Intermediaries as Information Aggregators: An Application to U.S. Treasury Auctions*

Nina Boyarchenko, David O. Lucca, and Laura Veldkamp

Federal Reserve Bank of New York and
New York University, Stern School of Business

April 20, 2015

Abstract

According to most theories of financial intermediation, intermediaries diversify risk, transform maturity or liquidity, and screen or monitor borrowers. In U.S. Treasury auctions, none of these rationales apply. Intermediaries submit their customer bids without transforming liquidity or maturity, and they do not screen and monitor borrowers or diversify fiscal policy risk. Yet, most end investors place their Treasury auction bids through an intermediary rather than submit them directly. Motivated by this evidence, we explore a new information aggregation model of intermediation. Intermediaries observe each client's order flow, aggregate that information across clients, and use it to advise their clients as a group. In contrast to underwriting theories in which intermediaries, by acting as gatekeepers, extract rents but reduce revenue variance, information aggregators increase expected auction revenue, but also make the revenue more sensitive to changes in asset value. We use the model to examine current policy questions, such as the optimal number of intermediaries, the effect of non-intermediated bids, and minimum bidding requirements.

*The views expressed here are the authors' and are not representative of the views of the Federal Reserve Bank of New York or of the Federal Reserve System. We thank Ken Garbade, Luis Gonzalez, Richard Tang, Haoxiang Zhu and participants at the 2014 U.S. Treasury Roundtable on Treasury Markets for comments. Nic Kozeniauskas and Karen Shen provided excellent research assistance. Emails: nina.boyarchenko@ny.frb.org; david.lucca@ny.frb.org; lveldkam@stern.nyu.edu.

Investors often access financial markets through intermediaries. Sometimes these intermediaries have exclusive access to a trading venue. Other times, they lower risk by screening investments or monitoring borrowers' behavior. In U.S. Treasury auctions, which are the world's largest, intermediaries channel investors' bids, without diversifying or transforming risks. Further, in contrast to public offerings of private issuers (e.g., Beatty and Ritter, 1986; Hansen and Torregrosa, 1992), intermediaries in Treasury auctions cannot screen or monitor the issuer because they do not influence fiscal policy. Despite these limitations, and even though investors can bid directly through an electronic system, most bids are still placed through intermediaries. The prevalence of intermediation in a market where none of the typical rationales apply prompts us to examine a new role for intermediaries and its consequences for asset prices and auction revenue.

We present a new theory of financial intermediaries who collect information from order flow, use it to advise clients, and bid for their in-house account.[1] Existing work in the initial public offering (IPO) literature studies the effects of concentrated underwriting, which typically involves a single lead, or a handful of co-lead, underwriters. It finds that this structure lowers issuers' revenues but also revenue variance.[2] In Treasury auctions, there are many information intermediaries, investors have the option of bidding directly without an intermediary, and intermediaries are subject to minimum bidding requirements. Finally, intermediaries place very large bids for their own accounts. We show that in this setting, the conventional wisdom of underwriting is reversed: information intermediaries raise expected revenue but also revenue variance. By sharing valuable information with their clients, dealers lower clients' risk, which encourages them to bid more aggressively and boost expected auction revenue. At the same time, more precise information about the asset value makes beliefs and bids more sensitive to changes in that value. Therefore, auction revenue is also more sensitive to information about the future value and as a result, more variable. Thus, information aggregation intermediaries provide value both for investors and for the asset issuer, but their effects on auction revenue are exactly the opposite from those of a traditional security underwriter.

[1] Sovereign auction rules regarding the use of client information vary across countries. In the UK, the Debt Management Office explicitly sanctions that Gilt-edged Market Makers, which have exclusive access to the auction and route orders for all other bidders, "whilst not permitted to charge a fee for this service, may use the information content of that bid to its own benefit" (GEMM Guidebook, 2011). We are not aware of similar rules in the context of U.S. Treasury auctions. In the U.S., a financial intermediary's use of client information, including sharing such information with other clients or using the information for other benefit to such intermediary, may violate legal requirements, be they statutory, regulatory or contractual, and/or violate market best practices or standards. This paper does not take a view as to whether the described use of client information with respect to Treasury auction activity is legal or proper. The objective of the paper is to study the economic effects of order flow dissemination ahead of the auction as a mechanism to lower auction risk and raise revenues.

[2] In a "full commitment IPO," the underwriter generally earns a large first-day secondary market return, and stabilizes the market value by raising supply elasticity, either offering additional ("greenshoe option") or buying some of the securities being offered (Ritter and Welch, 2002).

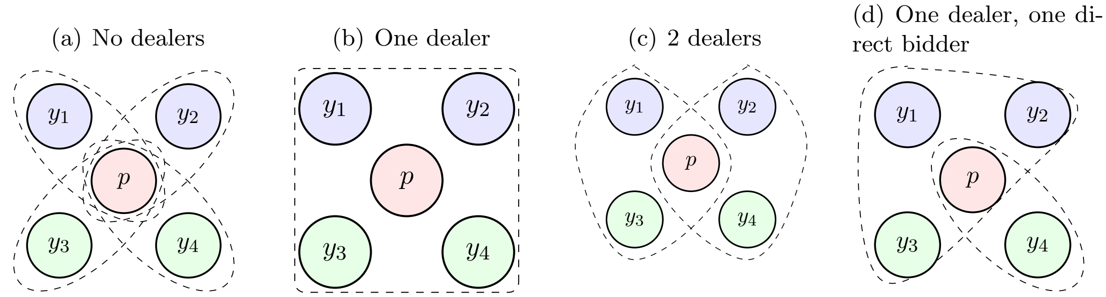
The mixed nature of Treasury auctions also introduces negative skewness into auction revenues. In a mixed auction, investors can choose to bid directly through an electronic system, or indirectly through the intermediary dealer. When an investor bids directly, no one observes their order flow and their information remains private. When they bid through an intermediary, they share their signal realization with the intermediary through their bidding behavior but also learn information about the overall order flow observed by the intermediary. An investor whose signal indicates a high future value for the security expects to take a large position, which will make his utility more sensitive to the auction-clearing price. Sharing his good news with others will increase the clearing price and negatively affect his expected utility. Thus, an investor with good news prefers not to share his information and bids directly. Conversely, when the news is bad, the investor expects to take a small position in the auction making his utility not as sensitive to the clearing price. With a low signal, the investor is less concerned about sharing his signal but also benefits from learning new information from other investors. Thus, low-signal investors are more likely to bid indirectly through the dealer. But when negative signals are shared, they affect bids of many investors and their price impact is amplified. Thus the mixed nature of the auction results in a new financial accelerator channel, which is characterized by the asymmetric diffusion of information. We use the model to run policy counterfactuals and determine whether minimum bidding requirements mitigate or amplify this source of asymmetry and fragility.

In many markets, intermediaries such as market makers or stock brokers collect information from order flow and use it to advise clients. We consider Treasury auctions in particular because of their importance (being central to funding the US federal debt) and because standard roles for intermediaries (as gatekeepers or monitors) do not apply. These auctions are, however, also complex and unique in their structure. In our modeling, we attempt to balance a detailed description with a tractable and transparent model which highlights insights that are broadly applicable. The basis for the model is a standard, common-value, uniform-price auction with heterogeneous information, and limit and market orders. We consider a large, finite number of competitive investors. On top of that foundation, we add five features that distinguish Treasury auctions from other settings.

Feature 1: Information Aggregation In Treasury auctions, primary dealers are the intermediaries that serve as information aggregators.[3] Primary dealers bid on behalf of

[3] As discussed on the New York Fed's website, primary dealers are "[... obligated] to: (i) participate consistently in open market operations to carry out U.S. monetary policy [...], and [to provide] the New York Fed's trading desk with market information and analysis helpful in the formulation and implementation of monetary policy. Primary dealers are also required to participate in all auctions of U.S. government debt and to make reasonable markets for the New York Fed when it transacts on behalf of its foreign official account-holders."

Figure 1: Information sets



Note: y_1 and y_2 are signals observed by small investors; y_3 and y_4 are signals observed by large investors; p is the equilibrium price. Dashed lines indicate common information sets.

clients and for their own account. They collect information from order flow and use it to advise clients. In effect, dealers are pooling the information sets of the clients they advise. Hortaçsu and Kastl (2012) find that this information pooling effect accounts for a large part of dealer surplus in Canadian Treasury Auctions. Dealers in US auctions acknowledge its central role:

> "[Observing overall patterns of buying in the Treasury market] can be one of the greatest benefits of being a primary dealer, since the service itself often doesn't pull in big profits directly." (Reuters, 2011)

Figure 1 illustrates how we model this information pooling. Without any dealers, investors only bid directly, keeping their private signals private. Each investor conditions their quantity demanded on the realized market-clearing price, which is a noisy signal about all agents' information. Thus, optimal bids use this price information (represented by p in the figure), making the information set of investor i effectively $\{y_i, p\}$. With one dealer and no direct bidding (panel b), the dealer observes all trades and shares all private signals with all agents. When there are multiple intermediaries, they pool the information of the subset of clients they advise. The more dealers, the less information pooling (panel c). Investors who bid independently from the intermediary keep their signal private (panel d).

Feature 2: Strategic Bidding The small number of primary dealers makes the U.S. Treasury primary market an imperfectly competitive one (Bikhchandani and Huang, 1993). The strategic aspect of primary dealers' bids is a central feature of our model. The model also includes other large strategic bidders that are non primary dealers. Varying the number of primary dealers in the model does not directly change the competitive structure in the auction, but it rather transforms a non-dealer strategic investor into an intermediary.

Feature 3: Non-Competitive Bidding A non-competitive bid is a purchase order for a given value of Treasuries at whatever the auction-clearing price turns out to be. In practice these bids are placed by retail investors through Treasury Direct as well as by foreign and international monetary authorities (FIMA) that hold securities in custody at the Federal Reserve Bank of New York. We model the demand of non-competitive bidders as an exogenous shock. These bids function as noise that prevents the market-clearing price from perfectly aggregating all private information. We estimate that bids from noise traders account for about 10% of the total.[4]

The first version of the model (Section 1) incorporates these three features and examines how the number of dealers affects the trade-off between the market power that an intermediary enjoys and the information they provide to others.

Feature 4: Direct and Indirect Bidding Treasury auctions are mixed auctions, meaning that investors have the option to bid indirectly, through a dealer, or directly, without any intermediary. While direct bidding has been historically allowed since 1992, electronic bidding systems and the elimination of deposit requirements for all bidders have facilitated direct bids. Direct bidding has grown from 2 percent of all bids in 2003 to 10 percent in 2014. While auction results do not disclose the number of direct bidders, public remarks of Treasury officials suggest there were about 1200 direct bidders in 2001, and 825 in 2004.

Section 4 examines a large, strategic bidder's choice between a direct or an indirect bid. The results uncover two valuable lessons. The first is about the optimal number of dealers. With few dealers who each aggregate the information of many clients, dealers provide a lot of information, making the benefit of indirect bidding high. However, indirect bidding is also costly because the investor signal is pooled with signals of other clients, lowering its value. Our results show that the latter effect dominates. As a result, reducing the number of dealers to improve information aggregation could be costly as dealer-intermediated bids could decline.

Second, we show that intermediation can amplify negative shocks to asset values with direct bidding. This new financial accelerator channel of intermediation arises because bad news are shared with dealers and thus with other investors, but good news are kept private. As a result, the distribution of auction revenues is negatively skewed.

Feature 5: Minimum bidding requirements Prior to 1992, being an active counterparty meant being a "consistent and meaningful participant" in Treasury auctions by submitting bids roughly commensurate with the dealer's capacity. A 2010 policy change

[4]Non-competitive bids for all FIMA accounts are capped at $1 billion, so many FIMA bids are placed competitively although official data do not exist. Because non-competitive bidders are in our model noise traders we group our estimate of FIMA competitive bids with other non-competitive bids.

strengthened these requirements. Today, primary dealers are expected to bid at all auctions an amount equal to the pro-rata share of the offered amount, with bids that are "reasonable" compared to the market. A dealer that consistently fails to bid for a large enough quantity at a high enough price could lose his primary dealer status.[5]

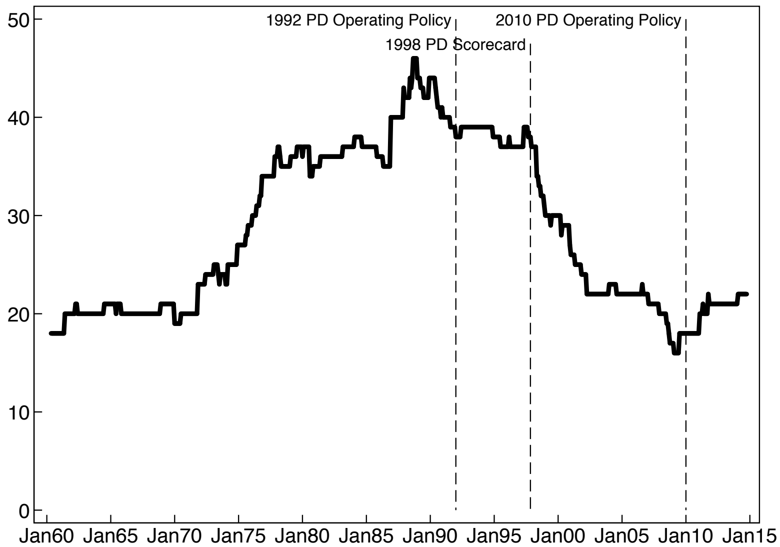
Section 5 models such a requirement as a shadow cost for low bids. The cost is based on the realized price times quantity awarded at auction. By encouraging aggressive bidding, this policy helps to counteract the monopsony power distortion of primary dealers. Less use of monopsony power both increases revenue and makes makes investors more willing to bid through the dealer, which reduces the fragility of the mixed auction. However we find that with less monopsony power and more price-taking behavior, auction revenues are not as smooth across states. Thus minimum bidding requirements raise expected revenue, but also revenue volatility.

We calibrate the full version of the model, using Treasury auction data, in order to do quantitative policy analysis. The centerpiece of this analysis is an active policy question: How does the number of dealers affects auction outcomes? This is an important question because entry into the primary dealer system is regulated. Policy makers have been debating the merits of a robust primary dealer system for decades (see the Brady, Breeden, and Greenspan (1992) report). In 1960, there were 18 primary dealers (Figure 2). Amid the rapid rise in federal debt and the sharp increase in interest rate volatility, the number of primary dealers rose throughout the 1970s to 46 in the mid-1980s. Subsequently, the population of primary dealers dwindled, reaching its nadir in 2009. The number of primary dealers has since increased to about 22 today. Our quantitative model reveals that increasing the number of dealers from 5 to 50 raises expected excess revenue by 15 bps for the baseline calibration and lowers volatility by 4 bps. Thus, for our baseline calibration, the revenue benefit of less dealer monopsony power outweighs the cost of diminished information aggregation.

A common tension at the core of each model is the trade-off between the dealer's ability to use market power and their ability to aggregate information in a way that lowers risk and raises revenue. This trade-off raises the question of whether one could improve information without the adverse effects of market power. Our model embodies the idea that intermediation and information aggregation are inextricably interlinked: it is the process of intermediating trades that reveals information to dealers and that empowers them. This is a tension that optimal auction design must confront.

[5]See Appendix E in Brady et al. (1992) for pre-1992 policies. In 1997, the New York Fed instituted an explicit counterparty performance scorecard and dealers were evaluated based on the volume of allotted securities. In 2010 the NY Fed clarified their primary dealer operating policies. See New York Fed website for the most recent rules.

Figure 2: Number of primary dealers



Note: Data are of the last of the month. Source: Federal Reserve Bank of New York

Contribution to the existing literature. In closely related work, Hortaçsu and Kastl (2012) also model dealers who observe order flow. Using a share auction model with discrete bids, they estimate to what extent client order flow is informative about demand versus asset value. Consistent with their results, we assume that both demand and asset fundamentals affect bids. We make no attempt to distinguish between these channels. Instead, we use a more tractable menu auction, which allows us to vary the number of dealers, understand how indirect/direct bidding choices contribute to fragility, and inform the debate on minimum bidding requirements.

The idea that multiple dealers fragment a market is similar to Babus and Parlatore (2015), where investor valuations are heterogeneous and fragmentation inhibits risk-sharing. Ours is a common-value auction where investor information is heterogeneous and fragmentation inhibits information-sharing. In addition, our model is tailored to Treasury auctions. It features competitive and non-competitive bidders, an option to bid directly, minimum bidding requirements and information that is partially revealed through the realized market-clearing price.

The assumptions of our model are informed by many empirical findings. First, we assume that bidders have private signals about future Treasury values. Indeed, bids contain information about the future value of the Treasury (e.g Coutinho, 2013).[6] Second, we assume that bids condition on realized prices, which is the rational response to the winner's curse.

[6] Economic conditions affect bidding behavior in non-U.S. markets as well. Using Finnish Treasury data, Keloharju et al. (2005) find strong uncertainty effects.

Nyborg et al. (2002) find that Swedish Treasury bidding behavior is consistent with auction models featuring private information and a winner's curse. Finally, we assume that dealers behave strategically and reap benefits from their privileged position. Lou et al. (2013) argue that "auction concessions"—the fact that Treasury prices decrease into an auction and recover afterwards— is evidence of dealer rents.

Finally, much of the literature studies how the format of sovereign auctions affects revenues. Theoretical work by Chari and Weber (1992), Bikhchandani and Huang (1989), Back and Zender (1993), and Wilson (1979) considers the merits of uniform-price auctions versus other possible alternatives. Empirical work by Nyborg and Sundaresan (1996), Malvey et al. (1995) and Malvey and Archibald (1998) compares revenues from 1992-1998 when the U.S. Treasury used both uniform and discriminatory price auctions. Armantier and Sbaï (2006) and Hortaçsu and McAdams (2010) use French and Turkish Treasury auction bids to structurally estimate the benefits of uniform price auctions. This literature complements our project, which fixes the auction format to a uniform-price menu auction and focuses on the effect of intermediation.

1 Baseline Auction Model with Primary Dealers

The model economy lasts for one period and agents can invest in a risky asset (the newly issued Treasury security) and a risk-less storage technology with zero net return. The risky asset is auctioned by Treasury in a fixed number of shares (normalized to 1) using a uniform-price auction with a market-clearing price p. The fundamental value of the newly issued asset is unknown to the agents and normally distributed: $f \sim N(\mu, \tau_f^{-1})$.

We consider four type of bidders to match key features of Treasury auction participation: small and large limit-order bidders, intermediaries (or dealers) and non-price contingent bidders. Limit-order bidders and intermediaries place price-contingent bids, which specify for each clearing price p, a price-quantity pair. Limit-order bidders can be small (price takers) or large (strategic bidders). We refer to large and small limit-order bidders as investors. Dealers are just like large limit-order bidders but they also intermediate bids from other limit-order bidders. Dealers place bids directly in the auction while small and other large limit-order bidders bid indirectly through the dealers.[7] To understand the role of intermediation, we study auction revenues as the number of dealers varies. In these experiments we keep the total number of large (and small) limit-order bidders fixed. Denoting by N_L (N_S) the number of large (small) investors, and by N_D the total

[7] We relax this assumption in Section 4.

number of dealers, when varying N_D or N_L, the total $N_D + N_L$ remains unchanged.[8] Non-price-contingent bidders are the fourth type of agent that places bids. Differently from other investors and dealers, these bidders only place market orders, which only specify a quantity but not a price (called non competitive bids). In practice, non competitive bidders are either small retail investors or foreign central banks who roll over expiring debt or accumulate liquid assets for exchange rate management. The net quantity of market orders, x, is unknown to other investors and normally distributed $x \sim N(\bar{x}, \tau_x^{-1})$.

We index investors (small and large) and dealers with $i = 1, \ldots, N$, where $N \equiv N_L + N_S + N_D$. Each small investor has initial wealth W_i, and chooses the quantity of the asset to hold, q_i (which could be negative) at price p per share, in order to maximize his expected utility, $\mathbb{E}[-\exp(-\rho_i W_i)]$, where ρ_i denotes agent i's coefficient of absolute risk aversion.[9] The budget constraint dictates that final wealth is $W_i = (W_0 - q_i p)r + q_i f$.

Before trading, each investor and dealer gets a signal about the payoff of the asset. These signals are unbiased, normally distributed and have private noise:

$$y_i = f + \varepsilon_i,$$

where $\varepsilon_i \sim N(0, \tau_{\varepsilon,i}^{-1})$. Dealer and large investors solve the same problem as small investors but they also internalize the effect they have on market prices. They maximize their final utility with risk aversion ρ_i subject to the same constraints as well as the market clearing condition. We assume that all large investors and dealers share the same risk aversion and signal precision: $\rho_i = \rho_l$ and $\tau_{\varepsilon,i} = \tau_{\varepsilon,l}$ for $i \in \{\mathcal{N}_L, \mathcal{N}_D\}$, where with $\mathcal{N}_j$ we denote the set of agents of type j. Similarly, all small investors are symmetric: $\rho_i = \rho_s$ and $\tau_{\varepsilon,i} = \tau_{\varepsilon,s}$ for $i \in \mathcal{N}_S$.

Describing Information Sets and Updating Beliefs with Correlated Signals Investors and dealers observe three types of information. They observe their private signals y_i. In addition, primary dealers, who place bids on behalf of their customers, observe their investors' bids and advise their clients as a whole by disclosing the average order to all of their customers. Each intermediary d receives orders from an equal number of investors. He observes the orders of $N_S/N_D \equiv \nu_s$ small and $N_L/N_D \equiv \nu_l$ large investors. Since bids

[8] When we change the number of intermediaries, we want to isolate the information and competition aspects on this change. We do not want to change the total demand for the asset by changing the number or size of market participants.

[9] We use exponential (CARA) utility here for all agents to keep the problem tractable. Of course, this rules out wealth effects on portfolio choices. At the same time, we want to capture the idea that investors with larger balance sheets naturally hold larger positions of risky assets. Therefore, we assign large investors and dealers a smaller absolute risk aversion. In other words, we are capturing wealth effects with differences in risk aversion.

will turn out to be linear functions of beliefs, revealing average bids and revealing average signals is equivalent. Thus, investment through a primary dealer d allows investors to observe an average of other investors' signals $\bar{y}_d$. The final piece of information is the auction-clearing price p. Of course, the agent does not know this price at the time he bids. However, the agent conditions his bid $q(p)$ on the realized auction price p. Thus, each quantity q demanded at each price p conditions on the information that would be conveyed if p were the realized price. Since p contains information about the signals that other investors received, the investor uses a signal derived from p to form his posterior beliefs about the asset payoff. We guess (and later verify) that the market-clearing (settle) price is a linear function of the average signals $\{\bar{y}_k\}_{k=1}^{N_D}$ and of the market orders x

$$p = A + B\, N_D^{-1} \sum_{d=1}^{N_D} \bar{y}_d + Cx, \tag{1}$$

where

$$\bar{y}_d = \frac{\tau_{\varepsilon,s} \sum_{k \in \mathcal{I}_d^s} y_k + \tau_{\varepsilon,l} \left(\sum_{j \in \mathcal{I}_d^l} y_j + y_d \right)}{\nu_s \tau_{\varepsilon,s} + (1+\nu_l)\tau_{\varepsilon,l}} \equiv \omega_s \sum_{k \in \mathcal{I}_d^s} y_k + \omega_l \sum_{k \in \mathcal{I}_d^l} y_k + \omega_l y_d,$$

and A, B and C are a function of the number of dealers in the market: $A \equiv A\,(N_D)$, $B \equiv B\,(N_D)$, $C \equiv C\,(N_D)$. The average signal $\bar{y}_d$ from dealer d to the subset of large $(\mathcal{I}_d^l)$ and small $(\mathcal{I}_d^s)$ investors that place bids through the dealer is a signal-precision-weighted average of all of the dealer's customer signals as well as the dealer's own signal (y_d).

It follows that $\frac{p-A-C\bar{x}}{B}$ provides an unbiased signal about f

$$\frac{p - A - C\bar{x}}{B} \sim N\left(f, \left(\nu_s \tau_{\varepsilon,s} + (\nu_l+1)\,\tau_{\varepsilon,l}\right)^{-1} + \tau_p^{-1}\right),$$

where $\tau_p = (\frac{B}{C})^2 \tau_x$. For every agent, we use Bayes' law to update beliefs about f. But because the signals agents observe have correlated signal errors, we need to use a procedure that adjusts for this correlation. The following general information structure is one we can use to solve all of the versions of the model that follow in the paper.

Let S be the vector of all signals (including price signals) available to any agent.

$$S = \left[\begin{array}{cccccccc} y_1 & \ldots & y_N & \bar{y}_1 & \ldots & \bar{y}_{N_D} & \frac{p-A-C\bar{x}}{B} \end{array}\right]'. \tag{2}$$

While all investors and dealers condition on the p information, each agent in the economy observes a subset of all signals. For example an investor j with dealer d will observe its own signal j and the mean signal from dealer d $\bar{y}_d$. Let X_j be an operator that selects the subset of all signals observed by agent j. (See appendix A for definition of this operator.) Then any investor j assigned to intermediary k conditions his beliefs on the signal vector $X_j S = [y_j, \bar{y}_k, (p - A - C\bar{x})/B]$, and similarly for the intermediary (who conditions on the the same mean signal as well as its own realization.) Note that $X_j S$ is a vector of all the signals known to agent j at the time when he invests. Of course, some of these signals are redundant. But our filtering algorithm will put zero weight on signals that provide no additional information.

We can now solve for the beliefs of any agent j (investor or dealer) who observes the vector of signals $X_j S$. Since all the uncertain quantities are normally distributed, we have the following optimal linear projections[10]

$$\mathbb{E}\left[f|X_j S\right] = (1 - \beta' 1_m)\mu + \beta' X_j S \qquad \text{where} \tag{3}$$

$$\beta \equiv \mathbb{V}\left(X_j S\right)^{-1} \mathbf{Cov}\left(f, X_j S\right) \tag{4}$$

$$\mathbb{V}\left[f|X_j S\right] = \mathbb{V}\left(f\right) - \mathbf{Cov}\left(f, X_j S\right)' \mathbb{V}\left(X_j S\right)^{-1} \mathbf{Cov}\left(f, X_j S\right) \equiv \hat{\tau}_j^{-1}, \tag{5}$$

where m is the number of signals selected by $X_j S$, the covariance vector is $\mathbf{Cov}\left(f, X_j S\right) = 1_m \tau_f^{-1}$ and the signal variance-covariance $\mathbb{V}\left(X_j S\right)$, is worked out in the appendix.

Although this matrix information structure may seem cumbersome for the simple problem at hand, it allows us to examine various forms of the model with minimal changes to the setup. The different versions of the model we explore in the following section will change two things about the model: The set of signals S and the information sets of each agent, summarized by a set of X_j's.

Equilibrium. A Nash equilibrium is

1. A menu of price-quantity pairs bid by each small investor i that solves

$$\max_{q_i(p)} \mathbb{E}[-\exp(-\rho W_i)|X_i S]$$
$$\text{s.t. } W_i = W_{0i} + q_i(f - p).$$

[10] Note that the following formulas are just like the OLS formulas in a context where all the means, variances and covariances of variables are known. The OLS additive constant α is $(1_N - \beta)' 1_N \mu$. β is the infinite sample version of $(X'X)^{-1}X'y$. The conditional mean here is analogous to the optimal linear estimate in the OLS problem. This equivalence holds because in linear systems, both OLS and Bayesian estimators are consistent. It also has the Bayesian interpretation as a weighted average of normal priors and signals, where each is weighted by their relative precision. Here that precision is adjusted to account for correlation.

The optimal bid function is the inverse function: $p(q_i)$.

2. A menu of price-quantity pairs bid by each strategic player (dealer and large investor) that maximizes

$$\max_{q_j,p} \mathbb{E}[-\exp(-\rho_l W_j)|X_j S] \tag{6}$$

$$\text{s.t. } W_j = W_{0,D} + q_j(f-p), \tag{7}$$

$$x + \sum_{i=1}^{N_S} q_i + \sum_{j=1}^{N_D} q_j + \sum_{k=1}^{N_L} q_k = 1. \tag{8}$$

The second constraint is the market clearing condition and reflects that the dealer must choose his quantity and the price so that the market clears.

3. An auction-clearing (settle) price that equates demand and supply: $x + \sum_{i=1}^{N_S} q_i + \sum_{j=1}^{N_D} q_j + \sum_{k=1}^{N_L} q_k = 1$.

1.1 Solution: Optimal Menu Bids

Since all investors' posterior beliefs about f are normally distributed, we can use the properties of a log-normal random variable to evaluate the expectation of each agent's objective function. It then follows that the FOC of the small investors' problem is to bid the following set of price-quantity pairs:

$$q_i(p) = \frac{1}{\rho} \mathbb{V}[f|X_i S]^{-1} \left(\mathbb{E}[f|X_i S] - p \right). \tag{9}$$

This is a standard portfolio expression in an exponential-normal portfolio problem. The fact that it is an auction rather than a competitive market doesn't change choices. The novelty of the model is in modeling the large, strategic players (dealers and large investors) and in thinking carefully about how primary dealers affect the information sets that determine the conditional mean and variance of the asset payoff.

For a large investor or dealer, we substitute the equation for W_j in the dealer's problem into the objective function, evaluating the expectation and taking the log, we can simplify the strategic investor problem to be $\max_{q_j,p} \quad q_j(\mathbb{E}[f|X_j S] - p) - \frac{1}{2}\rho_l q_j^2 \mathbb{V}[f|X_j S]$ s.t. the market clearing condition (8). Taking the first order condition with respect to q_j, we obtain

$$q_j(p) = \frac{\mathbb{E}[f|X_j S] - p}{\rho_l \mathbb{V}[f|X_j S] + dp/dq_j} \equiv M_L \left(\mathbb{E}[f|X_j S] - p \right). \tag{10}$$

The term dp/dq_d measures the price impact of a strategic investor bid. As the price impact increases, the dealer's demand becomes less sensitive to his beliefs about the value of the security.

1.2 Auction Revenue without Dealers

To understand the costs and benefits of having primary dealers, we now consider the benchmark setup without any primary dealers. In this model, each type of limit order investor submits bids on their own behalf rather than through intermediaries. The lack of intermediaries in turn implies absence of any information aggregation.

As a result, the investors' information set contains only their private signal y_i and the price information $(p - A - C\bar{x})/\tilde{B}$, but not the average signal that was previously provided by the intermediary. Thus, the set of possible signals will remain the same but the vector of signals observed by each investor i is now $X_i^C S = [y_i, (p - A - C\bar{x})/\tilde{B}]$. The signal vector $X_i^C S$ gives the investor only his own private signal and the price information.

To solve the model, we conjecture a linear auction-clearing (settle) price.

$$p = A\left(0\right) + B\left(0\right)\ N_S^{-1} \sum_{i=1}^{N} y_i + F\left(0\right)\ N_L^{-1} \sum_{j=1}^{N_L} y_j + C\left(0\right) x \tag{11}$$

Then, we use the price conjecture to determine what information agents' can extract by conditioning their quantity demanded on the realized price. That unbiased signal about the true payoff f is $(p - A(0) - C(0)\bar{x})/(B(0) + F(0))$. Using (3) and (5), we find that the posterior belief is a weighted average of the prior belief μ, the private signal y_i and the price signal. Let β_s be the vector of weights used by the small investors and β_l be the vector of weights used by the large investors, so that the beliefs of a small investor are

$$\mathbb{E}\left[f|X_i S_i\right] = \left(1 - \beta_s\left(1\right) - \beta_s\left(2\right)\right)\mu + \beta_s\left(1\right) y_i + \beta_s\left(2\right) \frac{p - A - C\bar{x}}{\tilde{B}}$$

with conditional variance $\mathbb{V}\left[f|X_i S_i\right] = \tau_f^{-1} - \tau_f^{-1}\left(\beta_s(1) + \beta_s(2)\right) \equiv \hat{\tau}_s^{-1}$. The beliefs of a large investor are

$$\mathbb{E}\left[f|X_i S_i\right] = \left(1 - \beta_l\left(1\right) - \beta_l\left(2\right)\right)\mu + \beta_l\left(1\right) y_i + \beta_l\left(2\right) \frac{p - A - C\bar{x}}{\tilde{B}}.$$

with conditional variance $\mathbb{V}\left[f|X_i S_i\right] = \tau_f^{-1} - \tau_f^{-1}\left(\beta_l(1) + \beta_l(2)\right) \equiv \hat{\tau}_l^{-1}$.

We substitute these beliefs into the optimal bid formulas (9) and (10), and use those bid functions to determine the price that clears the market. This result and all future results are proven in Appendix A.

Result 1. *Without primary dealers, the auction revenue is (11), where*

$$\begin{aligned}
A(0) &= -C(0)\left(1+\left(N_S\frac{\hat{\tau}_s}{\rho}\beta_s(2)+N_L M_L(0)\beta_l(2)\right)\frac{A+C\bar{x}}{\tilde{B}}\right)\\
&\quad + C(0)\left(N_S\frac{\hat{\tau}_s}{\rho}\left(1-\beta_s(1)-\beta_s(2)\right)+N_L M_L(0)\left(1-\beta_l(1)-\beta_l(2)\right)\right)\mu\\
B(0) &= C(0)\,N_S\rho^{-1}\hat{\tau}_s\beta_s(1)\\
F(0) &= C(0)\,N_L M_L(0)\,\beta_l(1)\\
C(0) &= -\tilde{B}\left[N_S\rho^{-1}\hat{\tau}_s\left(\beta_s(2)-\tilde{B}\right)+N_L M_L\left(\beta_l(2)-\tilde{B}\right)\right]^{-1},
\end{aligned}$$

and the elasticity of the large investor's demand to information is given by

$$M_L(0)^{-1} = \rho_l\hat{\tau}_l^{-1} - \tilde{B}\left[N_S\rho^{-1}\hat{\tau}_s\left(\beta_s(2)-\tilde{B}\right)+(N_L-1)\,M_L(0)\left(\beta_l(2)-\tilde{B}\right)\right]^{-1}.$$

The result provides an implicit solutions for the price coefficients. Normally, a competitive market model has simple closed form solutions for prices. The complication here is two-fold: 1) there are large strategic agents whose demand is not linear and 2) the average of investors' signals $\bar{y}$ is not equal to the true payoff. The latter effect implies that there is both public and private signal noise. Both sources of extra complexity are easy to resolve numerically and are essential in the following model to understand how the number of dealers affects information aggregation and auction revenue.

1.3 Auction Revenue with N_D Dealers

In this case, all the dealers in the market are symmetric, and the beliefs of any investor bidding through dealer d are the same as the dealer's beliefs. Using (3) and (5), we find that this posterior belief takes the form

$$\mathbb{E}\left[f|X_d S\right] = \left(1-\beta(1)-\beta(2)\right)\mu+\beta(1)\,\bar{y}_d+\beta(2)\,\frac{p-A-C\bar{x}}{B}, \tag{12}$$

with conditional variance $\mathbb{V}\left[f|X_d S\right] = \tau_f^{-1}-\tau_f^{-1}\left(\beta(1)+\beta(2)\right) \equiv \hat{\tau}^{-1}$. As before, we substitute these beliefs into the first-order conditions for optimal bids and then equate demand and supply to determine the auction settle price.

Result 2. *With N_D primary dealers, auction revenue is given by (1) where*

$$A = C\left[-1 + N_D\left(\nu_s\rho^{-1}\hat{\tau} + (\nu_l+1)M_L\right)\left((1-\beta(1)-\beta(2))\mu - \frac{\beta(2)}{B}(A + C\bar{x})\right)\right] \tag{13}$$

$$B = N_D C\left(\nu_s\rho^{-1}\hat{\tau} + (\nu_l+1)M_L\right)\beta(1) \tag{14}$$

$$C = -B\left[N_D\left(\nu_s\rho^{-1}\hat{\tau} + (\nu_l+1)M_L\right)(\beta(2) - B)\right]^{-1}. \tag{15}$$

Increasing the number of primary dealers, who bid on behalf of their clients, disperses information across a larger number of participants. In other words, dealer competition inhibits information-aggregation. On the other hand, a larger number of dealers increases competitive pressures and limits rent extraction. This second effect is shut down because every time a dealer is added, we subtract a large non-dealer investor. Recall that we hold the number of strategic agents ($N_L + N_D$) fixed when we change the number of dealers N_D. We make this assumption to isolate the information aggregation effect.

Auction Revenue Since we normalized the supply of the Treasuries to one, price and auction revenue are the same. Our objective is to determine what the expected revenue is, what the variance of that revenue is, and how this mean and variance compare to other primary dealer arrangements. In every auction, the unconditional expected revenue will be $A + B\mu + C\bar{x}$ and unconditional revenue variance will be $B^2\mathbb{V}[\bar{y}] + C^2\tau_x^{-1}$.

2 Mapping the Model to Data

To measure the impact of primary dealers on auction revenue, we calibrate the model parameters using Treasury auction result data and secondary market prices. Before detailing the mapping of the model to the data, we provide institutional details that are useful to understand that mapping.

2.1 Institutional Detail

In 2013 alone, Treasury issued nearly $8 trillion direct obligations in the form of marketable debt as bills, notes, bonds and inflation protected securities (TIPS), in about 270 separate auctions.[11] An auction begins with an announcement one day to one week ahead of the

[11] Treasury bills are auctioned at a discount from par, do not carry a coupon and have terms of not more than one year. Bonds and notes, instead, pay interest in the form of semi-annual coupons. The maturity of notes range between 1 and 10 years, while the term of bonds is above 10 year. For TIPS, the coupon is

auction. The announcement specifies CUSIPs (an alphanumeric code uniquely identifying an asset), offering amount, issue and maturity dates.

There are two types of bids: Competitive bids specify a quantity and a rate (a discount rate for bills, a nominal yield for notes/bonds, or a real yield for inflation protected securities). Non-competitive bids specify a total amount (value) to purchase at the market-clearing rate. Each bidder can only place a single non-competitive bid with a maximum size of $5 million.[12] Bids can be direct or indirect. To place a direct bid, investors submit electronic bids to Treasury's Department of the Public Debt or the Federal Reserve Bank of New York. Indirect bids are placed by a depository institution (banks that accept demand deposits), or dealers on behalf of their clients. On the auction day, bids are received prior to the auction close. The auction clears at a uniform price, which is determined by first accepting all non-competitive bids, and then competitive bids in ascending yield or discount rate order.[13]

The type of uncertainty faced by Treasury bidders is different from the risks faced by corporate bond investors. Because sovereign secondary markets are deep and liquid, Treasury investors can hedge issuer-specific risks by shorting already-issued securities. Newly issued government securities are, however, only imperfect substituted for the outstanding ones because of differential liquidity (Lou et al. (2013), Amihud and Mendelson (1991) and Krishnamurthy (2002)). Investors' demand for specific issues is the key determinant of this liquidity, and so the key underwriting risks are issue-specific rather than issuer-specific. In our model, each investor knows their own issue-specific demand for this set of Treasuries. They use this signal to forecast aggregate future demand. Taken together, the dispersed signals are informative about the future value of the asset.

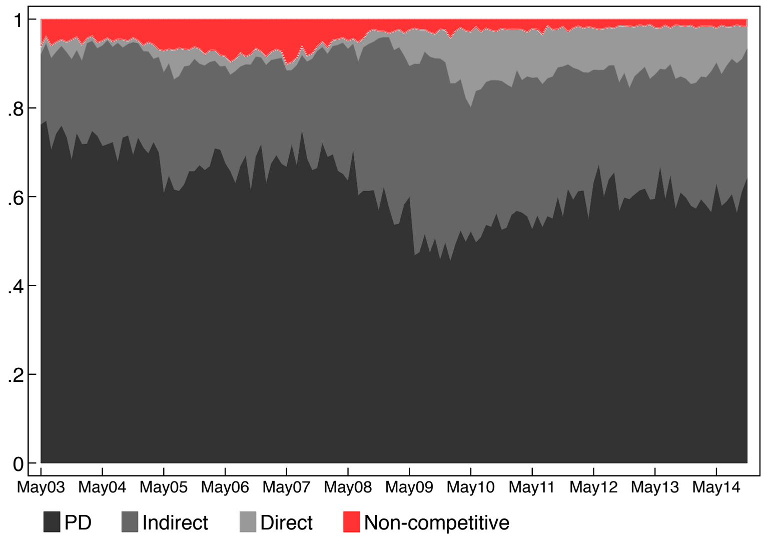
Primary dealers are the firms, commercial bank dealer departments or brokers and dealers not associated to banking organizations, with which the Federal Reserve conducts its open market operations. Figure 3 illustrates that primary dealers, bidding for their own account, are the largest bidder category at auctions (57 percent of allotted securities). Indirect bidders are the second largest at 32 percent. They may bid through depository institutions

applied to an inflation-adjusted principal, which also determines the maturity redeemable principal. TIPS maturities range between 1 and 30 years.

[12]Foreign and international monetary authorities (FIMA) that have accounts a the NY Fed can place bids up to $100 million per account and $1billion in total.

[13]The rate at the auction (or stop-out rate) is then equal to the interest rate that produces the price closest to, but not above, par when evaluated at the highest accepted discount rate or yield at which bids were accepted. The maximum auction award to a single bidder at a given rate is 35 percent of the offering, less the bidder's *reportable net long position* in the security. Competitive bidders can place multiple bids at different rates without an overall upper bound. This cap is meant to prevent a single bidder from acquiring a disproportionate amount of securities in the event of a proration, which is when, multiple investors place bids at the stop-out rate.

Figure 3: Allotted shares by bidders across all auctions



Note: Non-competitive bids in the chart exclude indirect FIMA bids. Source: Treasury auction results.

or other brokers or dealers, but primarily they bid through primary dealers.[14] The Federal Reserve Bank of New York (NY Fed) selects primary dealers. They are not unique to U.S. but are prevalent in most OECD countries.

2.2 Calibration

We calibrate model parameters using data from two main data sources: Treasury auction results and market prices. The target moments are shown in Table 1. Our sample starts in September 2004 and ends in June 2014.To study a comparable sample and estimate yield curves, we restrict attention to 2-, 3-, 5-, 7- and 10-year notes and exclude bills, bonds and TIPS.

For each maturity we compute the mean share of securities allotted to primary dealers, direct and indirect bidders.[15] As discussed above, the definition of indirect bidders from official auction results includes FIMA competitive bids placed through the NY Fed. Because these bids do not provide information to primary dealers, we attempt to abstract from them

[14]Brokers and dealers include all institutions registered according to Section 15C(a)(1) of the Securities Exchange Act. Indirect bidders also include foreign and international monetary authorities placing competitive bids through the New York Fed. These bids are not parsed out in the auction results and we attempt to estimate their size in the model calibration.

[15]We exclude amounts allotted to the Fed's own portfolio, the System Open Market Account or "SOMA," which are an add-on to the auction.

using a simple imputation.[16] We reclassify these bids as part of the noise trader group, as they do not reflect the short-term issue-specific value (more on this in the next paragraph) because foreign official investors have long investment horizons invest reflecting foreign exchange strategies.

To calibrate the auction price and fundamental value we first note that, up to rounding, the auction price clears at par. The stop-out coupon rate is, instead, a function of issue-specific value as well as the term structure of interest rates at the time of the auction, which depends on factors unrelated to the auction, such as monetary policy and inflation expectations. We focus on issue-specific fundamentals, or the "on-the-run" value of the issue, for two reasons. First, an investor can easily hedge interest rate risk into the auction by shorting a portfolio of currently outstanding securities. Second, from the issuer perspective, the stop-out rate could be very low because of low interest rates, but an issue could still be "expensive" relative to the rate environment due to auction features, which is what we are after. To strip out the aggregate interest-rate effects, we assume that the bidder enters the auction with an interest-rate-neutral portfolio, which holds one unit of the auctioned security and shorts a replicating portfolio of bonds trading in the secondary market. This portfolio is equivalent to the excess revenue on the current issue, relative to outstanding securities. Thus, price p in our model corresponds to *the auction price, minus the present value of the security's cash-flows*, where future cash flows are discounted using a yield curve.[17] The fundamental value f in the model corresponds to the *value of the interest-rate neutral portfolio on the date when the security is delivered to the winning bidders* (close of issue date). The issue date in our sample lags the auction date by an average of 5.5 days with a standard deviation of about 2.3 days. For example, in Table 1, the average revenue from selling a new coupon-bearing security is 37.18 basis points higher than the replicating portfolio formed using outstanding securities. Thus, we calibrate the model to have this average asset payoff. This excess revenue is positive across all maturities. This is the well-known "on-the-run" premium.

We fit the parameters of the full model (the model of Section 5) to aggregate moments. The full model differs from the one presented in the previous section for the inclusion of minimum bidding requirements for dealers, which, as we discuss in the next sections, are key features of Treasury auctions. The objective function matches a few moments from the

[16]From Treasury International Capital (TIC), as of August 2014, about $6 trillion of securities are held by foreign investors, while from the Fed Board's H.41 release, foreign and international monetary accounts (FIMA) at the NY Fed are about $3.4 trillion as of that time. Assuming that the portfolio composition and bidding strategy of FIMA and non-FIMA are similar then an estimate of FIMA's share of competitive bids reported as indirect ones is: 3.3/6 X all foreign bids (from investment allotment) less FIMA non-competitive bids that are reported separately.

[17] We estimate a Svensson yield curve following the implementation details of Gürkaynak, Sack, and Wright (2007) but using intraday price data as of 1pm, which is when the auction closes (data from Thomson Reuters TickHistory).

Table 1: Calibration targets and model-implied values.

	Data	Model
A	-17.01	-12.44
Price sensitivity to fundamental	0.97	0.88
C	124.38	114.67
Error Std. Dev.	29.72	25.97
Expected excess revenue	37.18	37.60
Volatility of excess revenue	72.64	69.81
Indirect share	0.25	0.51
Volatility of indirect share	0.09	0.57
Dealer share	0.53	0.42
Volatility of dealer share	0.14	0.20
Direct share	0.10	0.01
Volatility of direct share	0.09	0.01

Note: Prices and excess revenues are all expressed in basis points.

model to their empirical counterparts: the pricing coefficients A, B and C in equation (24), the mean and variance of the price of the interest-rate-risk-neutral portfolio p at auction, the mean and variance of the price of the portfolio f at issuance, the mean allotted share and variance of non-competitive bids x (including the FIMA trades, or market orders), the mean allotted share to primary dealers, $\sum_{d=1}^{D} q_d$, the mean allotted share to indirect bidders ($\sum_{i=1}^{N} q_i$), and the mean allotted share to direct bidders, q_L. We obtain sample estimates of A, B and C by regressing the stop-out-price at each auction on a constant, the end-of-day secondary price on the issue date of the auction security (data from Bloomberg LP) and the non-competitive bids. As shown in Table 1, consistent with the model, excess revenues are positively correlated to the fundamental value on issue date (positive $B+F$), and it also increases with the share of securities allocated to market orders (positive C). The model moments are computed by making 100000 draws of realization of the fundamental f, all the signals in the economy y_i and the non-competitive demand x from the model, and calculating the equilibrium outcomes.

Table 2: Calibrated parameters

μ	$\tau_f^{-\frac{1}{2}}$	$\tau_{\varepsilon,s}^{-\frac{1}{2}}$	$\tau_{\varepsilon,l}^{-\frac{1}{2}}$	$\tau_x^{-\frac{1}{2}}$	$\bar{x}$	ρ	ρ_L	χ	N_S	N_L	N_D
40.8	73.5	692.1	282.4	0.06	0.12	64542	908	0.07	240	40	20

Note: μ, $\tau_f^{-\frac{1}{2}}$, $\tau_{\varepsilon,s}^{-\frac{1}{2}}$ and $\tau_{\varepsilon,l}^{-\frac{1}{2}}$ are all expressed in basis points.

3 Quantitative Results: Primary Dealers and Auction Revenue

To understand the role of intermediation we first examine auction revenue with either one or no dealers. This comparison illustrates two competing effects: 1) dealers lower auction revenue through market power; and 2) dealers reduce uncertainty by pooling information from their own and their customers' demands. Lower uncertainty raises the price that risk-averse investors are willing to bid. Such information pooling increases revenue, particularly in high-uncertainty times when demand might otherwise collapse.

Since the quantity of Treasury securities sold is normalized to 1, the auction price and auction revenues are the same. Therefore, in the plots that follow, we report the expected price and the variance of that price, varying one exogenous parameter at a time. In each exercise, all parameters other than the one being varied are held at their calibrated values.

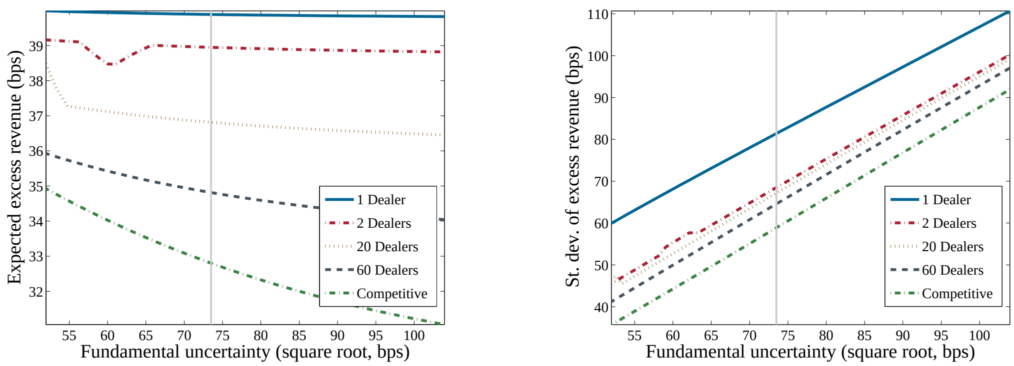
The left panel of Figure 4 plots expected auction revenues as a function of different levels of fundamental uncertainty as we vary the number of dealers, as well as in the case in which no dealer is present (line labelled competitive). The expected revenue gain from the introducing a single dealer into the auction varies from about 5 to about 8 basis points. The revenue gain comes from dealer information aggregation and is larger when the fundamental uncertainty about the future asset value (τ_f^{-1}) is large. The reason the dealer increases auction revenue is that he is providing his clients with information. This information makes Treasuries less risky to investors, eliciting stronger bids and increasing auction revenues.

The right panel of Figure 4 plots the conditional variance of auction revenue with a varying number of dealers as well as without dealers (line labelled competitive). Dealers increase the variance of auction revenue and this effect is largest when investors are most uncertain about the future value of Treasuries. The size of the increase in variance ranges from 30 to 40 basis points. This higher variance arises because dealers make investors better informed about the future value of Treasury securities. Absent any information about the future value of a security, bidders would always bid the same amount and revenue would be constant. With more precise information, bidders condition their bids on this information. When the fundamental value of the securities fluctuate, investors learn this information with a high degree of accuracy, and use this information in their bids leading to more volatile auction revenues.

We also examine how the effect of dealers varies as we change the variance of non-competitive bids. When these bids are less predictable, auction clearing prices are less

Figure 4: Dealers increase revenue in times of high uncertainty

(a) Expected auction revenue (b) Standard deviation of revenue



Note: Vertical line denotes the baseline calibration level of fundamental uncertainty.

clear signals about the true value of the asset. As a result, the value of information aggregation increases, which makes dealers more valuable in expected auction revenue terms–but the magnitude of this effect is, however, limited.

3.1 Dealers as Insurance Against Uncertainty Shocks

In standards models of IPO underwriting, dealers stabilize revenues. In our model, dealers increase the conditional variance of revenue, but they do provide revenue insurance in another way, by providing a hedge against uncertainty shocks.

Figure 4 shows that an auction with a dealer delivers lots more revenue than one without when fundamental uncertainty (τ_f^{-1}) is high. The result comes from the feedback of uncertainty through price information. In the absence of a dealer, when beliefs are uncertain, investors make less aggressive bids, meaning that their demand is less sensitive to the private information they have. However, when bids are less sensitive to private information, the market-clearing price contains less information, relative to the noise of the non-competitive bids. Thus, investors are all the more uncertain because the price does not reveal much information to them either.

The presence of a dealer avoids this problem. By providing investors signals about the average trade, without noise, investors do not need to only rely on prices to aggregate information. So, when prices become more noisy signals, the increase in uncertainty does not affect investors who trade through dealers because they already observe that average signal from other agents from the dealer. Since the difference between the price information and the dealer information is greatest when prices have the most noise, and prices have the most noise when uncertainty is high, auction revenue diverges in high-uncertainty times. This ability of dealers to raise revenue in times when it would otherwise be low by improving the information environment and reducing uncertainty is valuable to Treasury auctioneers.

Conditional on high uncertainty, revenue is still less predictable with the dealer than without. What we learn is that dealers make revenue more sensitive to the issue-specific value of the asset, because they aggregate and provide information about that value. But they make the auction revenue less sensitive to changes in uncertainty.

3.2 Quantitative Results: The Effect of More Dealers

We study mean and variance of auction revenues when increasing the number of dealers. Adding dealers reduces auction revenues (Figure 4). An auction with 20 dealers earns about 1.5 bps less on-the-run premium than as an auction with 1 dealer in the baseline calibration. When prior uncertainty about the future value of the asset is high (precision τ_f is low), the reduction in excess revenue doubles.

Increasing the number of dealers reduces revenue because more dealers disperse information. With two dealers, each dealer sees half of all the signals. With three dealers, each observes a third. Each time a new dealer is added, all dealers have less information. Since dealers disclose their information to their clients, all investors have less precise information sets, and are therefore more uncertain about the asset value, when the number of dealers rises. Because they are risk averse, more uncertain investors reduce their bids and average auction revenue falls.

The standard deviation of auction revenue also declines in the number of dealers (Figure 4 right panel) because dealers are segmenting information sets. When information about the true value of the asset is less precise, investors' bids respond less to changes in that true value. Since investors' bids are less sensitive to a random variable, they are also less variable themselves and create less variance in revenue. That effect shows up as a lower standard deviation of auction revenue from increasing the number of dealers.

4 Mixed Auctions: Choosing Direct or Indirect Bidding

A key distinguishing feature of US Treasury auctions is that they are mixed auctions, meaning that non-dealer investors can either place intermediated bids through primary dealers, or bid directly. As shown in this section, the intermediation choice affects auction revenue, lowers the revenue-maximizing number of dealers and amplifies the effect of low signals on auction revenue. To explore these effects in a simple setting, we now allow one investor to choose between bidding directly or indirectly through an intermediary. Without loss of generality, we assume that this choice is made by one of the large investors bidding through dealer 1.

4.1 A Model with Intermediation Choice

The large investor's choice to bid directly or indirectly (through an intermediary) only affects the information structure of that investor, its dealer, and other investors bidding with that same dealer. By affecting the information sets of these investors, the choice to bid directly affects the information content of the price y_p as well.

When the large investor chooses to bid directly on his own behalf, the first dealer's signal is the average of the first ν_s investors', the first $\nu_l - 1$ large investors' and the first dealer's signals:

$$\bar{y}_1 = \frac{\tau_{\varepsilon,s} \sum_{k \in \mathcal{I}_1^s} y_k + \tau_{\varepsilon,l} \left(\sum_{j \in \mathcal{I}_1^l} y_j + y_d \right)}{\nu_s \tau_{\varepsilon,s} + \nu_l \tau_{\varepsilon,l}}$$

As in the previous model, investor i who bids through intermediary d observes signals $X_i S = [y_i, \bar{y}_d, y_p]$. The large investor bidding directly observes only his own signal and the price information: $X_L S = [y_L, y_p]$.

Solution: Auction Outcomes If the large investor bids indirectly, the problem and the solution are the same as in the previous section. We solve now for the case with direct bidding. The following result shows that the auction price is a linear function of the dealer-level average individual investor signals, $\bar{y}_d$, the signal of the large investor, y_{N_S+1}, and of market orders x. Since the supply of the asset is 1, the price and the auction revenue are the same.

Result 3. *With N_D dealers and 1 large investor who bids directly, the auction revenue is*

$$p = A\left(N_D, 1\right) + B_1\left(N_D, 1\right)\bar{y}_1 + \frac{B_2\left(N_D, 1\right)}{N_D - 1}\sum_{d=2}^{N_D} \bar{y}_d + C\left(N_D, 1\right)x + F\left(N_D, 1\right)y_{N_S+1}, \quad (16)$$

where $B_1\left(N_D, 1\right) \equiv B\left(N_D, 1, 1\right)$, $B_2\left(N_D, 1\right) \equiv B\left(N_D, 1, 2\right)$, and the coefficient formulas are in the appendix.

4.2 Understanding Intermediation Choice

Given the solution to the auction outcomes, we study the large investor's choice of whether to bid directly or indirectly $l \in \{Ld, Li\}$. This decision, which determines the information of the large investor and of its dealer, will affect the distribution of the auction-clearing price and the signals that the investor will see. We computed the expected utility conditional on signals and a realized price. But when the investor chooses whether to invest through a dealer the only signal that he has seen is the realization of his private signal y_i. To solve for the large investor's choice to bid through a dealer we need to take an additional expectation over the information that has not yet realized. By computing the expectation over possible price realizations and what the investor might learn from the dealer, we find that expected utility takes the form

$$EU(l) = -\exp(\rho_L W_L)(1 + 2\theta_l \Delta\mathbb{V}_l)^{-\frac{1}{2}} \exp\left(-\frac{\mu_{rl}^2}{\theta_l^{-1} + 2\Delta\mathbb{V}_l} \right). \tag{17}$$

The intermediation decision affects utility in three ways: through the expected profit per unit μ_{rl}, the sensitivity of demand to expected profit θ_l, and through the ex-ante variance of expected profit $\Delta\mathbb{V}_l$. These three terms are:[18]

$$\mu_{rl} \equiv \mathbb{E}\{\mathbb{E}[f|X_l S] - p|y_i\}, \tag{18}$$

$$\theta_l \equiv \rho_L M_l \Big(1 - \frac{1}{2}\rho_L M_l \mathbb{V}[f|X_l S]\Big), \tag{19}$$

$$\Delta\mathbb{V}_l \equiv \mathbb{V}\{\mathbb{E}[f|X_l S] - p|y_i\} = \mathbb{V}[f - p|y_i] - \mathbb{V}[f|X_l S]. \tag{20}$$

Trading through a dealer has both costs and benefits for the large investor. We'll now discuss these and how they affect the expressions in (17)–(20).

The main cost of intermediation is that it reveals one's private information y_i to others. This effect shows up as a reduction in μ_{rl}, the ex-ante expectation today of expected profit per share, after all signals are observed. When the large investor bids through a dealer, their information set and their expectations will be the same as for every other agent who bids through that dealer $d1$: $E[f|X_l S] = E[f|X_{d1} S]$. Information sharing reduces μ_{rl} for two reasons. First, since many investors all condition their bids on this information, $E[f|X_{d1} S]$ has a large effect (closer to 1) on the auction-clearing price. Thus the difference $E[f|X_l S] - p$ is closer to zero with intermediation. Second, improving the precision of other investors' information raises the expected price p, which in turn, lowers μ_{rl} (see eq (18)). In equation (17), a decrease in μ_{rl} decreases expected utility because $\theta_l > 0$ (see appendix).

[18] See appendix for derivations of the following three equations and support for the analysis that follows.

The advantage of intermediation is that it allows the large investor to observe more information. Better information allows the large investor to make better trading decisions and increases his expected utility. This effect shows up through θ_l. In the appendix we show that $\theta_l > 0$ and is strictly decreasing in the posterior variance of the asset payoff $\mathbb{V}[f|X_l S]$. When the large investor trades through a dealer he receives more information, which decreases $\mathbb{V}[f|X_l S]$. This ensures that $\theta_{Li} > \theta_{Ld}$. Equation (17) shows that expected utility is increasing in θ_l. Thus, the reduction in uncertainty effect embodied in θ_l works in favor of dealer-intermediated trade.

Trading through a dealer also affects the ex-ante variance $\Delta\mathbb{V}_l$. As already discussed, when the large investor trades through a dealer the variance of her posterior beliefs $\mathbb{V}[f|X_l S]$ is lower. From equation (20) we can see that this increases the ex-ante variance of the expected profit $\Delta\mathbb{V}_l$. The reason for this is that when the agent trades through a dealer she will receive more information. This means that her beliefs about the profit will change more, which means a higher ex-ante variance. This change in $\Delta\mathbb{V}_l$ has two opposing effects on expected utility. First, an increase in $\Delta\mathbb{V}_l$ increases the exponential term in equation (17), which decreases $EU(l)$. This effect arises because the large investor is risk averse and higher $\Delta\mathbb{V}_l$ corresponds to more risk in continuation utility. The second effect is that an increase in $\Delta\mathbb{V}_l$ reduces $(1 + 2\theta_L \Delta\mathbb{V}_l)^{-\frac{1}{2}}$, which increases $EU(l)$. The intuition for this is that when the variance of the expected profit is larger, there are more realizations with large magnitude (more weight in the tails of the distribution). Since these are the states that generate high profit, this effect increases expected utility.

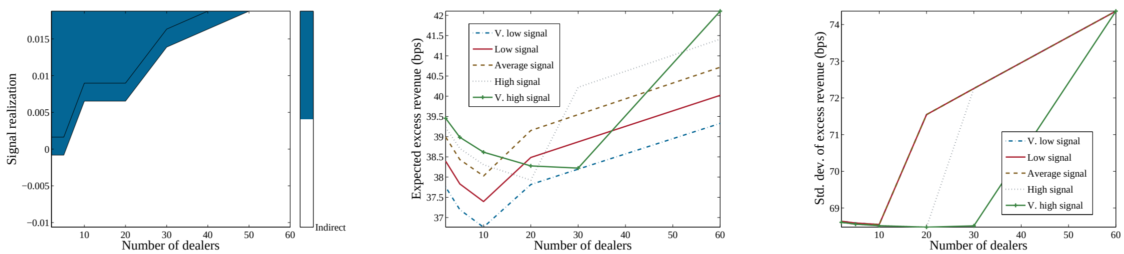
Note from equation (17) that as $\mu_{rl} \to 0$, the first effect disappears and an increase in the ex-ante variance of the profit will unambiguously increase expected utility. The reason for this is that the strength of the second effect depends on the mean of the expected profit, μ_{rl}. When $\mu_{rl} \neq 0$ the increase in the ex-ante variance $\Delta\mathbb{V}_l$ increases the probability that the expected profit $\mathbb{E}[f|X_l S] - p$ will be close to zero as well as increasing the probability of large observations. So intuitively the gains from the increase in ex-ante variance larger when μ_{rl} is closer to zero. We use these three effects to understand the intermediation choice results below.

4.3 Quantitative Results: Intermediation Choice

Figure 5 reveals that a large investor with a medium or high signal always bids directly, while an investor with a low signal may or may not choose to bid through a dealer. Why trade through a dealer when the signal is low? The magnitude of the cost of intermediation (represented by a decrease in μ_{rl} in (18)) depends on the large investor's private signal y_i. To see the effect of y_i, first note that on average, p will usually fall below $E[f|X_l S]$, reflecting

Figure 5: Mixed auctions.

(a) Direct bidding decision (b) Auction Revenue Mean (c) Auction Revenue St. Dev.



Note: Bidding directly is optimal in the blue region. "Very low signal": $y_{N_S+1} = \mu - 2\left(\tau_f + \tau_{\varepsilon,l}\right)^{-\frac{1}{2}}$; "low signal": $y_{N_S+1} = \mu - \left(\tau_f + \tau_{\varepsilon,l}\right)^{-\frac{1}{2}}$; "average signal": $y_{N_S+1} = \mu$; "high signal": $y_{N_S+1} = \mu + \left(\tau_f + \tau_{\varepsilon,l}\right)^{-\frac{1}{2}}$; "very high signal": $y_{N_S+1} = \mu + 2\left(\tau_f + \tau_{\varepsilon,l}\right)^{-\frac{1}{2}}$.

a risk premium in the return on the risky asset. If the large investor observes a signal y_i that is below the prior mean μ, it lowers the ex-anted expectation $\mathbb{E}\{E[f|X_l S] - p|y_i\}$ toward zero. The closer this expected value is to zero the lower the expected profit from trading, which means that that the cost of sharing one's information is low. The economic force here is that investors with low signals are awarded fewer securities, on average. (See low large investor allocation in Table 3.) If they do not expect to hold many Treasuries, they are neither helped or hurt (in expectation) if the auction price incorporates their private information. Thus, the cost of bidding through a dealer is very low when the investor's private signal is low. Therefore, in these low-signal states, investors choose intermediation. In summary, large investors who observe bad news bid through the dealer.

The benefit of bidding through a dealer is the ability to observe the dealer's information. This information benefit (represented by a higher θ_l in (17)) is invariant in the investor's private signal y_i. Since a low private signal reduces the cost of intermediation μ_{rl} and leaves the benefit θ_l unchanged, low private signals make intermediation more attractive.

Figure 5 also illustrates that with more dealers (moving to the right), the large investor is more likely to bid indirectly. There are two competing effects underlying this relationship. When the number of dealers is large, each dealer aggregates fewer signals and provides less precise information to their clients (θ_l is lower). Thus the benefit of intermediation is less. On the other hand, sharing the private signal is less costly with more dealers because the dealer will reveal the signal to fewer other clients (μ_{rl} is higher). Also, the dealer benefits less from investors' signals when more dealers reduce the competitive power of any one dealer. Our results teach us that for the calibrated parameter values, the expected return effect (μ_{rl}) dominates. Thus, for around 20 dealers (the current level), intermediation is more valuable when the number of dealers is large.

This result offers an important caveat to the discussion in the previous section about the optimal number of dealers. If policy-makers lower the number of dealers to improve information aggregation, this smaller set of dealers imply that intermediation may become less desirable. The net effect of limiting dealer entry could be to reduce intermediated trade, and with it, auction revenue.

As in the previous models, expected auction revenue and its standard deviation are higher when agents bid through dealers. When the number of dealers increases, revenue falls in cases where the bidders eschew intermediaries and bid directly.

4.4 A New Financial Accelerator Channel of Intermediation

This result reveal a new channel through which financial intermediaries can amplify shocks. When signals about the value of a financial asset are negative, this information is more likely to be shared with a dealer and his clients. Positive signals are less likely to be shared because an investor who receives a positive signal then expects to take a large portfolio position in the asset and faces a high expected cost from sharing his information. But sharing a bad signal places that signal in the information set of many more investors and causes a large number of investors to demand less of the asset. Thus, bad signals may affect the demand of more investors than good signals do and have a larger effect on asset prices.

This effect shows up as a distribution of auction revenues that has unconditional negative skewness. The first panel of Table 3 reveals that the unconditional skewness is -1.05. The fourth panel reveals that most of this skewness comes from states where the large investor receives a low signal. When shocks are good, they have a moderate effect on the asset price and the auction revenue. But with a bad realization of the asset's value, large investors observe negative signals. These investors choose to share their low signals with primary dealers, which in turn lowers the demands of other investors and has a significantly negative effect on auction revenues.

5 Minimum Bidding Requirements

The final feature of Treasury auctions that we examine is the requirements that primary dealers be consistent, active participants in Treasury auctions. A familiar tension reappears: Our analysis demonstrates how such a bidding requirement reduces market power, but also discourages investors' use of dealers. When more investors bid directly, dealers are less able to aggregate information.

In any given auction, there is no strict bidding requirement. But if on average, a dealer is not awarded a sufficient quantity of Treasuries, his primary dealer status could be revoked. To capture the essence of this dynamic requirement in a static model, we model the bidding requirement as a cost levied on a dealer who purchases too little. Conversely, a dealer who purchases a large dollar amount of Treasuries faces a relaxed bidding constraint in the future. We model this benefit as a current transfer. Thus, for a dealer who purchases a dollar amount qp of Treasuries through the auction, we assume a low-bid penalty of $\chi_0 - \chi qp$, or equivalently a benefit of $\chi qp - \chi_0$.

We introduce this minimum bidding policy in the model from the previous section. The investors' objectives are the same as before. But the dealers' problem becomes

$$\max_{q_D,p} \mathbb{E}[-\exp(-\rho_l W_D)|X_D S] \tag{21}$$

$$\text{s.t. } W_D = W_{0,D} + q_D(f-p) - \chi_0 + \chi q_D p, \tag{22}$$

and subject to the market-clearing condition (8). By substituting the equation for W_D in the dealer's problem into the objective function, evaluating the expectation and taking the log, and dropping the constant terms that do not affect optimization, we can simplify the dealer's bidding problem to be: $\max_{q_D,p} \; q_D(\mathbb{E}[f|X_D S] - p(1-\chi)) - \frac{1}{2}\rho_l q_D^2 \mathbb{V}[f|X_D S]$, s.t. the market clearing constraint (8). Taking the first order condition with respect to q_D, we obtain

$$q_D(p) = \frac{\mathbb{E}[f|X_D S] - p(1-\chi)}{\rho_l \mathbb{V}[f|X_D S] + (1-\chi) dp/dq_D} \equiv M_D \left(\mathbb{E}[f|X_D S] - p(1-\chi)\right). \tag{23}$$

Note that the bidding requirement shows up like a dealer price subsidy, encouraging the dealer to purchase more of the asset. It also mitigates the effect of dealer market power by multiplying the dp/dq_D term by a number less than one. It does not change expectations or the investors' problem, except through the change in equilibrium price coefficients. The following results show how the penalty χ affects the pricing coefficients. As before, there are two cases: one where the large bidder bids directly and another where they bid through dealer 1. We explore each in turn.

Result 4. *In an auction with N_D dealers and 1 large investor who bids directly, the auction revenue is*

$$p = A(N_D, 1, \chi) + B_1(N_D, 1, \chi)\bar{y}_1 + \frac{B_2(N_D, 1, \chi)}{N_D - 1} \sum_{d=2}^{N_D} \bar{y}_d + C(N_D, 1, \chi)x + F(N_D, 1, \chi) y_{N_S+1}. \tag{24}$$

where the coefficients A, $B_1(N_D, 1, \chi)$, $B_2(N_D, 1, \chi)$, $F(N_D, 1, \chi)$ and $C(N_D, 1, \chi)$ are given by (34)-(38) in the Appendix.

Note that the cost of bidding too little χ enters directly only through the coefficient C. However, the penalty χ also changes the dealers' demand coefficients $M_{D,1}$ and $M_{D,2}$; by making dealers less responsive to the price, it changes the large investor's demand coefficients $M_{L,1}$, $M_{L,2}$ and $M_{L,0}$ as well.

By examining the price impact of a dealer, we see that one clear effect of minimum bidding requirements is to help counteract the tendency of dealers to use their market power to hold price low. The proof of Result 4 shows that, for dealer 1,

$$\frac{dp}{dq_d} = -\Bigg\{\left(\nu_s\hat{\tau}_1\rho^{-1} + (\nu_l - 1)\, M_{L,1}\right)\left(\frac{\beta_1(2) - \tilde{B}}{\tilde{B}}\right) + \chi\,(N_D - 1)\, M_{D,2}$$
$$+ (N_D - 1)\left(\nu_s\hat{\tau}_2\rho^{-1} + \nu_l M_{L,2} + M_{D,2}\right)\left(\frac{\beta_2(2) - \tilde{B}}{\tilde{B}}\right) + M_{L,0}\left(\frac{\beta_L(2) - \tilde{B}}{\tilde{B}}\right)\Bigg\}^{-1}.$$

Notice that the bidding cost raises the term in brackets, which is raised to the power -1 and multiplied by -1, to make it positive. Thus, an increase in the penalty χ make the effect of dealer demand on the price smaller. This reduces the strategic power of dealers.

Next, we turn to the case where the large investor bids indirectly, through dealer 1.

Result 5. *In an auction with N dealers and 1 large investor who bids indirectly through dealer 1, the auction revenue is*

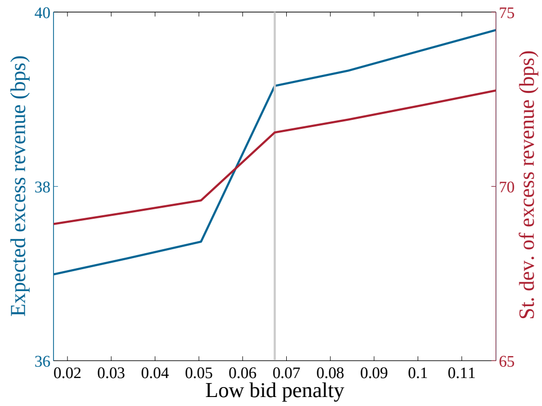
$$p = A(N_D, 0, \chi) + \frac{B(N_D, 0, \chi)}{N_D}\sum_{d=1}^{N_D}\bar{y}_d + C(N_D, 0, \chi)\, x. \tag{25}$$

where $A(N_D, 0, \chi)$, $B(N_D, 0, \chi)$ and $C(N_D, 0, \chi)$ are given by (39), (40) and (41) in the Appendix.

As before, the minimum bidding penalty χ affects directly the pricing coefficient C. It also affects A and B through the coefficients M_D and M_L that multiply expected profit to produce dealers' and the large investors' demands.

Results An increase in the penalty χ for low bidding does raise the expected revenues (Figure 6, left scale) by incentivizing dealers to make larger bids. This effect is intuitive. However, contrary to policy wisdom, low-bid penalties do not reduce auction revenue risk. Figure 6 (right scale) shows that more stringent bidding requirements increase the volatility of auction revenue. Variance rises because the greater quantity demanded by dealers makes

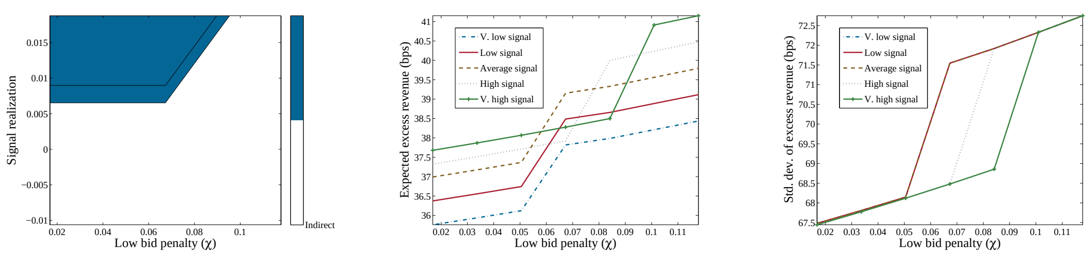
Figure 6: Auction with a Low-Bid Penalty.



Note: Auction outcomes as a function of the low-bid penalty χ, assuming that the large investor bids indirectly.

Figure 7: Intermediation Decision with a Low-Bid Penalty.

(a) Direct bidding decision (b) Auction Revenue Mean (c) Auction Revenue St. Dev.



Note: Bidding directly is optimal in the blue region. “Very low signal”: $y_{N_S+1} = \mu - 2\left(\tau_f + \tau_{\varepsilon,l}\right)^{-\frac{1}{2}}$; “low signal”: $y_{N_S+1} = \mu - \left(\tau_f + \tau_{\varepsilon,l}\right)^{-\frac{1}{2}}$; “average signal”: $y_{N_S+1} = \mu$; “high signal”: $y_{N_S+1} = \mu + \left(\tau_f + \tau_{\varepsilon,l}\right)^{-\frac{1}{2}}$; “very high signal”: $y_{N_S+1} = \mu + 2\left(\tau_f + \tau_{\varepsilon,l}\right)^{-\frac{1}{2}}$.

the price level more informative about the issue-specific value of the asset f. So, when investors condition their bids on the realized auction-clearing price, they end up having a demand that is more responsive to the random value. The more sensitive response to a random variable is the increase in revenue variance.

But minimum bidding requirements do induce investors to bid through the dealers. Figure 7 shows that for sufficiently high costs χ, investors almost always choose intermediated bidding. The high cost acts like a price subsidy and helps to offset their incentive to bid low that comes from market power. If dealers acts less strategically, they also use investors' information in less strategic ways. As a result sharing private information with a dealer is less costly and encourages bidding through the dealer.

6 Conclusions

The vast majority of US Treasury auction bids are placed by a small set of primary dealers for either their own or their customers' accounts. Using data from U.S. Treasury auctions, we estimate a structural auction model to quantify the costs and benefits of the primary dealer system. The model allows us to do counter-factuals and policy evaluation. We estimate the effects of increasing dealer entry, allowing investors to bid directly, and changing minimum bidding requirements. Overall we find that despite their pricing power, the primary dealer system sustains revenues by aggregating information. Policies aimed at reducing dealers' market power also adversely affect their ability to aggregate information. Better information aggregation is important because it increases expected auction revenue. However, information aggregation also increases auction revenue variance because the bids of better-informed investors are more sensitive to changes in Treasuries' value.

The common theme throughout the paper is a reversal of the common wisdom about dealers as underwriters. The prevailing thinking about underwriters is that they lower auction revenue, but also revenue risk. In the information model that we present, we find the exact opposite: when investors bid through dealers, both mean and variance of auction revenue increase. The stark difference in these predictions highlights how policy prescriptions may be heavily dependent on the role of intermediation. While many intermediaries perform roles other than information aggregation, this role is a key one in Treasury auctions and is likely to be present in some form in other markets as well. Thus the unique features of Treasury auctions makes them a useful laboratory to isolate and investigate this new facet of intermediation.

References

Amihud, Y. and H. Mendelson (1991): "Liquidity, maturity, and the yields on US Treasury securities," *The Journal of Finance*, 46, 1411–1425.

Armantier, O. and E. Sbaï (2006): "Estimation and comparison of Treasury auction formats when bidders are asymmetric," *Journal of Applied Econometrics*, 21, 745–779.

Babus, A. and C. Parlatore (2015): "Strategic Fragmented Markets," Working paper, New York University.

Back, K. and J. F. Zender (1993): "Auctions of divisible goods: on the rationale for the treasury experiment," *Review of Financial Studies*, 6, 733–764.

Beatty, R. P. and J. R. Ritter (1986): "Investment banking, reputation, and the underpricing of initial public offerings," *Journal of financial economics*, 15, 213–232.

Bikhchandani, S. and C.-f. Huang (1989): "Auctions with resale markets: An exploratory model of Treasury bill markets," *Review of Financial Studies*, 2, 311–339.

——— (1993): "The economics of Treasury securities markets," *The Journal of Economic Perspectives*, 117–134.

Brady, N., R. Breeden, and A. Greenspan (1992): "Joint report on the government securities market," *US Government Printing Office*.

Chari, V. V. and R. J. Weber (1992): "How the U.S. Treasury Should Auction Its Debt," *Federal Reserve Bank of Minneapolis*, 16, 1–8.

Coutinho, P. B. (2013): "When-Issued Markets and Treasury Auctions," Working paper, University of California at Los Angeles.

Flitter, E. (2011): "U.S. caught China buying more debt than disclosed," *Reuters*.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2007): "The US Treasury yield curve: 1961 to the present," *Journal of Monetary Economics*, 54, 2291–2304.

Hansen, R. S. and P. Torregrosa (1992): "Underwriter compensation and corporate monitoring," *The Journal of Finance*, 47, 1537–1555.

Hortaçsu, A. and J. Kastl (2012): "Valuing Dealers' Informational Advantage: A Study of Canadian Treasury Auctions," *Econometrica*, 80, 2511–2542.

Hortaçsu, A. and D. McAdams (2010): "Mechanism Choice and Strategic Bidding in Divisible Good Auctions: An Empirical Analysis of the Turkish Treasury Auction Market," *Journal of Political Economy*, 118, 833–865.

Keloharju, M., K. G. Nyborg, and K. Rydqvist (2005): "Strategic Behavior and Underpricing in Uniform Price Auctions: Evidence from Finnish Treasury Auctions," *The Journal of Finance*, 60, 1865–1902.

Krishnamurthy, A. (2002): "The bond/old-bond spread," *Journal of Financial Economics*, 66, 463–506.

Lou, D., H. Yan, and J. Zhang (2013): "Anticipated and Repeated Shocks in Liquid Markets," *Review of Financial Studies*, 26, 1891–1912.

Malvey, P. F. and C. M. Archibald (1998): "Uniform-price auctions: Update of the Treasury experience," *US Treasury*.

Malvey, P. F., C. M. Archibald, and S. T. Flynn (1995): "Uniform-Price Auctions: Evaluation of the Treasury Experience," *Office of Market Finance, US Treasury*.

Nyborg, K. G., K. Rydqvist, and S. M. Sundaresan (2002): "Bidder Behavior in Multiunit Auctions: Evidence from Swedish Treasury Auctions," *Journal of Political Economy*, 110, 394–424.

Nyborg, K. G. and S. M. Sundaresan (1996): "Discriminatory versus uniform Treasury auctions: Evidence from when-issued transactions," *Journal of Financial Economics*, 42, 63–104.

Ritter, J. R. and I. Welch (2002): "A Review of IPO Activity, Pricing, and Allocations," *The Journal of Finance*, 57, 1795–1828.

Svensson, L. E. (1994): "Estimating and interpreting forward interest rates: Sweden 1992-1994," Tech. rep., National Bureau of Economic Research.

Wilson, R. (1979): "Auctions of shares," *Quarterly Journal of Economics*, 93, 675–689.

A Technical Details

Let Z be the vector of orthogonal shocks.

$$Z = \begin{bmatrix} \varepsilon_1 & \ldots & \varepsilon_N & x \end{bmatrix}'. \tag{26}$$

Then, we can represent the vector of signals as

$$S = 1_{N+N_D+1} f + \Pi Z,$$

where 1_{N+N_D+1} is an $(N+N_D+1) \times 1$ vector of ones and the $(N+N_D+1) \times (N+N_D+1)$ matrix Π is given by

$$\Pi = \begin{bmatrix} & I_{N_S} & & & 0 & & & 0 & & 0 \\ & 0 & & & I_{N_L} & & & 0 & & 0 \\ & 0 & & & 0 & & & I_{N_D} & & 0 \\ \omega_s \cdot 1'_{\nu_s} & \cdots & 0'_{\nu_s} & \omega_l \cdot 1'_{\nu_l} & \cdots & 0'_{\nu_l} & \omega_l & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0'_{\nu_s} & \cdots & \omega_s \cdot 1'_{\nu_s} & 0'_{\nu_l} & \cdots & \omega_l \cdot 1'_{\nu_l} & 0 & \cdots & \omega_l & 0 \\ & \frac{\omega_s}{N_D} \cdot 1'_{N_S} & & & \frac{\omega_l}{N_D} \cdot 1'_{N_L} & & & \frac{\omega_l}{N_D} \cdot 1'_{N_D} & & \frac{C}{B} \end{bmatrix}.$$

where I_N denotes the $N \times N$ identity matrix. The matrix Π tells us how signals weight the orthogonal shocks. Each private signal y_i has a weight of one on its own signal noise. The price signal $(p - A - C\bar{x})/B$ is the average of all the $N+1$ signals, plus supply noise x. Thus, its weight on each small signal noise is ω_s, on large investor's noise ω_l, and ω_l on each dealer's signal noise.

This representation of signals as linear combinations of independent shocks allows us to express signal variance and covariances easily as

$$\begin{aligned} \mathbf{Cov}\left(f, X_j S\right) &= \mathbf{Cov}\left(f, X_j \left(1_{N+N_D+1} f + \Pi Z\right)\right) \\ &= X_j \left(\mathbb{E}\left[1_{N+1} f^2 + \Pi Z f\right] - \mathbb{E}\left[f\right] \mathbb{E}\left[1_{N+1} f + \Pi Z\right]\right) = 1_m \tau_f^{-1} \end{aligned} \tag{27}$$

$$\mathbb{V}\left(X_j S\right) = X_j \mathbb{V}\left(1_{N+N_D+1} f + \Pi Z\right) X_j' = X_j \left(\tau_f^{-1} 1_{N+N_D+1} 1'_{N+N_D+1} + \Pi \mathbb{V}\left[Z\right] \Pi'\right) X_j', \tag{28}$$

where the variance matrix of the shocks is $\mathbb{V}[Z] = diag([\tau_{\varepsilon,s}^{-1} 1_{N_S}, \tau_{\varepsilon,l}^{-1} 1_{N_L}, \tau_{\varepsilon,l}^{-1} 1_{N_D}, \tau_x^{-1}])$.

The X operators that select the relevant signals that each agent observes are as follows. In each case, the operator X is a linear matrix operation. It is described by a matrix of zeroes and ones that pre-multiplies the vector of all signals.

$$X_i = \begin{bmatrix} \mathbf{1}'_i & 0 & 0 \\ 0'_N & 1 & 0 \\ 0'_N & 0 & 1 \end{bmatrix}$$

where $\mathbf{1}_i$ is an $(N \times 1)$ vector of n zeros with a 1 in the ith position and 0_N is an $N \times 1$ vector of zeros. The matrix X_i is constructed so that it picks off from the vector of all signals S only the signals that agent i observes.

Competitive model without dealers In the model without intermediation, the X operators that select the relevant signals that each agent observes are as follows.

$$X_i^C = \begin{bmatrix} \mathbf{1}'_i & 0 & 0 \\ 0'_N & 0 & 1 \end{bmatrix}.$$

B Proofs

Proof of Result 1: Price in the no dealer model. Recall that the price conjecture in this model is

$$p = A(0) + B(0) N_S^{-1} \sum_{i=1}^{N_S} y_i + F(0) N_L^{-1} \sum_{j=1}^{N_L} y_j + C(0) x.$$

Then, an investor's unbiased price signal is

$$\frac{p - A - C\bar{x}}{\tilde{B}} = \frac{B}{\tilde{B}} N_S^{-1} \sum_{i=1}^{N_S} y_i + \frac{F}{\tilde{B}} N_L^{-1} \sum_{j=1}^{N_L} y_j + \frac{C}{\tilde{B}} (x - \bar{x}).$$

That is an unbiased signal about f. However, the price signal and the private signals have correlated signal errors. The β coefficients are formed using the optimal linear projection formulas correct for this covariance.

The small investors' demand functions (bids) in the competitive market model are given by equation (9) and the large investors' bids by equation (10). Substituting these bids into the market clearing condition $x + \sum_{i=1}^{N} q_i = 1$ yields

$$x + \rho^{-1} \hat{\tau}_s \sum_{i=1}^{N_S} (\mathbb{E}[f|X_i, S] - p) + M_L \sum_{j=1}^{N_L - 1} (\mathbb{E}[f|X_j, S] - p) + q_L = 1.$$

Substituting for $\mathbb{E}[f|X_i, S]$, we obtain

$$\begin{aligned} 1 = x + \rho^{-1} \hat{\tau}_s \sum_{i=1}^{N_S} \left((1 - \beta_s(1) - \beta_s(2)) \mu + \beta_s(1) y_i + \beta_s(2) \frac{p - A - C\bar{x}}{\tilde{B}} - p \right) \\ + M_L \sum_{j=1}^{N_L - 1} \left((1 - \beta_l(1) - \beta_l(2)) \mu + \beta_l(1) y_i + \beta_l(2) \frac{p - A - C\bar{x}}{\tilde{B}} - p \right) + q_L. \end{aligned}$$

Taking the derivative with respect to q_L, we obtain

$$0 = \left(N_S \hat{\tau}_s \rho^{-1} \frac{\beta_s(2) - \tilde{B}}{\tilde{B}} + (N_L - 1) M_L \frac{\beta_l(2) - \tilde{B}}{\tilde{B}} \right) \frac{dp}{dq_L} + 1,$$

so that M_L is given implicitly by

$$M_L^{-1} = \rho_L \hat{\tau}_l^{-1} - \left(N_S \hat{\tau}_s \rho^{-1} \frac{\beta_s(2) - \tilde{B}}{\tilde{B}} + (N_L - 1) M_L \frac{\beta_l(2) - \tilde{B}}{\tilde{B}} \right)^{-1}.$$

Solving

$$\begin{aligned} 1 = x + \rho^{-1} \hat{\tau}_s \sum_{i=1}^{N_S} \left((1 - \beta_s(1) - \beta_s(2)) \mu + \beta_s(1) y_i + \beta_s(2) \frac{p - A - C\bar{x}}{\tilde{B}} - p \right) \\ + M_L \sum_{j=1}^{N_L} \left((1 - \beta_l(1) - \beta_l(2)) \mu + \beta_l(1) y_i + \beta_l(2) \frac{p - A - C\bar{x}}{\tilde{B}} - p \right) \end{aligned}$$

for p and matching coefficients yields the implicit solution to the price equation in the result. The existence of a set of coefficients verifies the price conjecture.

Since the supply of the asset is one, auction revenue is the price of the asset.

Proof of Result 2: Price in the N dealer model with all investors bidding through a dealer. The small investors' demand functions (bids) in the competitive market model are given by equation (9) and the large investors' bids by equation (10). Substituting these bids into the market clearing condition $x + \sum_{i=1}^{N} q_i = 1$ yields

$$1 = x + \sum_{d=1}^{N_D} \left(\nu_s \hat{\tau} \rho^{-1} + (\nu_l + 1) M_L\right) \left(\mathbb{E}\left[f | X_d S\right] - p\right).$$

Substituting in (12) for the conditional expectation and then taking the derivative with respect to q_L, we obtain the price impact of a large investor

$$0 = 1 + \left(N_S \hat{\tau} \rho^{-1} + (N_L + N_D - 1) M_L\right) \left(\frac{\beta(2) - B}{B}\right) \frac{dp}{dq_L},$$

so that M_L is given implicitly by

$$M_L^{-1} = \rho_L \hat{\tau}^{-1} - B \left(\left(N_S \hat{\tau} \rho^{-1} + (N_L + N_D - 1) M_L\right) (\beta(2) - B)\right)^{-1}.$$

Finally, solving the market clearing condition for the settle price and equating the coefficients, we obtain the coefficient expressions in the result.

Proof of Result 3: Price when the large investor bids directly. In this setting, we can represent the vector of signals as

$$S = 1_{N+N_D+1} f + \Pi^D Z,$$

where the $(N + N_D + 1) \times (N + 1)$ matrix Π^D is given by

$$\Pi^D = \left[\begin{array}{ccccccccccc}
\multicolumn{3}{c}{I_{N_S}} & \multicolumn{4}{c}{0} & \multicolumn{3}{c}{0} & 0 \\
\multicolumn{3}{c}{0} & \multicolumn{4}{c}{I_{N_L}} & \multicolumn{3}{c}{0} & 0 \\
\multicolumn{3}{c}{0} & \multicolumn{4}{c}{0} & \multicolumn{3}{c}{I_{N_D}} & 0 \\
\omega_{s,2} \cdot 1'_{\nu_s} & \cdots & 0'_{\nu_s} & 0 & \omega_{l,2} \cdot 1'_{\nu_l - 1} & \cdots & 0'_{\nu_l} & \omega_{l,2} & \cdots & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0'_{\nu_s} & \cdots & \omega_s \cdot 1'_{\nu_s} & 0 & 0'_{\nu_l} & \cdots & \omega_l \cdot 1'_{\nu_l} & 0 & \cdots & \omega_l & 0 \\
\frac{B_1}{\tilde{B}} \omega_{s,2} \cdot 1'_{\nu_s} & \cdots & \frac{B_2}{\tilde{B}(N_D - 1)} \omega_s \cdot 1'_{\nu_s} & \omega_{l,2} \frac{F}{\tilde{B}} & \frac{B_1}{\tilde{B}} \omega_{l,2} \cdot 1'_{\nu_l - 1} & \cdots & \frac{B_2}{\tilde{B}(N_D - 1)} \omega_l \cdot 1'_{\nu_l} & \frac{B_1}{\tilde{B}} \omega_{l,2} & \cdots & \frac{B_2}{\tilde{B}} \omega_l & \frac{C}{\tilde{B}}
\end{array}\right].$$

As before, each agent in the economy observes a subset of these signals, given by X_j, a matrix of zeros and ones. This equilibrium has the following information structure.

- Individual investor $i \in \mathcal{I}_d$ bidding through dealer d observes his own private signal y_i. In addition, the dealer tells them the average of the signals of the dealer's customers, $\bar{y}_d$. Finally, the investor can condition on the settle price p. Thus, for an investor bidding through dealer d observes

$$X_i = \left[\begin{array}{cccc} \mathbf{1}'_i & 0 & 0'_D & 0 \\ 0'_N & 0 & \mathbf{1}'_d & 0 \\ 0'_N & 0 & 0'_D & 1 \end{array}\right], \quad i \in \mathcal{I}_d.$$

- The large investor bidding directly on the other hand observes only his own signal and conditions on the price, so that

$$X_L = \left[\begin{array}{cccc} 0'_N & 1 & 0'_D & 0 \\ 0'_N & 0 & 0'_D & 1 \end{array}\right].$$

The belief updating in this case is similar to the setting studied in the previous Section. The beliefs of an investor bidding through dealer d coincide with the beliefs of the dealer. All investors use the price to form a second unbiased signal about f, which is given by:

$$y_p = \frac{B_1}{\tilde{B}} \bar{y}_1 + \frac{B_2}{\tilde{B}} \frac{1}{N_D - 1} \sum_{d=2}^{N_D} \bar{y}_d + \frac{F}{\tilde{B}} y_L + \frac{C}{\tilde{B}} (x - \bar{x}).$$

As in the previous Section, the average signal of dealer d's clients is contained in the price signal y_p. Using equation (3), we find that the beliefs of the investors bidding through dealer 1 are

$$\mathbb{E}[f|X_1 S] = (1 - \beta_1(1) - \beta_1(2))\,\mu + \beta_1(1)\,\bar{y}_1 + \beta_1(2)\,y_p,$$

and the conditional variance of beliefs is

$$\mathbb{V}[f|X_1 S] = \tau_f^{-1}(1 - \beta_1(1) - \beta_1(2)) \equiv \hat{\tau}_1^{-1}.$$

Similarly, the beliefs of investors bidding through dealer 2 are

$$\mathbb{E}[f|X_d S] = (1 - \beta_2(1) - \beta_2(2))\,\mu + \beta_2(1)\,\bar{y}_d + \beta_2(2)\,y_p, \quad d = 2, \ldots, N_D,$$

and the conditional variance of beliefs is

$$\mathbb{V}[f|X_d S] = \tau_f^{-1}(1 - \beta_2(1) - \beta_2(2)) \equiv \hat{\tau}_2^{-1}.$$

Finally, the direct bidder averages his signal and the signal from the price to obtain

$$\mathbb{E}[f|X_L S] = (1 - \beta_L(1) - \beta_L(2))\,\mu + \beta_L(1)\,y_L + \beta_L(2)\,y_p$$

and

$$\mathbb{V}[f|X_L S] = \tau_f^{-1}(1 - \beta_L(1) - \beta_L(2)) \equiv \hat{\tau}_L^{-1}.$$

Turn now to solving for the price impact of large investors bidding through dealer 1. Substituting demand functions into the market clearing constraint, we obtain

$$\begin{aligned}
1 = x + q_L &+ \left(\nu_s \hat{\tau}_1 \rho^{-1} + (\nu_l - 1)\,M_{L,1}\right)\left((1 - \beta_1(1) - \beta_1(2))\,\mu + \beta_1(1)\,\bar{y}_1 + \beta_1(2)\,y_p - p\right) \\
&+ \sum_{d=2}^{N_D}\left(\nu_s \hat{\tau}_2 \rho^{-1} + (\nu_l + 1)\,M_{L,2}\right)\left((1 - \beta_2(1) - \beta_2(2))\,\mu + \beta_2(1)\,\bar{y}_d + \beta_2(2)\,y_p - p\right) \\
&+ M_{L,0}\left((1 - \beta_L(1) - \beta_L(2))\,\mu + \beta_L(1)\,y_{N_S+1} + \beta_L(2)\,y_p - p\right).
\end{aligned}$$

Taking the derivative with respect to the demand of the large investor, we obtain

$$\begin{aligned}
\frac{dp}{dq_L} = -\tilde{B}\Big\{&\left(\nu_s \hat{\tau}_1 \rho^{-1} + (\nu_l - 1)\,M_{L,1}\right)\left(\beta_1(2) - \tilde{B}\right) \\
&+ (N_D - 1)\left(\nu_s \hat{\tau}_2 \rho^{-1} + (\nu_l + 1)\,M_{L,2}\right)\left(\beta_2(2) - \tilde{B}\right) + M_{L,0}\left(\beta_L(2) - \tilde{B}\right)\Big\}^{-1}.
\end{aligned}$$

Thus, $M_{L,1}$ is given implicitly as the solution to

$$\begin{aligned}
M_{L,1}^{-1} = \rho_l \hat{\tau}_1^{-1} - \Bigg\{&\left(\nu_s \hat{\tau}_1 \rho^{-1} + (\nu_l - 1)\,M_{L,1}\right)\left(\frac{\beta_1(2) - \tilde{B}}{\tilde{B}}\right) \\
&+ (N_D - 1)\left(\nu_s \hat{\tau}_2 \rho^{-1} + (\nu_l + 1)\,M_{L,2}\right)\left(\frac{\beta_2(2) - \tilde{B}}{\tilde{B}}\right) + M_{L,0}\left(\frac{\beta_L(2) - \tilde{B}}{\tilde{B}}\right)\Bigg\}^{-1}.
\end{aligned}$$

Similarly, $M_{L,2}$ is given implicitly as the solution to

$$\begin{aligned}
M_{L,2}^{-1} = \rho_l \hat{\tau}_2^{-1} - \Bigg\{&\left(\nu_s \hat{\tau}_1 \rho^{-1} + \nu_l M_{L,1}\right)\left(\frac{\beta_1(2) - \tilde{B}}{\tilde{B}}\right) \\
&+ \left[(N_D - 1)\left(\nu_s \hat{\tau}_2 \rho^{-1} + (\nu_l + 1)\,M_{L,2}\right) - M_{L,2}\right]\left(\frac{\beta_2(2) - \tilde{B}}{\tilde{B}}\right) + M_{L,0}\left(\frac{\beta_L(2) - \tilde{B}}{\tilde{B}}\right)\Bigg\}^{-1},
\end{aligned}$$

and $M_{L,0}$ as the solution to

$$\begin{aligned}
M_{L,0}^{-1} = \rho_l \hat{\tau}_L^{-1} - \Bigg\{&\left(\nu_s \hat{\tau}_1 \rho^{-1} + \nu_l M_{L,1}\right)\left(\frac{\beta_1(2) - \tilde{B}}{\tilde{B}}\right) \\
&+ \left[(N_D - 1)\left(\nu_s \hat{\tau}_2 \rho^{-1} + (\nu_l + 1)\,M_{L,2}\right)\right]\left(\frac{\beta_2(2) - \tilde{B}}{\tilde{B}}\right)\Bigg\}^{-1}.
\end{aligned}$$

Finally, we use the market clearing condition

$$\begin{aligned}1 &= x + \left(\nu_s \hat{\tau}_1 \rho^{-1} + \nu_l M_{L,1}\right)\left(\left(1-\beta_1(1)-\beta_1(2)\right)\mu + \beta_1(1)\bar{y}_1 + \beta_1(2)y_p - p\right)\\ &+ \sum_{d=2}^{N_D}\left(\nu_s \hat{\tau}_2 \rho^{-1} + (\nu_l+1) M_{L,2}\right)\left(\left(1-\beta_2(1)-\beta_2(2)\right)\mu + \beta_2(1)\bar{y}_d + \beta_2(2)y_p - p\right)\\ &+ M_{L,0}\left(\left(1-\beta_L(1)-\beta_L(2)\right)\mu + \beta_L(1)y_{N_S+1} + \beta_L(2)y_p - p\right).\end{aligned}$$

Equating coefficients, we obtain

$$\begin{aligned}A &= -C\left(1-\left(\nu_s\rho^{-1}\hat{\tau}_1+\nu_l M_{L,1}\right)\left[\left(1-\beta_1(1)-\beta_1(2)\right)\mu - \beta_1(2)\frac{A+C\bar{x}}{\tilde{B}}\right]\right)\\ &\quad + C\left(N_D-1\right)\left(\nu_s\rho^{-1}\hat{\tau}_2+(\nu_l+1)M_{L,2}\right)\left[\left(1-\beta_2(1)-\beta_2(2)\right)\mu - \beta_2(2)\frac{A+C\bar{x}}{\tilde{B}}\right]\\ &\quad + CM_L\left[\left(1-\beta_L(1)-\beta_L(2)\right)\mu - \beta_L(2)\frac{A+C\bar{x}}{\tilde{B}}\right] \qquad (29)\\ B_1 &= C\left(\nu_s\rho^{-1}\hat{\tau}_1+\nu_l M_{L,1}\right)\beta_1(1) \qquad (30)\\ B_2 &= C\left(\nu_s\rho^{-1}\hat{\tau}_2+(\nu_l+1)M_{L,2}\right)\left(N_D-1\right)\beta_2(1) \qquad (31)\\ C &= -\tilde{B}\left[\left(\beta_1(2)-\tilde{B}\right)\left(\nu_s\rho^{-1}\hat{\tau}_1+\nu_l M_{L,1}\right) + \left(N_D-1\right)\left(\beta_2(2)-\tilde{B}\right)\left(\nu_s\rho^{-1}\hat{\tau}_2+(\nu_l+1)M_{L,2}\right)\right.\\ &\quad \left. + M_{L,0}\left(\beta_0(2)-\tilde{B}\right)\right]^{-1} \qquad (32)\\ F &= CM_{L,0}\beta_0(1), \qquad (33)\end{aligned}$$

and $\tilde{B} = B_1 + B_2 + F$. This verifies the price conjecture. Since supply of the asset is one, price and revenue are equal.

Proof of Result 4: A large investor who bids directly and a dealer with minimum bidding requirements. All the conditional expectations and variances are the same as in the previous model (proof of Result 3 holds up until this point). Recall that, when a large investor bids directly, the price impact of dealer 1 is different from that of other dealers since he has one fewer large investors.

Turn now to solving for the price impact of dealer 1. Substituting demand functions into the market clearing constraint, we obtain

$$\begin{aligned}1 &= x + q_{d_1} + \left(\nu_s \hat{\tau}_1 \rho^{-1} + (\nu_l-1) M_{L,1}\right)\left(\left(1-\beta_1(1)-\beta_1(2)\right)\mu + \beta_1(1)\bar{y}_1 + \beta_1(2)y_p - p\right)\\ &+ \sum_{d=2}^{N_D}\left(\nu_s \hat{\tau}_2 \rho^{-1} + \nu_l M_{L,2} + M_{D,2}\right)\left(\left(1-\beta_2(1)-\beta_2(2)\right)\mu + \beta_2(1)\bar{y}_d + \beta_2(2)y_p - p\right)\\ &+ \chi\left(N_D-1\right)M_{D,2}p + M_{L,0}\left(\left(1-\beta_L(1)-\beta_L(2)\right)\mu + \beta_L(1)y_{N_S+1} + \beta_L(2)y_p - p\right),\end{aligned}$$

where the first line is the demand of the small investors (unchanged), the second line is the other dealers (where the minimum bidding penalty χ enters), and the third line is the demand of the large investor (unchanged). Taking the derivative with respect to q_{d_1}, we obtain

$$\begin{aligned}\left(\frac{dp}{dq_{d_1}}\right)^{-1} &= -\left(\nu_s\hat{\tau}_1\rho^{-1}+(\nu_l-1)M_{L,1}\right)\left(\frac{\beta_1(2)-\tilde{B}}{\tilde{B}}\right) - \chi\left(N_D-1\right)M_{D,2}\\ &\quad - \left(N_D-1\right)\left(\nu_s\hat{\tau}_2\rho^{-1}+\nu_l M_{L,2}+M_{D,2}\right)\left(\frac{\beta_2(2)-\tilde{B}}{\tilde{B}}\right) - M_{L,0}\left(\frac{\beta_L(2)-\tilde{B}}{\tilde{B}}\right).\end{aligned}$$

Similarly, the sensitivity of the price to the demand of any other dealer d is given by:

$$\left(\frac{dp}{dq_d}\right)^{-1} = -\left(\nu_s\hat{\tau}_1\rho^{-1} + (\nu_l - 1)\,M_{L,1} + M_{D,1}\right)\left(\frac{\beta_1(2) - \tilde{B}}{\tilde{B}}\right) - \chi\,(N_D - 2)\,M_{D,2} - \chi M_{D,1}$$
$$- \left[(N_D - 1)\left(\nu_s\hat{\tau}_2\rho^{-1} + \nu_l M_{L,2} + M_{D,2}\right) - M_{D,2}\right]\left(\frac{\beta_2(2) - \tilde{B}}{\tilde{B}}\right) - M_{L,0}\left(\frac{\beta_L(2) - \tilde{B}}{\tilde{B}}\right).$$

Thus, the constant $(M_{D,1})$ that maps expected profit into demand for dealer 1 solves

$$M_{D,1}^{-1} = \rho_l\hat{\tau}_1^{-1} - (1 - \chi)\left\{\left(\nu_s\hat{\tau}_1\rho^{-1} + (\nu_l - 1)\,M_{L,1}\right)\left(\frac{\beta_1(2) - \tilde{B}}{\tilde{B}}\right) + \chi\,(N_D - 1)\,M_{D,2}\right.$$
$$\left. + (N_D - 1)\left(\nu_s\hat{\tau}_2\rho^{-1} + \nu_l M_{L,2} + M_{D,2}\right)\left(\frac{\beta_2(2) - \tilde{B}}{\tilde{B}}\right) + M_{L,0}\left(\frac{\beta_L(2) - \tilde{B}}{\tilde{B}}\right)\right\}^{-1},$$

and the constant $(M_{D,2})$ that maps expected profit into demand for any other dealer d solves

$$M_{D,2}^{-1} = \rho_l\hat{\tau}_2^{-1} - (1 - \chi)\left\{\left(\nu_s\hat{\tau}_1\rho^{-1} + (\nu_l - 1)\,M_{L,1} + M_{D,1}\right)\left(\frac{\beta_1(2) - \tilde{B}}{\tilde{B}}\right) + \chi M_{D,1} + \chi\,(N_D - 2)\,M_{D,2}\right.$$
$$\left. + \left[(N_D - 1)\left(\nu_s\hat{\tau}_2\rho^{-1} + \nu_l M_{L,2} + M_{D,2}\right) - M_{D,2}\right]\left(\frac{\beta_2(2) - \tilde{B}}{\tilde{B}}\right) + M_{L,0}\left(\frac{\beta_L(2) - \tilde{B}}{\tilde{B}}\right)\right\}^{-1}.$$

Similarly, in this case, the price sensitivity to the demand of a large investor bidding through dealer 1 is

$$\left(\frac{dp}{dq_{L,1}}\right)^{-1} = -\left(\nu_s\hat{\tau}_1\rho^{-1} + (\nu_l - 2)\,M_{L,1} + M_{D,1}\right)\left(\frac{\beta_1(2) - \tilde{B}}{\tilde{B}}\right) - \chi M_{D,1} - \chi\,(N_D - 1)\,M_{D,2}$$
$$- (N_D - 1)\left(\nu_s\hat{\tau}_2\rho^{-1} + \nu_l M_{L,2} + M_{D,2}\right)\left(\frac{\beta_2(2) - \tilde{B}}{\tilde{B}}\right) - M_{L,0}\left(\frac{\beta_L(2) - \tilde{B}}{\tilde{B}}\right),$$

the price sensitivity to the demand of a large investor bidding through any other dealer d is

$$\left(\frac{dp}{dq_{L,d}}\right)^{-1} = -\left(\nu_s\hat{\tau}_1\rho^{-1} + (\nu_l - 1)\,M_{L,1} + M_{D,1}\right)\left(\frac{\beta_1(2) - \tilde{B}}{\tilde{B}}\right) - \chi M_{D,1} - \chi\,(N_D - 1)\,M_{D,2}$$
$$- \left[(N_D - 1)\left(\nu_s\hat{\tau}_2\rho^{-1} + \nu_l M_{L,2} + M_{D,2}\right) - M_{L,2}\right]\left(\frac{\beta_2(2) - \tilde{B}}{\tilde{B}}\right) - M_{L,0}\left(\frac{\beta_L(2) - \tilde{B}}{\tilde{B}}\right),$$

and the price sensitivity to the demand of a large investor bidding directly is

$$\left(\frac{dp}{dq_{L,0}}\right)^{-1} = -\left(\nu_s\hat{\tau}_1\rho^{-1} + (\nu_l - 1)\,M_{L,1} + M_{D,1}\right)\left(\frac{\beta_1(2) - \tilde{B}}{\tilde{B}}\right) - \chi M_{D,1} - \chi\,(N_D - 1)\,M_{D,2}$$
$$- (N_D - 1)\left(\nu_s\hat{\tau}_2\rho^{-1} + \nu_l M_{L,2} + M_{D,2}\right)\left(\frac{\beta_2(2) - \tilde{B}}{\tilde{B}}\right).$$

Thus, the constant $M_{L,1}$ that maps expected profit into demand for a large investor bidding through dealer 1 solves

$$M_{L,1}^{-1} = \rho_l\hat{\tau}_1^{-1} - \left\{\left(\nu_s\hat{\tau}_1\rho^{-1} + (\nu_l - 2)\,M_{L,1} + M_{D,1}\right)\left(\frac{\beta_1(2) - \tilde{B}}{\tilde{B}}\right) + \chi M_{D,1} + \chi\,(N_D - 1)\,M_{D,2}\right.$$
$$\left. + \,(N_D - 1)\left(\nu_s\hat{\tau}_2\rho^{-1} + \nu_l M_{L,2} + M_{D,2}\right)\left(\frac{\beta_2(2) - \tilde{B}}{\tilde{B}}\right) + M_{L,0}\left(\frac{\beta_L(2) - \tilde{B}}{\tilde{B}}\right)\right\},$$

the constant $M_{L,2}$ that maps expected profit into demand for a large investor bidding through dealer 2 solves

$$M_{L,2}^{-1} = \rho_l\hat{\tau}_2^{-1} - \left\{\left(\nu_s\hat{\tau}_1\rho^{-1} + (\nu_l - 1)\,M_{L,1} + M_{D,1}\right)\left(\frac{\beta_1(2) - \tilde{B}}{\tilde{B}}\right) + \chi M_{D,1} + \chi\,(N_D - 1)\,M_{D,2}\right.$$
$$\left. + \left[(N_D - 1)\left(\nu_s\hat{\tau}_2\rho^{-1} + \nu_l M_{L,2} + M_{D,2}\right) - M_{L,2}\right]\left(\frac{\beta_2(2) - \tilde{B}}{\tilde{B}}\right) + M_{L,0}\left(\frac{\beta_L(2) - \tilde{B}}{\tilde{B}}\right)\right\},$$

and the constant $M_{L,0}$ that maps expected profit into demand for the direct bidder solves

$$M_{L,0}^{-1} = \rho_l \hat{\tau}_l^{-1} - \Bigg\{ \left(\nu_s \hat{\tau}_1 \rho^{-1} + (\nu_l - 1) M_{L,1} + M_{D,1}\right) \left(\frac{\beta_1(2) - \tilde{B}}{\tilde{B}}\right) + \chi M_{D,1} + \chi (N_D - 1) M_{D,2} + (N_D - 1) \left(\nu_s \hat{\tau}_2 \rho^{-1} + \nu_l M_{L,2} + M_{D,2}\right) \left(\frac{\beta_2(2) - \tilde{B}}{\tilde{B}}\right) \Bigg\}.$$

Finally, we use the market clearing condition again to solve for price p

$$\begin{aligned} 1 = x &+ \left(\nu_s \hat{\tau}_1 \rho^{-1} + (\nu_l - 1) M_{L,1} + M_{D,1}\right) \left((1 - \beta_1(1) - \beta_1(2)) \mu + \beta_1(1) \bar{y}_1 + \beta_1(2) \frac{p - A - C\bar{x}}{\tilde{B}} - p \right) \\ &+ \sum_{d=2}^{N_D} \left(\nu_s \hat{\tau}_2 \rho^{-1} + \nu_l M_{L,2} + M_{D,2}\right) \left((1 - \beta_2(1) - \beta_2(2)) \mu + \beta_2(1) \bar{y}_d + \beta_2(2) \frac{p - A - C\bar{x}}{\tilde{B}} - p \right) \\ &+ \chi M_{D,1} p + \chi (N_D - 1) M_{D,2} p + M_{L,0} \left((1 - \beta_L(1) - \beta_L(2)) \mu + \beta_L(1) y_{N_S+1} + \beta_L(2) \frac{p - A - C\bar{x}}{\tilde{B}} - p \right). \end{aligned}$$

Equating coefficients, we obtain the system of equations for the price coefficients

$$\begin{aligned} A = C &\left(-1 + \left(\nu_s \hat{\tau}_1 \rho^{-1} + (\nu_l - 1) M_{L,1} + M_{D,1}\right) \left((1 - \beta_1(1) - \beta_1(2)) \mu - \frac{\beta_1(2)}{\tilde{B}} (A + C\bar{x}) \right) \right) \\ &+ C (N_D - 1) \left(\nu_s \hat{\tau}_2 \rho^{-1} + \nu_l M_{L,2} + M_{D,2}\right) \left((1 - \beta_2(1) - \beta_2(2)) \mu - \frac{\beta_2(2)}{\tilde{B}} (A + C\bar{x}) \right) \\ &+ C M_{L,0} \left((1 - \beta_L(1) - \beta_L(2)) \mu - \frac{\beta_L(2)}{\tilde{B}} (A + C\bar{x}) \right) \end{aligned} \tag{34}$$

$$B_1 = C \left(\nu_s \hat{\tau}_1 \rho^{-1} + (\nu_l - 1) M_{L,1} + M_{D,1}\right) \beta_1(1) \tag{35}$$

$$B_2 = C (N_D - 1) \left(\nu_s \hat{\tau}_2 \rho^{-1} + \nu_l M_{L,2} + M_{D,2}\right) \beta_2(1) \tag{36}$$

$$F = C M_{L,0} \beta_L(1) \tag{37}$$

$$\begin{aligned} C^{-1} = &- M_{L,0} \left(\frac{\beta_L(2) - \tilde{B}}{\tilde{B}}\right) - \left(\nu_s \hat{\tau}_1 \rho^{-1} + (\nu_l - 1) M_{L,1} + M_{D,1}\right) \left(\frac{\beta_1(2) - \tilde{B}}{\tilde{B}}\right) - \chi M_{D,1} \\ &- (N_D - 1) \left(\nu_s \hat{\tau}_2 \rho^{-1} + \nu_l M_{L,2} + M_{D,2}\right) \left(\frac{\beta_2(2) - \tilde{B}}{\tilde{B}}\right) - (N_D - 1) \chi M_{D,2}. \end{aligned} \tag{38}$$

This verifies the price conjecture. Since supply of the asset is one, price and revenue are equal.

Proof of Result 5: Auction price with minimum bidding requirements and a large bidder who bids through a dealer. When the large investor bids indirectly through a dealer, all dealers have a symmetric market impact. In particular, from the market clearing constraint, we have:

$$\begin{aligned} 1 = x &+ \sum_{d=1}^{N_D} \left(\nu_s \hat{\tau} \rho^{-1} + \nu_l M_L + M_D\right) \left((1 - \beta(1) - \beta(2)) \mu + \beta(1) \bar{y}_d + \beta(2) \frac{p - A - C\bar{x}}{B} - p \right) \\ &+ \chi N_D M_D p, \end{aligned}$$

where the last term is the new term that arises because of the minimum bidding penalty χ.

Taking the derivative with respect to q_d, we obtain the price impact of a dealer d:

$$\frac{dp}{dq_d} = - \left[\left(N_D \left(\nu_s \hat{\tau} \rho^{-1} + \nu_l M_L + M_D\right) - M_D \right) \frac{\beta(2) - B}{B} + \chi (N_D - 1) M_D \right]^{-1}.$$

Using the dealer's first order condition, we find the constant (M_D) that maps expected profit into demand solves

$$M_D^{-1} = \rho_l \hat{\tau}^{-1} - (1-\chi)\left[\left(N_D\left(\nu_s\hat{\tau}\rho^{-1} + \nu_l M_L + M_D\right) - M_D\right)\frac{\beta(2) - B}{B} + \chi\left(N_D - 1\right)M_D\right]^{-1}.$$

Similarly, the price impact of any large investor is:

$$\frac{dp}{dq_d} = -\left[\left(N_D\left(\nu_s\hat{\tau}\rho^{-1} + \nu_l M_L + M_D\right) - M_L\right)\frac{\beta(2) - B}{B} + \chi N_D M_D\right]^{-1},$$

and M_L is given implicitly as the solution to

$$M_L^{-1} = \rho_l \hat{\tau}^{-1} - \left[\left(N_D\left(\nu_s\hat{\tau}\rho^{-1} + \nu_l M_L + M_D\right) - M_L\right)\frac{\beta(2) - B}{B} + \chi N_D M_D\right]^{-1}.$$

Finally, solving the market clearing condition for the settle price and equating the coefficients, we obtain the following coefficients

$$A = C\left(-1 + N_D\left(\nu_s\hat{\tau}\rho^{-1} + \nu_l M_L + M_D\right)\left(\left(1 - \beta\left(1\right) - \beta\left(2\right)\right)\mu - \frac{\beta(2)}{B}\left(A + C\bar{x}\right)\right)\right) \tag{39}$$

$$B = CN_D\left(\nu_s\hat{\tau}\rho^{-1} + \nu_l M_L + M_D\right)\beta(1) \tag{40}$$

$$C = -\left[N_D\left(\left(\nu_s\hat{\tau}\rho^{-1} + \nu_l M_L + M_D\right)\left(\frac{\beta(2) - B}{B}\right) + \chi M_D\right)\right]^{-1}. \tag{41}$$

C Analysis of the intermediation choice

Derivation of expected utility Whether the large investor invests directly or through the first intermediary we can write his expected utility conditional on his information set $X_l S$ as

$$-\exp(\rho_L W_L)\exp\big(-\rho_L q_i(\mathbb{E}[f|X_l S] - p) + \frac{1}{2}\rho_L^2 q_i^2 \mathbb{V}[f|X_l S]\big).$$

Substituting in the large investor's optimal portfolio and taking the expectation conditional on y_i gives

$$EU(l) = -\exp(\rho_L W_L)\mathbb{E}\Big\{\exp\Big(\rho_L M_l\Big(\frac{1}{2}\rho_L M_l \mathbb{V}[f|X_l S] - 1\Big)(\mathbb{E}[f|X_l S] - p)^2\Big)\Big|y_i\Big\}. \tag{42}$$

Given y_i, the only random variable in the expectation is $\mathbb{E}[f|X_l S] - p$. This is a sum of two normal random variables, so it is also normally distributed. Using the formula for the expected value of the exponential of a quadratic function of a normal random variable, the formula for expected utility in equations (17)–(20) follows from equation (42).

Effect of intermediation on $EU(l)$ through μ_{rl} First note that $\theta_l > 0$. This follows from equation (19) because $M_l > 0$ and

$$\rho_L M_L \mathbb{V}[f|X_L S] = \frac{\rho_L \mathbb{V}[f|X_L S]}{\rho_L \mathbb{V}[f|X_L S] + \frac{dp}{dq_L}} < 1.$$

Therefore, from equation (17), $EU(l)$ is strictly increasing in μ_{rl}.

When the large investor invests through an intermediary it causes μ_{rl} to decrease through two channels. This decrease expected utility. The first channel is that it causes the price to increase. This happens through the constant in the price function A.

$$\begin{aligned}A = &- C\Big(1-(\nu_s\rho^{-1}\hat{\tau}_1+\nu_l M_L)\Big[(1-\beta(1)-\beta(2))\mu-\beta(2)\frac{A+C\bar{x}}{B}\Big]\Big)\\&+C(N_D-1)(\nu_s\rho^{-1}\hat{\tau}_1+(\nu_l+1)M_L)\Big[(1-\beta(1)-\beta(2))\mu-\beta(2)\frac{A+C\bar{x}}{B}\Big]\\&+CM_L\Big[(1-\beta(1)-\beta(2))\mu-\beta(2)\frac{A+C\bar{x}}{B}\Big].\end{aligned}$$

The constant A will differ in the two cases because: 1) The coefficients in the price function differ—C differs for the two cases and B and $\tilde{B}$ differ; 2) The weights that the agents place on their prior beliefs and signals differ—β, β_1, β_2 and β_L; and 3) The posterior precisions and the stategic considerations of the agents differ—M_L, $M_{L,0}$, $M_{L,1}$ and $M_{L,2}$.

We are interested in the effect of changes in the precisions of the agents' posterior beliefs on A. To assess this affect we'll consider how changes in these precisions cause the M's to change and ignore the effects of other variables on A. To help with this note two features of the expressions for A. First, it should be the case that $C > 0$ because the price should increase in x (the demand of the agents who are not sensitive to the price). Second, if $\mu > 0$ and sufficiently large, then all the terms in square brackets in both expression will be strictly positive. These terms are the posterior expected payoff of the asset for investors when $\bar{s}_d = 0$ for all d and $p = 0$.

When these conditions on C and μ are satisfied, an increase in the posterior precision of all agents beliefs will cause A to increase. This is because M_L is increasing in the precision of the large investor's posterior beliefs and when precisions of the dealers' posterior beliefs increase it will cause $M_L > \max\{M_{L,0}, M_{L,1}, M_{L,2}\}$. An increase in A causes μ_{rL} to decrease.

The second channel is that the large investor's information has more influence on the price when he invests through a dealer and this reduces the ex-ante expected profit μ_{rl}. The expected profit in the two cases can be written as

$$\begin{aligned}\mathbb{E}[f|X_{Ld}]-p =\ & (1-\beta_L(1)-\beta_L(2))\mu+\beta_L(1)y_{N_S+1}+\beta_L(2)y_p\\&-\Big[A+B_1\bar{y}_1+\frac{B_2}{N_D-1}\sum_{d=2}^{N_D}\bar{y}_d+Cx+Fy_i\Big],\\\mathbb{E}[f|X_{Li}]-p =\ & (1-\beta(1)-\beta(2))\mu+\beta(1)\bar{y}_1+\beta(2)y_p-\Big[A+\frac{B}{N_D}\sum_{d=1}^{N_D}\bar{y}_d+Cx\Big].\end{aligned}$$

In the first equation the effect of the large investor's private signal, y_{N_S+1}, on the price depends on $\beta_L(1)-F$ and in the second case the effect of the agent's private signal, $\bar{y}_1$, on the price depends on $\beta(1)-\frac{B}{N_D}$. We show numerically that $\beta_L(1)-F > \beta(1)-\frac{B}{N_D}$, which shows that the large investor gains more from her private signal when she invests directly.

$$\begin{aligned}\mu_{rLi} &= \Big(\frac{\tau_f\mu+\tau_{\varepsilon l}y_i}{\tau_f+\tau_{\varepsilon,l}}\Big)-A-B\Big[\omega_s\nu_s+\omega_l\Big(\frac{N_D\nu_l-1}{N_D}\Big)+\omega_l\Big]\Big(\frac{\tau_f\mu+\tau_{\varepsilon l}y_i}{\tau_f+\tau_{\varepsilon,l}}\Big)\\&\qquad -BN_D^{-1}\omega_l y_i-C\bar{x},\\\mu_{rLd} &= \Big(\frac{\tau_f\mu+\tau_{\varepsilon l}y_i}{\tau_f+\tau_{\varepsilon,l}}\Big)-A-\Big[B_1(\omega_{s,2}\nu_s+\omega_{l,2}\nu_l)+B_2(\omega_s\nu_s+\omega_l(\nu_l+1))\Big]\Big(\frac{\tau_f\mu+\tau_{\varepsilon l}y_i}{\tau_f+\tau_{\varepsilon,l}}\Big)\\&\qquad -Fy_i-C\bar{x}.\end{aligned}$$

Effect of intermediation on $EU(l)$ through $\Delta\mathbb{V}_l$ When the large investor invests through the intermediary he receives the signal $\bar{y}_1$ in addition to y_i and y_p. This decreases the variance of his posterior beliefs, $\mathbb{V}[f|X_lS]$.[19] To see that θ_l is strictly decreasing in $\mathbb{V}[f|X_lS]$ it helps to write θ_l in the

[19]This statement is true holding fixed the information contained in p. When the large investor invests through a dealer the price will be a less precise signal of the asset payoff, which will offset the decrease in

following form:

$$\theta_l = \left(\frac{\rho_L}{\rho_L \mathbb{V}[f|X_l S] + \frac{dp}{dq_L}}\right)\left(1 - \frac{1}{2}\left[\frac{1}{1 + (\frac{dp}{dq_L})/(\rho_L \mathbb{V}[f|X_l S])}\right]\right).$$

Since $\theta_l > 0$ and the terms in both sets of large round parentheses are strictly decreasing in $\mathbb{V}[f|X_l S]$, θ_l is strictly decreasing in $\mathbb{V}[f|X_l S]$. We can see from equation (17) that $EU(l)$ is strictly increasing in θ_l. Therefore $EU(l)$ is strictly decreasing in $\mathbb{V}[f|X_l S]$.

To compute $\Delta\mathbb{V}_l$ we need to compute $\mathbb{V}[f-p|y_{N_S+1}]$, given by equation (20) in these notes (we have already discussed computing $\mathbb{V}[f|X_l S]$). The general formula for this is

$$\mathbb{V}\left[f-p|y_{N_S+1}\right] = \mathbb{V}\left(f-p\right) - \mathbf{Cov}\left(f-p, y_{N_S+1}\right)^2 \mathbb{V}\left(y_{N_S+1}\right)^{-1},$$

with $\mathbb{V}\left(y_{N_S+1}\right) = \tau_f^{-1} + \tau_{el}^{-1}$.

When the large investor invests through a dealer, we can represent the price as

$$p = A + Bf + C\bar{x} + B\Pi_{p,I}Z,$$

where $\Pi_{p,I}$ is the last row of the Π matrix in the main text. Then:

$$\begin{aligned}
\mathbb{V}\left(f-p\right) &= \mathbb{V}\left(\left(1-B\right)f - B\Pi_{p,I}Z\right) = \left(1-B\right)^2 \tau_f^{-1} + B^2\Pi_{p,I}\mathbb{V}\left(Z\right)\Pi_{p,I}' \\
\mathbf{Cov}\left(f-p, y_{N_S+1}\right) &= \mathbf{Cov}\left(\left(1-B\right)f - B\Pi_{p,I}Z, f + \epsilon_{N_S+1}\right) = \left(1-B\right)\tau_f^{-1} - \frac{B}{N_D}\omega_l\tau_{el}^{-1}.
\end{aligned}$$

When the large investor invests directly, we instead have

$$p = A + \tilde{B}f + C\bar{x} + \tilde{B}\Pi_{p,d}Z,$$

where $\tilde{B} = B_1 + B_2 + F$. Then:

$$\begin{aligned}
\mathbb{V}\left(f-p\right) &= \mathbb{V}\left(\left(1-\tilde{B}\right)f - \tilde{B}\Pi_{p,d}Z\right) = \left(1-\tilde{B}\right)^2 \tau_f^{-1} + \tilde{B}^2\Pi_{p,d}\mathbb{V}\left(Z\right)\Pi_{p,d}' \\
\mathbf{Cov}\left(f-p, y_{N_S+1}\right) &= \mathbf{Cov}\left(\left(1-\tilde{B}\right)f - \tilde{B}\Pi_{p,d}Z, f + \epsilon_{N_S+1}\right) = \left(1-\tilde{B}\right)\tau_f^{-1} - F\tau_{el}^{-1}.
\end{aligned}$$

$\mathbb{V}[f|X_l S]$. We conjecture that this effect is not large enough to fully offset the decrease.

Table 3: Descriptive statistics for the calibrated, simulated model with direct and indirect bidding and low-bid penalty. Revenue is in basis points, and allocations are in percent.

Panel A: Full Sample

	Revenue	Dealer allocation	Direct allocation	Indirect allocation
Mean	37.5957	42.0538	0.5868	50.8466
Std. Dev.	69.8128	20.1859	0.7675	57.0472
Skew	-0.0150	0.0072	0.9225	0.1887
Kurtosis	3.0101	3.0185	2.5776	4.6919

Panel B: Direct Bidding

	Revenue	Dealer allocation	Direct allocation	Indirect allocation
Mean	52.7388	41.3514	1.3302	45.0985
Std. Dev.	66.7869	20.3264	0.5887	21.0972
Skew	0.0019	0.0033	0.0870	0.0011
Kurtosis	3.0634	3.0590	3.0824	3.0663

Panel C: Indirect Bidding

	Revenue	Dealer allocation	Direct allocation	Indirect allocation
Mean	25.6443	42.6081	0.0000	55.3831
Std. Dev.	69.8299	20.0570	0.0000	73.6539
Skew	0.0126	0.0126	-	-0.0126
Kurtosis	2.9827	2.9827	-	2.9827

Panel D: High signal

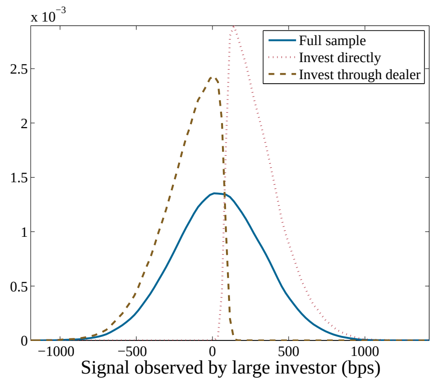
	Revenue	Dealer allocation	Direct allocation	Indirect allocation
Mean	56.6878	42.5339	1.4342	43.8340
Std. Dev.	66.6329	20.2831	0.5781	21.0393
Skew	0.0091	0.0089	0.0739	-0.0081
Kurtosis	3.0774	3.0716	3.1070	3.0837

Panel E: Low signal

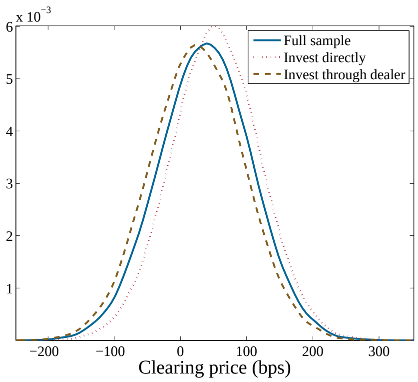
	Revenue	Dealer allocation	Direct allocation	Indirect allocation
Mean	17.3851	40.2359	0.0000	64.0946
Std. Dev.	69.5897	19.9880	0.0000	73.4005
Skew	0.0250	0.0250	-	-0.0250
Kurtosis	2.9751	2.9751	-	2.9751

Figure 8: Distribution of simulated outcomes

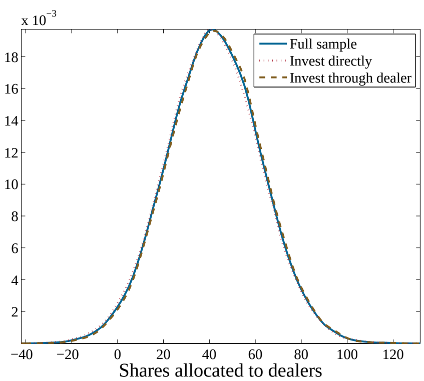
(a) Signals observed by large investor



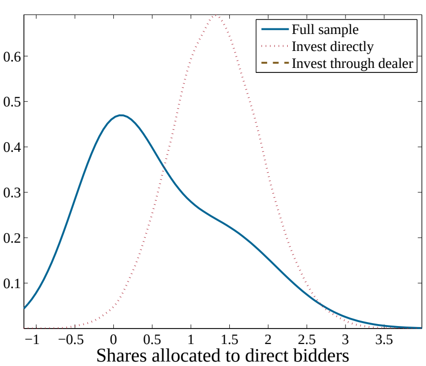
(b) Auction clearing price



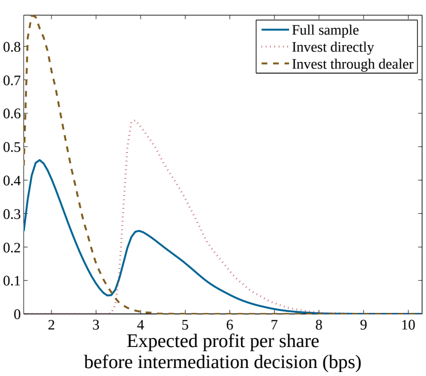
(c) Shares allocated to dealers



(d) Shares allocated to direct bidders



(e) Beliefs before intermediation decision



(f) Beliefs after intermediation decision