in the future becomes higher because the investor can acquire informational rents when he buys the asset in the future.

We calibrate our model to match features of the US municipal bond market, as studied in (Green et al., 2007). This data is ideal for our purposes as trade in this market is performed in a decentralized fashion, both in its primary and secondary markets. Interestingly, there is a large amount of dispersion in yields for assets issued by a given municipality, a feature that directly maps to our setting with search and information frictions. The municipal bond market in the US is large, with an average of $400 billion a year in new issuance, and features an active OTC secondary market that averages 40,000

[1] We focus on private information regarding individuals' opportunity costs in trade. The lack of complete information in our model has a similar flavor to models in which inefficiency arises from a lack of information about the *common value* component of the trade (e.g. Chiu and Koeppl, 2015). However, in our model, the inefficiency arises from lack of information about the *private value* component of the trade. As posed in Duffie (2012), both sources of private information are important frictions in decentralized asset markets, and both require attention. However, for the most part, the literature has focused only on the common component aspect of private information.

trades per day at an average par value of $14 million.[2]

With the calibrated version of the model, we find private information leads to a fall in aggregate welfare of 8% relative to the model with no private information. The effect on welfare is due to a sharp decline in trade volume and asset supply. Lower trade volume implies a larger extent of misallocation of assets and lower asset supply implies that fewer investors can satisfy their trading needs. Both these measures are depressed under private information because of the large informational rents implied by the model and the low value that owners of the asset obtain in the secondary market by attempting to sell the asset to investors that are non-owners (i.e. a low resale value). An interesting implication of these two forces is that chains of trade are shorter under incomplete information than under complete information, as evidenced by the low share of trade volume that involve investors with intermediate valuations. That is, the role of middlemen is severely undermined under private information relative to its complete information counterpart (Hugonnier et al., 2014).

Quantitatively, the aggregate welfare effects of private information crucially depend on the existence of private information in trade in the secondary market. If the secondary market were to run under complete information, the welfare losses in the economy would be small. If the primary market were to run under complete information, the welfare losses in the economy would continue to be large, close to those obtained when both markets operate under private information. However, this does not imply that the primary market is not an important channel in which private information affects welfare. As a result of private information in secondary markets, the resale value (and liquidity) of assets is lower, thereby reducing issuance in the primary market. This distinctive feature of our environment, not present in the OTC literature that largely abstracts from asset creation, allows us to uncover a prominent way by which informational frictions harm the economy: by generating an inefficiently low level of assets.

[2]Data on aggregate trading activity in the municipal bond market are available through the Municipal Securities Research Board's (MSRB) Electronic Municipal Market Access (EMMA) website, http://emma.msrb.org.

The asset supply channel is not only theoretically interesting, but also has policy implications. A policy that increases asset supply can undo some of the harm caused by the information friction. Indeed, we study a simple policy that taxes lump sum issuers and uses the revenue to subsidize asset issuance. We find that this policy reduces the welfare cost of private information from about 8% to close to 1%. Note that policies aimed at increasing asset issuance are not rare. During the last financial crisis, for example, the Federal Reserve created the Term Asset-Backed Securities Loan Facility (TALF) with the goal of spurring the issuance of asset-backed securities (ABS) collateralized by loans of various types to consumers and businesses.[3]

Our paper is related to both the micro and macro literature on trade with private information. At a micro-level, our trading mechanism follows early work by Myerson (1981) and Myerson and Satterthwaite (1983) that models trade of an indivisible asset under private information about agents' valuations. These papers, and the literature that followed, are largely interested in trade efficiency while taking the distribution of valuations as exogenous. Alternatively, we provide a framework where the distribution of valuations is endogenous and responds to frictions, both at a bilateral level (e.g. information frictions) and at an aggregate level (e.g. search frictions, supply/demand effects). The macro-finance literature on trade under private information has largely focused on uncertainty about common values, such as asymmetric information and adverse selection problems. Adverse selection has been used to understand asset market liquidity in both centralized markets (e.g. Eisfeldt, 2004; Chari et al., 2014; and Bigio, 2015) and decentralized markets (e.g. Guerrieri et al., 2010; Chang, 2014; and Chiu and Koeppl, 2015). Alternatively, we consider private information about private values, a relevant component of the valuation of assets in which the quality of the asset is easily verified, and analyze the effects on both secondary market liquidity and primary market issuance. Cujean and Praz (2015) study an environment with asymmetric information about private values in which asset holdings are unrestricted. Finally, Zhang (2016) also

[3]See https://www.federalreserve.gov/monetarypolicy/talf.htm

considers asymmetric information about privates values but in an environment in which traders can form long-term relationships. We abstract from long-term relationships.

The paper is structured as follows. Section 2 introduces the general environment, defines the steady state equilibrium, and proves its existence. Section 3 describe our benchmark with complete information. Section 4 describes the calibration and contains our main experiments. Section 5 provides a simple policy to mitigate the losses due to private information. Finally, Section 6 concludes.

2 Model

In this section, we introduce the environment, define an equilibrium, and provide a proof of equilibrium existence.

2.1 Environment

Time is continuous and infinite. There is a unit measure of infinitely lived and risk-neutral investors, and a unit measure of infinitely lived and risk-neutral issuers, who all discount the future at rate $r > 0$. Investors' asset holdings are discrete, either zero or one. We label investors holding the asset as *owners*, and those not holding the asset as *non-owners*. There is transferable utility across all agents. An investor receives utility flow ν from holding an asset. We refer to ν as the investor's type, which is private information and follows a distribution F. We assume that F has support $[\underline{\nu}, \bar{\nu}]$, density f, and $f(\nu) > 0$ for all ν in the support.[4] It is straightforward to extend the setup to allow for time-varying types, in the style of Duffie et al. (2005). In their setup, as in many in the literature, preference shocks are important in order to guarantee that trade occurs in the steady state. In our setup, since assets mature and are created, we do not need this

[4]This is a model of private values. Private values for an asset are distinguished from common values, in which the types of some agents convey information about the valuation of other agents. See Fundenberg and Tirole (1991), Chapter 7, for a discussion of how the information structure affects the mechanism design problem and Krishna (2010) for a detailed analysis of auctions under differing information structures.

assumption to generate trade in steady state. Investors meet each other with Poisson arrival rate $\lambda_b/2$, and they can trade assets.

Outstanding assets mature every period with Poisson arrival rate μ, and new assets are created by issuers. Issuers receive i.i.d. issuance opportunities (i.e. the opportunity to create an asset) at Poisson rate $\lambda_a > 0$. Creating an asset entails incurring in a cost c, which is drawn from a distribution G. A new issuance cost c is drawn each time an issuer has an issuance opportunity and is i.i.d. across issuers and over time. We assume that G has density g, support $[\underline{c}, \bar{c}]$, and $g(c) > 0$ for all costs c in the support. Upon the realization of an investment opportunity, an issuer contacts an investor randomly.[5] If the issuer decides to create the asset, they incur cost c. Issuers cannot hold assets. Note, however, that the issuers always have the option not to issue the asset, in which case no cost is incurred. For instance, this would happen if the issuer does not reach an agreement to sell the asset to the investor he is in contact with.

We restrict our attention to steady-state analysis so we do not include time t in the set of states of the economy. Let s denote the fraction of investors holding an asset, let $\Phi_o(\nu)$ denote the measure of owners with type ν or below, and $\Phi_n(\nu)$ denote the measure of non-owners of type ν or below. Let $V_o(\nu)$ and $V_n(\nu)$ be the value function of an *owner* and a *non-owner* of type ν, and let $\Delta(\nu) = V_o(\nu) - V_n(\nu)$ be his reservation value. The reservation value is the minimum/maximum price an investor is willing to sell/buy the asset for. Finally, let V_c denote the value function of an issuer before the issuance opportunity and before knowing his cost of producing the asset.

2.2 Primary-market trade mechanism

The primary market consists of trades between issuers and investors. We consider trading mechanisms that maximize the seller's (in this case the issuer's) expected gains from trade.[6] Without loss of generality with respect to characterizing payoffs and al-

[5] It is easy to extend this assumption so that the issuer can contact several investors at once, in which case he runs an auction instead of trading bilaterally.

[6] The model can easily be generalized in order to allow for different Pareto weights.

locations, we restrict attention to direct mechanisms. A direct mechanism is a pair of functions $m^a = (p^a, x^a) : [\underline{c}, \bar{c}] \times [\underline{v}, \bar{v}] \to \mathbb{R} \times [0, 1]$, where, for a given issuance cost c and non-owner type ν_n, $p^a(c, \nu_n)$ is the probability of the issuer transferring the asset to the non-owner, and $x^a(c, \nu_n)$ is the transfer of utility from the non-owner to the issuer. Define the functions

$$\bar{p}^a_n(\nu_n) = \int p^a(c, \nu_n) dG(c), \quad \bar{x}^a_n(\nu_n) = \int x^a(c, \nu_n) dG(c)$$
$$\bar{p}^a_c(c) = \int p^a(c, \nu_n) d\frac{\Phi_n(\nu_n)}{1-s}, \text{ and } \bar{x}^a_c(c) = \int x^a(c, \nu_n) d\frac{\Phi_n(\nu_n)}{1-s}.$$

The first term, $\bar{p}^a_n(\nu_n)$, is the investor's expected probability of receiving the asset if he announces type ν_n. The second term, $\bar{x}^a_n(\nu_n)$, is the investor's expected transfer if he announces type ν_n. The third and fourth terms, $\bar{p}^a_c(c)$ and $\bar{x}^a_c(c)$ are analogous but from the perspective of an issuer announcing cost c. Since the mechanism maximizes the issuer's gains from trade, it must solve

$$\max_{m^a} \int \int \left[x^a(c, \nu_n) - p^a(c, \nu_n) c \right] d\frac{\Phi_n(\nu_n)}{1-s} dG(c) \tag{1}$$

subject to

$$IR: \quad \bar{p}^a_n(\nu_n)\Delta(\nu_n) - \bar{x}^a_n(\nu_n) \geq 0 \tag{2}$$
$$IC: \quad \bar{p}^a_n(\nu_n)\Delta(\nu_n) - \bar{x}^a_n(\nu_n) \geq \bar{p}^a_n(\hat{\nu}_n)\Delta(\nu_n) - \bar{x}^a_n(\hat{\nu}_n) \tag{3}$$

for all c, ν_n, and $\hat{\nu}_n$. The first constraint, given by equation (2), is the non-owner's individual rationality constraint—it guarantees that he has an incentive to participate in the mechanism. The second constraint, given by equation (3), is the non-owner's incentive compatibility constraint—it guarantees that he has an incentive to truthfully reveal his type to the mechanism. Note that we do not state the individual rationality nor the incentive compatibility constraints of the issuer. One can show that these constraints are going to be satisfied because the mechanism maximizes the issuer gains from trade.

The expected gains from trade for an issuer and a non-owner in the meeting, for a given a direct mechanism $m^a = (p^a, x^a)$, are

$$\pi_c^a(c) = \bar{x}_c^a(c) - \bar{p}_c^a(c)c \quad \text{and} \tag{4}$$

$$\pi_n^a(\nu_n) = \bar{p}_n^a(\nu_n)\Delta(\nu_n) - \bar{x}_n^a(\nu_n). \tag{5}$$

The first equation is the expected gains from trade of an issuer with cost c conditional on meeting a non-owner. It takes expectation, with respect to ν_n, of the transfer from the non-owner, $x^a(c, \nu_n)$, minus the expected cost of issuance, $p^a(c, \nu_n)c$. The second equation is the expected gains from trade of a non-owner with type ν_n conditional on meeting with an issuer. It takes expectation, with respect to c, of the expected gain of obtaining the asset, $p^a(c, \nu_n)\Delta(\nu_n)$, minus the transfer to the issuer, $x^a(c, \nu_n)$. The gains from trade are evaluated under the assumption that issuers and non-owners truthfully reveal their types to the mechanism. The constraints (2) and (3) ensure that an equilibrium in truth-telling strategies exists. In order to keep the presentation simple, we assume that the truth-telling equilibrium is being played instead of explicitly writing the strategic game associated with the mechanism.

2.3 Solution to the primary-market trade mechanism

In this subsection, we assume that the distribution of valuations, $\frac{\Phi_n}{1-s}$, and the reservation value function, Δ, are differentiable, and Δ' is positive and bounded away from zero. Then we solve for the optimal trade mechanism following Myerson (1981) and discuss some of its properties. Note, however, that both $\frac{\Phi_n}{1-s}$ and Δ are endogenous objects—they depend on how people trade. As a result, one cannot make direct assumptions over them. When describing our existence results we discuss why $\frac{\Phi_n}{1-s}$ may not be differentiable in equilibrium and argue that the intuition derived here still holds. Moreover, we use the results in this subsection to prove the existence of a steady-state equilibrium in our economy, where a steady-state equilibrium is defined later in this section.

Let $haz(\nu_n) = \frac{\phi_n(\nu_n)}{1-s} / \left[1 - \frac{\Phi_n(\nu_n)}{1-s}\right]$ be the hazard function of the distribution of non-owners, $\frac{\Phi_n(\nu_n)}{1-s}$, where ϕ_n is the derivative of Φ_n. The following proposition simplifies problem (1)-(3). All the proofs are in the appendix.

Proposition 1. *Let the distribution of non-owners, $\frac{\Phi_n}{1-s}$, and the reservation value, Δ, be differentiable and Δ' strictly positive and bounded away from zero. Then, a direct mechanism $m^a = (p^a, x^a)$ solves problem (1)-(3) if, and only if, it solves*

$$\max_{m^a} \int \int p^a(c, \nu_n) \left[\Delta(\nu_n) - \frac{\Delta'(\nu_n)}{haz(\nu_n)} - c\right] d\frac{\Phi_n(\nu_n)}{1-s} dG(c) \tag{6}$$

subject to $\bar{p}^a_n(\nu_n)$ being increasing and $\pi^a_n(\nu_n) = \int_{\underline{\nu}}^{\nu_n} \bar{p}^a_n(\nu)\Delta'(\nu)d\nu$.

A proof of the proposition is available in the appendix. The probability of trade is determined by the investor's true utility gain from holding the asset, $\Delta(\nu_n)$, minus an informational rent, $\frac{\Delta'(\nu_n)}{haz(\nu_n)}$, minus the issuance cost, c. The term $\Delta(\nu_n) - \frac{\Delta'(\nu_n)}{haz(\nu_n)}$ is labeled in the mechanism design literature as the virtual valuation of the buyer. It measures how much the seller, in this case the issuer, values selling the asset to type ν_n investors.[7] The informational rent, $\frac{\Delta'(\nu_n)}{haz(\nu_n)}$, depends on the underlying distribution of non-owner investors $\frac{\Phi_n(\nu_n)}{1-s}$. A high hazard, $haz(\nu_n)$, at ν_n implies a high virtual valuation of type ν_n investors, which occurs because either the density of investors at ν_n, $\frac{\phi_n(\nu_n)}{1-s}$, is high, or because the measure of investors with valuation above ν_n, $1 - \frac{\Phi_n(\nu_n)}{1-s}$, is low. A high value of $\frac{\phi_n(\nu_n)}{1-s}$ leads to a high virtual valuation because it implies that there is a large amount of investors at ν_n, so that being able to sell to these investors is highly desirable for the issuer. However, selling to investors of type ν_n implies giving informational rents to investors of type above ν_n that have measure $1 - \frac{\Phi_n(\nu_n)}{1-s}$.

When the virtual valuation, $\Delta(\nu_n) - \frac{\Delta'(\nu_n)}{haz(\nu_n)}$, is increasing, we can define a cutoff $\nu^*_n(c)$

[7] With complete information there is no need for the issuer to give up informational rent, so we can think of the virtual valuation in a complete information model as being simply $\Delta(\nu_n)$. Note, however, that $\Delta(\nu_n)$ differ under complete versus incomplete information.

implicitly by the equation

$$\Delta(\nu_n^*(c)) - \frac{\Delta'(\nu_n^*(c))}{haz(\nu_n^*(c))} = c.$$

In this case, the solution to problem (6) is

$$p^a(c,\nu_n) = \begin{cases} 1 & \text{if } \nu_n \geq \nu_n^*(c) \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad x^a(c,\nu_n) = \begin{cases} \Delta(\nu_n^*(c)) & \text{if } \nu_n \geq \nu_n^*(c) \\ 0 & \text{otherwise.} \end{cases}.$$

This solution implies that (i) transfers only occur upon trade and (ii) every non-owner investor of type ν_n above the cutoff $\nu_n^*(c)$ buys the asset and pays the reservation value of the investor at the cutoff $\nu_n^*(c)$. Note that $\bar{p}_n^a(\nu_n)$ is weakly increasing because $\Delta(\nu_n) - \frac{\Delta'(\nu_n)}{haz(\nu_n)}$ is increasing, so the constraint is satisfied.

When the virtual valuation is not increasing, the solution involves a procedure to flatten the nonmonotone regions—producing an *adjusted* virtual valuation that is increasing. However, the same result applies to this *adjusted* virtual valuation. Namely, trade occurs whenever the adjusted virtual valuation is greater or equal to the issuance cost and investors above the associated cutoff pay the reservation value of the investor at the cutoff. In order to state the general solution, it is useful to define the functions

$$h(q) = \Delta[\Phi_n^{-1}((1-s)q)] - \frac{1-s-q}{\phi(\Phi_n^{-1}((1-s)q))}\Delta'[\Phi_n^{-1}((1-s)q)], \quad H(q) = \int_0^q h(r)dr$$

$$G(q) = \min\{\omega H(r_1) + (1-\omega)H(r_2);\ \omega r_1 + (1-\omega)r_2 = q\}, \quad \text{and} \quad g(q) = \frac{dG(q)}{dq}.$$

The following proposition characterizes the solution in general, whether or not the virtual valuation is increasing.

Proposition 2. *Let the distribution of non-owners,* $\frac{\Phi_n}{1-s}$*, and the reservation value,* Δ*, be both* *differentiable and Δ' strictly positive and bounded away from zero. Define the functions*

$$c_n(\nu) = g\left(\frac{\Phi_n(\nu)}{1-s}\right) \quad \textit{and} \quad c_n^{-1}(c) = \begin{cases} \inf\{\nu_n; c_n(\nu_n) \geq c\} & \textit{if } \; c_n(\nu_n) \geq c \textit{ for some } \nu_n \\ \bar{\nu} & \textit{otherwise} \end{cases}.$$

Then the direct mechanism $m^a = (p^a, x^a)$, defined as

$$p^a(c, \nu_n) = \begin{cases} 1 & \textit{if } \; c_n(\nu_n) \geq c \\ 0 & \textit{otherwise} \end{cases} \quad \textit{and} \quad x^a(c, \nu_n) = \begin{cases} \Delta(c_n^{-1}(c)) & \textit{if } \; c_n(\nu_n) \geq c \\ 0 & \textit{otherwise.} \end{cases},$$

achieves the maximum in problem (6).

There are two important aspects of this solution. First, issuance is always distorted within a meeting. It is efficient to issue an asset whenever the reservation value of the investor, $\Delta(\nu_n)$, is greater than the issuance cost c. However, due to the informational rent, $\frac{\Delta'(\nu_n)}{haz(\nu_n)}$, this policy is not implemented and some efficient issuance does not occur. The second aspect regards the split of the gains from trade among the issuer and the investor. In the complete information analogue of problem (1)-(3), it is optimal for the issuer to issue whenever $\Delta(\nu_n)$ is greater than the issuance cost c and charge exactly the reservation value of the investor, $\Delta(\nu_n)$. As a result, the issuer captures all the gains from trade. However, under incomplete information, the expected gains from trade for investors are

$$\mathbb{E}\left[p^a(c, \nu_n)\frac{\Delta'(\nu_n)}{haz(\nu_n)}\right] = \int\int p^a(c, \nu_n)\frac{\Delta'(\nu_n)}{haz(\nu_n)} d\frac{\Phi_n(\nu_n)}{1-s} dG(c), \tag{7}$$

which can be shown to be strictly positive. In the model, trade inefficiencies and the informational rent play an important role, shaping aggregate outcomes; we explore this in detail later once we calibrate our model. For now, to illustrate these two forces, it is worth studying a particular example. Below, we consider a static partial equilibrium example where $\frac{\Phi_n(\nu_n)}{1-s}$ is assumed to be Pareto, Δ is the identity, and G is degenerate.

Example 1. *Let* $\Delta(\nu) = \nu$, $\frac{\Phi_n(\nu_n)}{1-s} = 1 - \left(\frac{\underline{\nu}}{\nu_n}\right)^{\alpha}$ *and G be degenerated at* $c = \underline{\nu}$. $\underline{\nu} > 0$ *is the scale parameter (and also the lower bound of the support), and* $\alpha > 2$ *is the shape parameter. The assumption that* $\alpha > 2$ *guarantees that the results from this section hold even though the Pareto distribution is not defined in a compact support. The hazard function of* $\frac{\Phi_n(\nu_n)}{1-s}$ *is* $haz(\nu_n) = \frac{\alpha}{\nu_n}$, *and the investor's virtual valuation is* $\Delta(\nu_n) - \frac{\Delta'(\nu_n)}{haz(\nu_n)} = \frac{\alpha-1}{\alpha}\nu_n$, *which is linear in the investor's type. The cut-off value is* $\nu_n^*(c) = \frac{\alpha}{\alpha-1}c$. *Asset issuance is given by*

$$Issuance = \int\int \mathbb{1}_{\{\nu_n \geq \frac{\alpha}{\alpha-1}c\}} d\frac{\Phi_n(\nu_n)}{1-s} dG(c) = \left(\frac{\alpha-1}{\alpha}\right)^{\alpha} \in \left(\frac{1}{4}, \frac{1}{e}\right) ,$$

and the share of the total gains from trade accrued by issuers is given by

$$1 - \frac{\int\int p^a(c,\nu_n)\frac{\Delta'(\nu_n)}{haz(\nu_n)} d\frac{\Phi_n(\nu_n)}{1-s} dG(c)}{\int\int p^a(c,\nu_n)\left[\Delta(\nu_n) - c\right] d\frac{\Phi_n(\nu_n)}{1-s} dG(c)} = \frac{\alpha-1}{2\alpha-1} \in \left(\frac{1}{3}, \frac{1}{2}\right) .$$

From the example above we see that (i) issuance is always distorted and (ii) issuers give up a substantial fraction of the total gains from trade to the investor due to the informational rents. Regarding the distortion on issuance, note that under complete information every meeting ends up in issuance because investors' valuations are always above the issuance cost $c = \underline{\nu}$. Therefore, $Issuance^{CI} = \int d\frac{\Phi_n(\nu_n)}{1-s} = 1$. But with incomplete information, issuance is always smaller than $\frac{1}{e} \approx 0.37$. Regarding the informational rents, the share of gains from trade captured by issuers is one under complete information, but it is bounded by 0.5 under incomplete information.

2.4 Secondary-market trade mechanism

The secondary market consists of trades between owners and non-owners. As in the primary market, we consider mechanisms that maximize the expected gains from trade of owners in the secondary market. A direct mechanism in the secondary market is a pair of functions $m^b = (p^b, x^b) : [\underline{\nu}, \bar{\nu}]^2 \to [0,1] \times \mathbb{R}$, where $p^b(\nu_o, \nu_n)$ represents the probability of transferring the asset when the owner has type ν_o and the non-owner has type ν_n, and $x^b(\nu_o, \nu_n)$ represents the transfer from the non-owner to the owner when the owner has type $\bar{\nu}$ and the non-owner has type ν_n. Define the functions

$$\bar{p}_n^b(\nu_n) = \int p^b(\nu_o, \nu_n) d\frac{\Phi_o(\nu_o)}{s}, \quad \bar{x}_n^b(\nu_n) = \int x^b(\nu_o, \nu_n) \frac{\Phi_o(\nu_o)}{s}$$
$$\bar{p}_o^b(\nu_o) = \int p^b(\nu_o, \nu_n) d\frac{\Phi_n(\nu_n)}{1-s}, \quad \text{and} \quad \bar{x}_o^b(\nu_o) = \int x^b(\nu_o, \nu_n) d\frac{\Phi_n(\nu_n)}{1-s}.$$

The interpretations for $\bar{p}_n^b(\nu_n)$, $\bar{x}_n^b(\nu_n)$, $\bar{p}_o^b(\nu_o)$, and $\bar{x}_o^b(\nu_o)$ are the same as in the primary market. The mechanism that maximizes the owner's expected gains from trade solves

$$\max_{m^b} \int \int \left[x^b(\nu_o, \nu_n) - p^b(\nu_o, \nu_n)\Delta(\nu_o) \right] d\frac{\Phi_n(\nu_n)}{1-s} d\frac{\Phi_o(\nu_o)}{s} \tag{8}$$

subject to

$$IR: \quad \bar{p}_n^b(\nu_n)\Delta(\nu_n) - \bar{x}_n^b(\nu_n) \geq 0 \tag{9}$$
$$IC: \quad \bar{p}_n^b(\nu_n)\Delta(\nu_n) - \bar{x}_n^b(\nu_n) \geq \bar{p}_n^b(\hat{\nu}_n)\Delta(\nu_n) - \bar{x}_n^b(\hat{\nu}_n) \tag{10}$$

for all ν_o, ν_n, and $\hat{\nu}_n$. Problem (8)-(10) is closely related to problem (1)-(3) in the simple model and it can be solved in a similar fashion. However, it has two important differences. First, the seller's opportunity cost in the primary market is just the issuance cost, while in the secondary market, the opportunity cost reflects the value of holding onto the asset and potentially selling it later, which equals the reservation value Δ. Second, in the primary market, the distribution of sellers' types, G, is exogenous. While in the primary market it is $\frac{\Phi_o}{s}$, which is endogenous.

Given a mechanism $m^b = (x^b, p^b)$, the expected gains from trade of an owner and a non-owner in a meeting in the secondary market are given by

$$\pi_o^b(\nu_o) = \bar{x}_o^b(\nu_o) - \bar{p}_o^b(\nu_o)\Delta(\nu_o) \quad \text{and} \tag{11}$$
$$\pi_n^b(\nu_n) = \bar{p}_n^b(\nu_n)\Delta(\nu_n) - \bar{x}_n^a(\nu_n). \tag{12}$$

The first equation is the expected gains from trade of an owner with type ν_o conditional on meeting with a non-owner. It takes expectation with respect to ν_n of the transfer from the non-owner, $x^b(\nu_o, \nu_n)$, minus the expected opportunity cost of selling the asset, $p^b(\nu_o, \nu_n)\Delta(\nu_o)$, plus his continuation value if the asset is not sold, $\Delta(\nu_o)$. The second equation is the expected gains from trade of a non-owner with type ν_n conditional on meeting with an owner. It takes expectation with respect to ν_o of the expected gain of getting the asset, $p^b(\nu_o, \nu_n)\nu_n$, minus the transfer to the owner, $x^b(\nu_o, \nu_n)$. As before, the gains from trade are evaluated under the assumption that owners and non-owners are truthfully revealing their types to the mechanism.

2.5 Solution to the secondary-market trade mechanism

The secondary market can be solved in a similar fashion to the way we solve the mechanism in the primary market. We assume that the distribution of types of owner investors, $\frac{\Phi_o}{s}$, the distribution of types of non-owner investors, $\frac{\Phi_n}{1-s}$, and the reservation value function, Δ, are differentiable, and Δ' is positive and bounded away from zero. Define the functions

$$\bar{p}_n^b(\nu_n) = \int p^a(c, \nu_n) dG(c), \text{ and } \bar{x}_n^b(\nu_n) = \int x^a(c, \nu_n) dG(c).$$

The terms $\bar{p}_n^b(\nu_n)$ and $\bar{x}_n^b(\nu_n)$ are analogous to the primary market. Then we have the following characterization of the problem.

Proposition 3. *Let the distribution of owners, $\frac{\Phi_o}{s}$, of non-owners, $\frac{\Phi_n}{1-s}$, and the reservation value, Δ, be differentiable and Δ' strictly positive and bounded away from zero. Then a direct mechanism $m^b = (p^b, x^b)$ solves problem (8)-(10) if, and only if, it solves*

$$\max_{m^b} \int \int p^a(c, \nu_n) \left[\Delta(\nu_n) - \frac{\Delta'(\nu_n)}{haz(\nu_n)} - \Delta(\nu_o) \right] d\frac{\Phi_n(\nu_n)}{1-s} d\frac{\Phi_o(\nu_o)}{s} \tag{13}$$

subject to $\bar{p}_n^b(\nu_n)$ being weakly increasing in ν_n, and $\pi_n^b(\nu_n) = \int_{\underline{\nu}}^{\nu_n} \bar{p}_n^b(\nu) \Delta'(\nu) d\nu$.

The proof of the proposition is analogous to the proof of Proposition 1, and thus we omit it. The following proposition characterizes the solution.

Proposition 4. *Let the distribution of non-owners, $\frac{\Phi_n}{1-s}$, and the reservation value, Δ, be both differentiable and Δ' strictly positive and bounded away from zero. Define the functions*

$$c_n(\nu) = g\left(\frac{\Phi_n(\nu)}{1-s}\right) \quad \textit{and}$$
$$c_n^{-1}(\nu) = \begin{cases} \inf\{\Delta(\nu_n); c_n(\nu_n) \geq \Delta(\nu)\} & \textit{if } c_n(\nu_n) \geq \Delta(\nu) \textit{ for some } \nu_n \\ \bar{\nu} & \textit{otherwise} \end{cases} .$$

Then the direct mechanism $m^a = (p^a, x^a)$, defined as

$$p^b(\nu_o, \nu_n) = \begin{cases} 1 & \textit{if } c_n(\nu_n) \geq \Delta(\nu_o) \\ 0 & \textit{otherwise} \end{cases} \quad \textit{and} \quad x^b(\nu_o, \nu_n) = \begin{cases} c_n^{-1}(\nu_o) & \textit{if } c_n(\nu_n) \geq \Delta(\nu_o) \\ 0 & \textit{otherwise.} \end{cases} ,$$

achieves the maximum in problem (13).

The proof of the proposition is analogous to the proof of proposition 2, and thus we do not provide it.

2.6 The value functions of investors and issuers

The value function of an owner is given by

$$rV_o(\nu) = \nu + \mu\left[V_n(\nu) - V_o(\nu)\right] + \lambda_b(1-s)\pi_o^b(\nu) . \tag{14}$$

This equation states that the value of owning an asset, discounted at rate r, equals the sum of three terms. The first term accounts for the flow utility of holding the asset, ν. The second term is the change in value when the asset matures and the owner becomes a non-owner, which happens at rate μ. The third term accounts for the net gains of an owner meeting a non-owner. An owner meets a non-owner at rate $\lambda_b(1-s)$, where $1-s$ is the fraction of investors that are non-owners. The value function of a non-owner follows similarly,

$$rV_n(\nu) = \lambda_b s \pi_n^b(\nu) + \lambda_a \pi_n^a(\nu) \ . \tag{15}$$

The value for a non-owner of type ν is the sum of two terms. The first term accounts for the net gains for the non-owner meeting an owner. A non-owner meets an owner at rate $\lambda_b s$, where s is the fraction of investors that are owners. The second term accounts for the net gains accrued by the non-owner when meeting an issuer, an event that occurs at rate λ_a. Using equations (14)-(15), we can compute the reservation value function for investors, $\Delta(\nu) = V_o(\nu) - V_n(\nu)$, according to

$$r\Delta(\nu) = \nu - \mu\Delta(\nu) + \lambda_b(1-s)\pi_o^b(\nu) - \lambda_b s \pi_n^b(\nu) - \lambda_a \pi_n^a(\nu) \ . \tag{16}$$

Finally, the value function for an issuer is given by

$$rV_c = \lambda_a(1-s) \int \pi_c^a(c) dG(c) \ . \tag{17}$$

This equation provides that the flow value of being an issuer equals the arrival rate of an issuance opportunity times the probability of finding a non-owner investor times the expected profits that the issuer would get if he produces and sells the asset to a non-owner investor.

2.7 The distribution of assets

In this section, we provide equilibrium conditions for the distribution of owners' types, $\Phi_o(\nu)$, the distribution of non-owners' types, $\Phi_n(\nu)$, and for the fraction of agents holding assets in the secondary market, s. The law of motion for $\Phi_o(\nu)$ is given by

$$\dot{\Phi}_o(\nu) = -\mu\Phi_o(\nu) - \lambda_b \int_{\underline{\nu}}^{\nu} \int_{\nu}^{\bar{\nu}} p^b(\nu_o,\nu_n) d\Phi_n(\nu_n) d\Phi_o(\nu_o) + \lambda_a \int_{\underline{c}}^{\bar{c}} \int_{\underline{\nu}}^{\nu} p^a(c,\nu_n) d\Phi_n(\nu_n) dG(c) \ . \tag{18}$$

This equation considers the change in the mass of owners with valuation below ν. The first term on the righthand-side of the equation accounts for those owners with valuation below ν that become non-owners because the asset matures. The second term accounts for meetings in the secondary market where owners with valuation below ν sell the asset to buyers with valuation above ν. Finally, the third term accounts for meetings where issuers sell the asset to non-owners with valuation below ν. Given the law of motion for $\Phi_o(\nu)$, we can easily obtain an expression for $\Phi_n(\nu)$ using that

$$\Phi_o(\nu) + \Phi_n(\nu) = F(\nu) \ , \tag{19}$$

which simply states that the measure of owners of type below ν plus the measure of non-owners of type below ν has to be equal to the total measure of investors of type below ν. Finally, by definition, all issued assets must be held by owners, or

$$\Phi_o(\bar{\nu}) = s. \tag{20}$$

2.8 Equilibrium

In this section, we define an equilibrium and provide an existence result. We stress that we do not impose conditions on the distributions Φ_o and Φ_n, nor on the reservation value Δ, as we did in Propositions 2 and 4. We start with the definition of a symmetric, steady-state equilibrium.

Definition 1. *A symmetric, steady-state equilibrium is a family* $\{\Delta, V_c, \Phi_o, \Phi_n, s, x^a, p^a, x^b, p^b\}$ *satisfying the following conditions:*

(i) the value function of an issuer V_c satisfies equation (17) where π_c^a is given by equation (4), and the reservation value of investors $\Delta(\nu)$ satisfies equation (16) where π_n^a, π_o^b and π_n^b are given by equations (5), (11), and (12);

(ii) the measure Φ_o is such that $\dot{\Phi}_o = 0$, where $\dot{\Phi}_o$ is defined in equation (18), Φ_n satisfies equation (19), and the amount of assets in the economy s satisfy (20); and

(iii) the mechanism in the primary market, x^a and p^a, solves problem (1)-(3), while the mechanism in the secondary market, x^b and p^b, solves problem (8)-(10).

Notice that the definition of equilibrium does not include the value function V_o and V_n. We do not include these two terms because we can solve for the equilibrium solely with the reservation value function Δ and then back out V_o and V_n using (14) and (15).

The existence of a solution to the problem that we analyze here can be complicated. In a partial equilibrium sense, solving for the optimal trade mechanism reduces to finding a solution to problem (6) and (13) but takes as given the distribution of valuations. However, because the distributions, Φ_o and Φ_n, and reservation value, Δ, are endogenous, existence also involves a fixed point in these objects.

Proposition 5. *There exists a symmetric steady-state equilibrium.*

The idea of the proof of the proposition is the following. We begin by constructing a sequence of distribution functions, valuations, and mechanisms. The operator we use to construct this sequence updates the measures Φ_o^k and Φ_n^k using polynomials (for example, Chebyshev polynomials) of degree k for the associated densities. This guarantees that the distributions are well behaved through the whole sequence, and thus we can apply Proposition 2 and 4 to generate the $k+1$ term from the k term in the sequence. The next step is to show that the sequence $\{\Delta^k, \Phi_o^k, \Phi_n^k\}_k$ is equi-continuous and then, by the Arzelà-Ascoli theorem, it has a subsequence that converges uniformly.[8] The convergence of a subsequence of $\{\Delta^k, x^{ka}, p^{ka}, x^{kb}, p^{kb}\}_k$ follows similar arguments. Hence, we

[8]Note, however, that even though for the entire sequence we have that Φ_o^k and Φ_n^k are continuously differentiable, the limits do not have to be since the space $\mathscr{C}^1$ is not closed.

conclude that, passing to a subsequence if necessary, our original sequence has a limit $\{\Phi_o^*, \Phi_n^*, s^*, \Delta^*, x^{*a}, p^{*a}, x^{*b}, p^{*b}\}$. The last step in the proof is to show that this limit satisfies the equilibrium conditions. For that, we use the Stone-Weierstrass theorem to argue that the polynomials are a good approximation of the limit so that optimality must hold.

3 Complete information

An important benchmark for our model is a version where, in any meeting, investors' valuations, ν, are observed. That is, there is complete information. In this case, the seller can extract full surplus from the trade. The trade mechanisms are easy to characterize, given by

$$p^{a,CI}(c,\nu_n) = \begin{cases} 1 & \text{if } \Delta^{CI}(\nu_n) \geq c \\ 0 & \text{otherwise} \end{cases}, \quad x^{a,CI}(c,\nu_n) = \begin{cases} \Delta^{CI}(\nu_n) & \text{if } \Delta^{CI}(\nu_n) \geq c \\ 0 & \text{otherwise} \end{cases}, \tag{21}$$

$$p^{bCI}(\nu_o,\nu_n) = \begin{cases} 1 & \text{if } \nu_n \geq \nu_o \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad x^{bCI}(\nu_o,\nu_n) = \begin{cases} \Delta^{CI}(\nu_n) & \text{if } \nu_n \geq \nu_o \\ 0 & \text{otherwise} \end{cases}. \tag{22}$$

Notice, under complete information, trade is always bilaterally ex-post efficient. Issuers create assets if the seller they match with has a larger valuation than their cost and investors trade in the secondary market whenever the non-owner values the asset more than the owner. The equations for the reservation value and distribution of valuations are the same as in the incomplete information model but replacing the mechanisms (x^a, p^a, x^b, p^b) with $(x^{aCI}, p^{aCI}, x^{bCI}, p^{bCI})$. That is,

$$\begin{aligned} \dot{\Phi}_o^{CI}(\nu) = & -\mu\Phi_o^{CI}(\nu) \\ & -\lambda_b \int_{\underline{\nu}}^{\nu} \int_{\nu}^{\bar{\nu}} p^{bCI}(\nu_o,\nu_n) d\Phi_n^{CI}(\nu_n) d\Phi_o^{CI}(\nu_o) \qquad (23) \\ & +\lambda_a \int_{\underline{c}}^{\bar{c}} \int_{\underline{\nu}}^{\nu} p^{aCI}(c,\nu_n) d\Phi_n^{CI}(\nu_n) dG(c) \text{ , and} \qquad (24) \end{aligned}$$

$$r\Delta^{CI}(\nu) = \nu - \mu\Delta^{CI}(\nu) + \lambda_b(1-s)\pi_o^{bCI}(\nu) - \lambda_b s\pi_n^{bCI}(\nu) - \lambda_a\pi_n^{aCI}(\nu)\ . \tag{25}$$

The equilibrium definition also follows in the same fashion. We conjecture that a complete information equilibrium always exists, as in the incomplete information case, and it is also unique and implements the first best allocation (that is, the allocation that maximizes aggregate welfare subject to the search frictions). We have partial results suggesting this is the case, but the proof is still incomplete.

4 Quantitative Analysis

In this section, we quantitatively study the implications of private information in the US municipal bond market. The US municipal bond market is (i) large, with nearly 40,000 transactions per day at an average par of $14 million, (ii) decentralized – all trades occur bilaterally – and (iii) features low default risk. Low default risk is important in that private information about the payoffs of municipal bonds are not a major friction in this market.

4.1 MSRB data and sample description

We use data on municipal bond transactions collected and provided by the Municipal Securities Rulemaking Board (MSRB). The MSRB requires securities dealers, issuers, and those acting on their behalf to submit information on municipal bond trades and disclosure documents for all transactions within 15 minutes of the time of trade. The data are then released publically, with a lag, through the MSRB's Electronic Municipal Market Access (EMMA) portal, however we obtain historical data through Wharton Research Data Services (WRDS). The dataset includes transaction-level information for all trades of municipal bonds involving a securities dealer in which the municipal bond was assigned a unique CUSIP identification number. This nearly covers the universe of municipal bond trades. Our sample includes transactions from January 3, 2005, through

December 31, 2014.

For each transaction, we observe characteristics about the issue being traded, such as the unique CUSIP identifier, maturity date, coupon rate, and the date when interest began to accrue as well as trade characteristics such as the date and time the trade took place, the price, the par value, and whether or not the trade was a purchase by a dealer from an investor, a sale from a dealer to an investor, or an inter-dealer transaction. In some instances, the yield is reported. We obtain other bond and issuer characteristics from the CUSIP Global Services Master File including the geography and type of the issuer (e.g. country, school district, development authority, etc.) as well as the bond's type (e.g. tax revenue or general obligation), its tax-exempt status, and whether or not the bond is callable. Since the yield on the transaction is not reported for a significant portion of the sample, we infer the yield on missing observations using information on the coupon rate, price, and time to maturity. We further describe our process as well as our sampling procedure in Appendix B.

Table 1 reports the descriptive statistics for our sample. Following Green et al. (2007), we label all trades that take place within the first 90 days from issuance as 'primary market' transactions and those occurring after as 'secondary market' transactions.[9]

Municipal bonds are actively traded after their initial issuance. From 2005-2014, we observe 86.4 million transactions of 1.89 million bond issues, 78% of which occur in the secondary market. The total value of all trades was $17.7 quadrillion (or $1.77 trillion per year), of which 62% occur in the secondary market. Hence, primary market trades tend to be larger; the average par on transactions in the primary market was $370,000 while only $160,000 in the secondary market. The average maturity of issues traded was 17.83 years at an average yield of 3.9%. There is not a significant difference between the yield on trades in the primary market versus the secondary market, 3.97% and 3.90%, respectively. Table 1 reports descriptive statistics.

[9]Alternatively, we could follow the MSRB in denoting all "new issue trades" as those occurring within the first 30 days. However, we chose to be consistent with the literature for comparison.

Table 1: Descriptive Statistics on Municipal Bond Transactions, 2005-14

	All Transactions	Primary Market	Secondary Market
Observations (millions)	86.40	18.50	67.90
Issues (millions)	1.89	1.26	1.25
Total value (trillion $)	17,701.69	6,840.02	10,861.67
Average par (million $)	0.20	0.37	0.16
Coupon rate (%)	4.52	4.20	4.60
Yield-to-Worst (%)	3.91	3.97	3.90
Maturity (years)	17.83	15.34	18.51
Years until maturity	13.66	15.33	13.20

Note: Sample is all municipal bond transactions reported to the MSRB from 2005-14. We drop observations on variable rate securities, if missing price, or if the par value traded is less than $1,000. We winsorize the price and yield at 99%. If the yield-to-worst is not reported on the transaction, we use the estimated yield-to-maturity.

4.2 Calibration

Using the moments above, we calibrate the model as follows. We assume that the distribution of investors' valuations, F, is a truncated log-normal, with parameters μ_F and σ_F. The lower limit of the truncation, $\underline{v}$, is kept at zero and the upper limit, $\bar{v}$, is set at $\bar{v} = F^{-1}(0.999)$. The issuance cost is $c = \frac{\tilde{c}}{r+\mu}$, where $\tilde{c}$ is also truncated log-normal, with parameters μ_G and σ_G, and the limits are the same as in the F distribution (that is, $\underline{c} = \underline{v}$ and $\bar{c} = \bar{v}$). The above specification gives eight parameters: r, μ, λ_a, λ_b, μ_F, σ_F, μ_G, and σ_G. We calibrate to an annual frequency and set the discount rate $r = 0.05$. We set the maturity parameter so that the average maturity in the model matches the one in the data, $\mu = 1.0/17.83$.

This leaves us with six parameters to calibrate, the frequency of trade opportunities in the primary and secondary market, λ_a and λ_b, and the mean and variance of the value and cost distributions. We use the following moments in the data to discipline these six parameters. In order to capture the importance of the secondary market for municipal bonds, we the target the relative size of this market in the total value of trade of 62%. This helps discipline the relative size of λ_b versus λ_a. As we illustrated using the simple version of the model, the shape of the distribution of investor valuations and issuer costs are important in understanding the scope of informational rents that arise from private information. To discipline the parameters of these distribution, we target moments from the distribution of yields on transactions in the MSRB data. Specifically, we target the mean and dispersion of yields across transactions for both the primary and secondary market. The average yield on transactions in the primary and secondary market is taken directly from Table 1. Yield dispersion in the model arises only due to heterogeneity in valuations and costs for bonds with a known dividend (or coupon rate) and maturity. However, in the data, yield dispersion arises for many reasons. Yields vary within time and across issues because of differing observable characteristics, such as coupon rates or taxable status. Yields vary within issues and across time due to both aggregate effects, such as changing market conditions, as well as idiosyncratic effects. The dispersion in yields that is most closely related to the one in our model is dispersion within issues across time controlling for aggregate time effects. In order to capture time-variant aggregate volatility, we de-mean the yield on each transaction by a group-specific, daily average yield. We categorize issues into groups according tax-exempt status, callable status, and the source of repayment (general obligation or tax revenue). Then, we calculate the average, unweighted within group variance of the yield over our entire sample, 2005-14. We do this separately for the primary and secondary market. This gives us five moments to match six parameters. For the last moment, we use data on the expected number of times an issue is traded on the secondary market as a function of the amount of time since issuance. In the data, we observe that as issues mature, they are less likely to be traded. The same is true in the model since, on average, issues are allocated from low-valuation investors to high-valuation investors. Hence, the path of the frequency of trade also helps us discipline the meeting rate in the secondary market and the distribution of valuations.

Table 2 reports the calibrated parameters of the model, while Table 3 and Figure 1 depict the outcomes of the model in comparison with the data. In the general, the

Table 2: Calibrated parameters

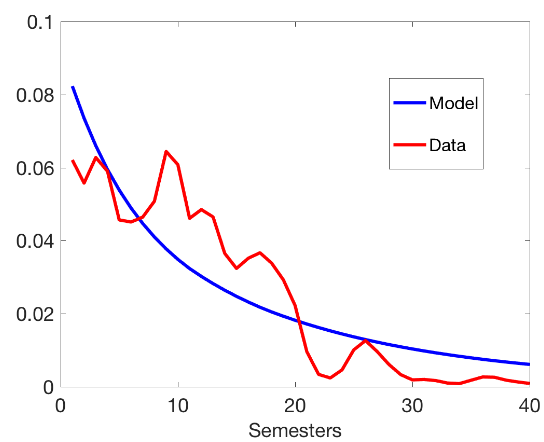
r	μ	λ_a	λ_b	μ_F	σ_F	μ_G	σ_G
0.050	0.056	0.100	4.650	-0.001	0.970	-0.097	0.688

model does well matching the data, with the exception being the difference in average yields between the primary and secondary market. The reason is that the model always generates prices in the secondary market higher than those in the primary market since investors who sell assets in the secondary market do not change valuation from when they originally purchased the issue in the primary market. Hence, if they re-sell, it must be at a higher price or lower yield. In the data, while the yield is lower in the secondary market, quantitatively they are very close, and as a result the model cannot match it exactly. However, the model matches the average yield well across all transactions; in the data, average yield in both markets is 3.91% while in the model it is 4.00%.

Table 3: Moments

Variables	Data	Model
Average maturity	14.96	14.96
Primary/Total market	0.38	0.39
Primary market		
Average yield	3.96	5.66
Yield std. deviation	0.30	0.31
Secondary market		
Average yield	3.90	2.96
Yield std. deviation	1.10	1.10
Average yield	3.91	4.00

Figure 1: Trade after issuance



4.3 The pattern of trades

In the secondary market, trade occurs when two investors meet, one who owns the asset and one who does not, and the virtual valuation of the non-owner of the asset is above the reservation value of the asset's owner. Figure 2 depicts the reservation value and the virtual valuation of investors given their type. The virtual valuation is strictly below the reservation value, which implies that trade is inefficient. That is, for some meetings between an investor who owns an asset and one who does not, trade will not occur in situations where the reservation value of the non-owner of the asset is higher than the one of the owner. Figure 3 illustrates this result. It depicts the region of asset owner valuations, ν_o, and non-owner valuations, ν_n, in which meetings result in trade. For all pairs of valuations (ν_o, ν_n) below the 45° line, trade is the efficient outcome. However, trade only occurs in the purple trade region. The pink region depicts the loss of efficient trades caused by private information.

Figure 2: Virtual valuation

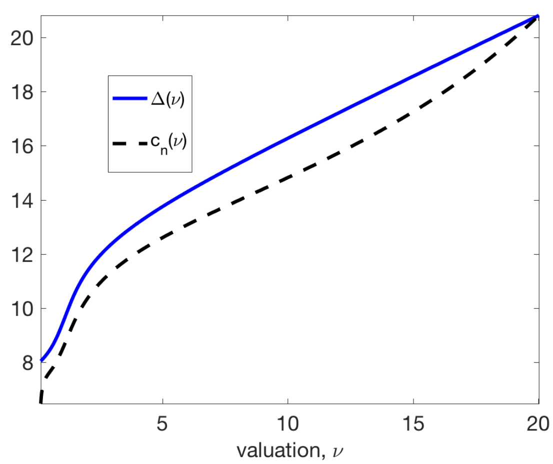
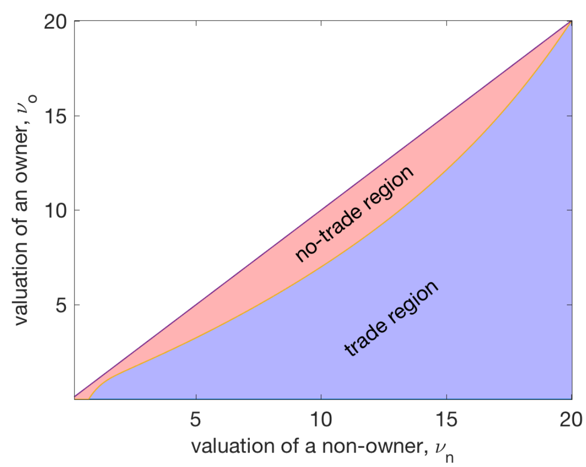


Figure 3: Trade region



Trade occurs in the primary market in the same way as the secondary market. In a meeting between an issuer and an investor who does not own an asset, the issuer issues an asset to the investor if the investor's virtual valuation is above the issuing cost. The fact that trade occurs only when virtual valuations are above the issuance cost, as opposed to needing the reservation value to be above the issuance cost, generates an inefficiency similar to the one in the primary market.

4.4 Asset supply, trade volume, and welfare

There are two important channels in which the private information affects welfare. The first one is a direct channel. Since virtual valuations are lower than reservation values, some meetings do not end up in trade/issuance even though the reservation value of the investor is higher than the reservation value/issuing cost of the potential buyer for the asset. As a consequence, the total surplus in a meeting is not realized. This channel is well explained in our simple version of the model.

The second channel in which private information affects welfare is an indirect channel due to the dynamic nature of asset markets. How much an investor is willing to pay for an asset (the reservation value) depends not only on his own valuation of the asset, ν, but also on the expected surplus from selling the asset, $\pi_o^b(\nu)$, his outside option of buying the asset from another investor, $\pi_n^b(\nu)$, and his outside option of buying the asset from an issuer, $\pi_n^a(\nu)$. The resale value of the asset goes down with private information because the potential buyer can extract informational rents from the investor trying to sell an asset. Moreover, the outside options go up because the investor can extract informational rents when buying the asset from another investor or from an issuer. The lower reservation value under incomplete information implies that investors are less willing to pay issuers to issue an asset. As a result, issuance and asset level are both inefficiently low due to private information.

To better understand what is going on in the model, it is worth comparing it with a complete information benchmark. The difference between reservation values in the complete and incomplete information models, using (14) and (25), is given by

$$\Delta^{CI}(\nu) - \Delta(\nu) = \underbrace{\frac{\lambda_b[(1-s^{CI})\pi_o^{bCI}(\nu) - (1-s)\pi_o^b(\nu)]}{r+\mu}}_{resale\ profits} + \underbrace{\frac{\lambda_a \pi_n^a(\nu)}{r+\mu}}_{\substack{primary\\ market\\ rents}} + \underbrace{\frac{\lambda_b s \pi_n^b(\nu)}{r+\mu}}_{\substack{secondary\\ market\\ rents}}, \tag{26}$$

noting that $\pi_n^{iCI} = 0$ for $i = a, b$ or that non-owners receive zero gains from trade under complete information. Equation (26) provides a way to decompose the difference in reservation values between the complete and the incomplete information models. Figure 4 illustrates this difference in the calibrated version of the model, and Figure 5 decomposes this difference. For investors with low types, ν, the difference is mostly due to losses in resale value. Resale value captures the expected gains from trade of an owner reselling the asset in the secondary market. Under incomplete information, a low-valuation owner is not able to seize all the gains from trade associated with meeting a high-valuation non-owner, which implies that π_o^{bCI} is larger than π_o^b. The effect of private information on resale values diminishes for larger investor types. For investors with higher values of ν, the difference in reservation values is largely due to the outside option of not owning the asset and collecting informational rents upon buying in the future. The last two terms correspond to this option for the primary and secondary market, respectively. Higher type investors collect larger informational rents and so posses a larger outside option under complete information. This creates a hold-up problem, leading to an inefficiently low reservation values for assets and reductions in liquidity. In the calibrated model, the hold-up problem in the secondary market is more severe than in the primary market, a feature we quantify later.

Figure 4: Reservation value

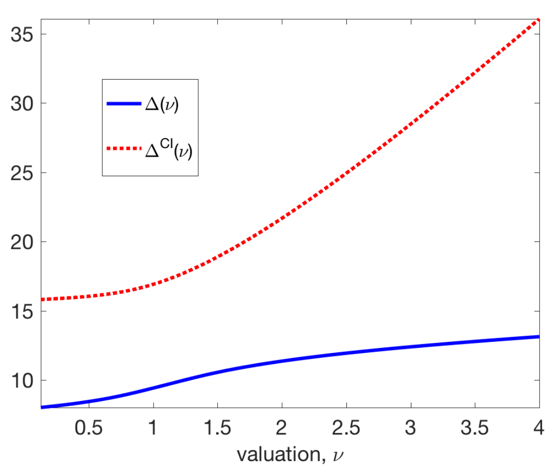
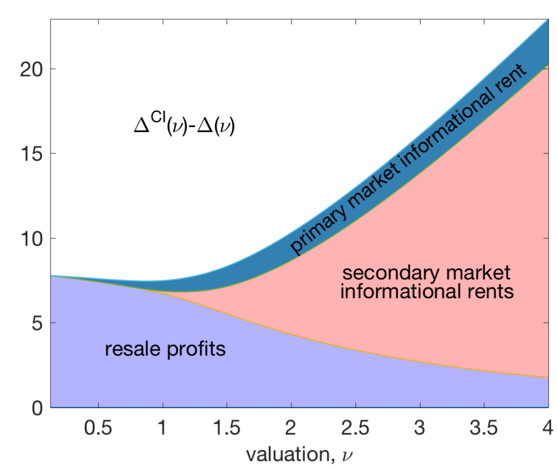


Figure 5: Difference in Δ



As illustrated in Figure 3, assets are not well allocated primarily because trade in the secondary market is not efficient (affecting resale profits and secondary market outside options). Additionally, the total level of assets is not efficient because trade in the primary market is also inefficient. This results in less issuance, leading to a lower asset supply, lower trade volume, and lower total welfare.

Table 4: Asset level, trade volume, and total welfare

	Asset level	Trade volume	Total welfare	Avg. Yield
Complete information	0.56	2.68	24.19	-1.23%
Incomplete information	0.44	0.60	22.36	4.00%
Difference	21.43%	77.61%	7.56%	5.23%

Table 4 illustrates the magnitude of the loss in asset supply, trade volume, and welfare for the US municipal bond market. Asset supply is 22% lower under incomplete information, and secondary market trade volume is 77% lower. Note that volume falls more than the level of assets because, besides a reduction in trade opportunities due to less assets in the economy, there are also trades which are lost due to the information friction. In terms of welfare, agents are nearly 8% worse off in the incomplete information economy than with complete information. Consistent with the fall in asset values, the average yield of assets in our economy is significantly higher under private information.

4.5 Misallocation and the distribution of assets

The inefficiencies generated by the information friction affect the distribution of asset holdings among owners and non-owners. Figures 6 and 7 depict these measures in comparison with the analogous measures associated with complete information. Total assets in the model with incomplete information is $s = 0.44$, while in the benchmark with complete information it is $s^{CI} = 0.56$. The difference in total assets, $s^{CI} - s$, is associated with the area between the curves in Figures 6 and 7.

The welfare cost associated with complete information is mostly due to investors

Figure 6: Measure of owners

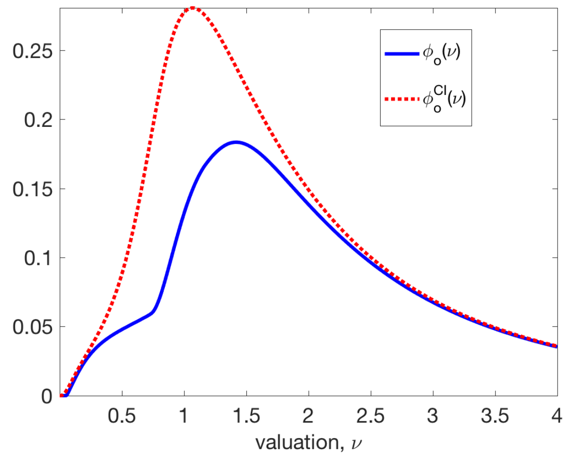
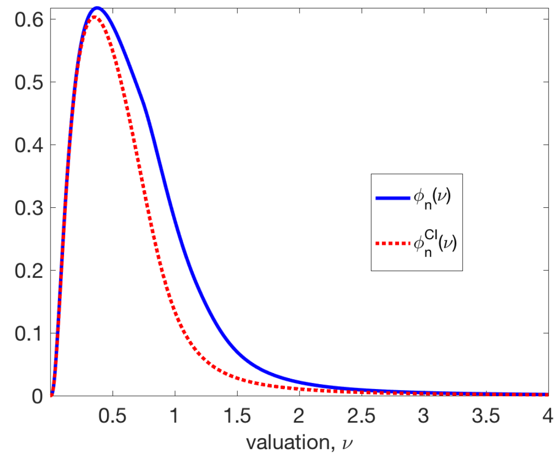


Figure 7: Measure of non-owners



with intermediate valuations (between 0.5 and 2.5) holding less assets than in the complete information benchmark. Investors with low valuations (ν close to zero) are not likely to hold assets in either economy. Similarly, investors with high valuation acquire assets in both economies. This pattern reflects a key feature of the trade mechanism: in order to capture rents from high-valuation investors, owners and issuers sacrifice trade with investors with average valuations, so that there fewer trades where investors with intermediate valuations participate. Private information skews the distribution of non-owners to the left resulting more high-valuation non-owners – both because of reduced asset supply and because of the no-trade region. Likewise, the distribution of owners is skewed to the right resulting in relatively more low-valuation owners.

Figures 8 and 9 illustrate the distribution of trades in the primary market between issuers and non-owner investors (left panel) and in the secondary market between owners and non-owners (right panel). The color indicates the mass of trades occurring for a given issuer or investor type. Although most trade in the primary market involves investors with low valuations in either the complete and incomplete information economies, the way trade is handled by these investors in the secondary market is very different between the two economies. Under complete information, owners sell their

Figure 8: Complete information

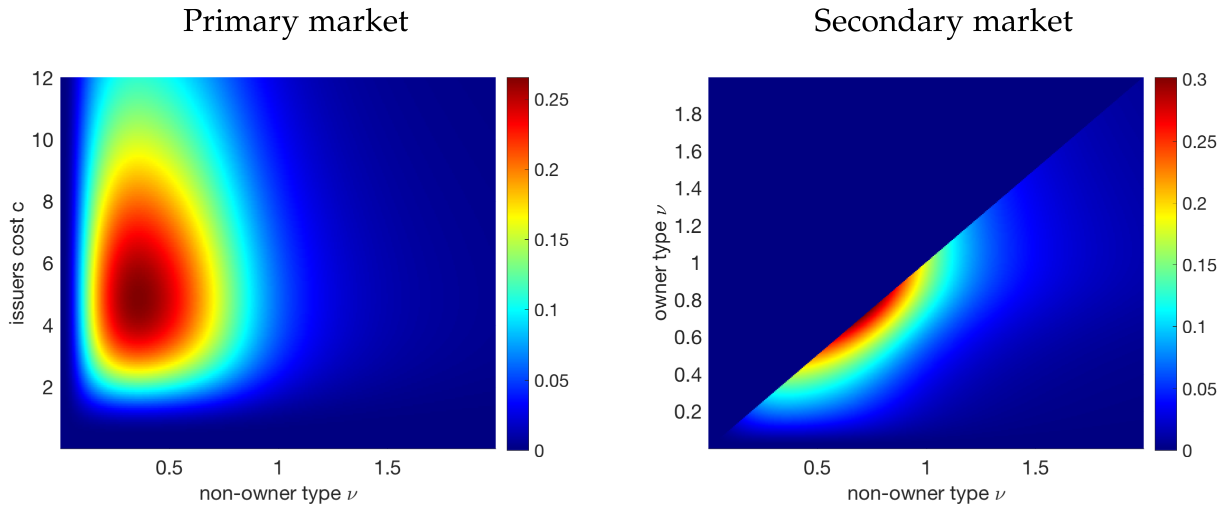
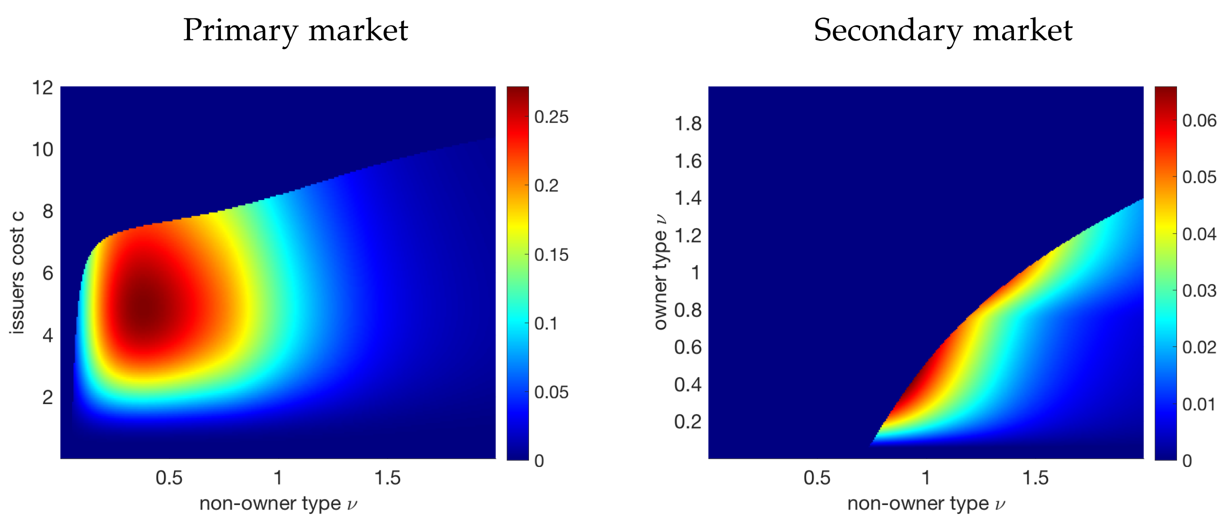


Figure 9: Incomplete information



assets to non-owners with valuation close to theirs, so that assets flow to high valuation investors slowly, involving intermediate valuation investors as middlemen—a feature identified by Hugonnier et al. (2014). Under incomplete information, because an owner wants to capture rents from trade from high-valuation investors, the role as middlemen of intermediate valuation investors is hindered, so that assets flow to high-valuation investors with fewer trades. Here, intermediate valuation investors have a more passive role in trade, and thus a smaller role as middlemen.

4.6 Primary *vs.* secondary markets

Private information distorts trade in both primary and secondary markets. A natural question is in which market are information frictions more severe. To answer this question, we solve two alternative models. In the first one, there is complete information in all meetings in the primary market but not in the secondary market. In the second one, there is complete information in all meetings in the secondary market but not in the primary market. Table 5 repeats the results from Table 4, decomposing the informational effects between the two markets.

Table 5: Decomposition: primary *vs* secondary market

	Asset level	Trade volume	Welfare	Welfare loss
Incomplete	0.44	0.60	22.36	7.59%
Complete—primary	0.45	0.64	22.58	6.64%
Complete—secondary	0.53	1.92	24.02	0.07%
Complete—both markets	0.56	2.68	24.19	-

The effect of private information in the primary market for municipal bonds is small. In this case, the welfare loss with respect to the complete information benchmark is 6.64%, while in the incomplete information version the welfare loss is 7.59%. Solving informational problems in the primary market only increases welfare by 0.95%. On the other hand, private information in the secondary market has substantial effects. Eliminating private information in secondary market trades nearly implements the complete information allocation. The welfare loss of only having private information in asset issuance is 0.07%.

The fact that the welfare cost of private information comes mostly from the information friction in the secondary market may lead some to think that primary market issuance has no role in generating the inefficiency, but that is not the case. The reason is that the channel in which the information friction in the secondary market reduces welfare works through asset supply, determined in the primary market. The information friction in the secondary market allows non-owners of assets to extract informational rents from owners. This reduces the value of having an asset, and as a result, issuance goes down because the value of buying an asset is lower. We can see this by noticing that even when there is complete information only in the primary market, asset supply is still 19.49% below the level under complete information in both markets. However, when there is complete information only in the secondary market, asset supply is only 5.61% lower.

5 Policy

In this section, we study a policy intervention with the aim of improving welfare in the economy with private information. The type of policy intervention that we consider is simple and restrictive. In particular, we restrict attention to a policy that subsidizes trade in the primary market, financed by charging a participation fee to issuers that intend to issue an asset in a given period. The restriction to only intervene in the primary market is motivated by the fact that in many asset markets, secondary market intervention is difficult given the market's opacity and the anonymity of participants. For instance in the municipal bond market, the MSRB requires dealers to report transactions (either as the buyer or the seller) but does not require them to report the identities of their counter-parties.[10] Even if taxation of individual trades were introduced, it is unclear if revealing the participants' identity would be individually rational. Our policy, while different, has a similar spirit to granting tax-exempt status to the interest payments of municipal bonds, a feature that the municipal finance literature finds encourages

[10]See MSRB Rule G-14 for dealer guidelines in reporting transactions and the Real Time Reporting System's (RTRS) user guide for an explanation of the information required when reporting, both available at www.msrb.org.

issuance.[11] Although the policy is simple, we show how the policy intervention in the primary market greatly improves welfare, even though the main distortion, as show in Section 4.6, occurs in the secondary market.

Let f denote the participation fee charged to issuers that intend to issue, and let τ denote the subsidy provided to those issuers that end up issuing an asset. We present the problem under the conjecture that $\{f, \tau\}$ are such that issuers are willing to participate in the program (and we then check ex-post that this is the case). With a slight abuse of notation, the value function for an issuer is given by

$$rV_c(\tau, f) = \lambda_a \left[(1-s) \int \pi_c^a(c; \tau)\, dG(c) - f \right]$$

where

$$\pi_c^a(c; \tau) = \int [x^a(c, \nu_n) - p^a(c, \nu_n; \tau)(c - \tau)]\, d\frac{\Phi_n(\nu_n)}{1-s}$$

are the expected gains from trade for an issuer with cost c in face of an issuance opportunity. Notice that all other objects in the economy remain unchanged, as they only depend on f and τ indirectly. The objective of the planner is

$$\max_{\tau} V_c(\tau, f) + \int V_o(\nu_o)\, d\Phi_o(\nu_o) + \int V_n(\nu_n)\, d\Phi_n(\nu_n)$$

subject to

$$\begin{aligned} \frac{f}{\lambda_a} &= (1-s) \int \int \tau p^a(c, \nu_n; \tau)\, d\frac{\Phi_n(\nu_n)}{1-s}\, dG(c)\, , \\ 0 &\leq V_c(\tau, f)\, . \end{aligned}$$

The first restriction states that the budget constraint of the planner needs to be satisfied, while the second restriction states that issuers need to be willing to participate in the market. Using the budget constraint of the planner together with the expression for

[11] See for example Gordon and Slemrod (1983) and Fortune (1998).

$V_c(\tau, f)$, we obtain that the participation restriction reduces to

$$\int [x^a(c, \nu_n) - p^a(c, \nu_n; \tau)c] \, d\frac{\Phi_n(\nu_n)}{1-s} \geq 0 \, .$$

If τ were to be zero (so that there is no intervention by the planner), this equation is immediately satisfied as this equation is simply the profits of an issuer with cost c, which we know are positive. However, for $\tau > 0$ the trade probability is distorted upwards, and thus this equation might not be satisfied for some $\tau > 0$.

We look for the optimal policy by a grid search method. Although slow, this method is reliable and lets us circumnavigate potential issues with local maxima. We find that the optimal subsidy is $\tau^* = 3.12$ with an implied participation fee $f^* = 1.00$. To put things in perspective, the expected production cost is around 9.8, so that the optimal subsidy accounts for about 1/3 of the production cost of the average issuer.

Table 6: Outcomes of policy

	Asset level	Trade volume	Welfare	Welfare loss
Incomplete	0.44	0.60	22.36	7.59%
Complete—primary	0.45	0.64	22.58	6.64%
Incomplete with policy	0.49	0.58	23.87	1.32%
Complete—secondary	0.53	1.92	24.02	0.70%
Complete—both markets	0.56	2.68	24.19	-

We learn an important result from this exercise. At least under our calibration, the main cost of private information is not due to the inefficiency of individual trades—though this cost is obviously still there. The main cost comes from the inefficiency in issuance that is caused by the private information. So a simple tax/subsidy scheme that spurs new issuance is able to reduce the welfare cost of private information from 7.59% to just 1.32%, as we can see in Table 6.

6 Conclusion

In this paper, we analyzed the effects of private information about idiosyncratic valuations for assets. The environment nests workhorse models of trade in decentralized OTC markets. Despite the fact that private information can easily lead to tractability problems, we can allow for a general distribution of valuations and prove that an equilibrium exists. Further, our model lends itself easily to quantitative analysis. We calibrate to the US municipal bond market and illustrate that private valuations are a significant source of inefficiency; restoring information leads to a welfare gain of 7.6%. Additionally, private information affects aggregate trade patterns. Assets are no longer reallocated by small surplus, infra-marginal trades. It is precisely these trades that private information eliminates.

In summary, we show that private value problems are a relevant friction in decentralized OTC markets. These problems are magnified in times of increased risk (dispersion of valuations) or liquidity events (times of decreased demand) and can lead to large welfare losses. Policies designed to address information problems in financial markets are often targeted to eliminate private information about common payoffs, seeking to address such problems as adverse selection. We emphasize that an environment with no uncertainty about common payoffs can still suffer from the inefficiency caused by private information.

References

Bigio, S. (2015). Endogenous liquidity and the business cycle. *American Economic Review*, 105(6):1883–1927.

Chang, B. (2014). Adverse selection and liqudity distortion.

Chari, V., Shourideh, A., and Zetlin-Jones, A. (2014). Reputation and persistence of adverse selection in secondary loan markets. *American Economic Review*, 104(12):4027–4070.

Chiu, J. and Koeppl, T. (2015). Trading dynamics with adverse selection and search: Market freeze, intervention and recovery. *Review of Economic Studies*, 00:1–35.

Cujean, J. and Praz, R. (2015). Asymmetric information and inventory concerns in over-the-counter markets.

Duffie, D. (2012). *Dark Markets*. Princeton University Press.

Duffie, D., Garleanu, N., and Pedersen, L. H. (2005). Over-the-counter markets. *Econometrica*, 73:1815–1847.

Eisfeldt, A. (2004). Endogenous liquidity in asset markets. *The Journal of Finance*, 59(1):1–30.

Fortune, P. (1998). Tax-exempt bonds really do subsidize municipal capital! *National Tax Journal*, 51(1):43–54.

Fundenberg, D. and Tirole, J. (1991). *Game Theory*. MIT Press.

Gordon, R. and Slemrod, J. (1983). A general equilibrium simulation of subsidies to municipal expenditures. *The Journal of Finance*, 38(2):585–594.

Green, R., Hollifield, B., and Schurhoff, N. (2007). Financial intermediation and the costs of trading in an opaque market. *The Review of Financial Studies*, 20(2):275–314.

Guerrieri, V., Shimer, R., and Wright, R. (2010). Adverse selection in competitive search equilibrium. *Econometrica*, 78(6):1823–1862.

Hugonnier, J., Lestera, B., and Weill, P.-O. (2014). Hetergeneity in decentralized asset markets.

Krishna, V. (2010). *Auction Theory*. Academic Press, 2nd edition.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73.

Myerson, R. B. and Satterthwaite, M. A. (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 29:265–281.

Zhang, S. (2016). Liqudity misallocation in an over-the-counter market.

A Appendix: Proofs

A.1 Proposition 1

It is useful to begin by stating the following lemma.

Lemma 1. *A mechanism $m^a = (p^a, x^a)$ satisfies the constraints*

$$
\begin{aligned}
\textit{IR(issuer)}: &\quad \bar{x}^a_c(c) - \bar{p}^a_c(c)c \geq 0 \\
\textit{IC(issuer)}: &\quad \bar{x}^a_c(c) - \bar{p}^a_c(c)c \geq \bar{x}^a_c(\hat{c}) - \bar{p}^a_c(\hat{c})c \\
\textit{IR(investor)}: &\quad \bar{p}^a_n(\nu_n)\Delta(\nu_n) - \bar{x}^a_n(\nu_n) \geq 0 \\
\textit{IC(investor)}: &\quad \bar{p}^a_n(\nu_n)\Delta(\nu_n) - \bar{x}^a_n(\nu_n) \geq \bar{p}^a_n(\hat{\nu}_n)\Delta(\nu_n) - \bar{x}^a_n(\hat{\nu}_n)
\end{aligned}
$$

if, and only if, $\bar{p}^a_c(c)$ is decreasing, $\bar{p}^a_n(\nu_n)$ is increasing,

$$\pi^a_c(c) = \pi^a_c(\bar{c}) + \int_c^{\bar{c}} \bar{p}^a_c(\tilde{c})d\tilde{c}, \tag{27}$$

$$\pi^a_n(\nu_n) = \pi^a_n(\underline{\nu}) + \int_{\underline{\nu}}^{\nu_n} \bar{p}^a_n(\nu)\Delta'(\nu)d\nu, \tag{28}$$

$\pi^a_c(\bar{c})$ and $\pi^a_n(\underline{\nu})$ are greater or equal to zero.

Proof. Let us start showing the necessity part. We show the result for the IR(investor) and IC(investor) of the investor since IR(issuer) and IC(issuer) follow the exact same steps. Note that we can write π_n^a as

$$\pi_n^a(\nu_n) = \bar{p}_n^a(\nu_n)\Delta(\nu_n) - \bar{x}_n^a(\nu_n).$$

The IC(investor) constraint implies that

$$\begin{aligned}\pi_n^a(\nu_n) \geq \bar{p}_n^a(\hat{\nu}_n)\Delta(\nu_n) - \bar{x}_n^a(\hat{\nu}_n) = \pi_n^a(\hat{\nu}_n) + \bar{p}_n^a(\hat{\nu}_n)[\Delta(\nu_n) - \Delta(\hat{\nu}_n)] \\ \implies \pi_n^a(\nu_n) - \pi_n^a(\hat{\nu}_n) \geq \bar{p}_n^a(\hat{\nu}_n)[\Delta(\nu_n) - \Delta(\hat{\nu}_n)]\end{aligned}$$

Analogously, the IC(investor) constraint implies that $\pi_n^a(\hat{\nu}_n) - \pi_n^a(\nu_n) \geq \bar{p}_n^a(\nu_n)[\Delta(\hat{\nu}_n) - \Delta(\nu_n)]$. Reorganizing these equations, we have that

$$\bar{p}_n^a(\nu_n)[\Delta(\nu_n) - \Delta(\hat{\nu}_n)] \geq \pi_n^a(\nu_n) - \pi_n^a(\hat{\nu}_n) \geq \bar{p}_n^a(\hat{\nu}_n)[\Delta(\nu_n) - \Delta(\hat{\nu}_n)].$$

And we can conclude that $\bar{p}_n^a$ is weakly increasing. Moreover, because $\bar{p}_n^a$ is monotone, it has at most countable many discontinuities. Therefore, we can use that

$$\bar{p}_n^a(\nu_n)\frac{\Delta(\nu_n) - \Delta(\hat{\nu}_n)}{\nu_n - \hat{\nu}_n} \geq \frac{\pi_n^a(\nu_n) - \pi_n^a(\hat{\nu}_n)}{\nu_n - \hat{\nu}_n} \geq \bar{p}_n^a(\hat{\nu}_n)\frac{\Delta(\nu_n) - \Delta(\hat{\nu}_n)}{\nu_n - \hat{\nu}_n}.$$

and that $\bar{p}_n^a$ is continuous almost everywhere to conclude that π_n^a is differentiable almost everywhere and it must satisfy

$$\frac{\partial \pi_n^a(\nu_n)}{\partial \nu_n} = \bar{p}_n^a(\nu_n)\Delta'(\nu_n) \quad \implies \quad \pi_n^a(\nu_n) = \pi_n^a(\underline{\nu}) + \int_{\underline{\nu}}^{\nu_n} \bar{p}_n^a(\nu)\Delta'(\nu)d\nu.$$

Moreover, because $m^a = (p^a, x^a)$ satisfies the IR constraint, $\pi_n^a(\nu_n) \geq 0$ for all ν_n. Hence, we must have that $\pi_n^a(\underline{\nu}) \geq 0$. The necessity part for issuers follow the same steps.

For the sufficient part we again only show it for investors, but the same logic applies for issuers. Note first that, if $\pi_n^a(\nu_n)$ satisfies (27) and $\pi_n^a(\underline{\nu}) \geq 0$, then $\pi_n^a(\nu_n) \geq 0$ for all ν_n since (27) implies that $\pi_n^a(\nu_n)$ is weakly increasing. Hence, the mechanism satisfies individual rationality. For incentive compatibility note that, after changing variables as

$\tilde{\Delta} = \Delta(\nu)$, we have that

$$\pi_n^a(\nu_n) - \pi_n^a(\hat{\nu}_n) = \int_{\hat{\nu}_n}^{\nu_n} \bar{p}_n^a(\nu)\Delta'(\nu)d\nu = \int_{\Delta(\hat{\nu}_n)}^{\Delta(\nu_n)} \bar{p}_n^a(\Delta^{-1}(\tilde{\Delta}))d\tilde{\Delta}.$$

Since $\bar{p}_n^a$ is weakly increasing and Δ is strictly increasing, we have that $\bar{p}_n^a \circ \Delta^{-1}$ is weakly increasing. Therefore,

$$\pi_n^a(\nu_n) - \pi_n^a(\hat{\nu}_n) \geq \bar{p}_n^a(\Delta^{-1}(\Delta(\hat{\nu}_n)))[\Delta(\nu) - \Delta(\hat{\nu}_n)] = \bar{p}_n^a(\hat{\nu}_n)[\Delta(\nu_n) - \Delta(\hat{\nu}_n)]$$
$$\implies \pi_n^a(\nu_n) \geq \bar{p}_n^a(\hat{\nu}_n)\Delta(\nu_n) - \bar{x}_n^a(\hat{\nu}_n).$$

That is, the incentive compatibility constraint is satisfied. This concludes the proof of the lemma. □

Now we can prove proposition 1.

Proof. Using Lemma 1, we can rewrite the objective function in (1) as

$$\begin{aligned}
&\int\int [x^a(c,\nu_n) - p^a(c,\nu_n)c]\, d\frac{\Phi_n(\nu_n)}{1-s}dG(c) = \\
&\int\int p^a(c,\nu_n)\,[\Delta(\nu_n) - c]\, d\frac{\Phi_n(\nu_n)}{1-s}dG(c) - \int \pi_n^a(\nu_n)d\frac{\Phi_n(\nu_n)}{1-s} = \\
&\int\int p^a(c,\nu_n)\,[\Delta(\nu_n) - c]\, d\frac{\Phi_n(\nu_n)}{1-s}dG(c) - \int\int_{\underline{\nu}}^{\nu_n} \bar{p}_n^a(\nu)\Delta'(\nu)d\nu d\frac{\Phi_n(\nu_n)}{1-s} - \pi_n^a(\underline{\nu}).
\end{aligned}$$

We can then apply integration by parts in the following term

$$\begin{aligned}
\int\int_{\underline{\nu}}^{\nu_n} \bar{p}_n^a(\nu)\Delta'(\nu)d\nu d\frac{\Phi_n(\nu_n)}{1-s} &= \int_{\underline{\nu}}^{\nu_n} \bar{p}_n^a(\nu)\Delta'(\nu)d\nu \frac{\Phi_n(\nu_n)}{1-s}\bigg|_{\underline{\nu}}^{\bar{\nu}} \\
&- \int \bar{p}_n^a(\nu_n)\Delta'(\nu_n)\frac{\Phi_n(\nu_n)}{1-s}d\nu_n = \int \bar{p}_n^a(\nu_n)\Delta'(\nu_n)\left[1 - \frac{\Phi_n(\nu_n)}{1-s}\right]d\nu_n.
\end{aligned}$$

Combining the above equations, we have that the objective function is

$$\int\int p^a(c,\nu_n)\left[\Delta(\nu_n)-\frac{\Delta'(\nu_n)}{haz(\nu_n)}-c\right]d\frac{\Phi_n(\nu_n)}{1-s}dG(c)-\pi_n^a(\underline{\nu}),$$

For the last, by Lemma 1, for the mechanism to satisfy the IC constraint we must have that $\bar{p}_n^a(\nu_n)$ is weakly increasing in ν_n. This concludes the proof of the proposition. □

A.2 Proposition 2

Proof. Using proposition 1, we can write the objective function as

$$\int\int p^a(c,\nu_n)\left[\Delta(\nu_n)-\frac{\Delta'(\nu_n)}{haz(\nu_n)}-c\right]d\frac{\Phi_n(\nu_n)}{1-s}dG(c)=$$
$$\int\int p^a(c,\nu_n)\left[c_n(\nu_n)-c\right]d\frac{\Phi_n(\nu_n)}{1-s}dG(c)+\int\bar{p}_n^a(\nu_n)\left[h\left(\frac{\Phi_n(\nu_n)}{1-s}\right)-g\left(\frac{\Phi_n(\nu_n)}{1-s}\right)\right]d\frac{\Phi_n(\nu_n)}{1-s}$$

Let us consider the last term of the above equation. Using integration by parts, we have

$$\int\bar{p}_n^a(\nu_n)\left[h\left(\frac{\Phi_n(\nu_n)}{1-s}\right)-g\left(\frac{\Phi_n(\nu_n)}{1-s}\right)\right]d\frac{\Phi_n(\nu_n)}{1-s}=$$
$$\left[H\left(\frac{\Phi_n(\nu_n)}{1-s}\right)-G\left(\frac{\Phi_n(\nu_n)}{1-s}\right)\right]\bar{p}_n^a(\nu_n)\Big|_{\underline{\nu}}^{\bar{\nu}}-\int\left[H\left(\frac{\Phi_n(\nu_n)}{1-s}\right)-G\left(\frac{\Phi_n(\nu_n)}{1-s}\right)\right]d\bar{p}_n^a(\nu_n)$$

Since G is the convex-hull of H, they coincide at the boundary points. Hence, the first term of the final expression is equal to 0. Therefore, the objective function equals

$$\int\int p^a(c,\nu_n)\left[c_n(\nu_n)-c\right]d\frac{\Phi_n(\nu_n)}{1-s}dG(c)-\int\left[H\left(\frac{\Phi_n(\nu_n)}{1-s}\right)-G\left(\frac{\Phi_n(\nu_n)}{1-s}\right)\right]d\bar{p}_n^a(\nu_n).$$

It is easy to see that our proposed mechanism maximizes the first term since, by construction, $p^a(c,\nu_n)=1$ whenever $c_n(\nu_n)\geq c$. Also, the proposed mechanism maximizes the second term. To see this, note that the second term is nonpositive for any p^a that satisfies the constraint that $\bar{p}_n^a$ is weakly increasing. Moreover, for our proposed mechanism, this term is exactly zero because whenever $H\left(\frac{\Phi_n(\nu_n)}{1-s}\right)>G\left(\frac{\Phi_n(\nu_n)}{1-s}\right)$ the derivative

$g(q) = \frac{dG(q)}{dq}$ is constant and, as a result, $d\bar{p}_n^a(\nu_n)$ is zero. Thus, we can conclude that the proposed mechanism achieves the maximum in problem (6). $\square$

A.3 Proposition 5

Proof. Given two direct mechanisms $m = (m^a, m^b) = (p^a, x^a, p^b, x^b)$, define the operator $(\Phi, \Delta) \xmapsto{T_D(m)} \hat{\Delta}$ as

$$\hat{\Delta}(\nu) = \frac{\nu + (\lambda_a + \lambda_b)\Delta(\nu) + \lambda_b[(1-s)\pi_o^b(\nu) - s\pi_n^b(\nu)] - \lambda_a \pi_n^a(\nu)}{\lambda_a + \lambda_b + \mu + r}. \tag{29}$$

In the same fashion, given two direct mechanisms $m = (m^a, m^b) = (p^a, x^a, p^b, x^b)$, define the operator $(\Phi, \Delta) \xmapsto{T_P(m)} \hat{\Phi}$ as

$$\hat{\Phi}(\nu) = \int_{\underline{\nu}}^{\nu} \hat{\phi}(\nu) d\nu, \;\; \text{where} \;\; \hat{\phi}(\nu) = \frac{\lambda_a \bar{p}_n^a(\nu) + \lambda_b s \bar{p}_n^b(\nu)}{\lambda_a[1 + \bar{p}_n^a(\nu)] + \lambda_b[s\bar{p}_n^b(\nu) + (1-s)\bar{p}_o^b(\nu)]} f(\nu). \tag{30}$$

Our plan for this proof is to find (Φ, Δ, m) such that $m = (m^a, m^b)$ solves the primary and secondary market problems for (Φ, Δ), $(\Phi, \Delta) \xmapsto{T_D(m)} \Delta$ and $(\Phi, \Delta) \xmapsto{T_P(m)} \Phi$. Then we use these three functions to backup the other equilibrium objects.

Let $a = \frac{\lambda_a + \lambda_b}{\lambda_a + \lambda_b + \mu + r}$, $b = \frac{1}{\lambda_a + \lambda_b + \mu + r}$, $\kappa = \frac{1}{1-a}$ and $\epsilon > 0$ a small constant. Let $\mathscr{P}$ and $\mathscr{D}$ be spaces of differentiable functions defined as

$$\mathscr{P} = \left\{ \Phi \in \mathscr{C}^1[\underline{\nu}, \bar{\nu}] \;\middle|\; \forall\, \nu: \; 0 \leq \Phi(\nu) \leq F(\nu) \text{ and } 0 \leq \phi(\nu) \leq f(\nu) + \epsilon \right\} \quad \text{and}$$
$$\mathscr{D} = \left\{ \Delta \in \mathscr{C}^1[\underline{\nu}, \bar{\nu}] \;\middle|\; \forall\, \nu: \; 0 \leq \Delta(\nu) \leq [\kappa b + \epsilon]\, \bar{\nu} \text{ and } b - \epsilon \leq \Delta'(\nu) \leq \kappa b + \epsilon \right\}$$

where $\phi(\nu) = d\Phi(\nu)/d\nu$. Construct the two sequences $\{\Delta^k, \Phi^k, H^k\}_k$ and $\{\hat{\Delta}^k, \hat{\Phi}^k\}_k$ in the following way. Start with any $(\Delta^0, \Phi^0) = (\hat{\Delta}^0, \hat{\Phi}^0) \in \mathscr{D} \times \mathscr{P}$ and let H^0 be any function. Then construct the sequence recursively in the following way. For given $\{\Delta^k, \Phi^k, H^k\}_k$, let $m^k = (m^{ak}, m^{bk})$ be the solution described in propositions (2) and (4), and H^{k+1} be the H associated with those mechanisms, where $\Delta = \Delta^k$, $\Phi_o = \Phi^k$, $\Phi_n = F - \Phi_o$ and $s = \Phi_o(\bar{v})$. Let $\hat{\Delta}^{k+1} = T_D(m^k) \cdot (\Delta^k, \Phi^k)$ and $\hat{\Phi}^{k+1} = T_P(m^k) \cdot (\Delta^k, \Phi^k)$. Let Δ^{k+1} be a solution to the problem

$$\min_{p^{k+1} \in P^{k+1} \cap \mathscr{D}} \sup_{v} |\hat{\Delta}^{k+1}(v) - p^{k+1}(v)|,$$

where P^{k+1} is the space of non-negative polynomials of degree $k+1$. Similarly, let the measure Φ^{k+1} be a solution to the problem

$$\min_{p^k \in P^k \cap \mathscr{P}} \sup_{v} |\hat{\Phi}^1(v) - p^k(v)|.$$

Note that $\{\Delta^k, \Phi^k, H^k\}_k$ is differentiable with uniformly bounded derivatives. Therefore, the sequence is equicontinuous and, by the Arzelà-Ascoli theorem, it has a convergence sub-sequence. Passing to a sub-sequence if necessary, let $\{\Delta^*, \Phi^*, H^*\}$ be the limit of $\{\Delta^k, \Phi^k, H^k\}_k$. We want to show that $\hat{\Delta}^k \to \Delta^*$ and $\hat{\Phi}^k \to \Phi^*$. To show that $\hat{\Delta}^k \to \Delta^*$, it suffices to show that, for k high enough,

$$\min_{p^k \in P^k \cap \mathscr{D}} \sup_{v} |\hat{\Delta}^k(v) - p^k(v)| = \min_{p^k \in P^k} \sup_{v} |\hat{\Delta}^k(v) - p^k(v)|.$$

That is, the constraint $p^k \in \mathscr{D}$ does not bind in the limit. Note that

$$\begin{aligned}\hat{\Delta}^{k+1}(v) = & \frac{v + (\lambda_a + \lambda_b)\Delta^k(v)}{\lambda_a + \lambda_b + \mu + r} \\ & + \frac{\lambda_b[(1-s)\int_v^{\bar{v}} \bar{p}_o^b(v)\Delta'^k(v)dv - s\int_{\underline{v}}^{v} \bar{p}_n^b(v)\Delta'^k(v)dv] - \lambda_a \int_{\underline{v}}^{v} \bar{p}_n^a(v)\Delta'^k(v)dv}{\lambda_a + \lambda_b + \mu + r}.\end{aligned}$$

Because $\bar{p}_o^b(v)$, $\bar{p}_n^b(v)$, and $\bar{p}_n^a(v)$ and monotone, they are continuous almost everywhere.

Hence, $\hat{\Delta}^{k+1}(\nu)$ is differentiable almost everywhere with

$$\hat{\Delta}'^{k+1}(\nu) = \frac{1 + (\lambda_a + \lambda_b)\Delta'^k(\nu) - \lambda_b[(1-s)\bar{p}_o^b(\nu) + s\bar{p}_n^b(\nu)]\Delta'^k(\nu) - \lambda_a \bar{p}_n^a(\nu)\Delta'^k(\nu)}{\lambda_a + \lambda_b + \mu + r}.$$

From the above formula, we can see that $\hat{\Delta}'^{k+1}(\nu) \geq \frac{1}{\lambda_a+\lambda_b+\mu+r} = b > b - \epsilon$. Moreover,

$$\begin{aligned}\hat{\Delta}'^{k+1}(\nu) &\leq \frac{1 + (\lambda_a + \lambda_b)\Delta'^k(\nu)}{\lambda_a + \lambda_b + \mu + r} = b + a\Delta'^k(\nu) \\ &\leq b + a\left(\frac{b}{1-a} + \epsilon\right) = \frac{b}{1-a} + a\epsilon < \frac{b}{1-a} + \epsilon.\end{aligned}$$

This shows that the constraint $b - \epsilon \leq \Delta'(\nu) \leq \kappa b + \epsilon$ is not binding. For the constraint $\Delta(\nu) \leq [\kappa b + \epsilon]\,\bar{\nu}$, note that

$$\begin{aligned}\hat{\Delta}^{k+1}(\bar{\nu}) &= \frac{\bar{\nu} + (\lambda_a + \lambda_b)\Delta^k(\bar{\nu})}{\lambda_a + \lambda_b + \mu + r} \\ &\quad + \frac{\lambda_b[(1-s)\int_{\bar{\nu}}^{\bar{\nu}} \bar{p}_o^b(\nu)\Delta'^k(\nu)d\nu - s\int_{\underline{\nu}}^{\bar{\nu}} \bar{p}_n^b(\nu)\Delta'^k(\nu)d\nu] - \lambda_a \int_{\underline{\nu}}^{\bar{\nu}} \bar{p}_n^a(\nu)\Delta'^k(\nu)d\nu}{\lambda_a + \lambda_b + \mu + r} \\ &\leq \frac{\bar{\nu} + (\lambda_a + \lambda_b)\Delta^k(\bar{\nu})}{\lambda_a + \lambda_b + \mu + r} = [b + a\Delta^k(\bar{\nu})]\bar{\nu} \leq \left[b + a\left(\frac{b}{1-a} + \epsilon\right)\right]\bar{\nu} < \left[\frac{b}{1-a} + \epsilon\right]\bar{\nu}.\end{aligned}$$

Again the constraint is not binding. Thus, since by the Stone-Weierstrass theorem the space of polynomials is dense in $\mathscr{C}^0$, if k is high enough

$$\min_{p^k \in P^k \cap \mathscr{D}} \sup_{\nu} |\hat{\Delta}^k(\nu) - p^k(\nu)| = \min_{p^k \in P^k} \sup_{\nu} |\hat{\Delta}^k(\nu) - p^k(\nu)|$$

and we can conclude that $\hat{\Delta}^k \to \Delta^*$. The argument to show that $\hat{\Phi}^k \to \Phi^*$ is analogous. The limit H^* also define a mechanism as

$$G^*(q) = \min\{\omega H^*(r_1) + (1-\omega)H^*(r_2);\ \omega r_1 + (1-\omega)r_2 = q\}, \quad \text{and} \quad g(q) = \frac{dG(q)}{dq^+}.$$

Note that H^* may not be continuously differentiable. Now we define the direct mechanisms $m^* = (m^{*a}, m^{*b})$ in propositions (2) and (4). By continuity of the object functions, the mechanisms $m^* = (m^{*a}, m^{*b})$ solve the primary and secondary market problems for (Φ^*, Δ^*). Moreover, $(\Phi^*, \Delta^*) \xmapsto{T_D(m^*)} \Delta^*$ and $(\Phi^*, \Delta^*) \xmapsto{T_P(m^*)} \Phi^*$. From Δ^*, Φ^* and m^* is standard to construct the equilibrium $\{\Delta^*, V_c^*, \Phi_o^*, \Phi_n^*, s^*, m^*\}$ with $\Phi_o^* = \Phi^*$, $\Phi_n^* = F - \Phi^*$, $s^* = \Phi^*(\bar{v})$ and

$$V_c = \frac{1}{r}\lambda_a(1-s)\int \pi_c^{*a}(c)dG(c).$$

Finally, it is easy to see that $\{\Delta^*, V_c^*, \Phi_o^*, \Phi_n^*, s^*, m^*\}$ must satisfy the equilibrium conditions since (29) and (30) are just rewriting of the equilibrium equations (4) and (18). $\square$

B Appendix: Municipal Bond Data

MSRB Data and Sample Description. The Municipal Securities Rulemaking Board (MSRB) requires securities dealers, issuers, and those acting on their behalf to submit information on municipal bond trades and disclosure documents for all transactions of municipal bonds within 15 minutes of the time of trade. The data are publicly available through the MSRB's Electronic Municipal Market Access (EMMA) portal, however we obtain historical data through Wharton Research Data Services. Our dataset includes transaction-level information for all trades of municipal bonds involving a securities dealer in which the municipal bond was assigned a unique CUSIP identification number. This nearly covers the universe of municipal bond trades.[12] Our sample includes transactions from January 3, 2005, through December 31, 2014.

An observation in our dataset is a unique transaction of a municipal bond either between a securities dealer and a customer (defined as any person or institution other than a securities dealer) or between two dealers. We are able to observe characteristics

[12] For example, trades in 529 college savings plans that include municipal bonds, municipal fund securities, or municipal derivatives are not reported on EMMA. These make up only a fraction of all municipal bond transactions.

of the transaction, including the date and time of trade, trade price, par-value traded, settlement date, and if the transaction was reported as a primary market sale executed on the first day of trading a new issue. For some transactions the yield-to-worst is reported. Additionally, we are able to observe characteristics about the bond being traded, including its unique CUSIP identifier, a text description of the bond, the date interest began to accrue, the maturity date, and the coupon rate (if applicable).[13] In addition to these variables reported through the MSRB, we merge bond information from CUSIP Global Services' Master File on the geography of the issuer, type of issuer (county, school district, development authority, financial authority, housing authority, sewer authority, or redevelopment authority), tax-exempt status, callable status, if the bond is refunding existing debt, and the type of municipal bond (tax revenue bond or general obligation, etc.).

In our main sample, we drop transactions that are missing a par value or price. We also drop all variable rate securities since we do not see any information about the current interest rate at the time of the transaction. In our data, 70% of all transactions are fixed-rate or discount bonds (0-coupon bonds). We also drop transactions whose par value is less than $1,000. In this sample selection, we are left with 86.4 million observations on 1.89 million unique bond issues where we identify a bond issue by its unique CUSIP identifier.[14]

A key variable in our analysis is the yield on the transaction. Seventy percent of transactions in the data report the yield-to-worst, defined as the lowest of the yield calculated to the call option, par option, or maturity. The relevant yield in terms of our model is the yield-to-expected-redemption, or a weighted average of the yield-to-call, yield-to-par, and yield-to-maturity, weighting by the respective probabilities of each event occurring. Since we do not observe this yield in the data, our procedure for calculating the yield is as follows. If the yield-to-worst is reported, we take that to

[13]Some municipal bonds do not pay regular interest payments.

[14]New issuance of municipal debt is typically done through a series of bonds with differing maturity dates and coupon rates. Each bond within the series is given a unique CUSIP identification number which serves as our definition of 'issue'.

be the yield generated in the model. When no yield is reported, we calculate a yield-to-maturity. For discount bonds, the yield-to-maturity, i, is given by the formula $price = par/(1+i)^T$, where T is the time until maturity. For other fixed-rate bonds, we use the console formula $i = c/(price - c)$, where c is the coupon rate.

Finally, we winsorize the price and yield in the estimation at the 99% level. The descriptive statistics of the sample are reported in Table 1.