FINANCE RESEARCH SEMINAR
SUPPORTED BY UNIGESTION

“Will Intermediate for Data: Information Sharing in U.S. Treasury Auctions”

Nina BOYARCHENKO
Federal Reserve Bank of New York

Abstract

According to most financial intermediation theories, the prevalence of primary dealers is puzzling. They do not transform liquidity or maturity; 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 model of intermediaries as advisors. Intermediaries observe each client’s order flow, aggregate that information across clients, and use it to advise their clients as a group. Although information-sharing is often maligned, it typically increases treasury revenue, but also makes revenue more variable. When bidders can choose whether or not to bid through a dealer, a new financial accelerator emerges. 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.

Friday, October 2, 2015, 10:30-12:00
Room 126, Extranef building at the University of Lausanne
Will Intermediate for Data:
Information Sharing in 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

August 18, 2015

Abstract

According to most financial intermediation theories, the prevalence of primary dealers is puzzling. They do not transform liquidity or maturity; 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 model of intermediaries as advisors. Intermediaries observe each client’s order flow, aggregate that information across clients, and use it to advise their clients as a group. Although information-sharing is often maligned, it typically increases treasury revenue, but also makes revenue more variable. When bidders can choose whether or not to bid through a dealer, a new financial accelerator emerges. 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 Bruno Biais, Ken Garbade, Luis Gonzalez, Jeremy Greenwood, Pablo Kurlat, Richard Tang, Haoxiang Zhu and participants at the 2014 U.S. Treasury Roundtable on Treasury Markets, NYU, the Macro-finance society meetings, SED, and Gerzensee summer institute for comments. Nic Kozeniauskas, Karen Shen and Peifan Wu provided excellent research assistance. Emails: nina.boyarchenko@ny.frb.org; david.lucca@ny.frb.org; lveldkam@stern.nyu.edu.
Hundreds of trillion of dollars are traded each day in global fixed income, currency and commodities (FICC) markets. In recent years, centralized exchanges have become increasingly more important in clearing these activities, but most of FICC trading still takes place in relatively opaque and fragmented over-the-counter (OTC) markets where financial intermediaries act as principals for their customers and trade on their own account. FICC trading is also very informationally intensive leading to high levels of concentration due to economies of informational scale and the ensuing barriers of entry. Partly owing to this concentration, FICC trading has been extremely profitable for the largest financial institutions and accounted for significant shares of their revenues.1 Given the economic importance of FICC markets in allocating financial resources, a key question is whether such a seemingly archaic intermediated trading structure appropriately serves client investors and issuers alike. This question has come to the forefront of the policy discussion in the wake of a number of misconducts in FICC markets ranging from collusive manipulations of key benchmarks, such as the London Interbank Offered Rate (Libor), to the misuse of confidential information. These misconducts have not only led to record fines and reputational damage to many large financial institutions, but have also called for a more widespread use of transparent, anonymized, all-to-all electronic trading venues, and have called into question the role of financial institutions more generally (for example, Zingales, 2015, in his presidential address).2 In this paper we study this question in the context of US Treasury auctions and show how financial intermediaries can benefit customers and issuers over a decentralized market allocation even in the presence of concentration and imperfect sharing of order flow information with clients.

We focus on Treasury auctions for three reasons. First, they are among the most important FICC venues, with X trillion of issues auctioned every year, and are also central in funding US federal debt. Second, the role of intermediaries in Treasury auctions are one of the most challenging to understand, as auction 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 issuers because they cannot influence fiscal policy. Despite these limitations, and even though investors can bid directly through an electronic system, most client bids are placed through intermediaries. Finally, we access detailed auction and pricing data to calibrate the model and study welfare implications of

---

1For example, as of X FICC trading has accounted for Y share of total revenues for GS, ...

intermediation for clients and auction revenues.

To understand the prevalence of intermediation in Treasury auctions, we present a new theory of intermediation. In our theory financial intermediaries collect information from order flow, use it to advise clients, and bid for their in-house account. This advising practice is referred providing “market color” in the market place. In the model we parametrize the extent to which intermediaries share order information with their clients ranging from an ideal perfect-information world to no sharing at all.  

We first study the benchmark of perfect information sharing between intermediaries and their clients. 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. 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.

We next study how these results change as information is only imperfectly shared to clients. The welfare implications in the case of imperfect information sharing are particularly important as a number of policy initiatives and change in market practices are under way. For example, some firms have introduced physical separation of certain functions to help control

---

3Sovereign 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.

4In 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 flow of sensitive information and manage potential conflicts of interest. The Financial Stability Board (FSB) report on FX benchmarks recommended that banks establish and enforce internal systems and controls to address potential conflicts of interest arising from managing customer flow. While these initiatives may be beneficial in reducing possible abuses, these change in practices also destroy the value of information aggregation.

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.

We consider Treasury auctions because of their importance, data availability 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.\textsuperscript{5} Primary dealers bid on behalf of

\textsuperscript{5}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
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 ??? 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.\(^6\)

---

\(^6\)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.
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 A.3 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 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.7

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

---

7See 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.
use of monopsony power both increases revenue and 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 ??). 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.

**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). Second, we assume that bids condition on realized prices, which is the rational response to the winner’s curse. 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

“Bankers have often shared broad guidance, both internally and to clients, on whether demand is slack or strong before auctions.” (Bloomberg, June 24, 2015)

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

---

8Economic conditions affect bidding behavior in non-U.S. markets as well. Using Finnish Treasury data, Keloharju et al. (2005) find strong uncertainty effects.
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 types 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. We relax this assumption in Section A.3.

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 \( N_L \) (\( N_I \)) the number of large (small) investors, and by \( N_D \) the total number of dealers, when varying \( N_D \) or \( N_L \), the total \( N_D + N_L \) remains unchanged.\(^{10} \) 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 \( N \equiv N_L + N_I + N_D \). 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 \( \rho_i \) denotes agent \( i \)'s coefficient of absolute risk aversion.\(^{11} \) 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:

\[ s_i = f + \varepsilon_i, \]

\(^{9} \)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.

\(^{11} \)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.
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 \( \rho_i \) 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 \( i \in \{N_L, N_D\} \), where with \( N_j \) we denote the set of agents of type \( j \). Similarly, all small investors are symmetric: \( \rho_i = \rho_I \) and \( \tau_{\varepsilon,i} = \tau_{\varepsilon,s} \) for \( i \in N_I \).

**Describing Information Sets and Updating Beliefs with Correlated Signals**

Investors and dealers observe three types of information. They observe their private signals \( s_i \). In addition, primary dealers, who place bids on behalf of their customers, observe their investors’ bids and advise their clients as a whole by giving a signal about the average order to all of their customers. Each dealer \( d \) receives orders from an equal number of investors. He observes the orders of \( N_I/N_D \equiv \nu_I \) small and \( N_L/N_D \equiv \nu_l \) large investors. Since bids will turn out to be linear functions of beliefs, observing average bids and observing average signals is equivalent. Thus, when a group of investors bid through a primary dealer \( d \), the dealer can construct \( \bar{s}_d \), an optimal signal-precision-weighted average of his and his clients’ private signals:

\[
\bar{s}_d = \frac{\tau_{\varepsilon,s} \sum_{k \in \mathcal{I}_d} s_k + \tau_{\varepsilon,l} \left( \sum_{j \in \mathcal{I}_d} s_j + s_d \right)}{\nu_I \tau_{\varepsilon,s} + (1 + \nu_l) \tau_{\varepsilon,l}}.
\]  

(1)

The dealers transmit some of this information back to their clients in the form of a noisy signal \( \bar{s}_d + \xi_d \) where \( \xi_d \sim N(0, \tau_{\xi}^{-1}) \) is the noise in the dealers’ advice, which is iid across dealers \( d \). Our model can capture noisy dealer advice, as well as two extreme cases: perfect information-sharing between dealers and clients (\( \tau_{\xi} = \infty \)) and no information-sharing (\( \tau_{\xi} = 0 \)).

The final piece of information that all agents observe 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.

Let \( s(p) \) denote the unbiased signal constructed from auction-clearing price. We guess and verify that \( (s(p) - f) \sim N(0, \tau_p^{-1}) \), where \( \tau_p \) is a measure of the informativeness of the auction-clearing settle price.

Thus the vector of signals observed by an investor \( j \) assigned to intermediary \( k \) is \( S_j = [s_j, \bar{s}_{\xi d(j)}, s(p)] \), where \( d(j) \) denotes the dealer that investor \( j \) is assigned to. This is the
same vector for large and small investors. The only difference is that large investors’ private signals \( s_j \) are more precise and have a different (higher) covariance with price information. A dealer observes the same signals, except that he sees the exact order flows, instead of a noisy signal of them. For dealer \( d \), \( S_d = [s_d, \bar{s}_d, s(p)] \). 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. To update beliefs, we use the following optimal linear projection formulas\(^{12}\)

\[
\mathbb{E}[f|S_j] = (1 - \beta'1_m)\mu + \beta' S_j \quad \text{where}
\]

\[
\beta_j \equiv \mathbb{V}(X_jS)^{-1} \text{Cov}(f, S_j) \tag{3}
\]

\[
\mathbb{V}[f|S_j] = \mathbb{V}(f) - \text{Cov}(f, S_j)' \mathbb{V}(S_j)^{-1} \text{Cov}(f, S_j) \equiv \hat{\tau}_j^{-1} \tag{4}
\]

where \( m \) is the number of signals in the vector \( S_j \), the covariance vector is \( \text{Cov}(f, S_j) = 1_m \tau_f^{-1} \) and the signal variance-covariance \( \mathbb{V}(S_j) \), is worked out in the appendix. The vector \( \beta_j = [\beta_{s_j}, \beta_{\xi_j}, \beta_{p_j}] \) dictates how much weight an agent puts on his signals \( [s_j, \xi_j, s(p)] \) in his posterior expectation. In a Kalman filtering problem, \( \beta \) is like the Kalman gain. In an econometric forecasting problem, these are the OLS coefficients that multiply the independent variables to forecast the dependent variable – in this case, the payoff \( f \).

The different versions of the model we explore in the following section will change two things about this information structure: The set of signals \( S_j \) that an agent \( j \) observes and covariance between those signals.

**Equilibrium.** A Nash equilibrium is

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

\[
\max_{q_i(p)} \mathbb{E}[- \exp(-\rho W_i)|S_i]
\]

\[
\text{s.t. } W_i = W_{0i} + q_i(f - p).
\]

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)

\(^{12}\)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 \( \alpha = (1_N - \beta)'1_{NH} \). \( \beta \) 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.
that maximizes

\[
\max_{q_j,p} \mathbb{E}[-\exp(-\rho_l W_j)|S_j] \tag{5}
\]

subject to

\[
W_j = W_{0,D} + q_j(f - p), \tag{6}
\]

\[
x + \sum_{i=1}^{N_I} q_i + \sum_{j=1}^{N_D} q_j + \sum_{k=1}^{N_L} q_k = 1. \tag{7}
\]

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: \( x + \sum_{i=1}^{N_I} q_i + \sum_{j=1}^{N_D} q_j + \sum_{k=1}^{N_L} q_k = 1. \)

### 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:

\[
q_i(p) = \frac{1}{\rho} \mathbb{V}[f|S_i]^{-1} (\mathbb{E}[f|S_i] - p). \tag{8}
\]

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 \( \max_{q_j,p} q_j(\mathbb{E}[f|S_j] - p) - \frac{1}{2} \rho q_j^2 \mathbb{V}[f|S_j] \) s.t. the market clearing condition (7). Taking the first order condition with respect to \( q_j \), we obtain

\[
q_j(p) = \frac{\mathbb{E}[f|S_j] - p}{\rho \mathbb{V}[f|S_j] + dp/dq_j}. \tag{9}
\]

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 $s_i$ and the price information $s(p)$, but not any signal from the dealer: $S_i = [s_i, s(p)]$.

**Result 1.** Without primary dealers, the auction revenue is a linear function of signals $s_i$, $s_j$ and the supply noise $x$:

$$p = A + B_I N_I^{-1} \sum_{i=1}^{N} s_i + B_L N_L^{-1} \sum_{j=1}^{N_L} s_j + C (0) x$$

where the coefficients $A, B_I, B_L$ and $C$ are reported in the appendix.

We use this 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 $s(p) = (p - A - C \bar{x})/(B_I + B_L)$. Using (2) and (4), we find that the posterior belief is a weighted average of the prior belief $\mu$, the private signal $s_i$ and the price signal. We substitute these beliefs into the optimal bid formulas (8) and (22), and use those bid functions to determine the price that clears the market. This result and all future results are proven in Appendix A.

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) with a finite number of bidders, the average of investors’ signals $\bar{s}$ is not equal to the true payoff. Both sources of extra complexity are essential in the following model to understand how the number of dealers and their information sharing affects auction revenue.

1.3 Auction Revenue with $N_D$ Dealers

When we add dealers, they inject new noise in the prices from the noise in their “advice” signal to clients. But they also leave all investors better informed. 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 causes information to be less aggregated because each dealer-generated signal $s_{d\xi}$ contains the average of a smaller number of private investor signals. These effects will show up in the numerical results on auction revenue, which are computed using the following result:

**Result 2.** When dealers share information imperfectly with clients and all investors bid through dealers, the auction revenue is a linear function of signals $s_i$, non-competitive
trades \( x \), and dealer signal noise \( \xi_d \):

\[
p = A + B N_{I}^{-1} \sum_{i=1}^{N_I} s_i + B L N_{L}^{-1} \sum_{j=N_I+1}^{N_I+N_L} s_j + B D N_{D}^{-1} \sum_{d=N_I+N_L+1}^{N} s_d + C x + D \sum_{d=1}^{N_d} \xi_d
\]

where the coefficients \( A, B, F \) and \( C \) are reported in the appendix.

On the other hand, one might think that a larger number of dealers increases competitive pressures and limits rent extraction. But 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 \( N_D \). Thus, this assumption isolates the information aggregation effect. More importantly, fixing the number of large players captures the idea that agents who become primary dealers were probably already trading in treasury markets. It is unlikely that increasing the number of dealers would attract more money to treasury markets and mechanically increase the demand for treasuries. Fixing the number of large players rules out this mechanical 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 vary with information sharing and compare to other primary dealer arrangements. The unconditional expected revenue will be \( A + B \mu + C \bar{x} \) and unconditional revenue variance will be \( B^2 \nu[s] + C^2 \tau_x^{-1} \).

## 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.\(^{13}\) An auction begins with an announcement one day to one week ahead of the...
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. 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.

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 1 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 or other brokers or dealers, but primarily they bid through primary dealers. The Federal

14Foreign and international monetary authorities (FIMA) that have accounts a the NY Fed can place bids up to $100 million per account and $1 billion in total.

15The 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.

16Brokers and dealers include all institutions registered according to Section 15C(a)(1) of the Securities
Note: Non-competitive bids in the chart exclude indirect FIMA bids. Source: Treasury auction results.

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.\textsuperscript{17} 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 using a simple imputation.\textsuperscript{18} We reclassify these bids as part of the noise trader group, as

\textsuperscript{17}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.

\textsuperscript{18}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
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. 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 model to their empirical counterparts: the pricing coefficients $A$, $B$ and $C$ in equation $??$, the mean and variance of the price of the interest-rate-risk-neutral portfolio $p$ at auction,

(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.

19 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.

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

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

Table 2: Calibrated parameters

<table>
<thead>
<tr>
<th>( \mu )</th>
<th>( \tau_f^{-\frac{1}{2}} )</th>
<th>( \tau_{\varepsilon,s}^{-\frac{1}{2}} )</th>
<th>( \tau_{\varepsilon,d}^{-\frac{1}{2}} )</th>
<th>( \tilde{x} )</th>
<th>( \rho )</th>
<th>( \rho_L )</th>
<th>( \chi )</th>
<th>( N_S )</th>
<th>( N_L )</th>
<th>( N_D )</th>
</tr>
</thead>
<tbody>
<tr>
<td>40.8</td>
<td>73.5</td>
<td>692.1</td>
<td>282.4</td>
<td>0.06</td>
<td>0.12</td>
<td>64542</td>
<td>908</td>
<td>0.07</td>
<td>240</td>
<td>40</td>
</tr>
</tbody>
</table>

Note: \( \mu \), \( \tau_f^{-\frac{1}{2}} \), \( \tau_{\varepsilon,s}^{-\frac{1}{2}} \) and \( \tau_{\varepsilon,d}^{-\frac{1}{2}} \) are all expressed in basis points.

The final parameter, \( \tau_\xi \) regulates how much information dealers share with clients. Without micro data, we cannot infer this value. In reality, different dealers probably engage in different degrees of information-sharing. Instead, we show results from the whole spectrum of potential values, from \( \tau_\xi = 0 \) (no information sharing) to \( \tau_\xi = \infty \) (perfect information sharing).
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 ?? 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 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 ($\tau^{-1}_f$) 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 ?? 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.

The effect of dealers also varies in the variance of non-competitive bids. When these bids are less predictable, auction clearing prices are less 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 ?? shows that an auction with a dealer delivers lots more revenue than one without when fundamental uncertainty \( \tau^{-1} \) 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 ??). 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 $\tau_f$ 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 ?? 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.

3.3 Test 1: The Share of Informed Traders Forecasts Profits.

An essential feature of the model is that a subset of agents have private information about the future resale value of treasuries. Before proceeding with more nuanced versions of the model, it is useful to stop and check whether this basic premise is supported by the data. The hallmark of informed trading is that the trades forecast future profits. An agent who has no information, beyond that incorporated in equilibrium prices, cannot systematically buy more when profits ($f - p$) are high and sell when they are low. Without information about the difference between $f$ and $p$, such an investment strategy is not a measurable strategy. Thus, $\text{cov}(q_i, f - p) > 0$ is evidence of informed trading.

The data counterpart to profit $f - p$ is post-auction appreciation. That’s the difference between the resale price of the asset in the secondary market, minus the price paid at auction. Since we assumed that all competitive bidders have private information and all non-competitive bidders are uninformed, the data counterpart to $q_i$ of an informed trader is the share of the auction awarded to competitive bidders. We use share of the auction, rather than face value because the size of auctions varies and this introduces noise in our regression and, if auction size is correlated with asset characteristics, could potentially bias our estimates.

**Corollary 1.** High competitive share predicts high post-auction appreciation: $\partial E[f - $
Using our auction data, we test prediction 1. Table 3.3 reports estimates of a regression of the price appreciation of the hedged portfolio from the time of the auction close to the issue date, or $f - p$. As shown in the first column of the table, the value of the newly issued security appreciates on average by about 3 basis points between the auction date and the issue date, consistent with the findings of Lou et al. (2013). As shown in the second column, this appreciation is higher the higher the share of competitive bids into the auction (column 2) with a highly statistically significant coefficient. This empirical result is robust to the inclusion of month and tenor fixed effects. This effect is consistent with non-competitive being noise traders and competitive bidders being informed. It is not a mechanical result from high demand. When informed traders demand is high, the price is lower on average, relative to the payoff. It is that low price relative to fundamental value that induces informed investors to buy more. We run analogous regressions using data generated from the model and find similar results. This result does not prove that primary dealers aggregate this information. But it does support the notion that the dealers’ clients have some private information to be aggregated.

### 4 Mixed Auctions: Choosing Direct or Indirect Bidding

“The primary dealers’ customers, for their part, channel orders to them in part as a thank you for providing services such as research, people in the market have long said. The open secret part of the equation is that some people at these banks can see orders flowing in.” (Bloomberg, June 24, 2015)
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. Since intermediation choice makes the model significantly more complex, we simplify along another dimension to keep the problem tractable. From here on, we consider the special case of full information-sharing by dealers by setting $\tau_\xi = \infty$.

### 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 $s(p)$ 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 $\nu_I$ investors’, the first $\nu_I - 1$ large investors’ and the first dealer’s signals:

$$\bar{s}_1 = \frac{\tau_{e,s} \sum_{k \in I} s_k + \tau_{e,l} \left( \sum_{j \in I} s_j + s_d \right)}{\nu_I \tau_{e,s} + \nu_l \tau_{e,l}}$$

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

**Solution: Auction Outcomes** If the large investor bids indirectly, the problem and the solution are the same as in the previous section. With direct bidding, the auction price (revenue) is a linear function of the dealer-level average individual investor signals, $\bar{s}_d$, the signal of the large direct bidder, $s_L$, and of market orders $x$.

**Result 3.** With $N_D$ dealers and 1 large investor who bids directly, the auction revenue is $p = A + B_d \bar{s}_1 + B_{d\neq1}/(N_D - 1) \sum_{d=2}^{N_D} \bar{s}_d + C (N_D, 1) x + B_L s_L$, where 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 \( l \in \{L_d, L_i\} \). 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 \( s_i \). 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

\[
EU(l) = - \exp(\rho L W_L (1 + 2 \theta_l \Delta V_l)) \cdot \frac{1}{2} \exp \left( - \frac{\mu_{rl}^2}{\theta_l^{-1} + 2 \Delta V_l} \right). \tag{10}
\]

The intermediation decision affects utility in three ways: through the expected profit per unit \( \mu_{rl} \), the sensitivity of demand to expected profit \( \theta_l \), and through the ex-ante variance of expected profit \( \Delta V_l \). These three terms are:

\[
\mu_{rl} \equiv \mathbb{E}\{\mathbb{E}[f|X_lS] - p|s_i]\}, \tag{11}
\]

\[
\theta_l \equiv \rho_L \left[ \rho_L \mathbb{V}[f|S_L] + \frac{dp/dq_L}{\rho_L} \right]^{-1} \left( 1 - \frac{1}{2} \rho_L \left[ \rho_L \mathbb{V}[f|S_L] + \frac{dp/dq_L}{\rho_L} \right] \mathbb{V}[f|S_L] \right), \tag{12}
\]

\[
\Delta V_l \equiv \mathbb{V}\{\mathbb{E}[f|X_lS] - p|s_i\} = \mathbb{V}[f - p|s_i] - \mathbb{V}[f|X_lS]. \tag{13}
\]

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 (10)–(13).

The main cost of intermediation is that it reveals one’s private information \( s_i \) to others. This effect shows up as a reduction in \( \mu_{rl} \), 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 \( d1 \): \( E[f|X_lS] = E[f|X_{d1}S] \). Information sharing reduces \( \mu_{rl} \) 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 \( E[f|X_lS] - p \) is closer to zero with intermediation. Second, improving the precision of other investors’ information raises the expected price \( p \), which in turn, lowers \( \mu_{rl} \) (see eq (11)). In equation (10), a decrease in \( \mu_{rl} \) decreases expected utility because \( \theta_l > 0 \) (see appendix).

\[\text{See appendix for derivations of the following three equations and support for the analysis that follows.}\]
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 $\theta_l$. In the appendix we show that $\theta_l > 0$ and is strictly decreasing in the posterior variance of the asset payoff $\mathbb{V}[f|X_lS]$. When the large investor trades through a dealer he receives more information, which decreases $\mathbb{V}[f|X_lS]$. This ensures that $\theta_{Ld} > \theta_{Li}$. Equation (10) shows that expected utility is increasing in $\theta_l$. Thus, the reduction in uncertainty effect embodied in $\theta_l$ 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 $\mathbb{V}[f|X_lS]$ is lower. From equation (13) 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 (10), 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 (10) 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, $\mu_{rl}$. When $\mu_{rl} \neq 0$ the increase in the ex-ante variance $\Delta \mathbb{V}_l$ increases the probability that the expected profit $\mathbb{E}[f|X_lS] - p$ 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 $\mu_{rl}$ is closer to zero. We use these three effects to understand the intermediation choice results below.

4.3 Quantitative Results: Intermediation Choice

Figure 2 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 $\mu_{rl}$ in (11)) depends on the large investor’s private signal $s_i$. To see the effect of $s_i$, first note that on average, $p$ will usually fall below $\mathbb{E}[f|X_lS]$, reflecting
Note: Bidding directly is optimal in the blue region. “Very low signal”: \( s_{N_{i+1}} = \mu - 2(\tau_f + \tau_{e,l})^{-\frac{1}{2}} \); “low signal”: \( s_{N_{i+1}} = \mu - (\tau_f + \tau_{e,l})^{-\frac{1}{2}} \); “average signal”: \( s_{N_{i+1}} = \mu \); “high signal”: \( s_{N_{i+1}} = \mu + (\tau_f + \tau_{e,l})^{-\frac{1}{2}} \); “very high signal”: \( s_{N_{i+1}} = \mu + 2(\tau_f + \tau_{e,l})^{-\frac{1}{2}} \).

a risk premium in the return on the risky asset. If the large investor observes a signal \( s_i \) that is below the prior mean \( \mu \), it lowers the ex-anted expectation \( \mathbb{E}\{E[f|X_lS] - p|s_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 4.) 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 \( \theta_l \) in (10)) is invariant in the investor’s private signal \( s_i \). Since a low private signal reduces the cost of intermediation \( \mu_{rl} \) and leaves the benefit \( \theta_l \) unchanged, low private signals make intermediation more attractive.

Figure 2 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 (\( \theta_l \) 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 (\( \mu_{rl} \) 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 (\( \mu_{rl} \)) 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.

4.5 Testable Prediction 2: Positive Skewness in Post-Auction Appreciation

The financial accelerator effect shows up as a distribution of auction-clearing prices that has unconditional negative skewness. Negative skewness in the price $p$ translates into positive skewness in the post-auction appreciation $f - p$. The first panel of Table 4 reveals that the unconditional skewness in $p$ is $-0.015$. This translates into positive skewness in appreciation of $0.013$. 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. The skewness in the model is a consistent feature of our simulations with intermediation choice. But it is small. The reason is that we only consider one bidder choosing to bid directly or indirectly. The one bidder’s choice introduces skewness. But amidst the many other agents in the model, the amount of skewness that one agent can
introduce is limited. If we could compute a model with many agents making such an intermediation choice, this effect would likely be amplified.

The skewness prediction is consistent with the empirical distribution of post-auction appreciation \( f - p \), which presents a strong positive skew of about 3.7, as evidenced by the histogram in Figure 3.

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

\[
\max_{q_D, p} \mathbb{E}[-\exp(-\rho W_D) | X_D S] \tag{14}
\]

\[
\text{s.t. } W_D = W_{0,D} + q_D (f - p) - \chi_0 + \chi q_D p, \tag{15}
\]

and subject to the market-clearing condition (7). 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: \( \max_{q_D, p} q_D (\mathbb{E}[f | X_D S] - p(1 - \chi)) - \frac{1}{2} \rho q_D^2 \mathbb{V}[f | X_D S] \), s.t. the market clearing constraint (7). Taking the first order condition with respect to \( q_D \), we
obtain

\[
q_D (p) = \frac{\mathbb{E}[f | X_D S] - p(1 - \chi)}{\rho \mathbb{V}[f | X_D S] + (1 - \chi) dp/dq_D} \tag{16}
\]

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 \( \chi \) 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 minimum bidding penalties and 1 large investor who bids
directly, the auction revenue is \( p = A + B_1 \bar{s}_1 + B_2 / (N_D - 1) \sum_{d=2}^{N_D} \bar{s}_d + Cx + F s_{N_I + 1} \), where
the coefficients are given in the Appendix.

**Result 5.** In an auction with minimum bidding penalties where all investors bid indirectly
through dealers, the auction revenue is \( p = A + B / N_D \sum_{d=1}^{N_D} \bar{s}_d + Cx \), where \( A, B \) and \( C \)
are reported in the Appendix.

**Quantitative Results: Effects of Minimum Bidding Penalties** An increase in
the penalty \( \chi \) for low bidding does raise the expected revenues (Figure 4, 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 4 (right scale)
shows that more stringent bidding requirements increase the volatility of auction revenue.
Figure 4: Auction with a Low-Bid Penalty.

Figure 5: Intermediation Decision with a Low-Bid Penalty.

(a) Direct bidding decision

(b) Auction Revenue Mean

(c) Auction Revenue St. Dev.

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

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

Variance rises because the greater quantity demanded by dealers makes 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 5 shows that for sufficiently high costs $\chi$, 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


A Appendix: Derivations and Proofs

A.1 Auction without Dealers

In this model, each type of limit order investor submits bids on their own behalf rather than through intermediaries. The investors’ information set contains only their private signal $s_i$ and the price information $s(p)$, but not the average signal that was previously provided by the intermediary. Thus, the vector of signals observed by each investor $i$ is now $X_i^C S = [s_i, s(p)]$.

Given the price formula in Result 1, an investor’s unbiased price signal is

$$s(p) = \frac{p - A - C\bar{x}}{B} = \frac{B_I}{B} N_I^{-1} \sum_{i=1}^{N_I} s_i + \frac{B_L}{B} N_L^{-1} \sum_{j=1}^{N_L} s_j + \frac{C}{B} (x - \bar{x})$$

where $\bar{B} \equiv B_i + B_L$. That $s(p)$ is an unbiased signal about $f$. However, the price signal and the private signals have correlated signal errors. The $\beta$ coefficients are formed using the optimal linear projection formulas correct for this covariance.

Let $Z$ be the vector of orthogonal shocks.

$$Z = [\varepsilon_1 \ldots \varepsilon_N \ x]'.$$(17)

where we order agents putting small, individual investors first, then large investors, then dealers. The variance matrix of this vector is

$$\mathbb{V}[\tilde{Z}] = \text{diag}([\tau_{s,-1} 1_{N_I}, \tau_{l,-1} 1_{N_L + N_D}, \tau_{x,-1}]).$$

Then, we can represent the vector of all $m$ signals observed by investor $i$ as

$$S_i = 1_m f + \Pi_i Z,$$

where $1_m$ is an $m \times 1$ vector of ones. The loading matrix $\Pi_i$ of signals on orthogonal shocks is

$$\Pi_i = \left[ \begin{array}{cc} (B_I/\bar{B}) 1_{N_I} & \tau_i \\ (B_L/\bar{B}) 1_{N_L + N_D} & C/\bar{B} \end{array} \right]$$

(19)

where $\tau_i$ is a vector of zeros with a 1 in the $i$th position, $N$ is the total number of agents ($N = N_I + N_L + N_D$, all exogenous), and $B_I$, $B_L$ and $C$ are equilibrium pricing coefficients (endogenous).

With this $\Pi_i$ matrix, we can express the Bayesian updating weights $\beta$ that small and large investors $s, l$ put on their private signal and the price signal. With the $\beta$’s, we can then easily compute posterior precision $\tilde{\tau}$. The variance-covariance matrix of signals is

$$\mathbb{V}(S_j) = \tau_f^{-1} + \Pi_j \mathbb{V}[Z] \Pi_j'$$

(20)

The $2 \times 1$ vector of Bayesian updating weights $\beta_j = [\beta_{sj}, \beta_{pj}]'$ is

$$\beta_j = \mathbb{V}(S_j)^{-1} 1_m \tau_f^{-1}.$$
Let $\beta_I = [\beta_{sI}, \beta_{pI}]$ be the vector of weights used by the small, individual investors and $\beta_L = [\beta_{sL}, \beta_{pL}]$ be the vector of weights used by the large investors, so that the posterior beliefs of a small investor are

$$E[f|S_i] = (1 - \beta_{sI} - \beta_{pI}) \mu + \beta_{sI}s_i + \beta_{pI}\frac{p - A - C\tilde{x}}{B}$$

with conditional variance $V[f|S_i] = \tau^{-1}_f - \tau^{-1}_I (\beta_{sI} + \beta_{pI}) \equiv \hat{\tau}^{-1}_I$. The beliefs of a large investor are

$$E[f|S_i] = (1 - \beta_{sL} - \beta_{pL}) \mu + \beta_{sL}s_i + \beta_{pL}\frac{p - A - C\tilde{x}}{B}$$

with conditional variance $V[f|S_i] = \tau^{-1}_f - \tau^{-1}_L (\beta_{sL} + \beta_{pL}) \equiv \hat{\tau}^{-1}_L$.

**Proof of Result 1: Price in the no-dealer auction.** The small investors’ demand functions (bids) in the competitive market model are given by equation (8) and the large investors’ bids by equation (22). In this and future results, it is convenient to re-express the large investors’ and dealers’ demand as

$$q_j(p) = M_L (E[f|X_jS] - p)$$

where

$$M_l \equiv [\rho_l V[f|S_l] + dp/dq_l]^{-1} \quad \text{for any large player } l$$

where $M_l$ describes the sensitivity of $l$’s treasury demand to changes in expected profit.

Substituting these bids into the market clearing condition $x + \sum_{i=1}^{N} q_i = 1$ yields

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

Substituting for $E[f|X_i, S]$, we obtain

$$1 = x + \rho^{-1}\hat{\tau}_I \sum_{i=1}^{N_l} \left((1 - \beta_{sI} - \beta_{pI}) \mu + \beta_{sI}s_i + \beta_{pI}\frac{p - A - C\tilde{x}}{B} - p\right) + M_L \sum_{j=1}^{N_L-1} \left((1 - \beta_{sL} - \beta_{pL}) \mu + \beta_{sL}s_i + \beta_{pL}\frac{p - A - C\tilde{x}}{B} - p\right) + q_L.$$ 

Taking the derivative with respect to $q_L$, we obtain

$$0 = \left(N_l\hat{\tau}_L\rho^{-1}\frac{\beta_{pL} - \tilde{B}}{B} + (N_L - 1) M_L \beta_{pL} - \tilde{B} \right) \frac{dp}{dq_L} + 1,$$

so that $M_L$ (defined as in (23)) is given implicitly by

$$M_L^{-1} = \rho_L\hat{\tau}_L^{-1} - \left(N_l\hat{\tau}_L\rho^{-1}\frac{\beta_{pL} - \tilde{B}}{B} + (N_L - 1) M_L \beta_{pL} - \tilde{B}\right)^{-1}.$$ 

(24)
Solving

\[ 1 = x + \rho^{-1} \tau_1 \sum_{i=1}^{N_I} \left( (1 - \beta_{sI} - \beta_{pI}) \mu + \beta_{sI}s_i + \beta_{pI} \frac{p - A - C\bar{x}}{B} - p \right) \]

\[ + M_L \sum_{j=1}^{N_L} \left( (1 - \beta_{sL} - \beta_{pL}) \mu + \beta_{sL}s_i + \beta_{pL} \frac{p - A - C\bar{x}}{B} - p \right) \]

for \( p \) and matching coefficients yields the implicit solution

\[ A = -C \left( 1 + \left( N_I \frac{\tilde{\tau}_I}{\rho} \beta_{pI} + N_L M_L \beta_{pL} \right) \frac{A + C\bar{x}}{B} \right) \]

(25)

\[ + C \left( N_I \frac{\tilde{\tau}_I}{\rho} (1 - \beta_{sI} - \beta_{pI}) + N_L M_L (1 - \beta_{sL} - \beta_{pL}) \right) \mu \] (26)

\[ B_I = C N_I \rho^{-1} \tilde{\tau}_I \beta_{sI} \]

(27)

\[ B_L = C N_L M_L \beta_{sL} \]

(28)

\[ C = -\tilde{B} \left[ N_I \rho^{-1} \tilde{\tau}_I \left( \beta_{pI} - \tilde{B} \right) + N_L M_L \left( \beta_{pL} - \tilde{B} \right) \right]^{-1}, \]

(29)

and the elasticity of the large investor’s demand to information is given by (24). 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.

The solution to this problem is a joint solution to (7) - (11) and (13)- (18).

A.2 Auction with Dealers and Imperfect Information-Sharing

When dealers reveal a noisy signal of average order flow \( \tilde{s}_d + \xi_d \), their clients’ private signal becomes relevant again because it is not contaminated with the noise \( \xi_d \). Therefore, for investor \( i \), \( S_i = [s_i, \tilde{s}_d + \xi_i, s(p)] \).

These three signals for investor \( j \) load on the following orthogonal shocks

\[ Z = [\varepsilon_1 \ldots \varepsilon_N \xi_1 \ldots \xi_N x]' \] (30)

with loadings \( \Pi_i \) for any non-dealer investor \( i \):

\[ \Pi_i = \begin{bmatrix} \iota_i' & 0 & 0 \\ \iota_{d(i)} \otimes 1_{N/N_D} N_D/N & \iota_i & 0 \\ B/(B + F)1_{N_I} & F/(B + F)1_{N_L + N_D} & C/B \end{bmatrix} \]

(31)

where \( \iota_i \) is an \( N \times 1 \) vector of zeros with a 1 in the \( i \)th position, \( \iota_{d(i)} \) is an \( N_D \times 1 \) vector of zeros with a 1 in the \( d(i) \)th position, and \( d(i) \) is the dealer of investor \( i \).

The variance matrix of the \( Z \) shocks is

\[ V[Z] = \text{diag}([\tau_{\varepsilon_1}^{-1} N_I, \tau_{\varepsilon_d}^{-1} 1_{N_I}, \tau_{\xi_1}^{-1} 1_{N_L + N_D}, \tau_{x}^{-1}]) \].

(32)

Given this \( \Pi \) matrix and \( V[Z] \), we can construct \( V(S_i) = \tau_f^{-1} + \Pi_i V[Z] \Pi_i' \).
Given the signal variance-covariance matrix for dealers and investors, the Bayesian updating weights are given by (21). For non-dealers, there are three Bayesian updating weights \( \beta = [\beta_s, \beta_s, \beta_p]' \). Using (2) and (4), we find posterior beliefs

\[
\mathbb{E}[f|S_j] = (1 - \beta_{sj} - \beta_{pj}) \mu + \beta_{sj} \bar{s}_d(j) + \beta_{pj} \frac{p - A - C\bar{x}}{B},
\]

(33)

with conditional variance \( V[f|S_j] = \tau_f^{-1} - \tau_f^{-1}(\beta_{sj} + \beta_{sj} + \beta_{pj}) \equiv \hat{\tau}_j^{-1} \).

The dealers in this setting are slightly different because they are the one set of agents who observe their order flow noiselessly. For dealers, there are only two relevant signals \( S_d[\bar{s}_d, s(p)] \) where

\[
\Pi_d = \left[ \begin{array}{cc}
\tau_d \otimes 1_{N_D} & 0 \\
B/(B + F)1_{N_I} & F/(B + F)1_{N_I + N_D} \\
C/B & C/B
\end{array} \right]
\]

(34)

are the loadings on the same \( Z \) shock vector as for investors. Thus, for dealers, \( V(S_d) = \tau_f^{-1} + \Pi_d V[Z]\Pi_d' \), which allows us to compute updating weights \( \beta_d \) and posterior means and variances. For dealers, since there are only two relevant signals, there are only two Bayesian weights: \( \beta_D = [\beta_{sd}, \beta_{pd}]' \). Given these weights, conditional expectations and variances are computed with the same formula as above.

**Proof of Result 2** For large investors and dealers, let the sensitivity of asset demand to changes in the expected profit per share be denoted \( M_L = V[f|S_L] + dp/dq_L \) and \( M_D = V[f|S_D] + dp/dq_D \).

The asset demand for each small investors is (8) and each large investor is (22). We equate total demand to supply (1) \( \int q_i di + x = 1 \) and substitute in for all demands, except for that of one large investor \( q_L \) to get

\[
x + \rho^{-1}\hat{\tau}_I \sum_{i=1}^{N_I} (\mathbb{E}[f|S_i] - p) + M_L \sum_{j=1}^{N_L-1} (\mathbb{E}[f|S_j] - p) + q_L + M_D \sum_{d=1}^{N_D} (\mathbb{E}[f|S_d] - p) = 1.
\]

Then, we substitute in for the conditional expectations and use the implicit function theorem to calculate \( dp/dq_L \), which we need to determine \( M_L \). Substituting for the expectations yields

\[
1 = x + \frac{\hat{\tau}_I}{\rho} \sum_{i=1}^{N_I} \left( (1 - \beta_{si} - \beta_{pi}) \mu + \beta_{si} \bar{s}_i + \beta_{si}(\bar{s}_d(i) + \xi_d(i)) + \beta_{pi} \frac{p - A - C\bar{x}}{B} - p \right) + M_L \sum_{j=1}^{N_L-1} \left( (1 - \beta_{sj} - \beta_{pj}) \mu + \beta_{sj} \bar{s}_j + \beta_{sj}(\bar{s}_d(j) + \xi_d(j)) + \beta_{pj} \frac{p - A - C\bar{x}}{B} - p \right) + q_L + M_D \sum_{d=1}^{N_D} \left( (1 - \beta_{sd} - \beta_{pd}) \mu + \beta_{sd} \bar{s}_d + \beta_{pd} \frac{p - A - C\bar{x}}{B} - p \right).
\]
Taking the derivative with respect to $q_L$, we obtain

$$0 = \left( N_I \tau_I \rho^{-1} \frac{\beta_{pi} - \hat{B}}{B} + (N_L - 1) M_L \frac{\beta_{pL} - \hat{B}}{B} + N_D M_D \frac{\beta_{pd} - \hat{B}}{B} \right) \frac{dp}{dq_L} + 1,$$

so that we can solve for $dp/dq_L$ and express $M_L$ implicitly as

$$M_L^{-1} = \rho_L \tilde{\tau}_L^{-1} - \left( N_I \tau_I \rho^{-1} \frac{\beta_{pi} - \hat{B}}{B} + (N_L - 1) M_L \frac{\beta_{pL} - \hat{B}}{B} + N_D M_D \frac{\beta_{pd} - \hat{B}}{B} \right)^{-1}. \quad (35)$$

To determine dealers’ demand $q_d = M_D (\mathbb{E}[f | S_d] - p)$, we need to solve for $M_D$ in the same fashion. The solution is identical, except that the dealers’ signal precision is higher and therefore, his posterior variance $\tilde{\tau}^{-1}$ is lower. Following the same steps, we find that

$$M_D^{-1} = \rho_L \tilde{\tau}_D^{-1} - \left( N_I \tau_I \rho^{-1} \frac{\beta_{pi} - \hat{B}}{B} + N_L M_L \frac{\beta_{pL} - \hat{B}}{B} + (N_D - 1) M_D \frac{\beta_{pd} - \hat{B}}{B} \right)^{-1}. \quad (36)$$

Finally, we can express the price coefficients as functions of these $M$’s. Do to this, we substitute all the investors’ demands in the market-clearing condition (same expression as before, without the $q_L$ term and summing from $j = 1$ to $N_L$, instead of $N_L - 1$). Coefficient matching yields the following price coefficients:

$$A = C \left[ N_I \tau_I \rho^{-1} (1 - \beta_{sI} - \beta_{sI} - \beta_{pI}) \mu + N_L M_L (1 - \beta_{sL} - \beta_{sL} - \beta_{pL}) \mu \right. + \left. N_D M_D (1 - \beta_{sD} - \beta_{sD} - \beta_{pD}) \mu - \hat{B}^{-1} (A + C \bar{x}) (\beta_{pI} + \beta_{pL} + \beta_{pD}) - 1 \right] \quad (37)$$

$$B_I = C \left( \tau_I \rho^{-1} \beta_{sI} + M_D \beta_{sd} \omega_I \right) + D \omega_I \quad (38)$$

$$B_L = C \left( M_L \beta_{sL} + M_D \beta_{sd} \omega_L \right) + D \omega_L \quad (39)$$

$$B_D = C M_D \beta_{sd} \omega_D + D \omega_D \quad (40)$$

$$C = - \left( N_I \tau_I \rho^{-1} \frac{\beta_{pI} - \hat{B}}{B} + N_L M_L \frac{\beta_{pL} - \hat{B}}{B} + N_D M_D \frac{\beta_{pd} - \hat{B}}{B} \right)^{-1} \quad (41)$$

$$D = C \left( \tau_I \rho^{-1} \frac{N_I}{N_D} \beta_{sI} + M_L \frac{N_L}{N_D} \beta_{sL} \right) \quad (42)$$

The solution to this model is the set of $\beta$’s for each type of agent, price coefficients, and $M$ coefficients that jointly solve (21), (37)-(42), (35) and (36).

**Proof of Corollary 1: Noncompetitive share reduces expected profit.** We can break $\partial E[f - p]/\partial x$ into the difference $\partial E[f]/\partial x - \partial E[p]/\partial x$. The first term, $\partial E[f]/\partial x = 0$, by assumption. Both $f$ and $x$ are assumed to be exogenous, independent variables. Using ??, the second term is $\partial E[A + B f + C x]/\partial x = C$. Finally, because $C > 0$, $\partial E[f - p]/\partial x > 0$ and the result holds.
A.3 Direct Bidding and Minimum Bidding Penalties

We solve the model generally with a minimum bidding penalty. Then we can set $\chi = 0$ to get the results for the no-penalty model.

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

$$\bar{s}_1 = \frac{\tau_{\epsilon,s} \sum_{k \in I^s} s_k + \tau_{\epsilon,l} \left( \sum_{j \in I^l} s_j + s_d \right)}{\nu_s \tau_{\epsilon,s} + \nu_l \tau_{\epsilon,l}}$$

As in the previous model, investor $i$ who bids through intermediary $d$ observes signals $\tilde{S}_i = [s_i, \bar{s}_d, s(p)]$. Since we are considering the case with perfect dealer information sharing, the investor’s own signal is redundant because it is included in the dealer’s average signal $\bar{s}_d$. Thus, for notational convenience, we drop the coefficient on $s_i$ and continue as if $\tilde{S}_i = [\bar{s}_d(i), s(p)]$, $\forall i \neq L$. The large investor bidding directly observes only his own signal and the price information: $\tilde{S}_L = [s_L, s(p)]$.

**Variance-covariance matrix of signals** The only difference in the signal construction is that we assume that the first large bidder does not trade through the large dealer. Thus, the large bidder’s signals are only his own private signal and the price. The first dealer’s information (and his clients’) are less precise because they miss the one large investor. And finally, the price information has a slightly different composition.

We can construct a condensed $\Pi$ matrix as before. Let

$$Z^{direct} = \left[ \begin{array}{cccc} \epsilon_{L1} & \bar{\epsilon}_1 & \bar{\epsilon}_2 & \ldots & \bar{\epsilon}_{Nd} & (x - \bar{x}) \end{array} \right]'$$

where $\epsilon_t = S_L - f$ and $\bar{\epsilon}_d = \bar{s}_d - f$ are the noise the the signals of the large investors and the dealer and $x$ is the non-competitive bids. Note that the information from the intermediation decision of the direct bidder in incorporated in the first dealer’s signal, and thus in $\bar{\epsilon}_1$. The variances of these shocks are:

$$\forall \left[ Z^{direct} \right] = diag((\tau_{L}^{-1}, (\tau_d - \tau_L)^{-1}, \tau_d^{-1}1_{N_d-1}, \tau_x^{-1})).$$  \hfill (43)

Then, we can express signals as the true payoff $f$ plus the following loadings on each of the orthogonal $Z$ shocks:

$$\Pi_L = \left[ \begin{array}{cccc} 1 & 0 & \ldots & 0 \\ B_L/\bar{B} & B_{d1}/\bar{B} & \ldots & C/\bar{B} \end{array} \right]$$  \hfill (44)

$$\Pi_{D1} = \left[ \begin{array}{cccc} 0 & 1 & \ldots & 0 \\ B_L/\bar{B} & B_{d1}/\bar{B} & \ldots & C/\bar{B} \end{array} \right]$$  \hfill (45)

$$\Pi_{d \neq 1} = \left[ \begin{array}{cccc} 1 & 0 & \ldots & 0 \\ B_L/\bar{B} & B_{d1}/\bar{B} & \ldots & C/\bar{B} \end{array} \right]$$  \hfill (46)
where \( \tilde{B} \equiv B_{d1} + B_{d \neq 1} + B_L \) and \( d \neq 1 \) stands for all dealers, besides dealer 1.

Then, for all agents \( j \) who are clients of dealer \( d(j) \), or for the large investor \( d(j) = L \), the signal variance-covariance is

\[
\forall(S_j) = \tau_j^{-1} + \Pi_d(j)\mathbb{P}(Z_{\text{direct}}^d)\Pi_d(j)' \quad \forall j.
\]

(47)

**Bayesian updating weights \( \beta \) and posterior precisions \( \tau \)** For each agent, there are two signals: (1) the signal \( s_i \) provided by the dealer or the large investor’s own signal, in the case of direct bidding; and (2) the signal conveyed by the price \( (p - A - Cx)/B \). The posterior expectation of the asset value is a linear combination of the prior and each of these two signals. \( \beta_i \) and \( \beta_{pi} \) are the weights the agent places on the two signals. According to Bayes’ law, the \( \beta \)'s are the precision of the signal, divided by the sum of the precisions of the prior and signals. But in this case, the signals are correlated. We need to use optimal linear projection formulas to adjust for this correlation. The \( 2 \times 1 \) vector \( \beta = [\beta_s, \beta_p]' \) is given by (21). The conditional expectation is then (2). The posterior precision is (4), which can always be expressed as a simple function of the Bayesian updating weights \( \beta \):

\[
\hat{\tau}_l^{-1} = \tau_l^{-1}(1 - \beta_s - \beta_p)
\]

(48)

where \( \tau_l \) is the prior variance of the asset payoff, and is exogenous.

**Proof of Results 3 and 5: Price when the large investor bids directly (with and without minimum bidding).** In order to solve for equilibrium price, we need to determine asset demands of the large investors. That in turn depends on their price impact. To determine price impact \( dp/dq \) for an agent \( j \), we need to know what the equilibrium price function looks like. The following steps derive a set of equations that jointly solve this fixed point problem.

Substitute demand functions of all agents, except one (dealer 1 in this case) into the market clearing constraint:

\[
1 = x + q_{d1} + (\nu_1 \hat{\tau}_{d1}^{-1} \mu + (\nu_1 - 1) M_{L(d=1)}) (1 - \beta_{s,d1} - \beta_{p,d1}) \mu + \beta_{s,d1} \hat{s}_1 + \beta_{p,d1} s(p) - p
\]

\[
+ \sum_{d=2}^{ND} \left( \nu_d \hat{\tau}_{d \neq 1}^{-1} + \nu_M M(d \neq 1) + M \right) (1 - \beta_{s,d \neq 1} - \beta_{p,d \neq 1}) \mu + \beta_{s,d \neq 1} \hat{s}_d + \beta_{p,d \neq 1} s(p) - p
\]

\[
+ \chi (N_D - 1) M_{d \neq 1} p + M_L (1 - \beta_{sL} - \beta_{pL}) \mu + \beta_{sL} s_{N_L + 1} + \beta_{pL} s(p) - p,
\]

where \( M \)'s are defined as in (23). The first line equates the supply (1) to the demands of non-competitive investors (\( x \)), dealer 1 (\( q_{d1} \)) and the demand of all the investors who bid through dealer 1, the second line and first term of the third line is the demand of other dealers and their investors, and the rest of the third line is the large direct bidder.

For each type of agent \( j \), the variable \( M_j \) is the sensitivity of \( j \)'s treasury demand to \( j \)'s expected profit. For small investor \( i \), this sensitivity is \( \hat{\tau}_i \rho^{-1} \) and is simply expressed as such in the formula above. For large investors and dealers, this object complicated by their market power. For each large, non-dealer investor \( l \), the first-order condition tells us that this sensitivity is given by (23):

\[
M_l \equiv \rho \hat{\tau}_l dp/dq_l^{-1}.
\]

The price sensitivity to demand...
\(dp/dq\) depends on whether the large investor is a client of dealer 1 \(dp/dq_{L(d=1)}\), a client of some other dealer \(dp/dq_{L(d\neq1)}\), or the direct bidder \(dp/dq_{L}\). Taking the market-clearing condition with all demand substituted in, except \(q_L\), and differentiating with respect to \(q_L\) and rearranging, we obtain the sensitivity of price to the demand of the large direct bidder:

\[
\left(\frac{dp}{dq_L}\right)^{-1} = -\left(\nu_I\hat{\tau}_1\rho^{-1} + (\nu_l - 1) M_{L(d=1)} + M_{d=1}\right) \left(\frac{\beta_{p,d1} - \tilde{B}}{B}\right) - \chi M_{d=1} - \chi (N_D - 1) M_{d\neq1} - \left(\nu_I\hat{\tau}_2\rho^{-1} + \nu_l M_{L(d\neq1)} + M_{d\neq1}\right) \left(\frac{\beta_{p,d\neq1} - \tilde{B}}{B}\right). \tag{49}
\]

Following the same steps, we obtain the sensitivity of price to the demand of the large investors who bid through dealer 1 (the dealer who lost the business of the large bidder)

\[
\left(\frac{dp}{dq_{L(d=1)}}\right)^{-1} = \left(\frac{dp}{dq_L}\right)^{-1} - M_L \left(\frac{\beta_{pL} - \tilde{B}}{B}\right) + M_{L(d=1)} \left(\frac{\beta_{p,d1} - \tilde{B}}{B}\right), \tag{50}
\]

and the sensitivity of price to the demand of the large investors who bid through other dealers

\[
\left(\frac{dp}{dq_{L(d\neq1)}}\right)^{-1} = \left(\frac{dp}{dq_L}\right)^{-1} - M_L \left(\frac{\beta_{pL} - \tilde{B}}{B}\right) + M_{L(d\neq1)} \left(\frac{\beta_{p,d\neq1} - \tilde{B}}{B}\right). \tag{51}
\]

For each dealer \(d\), the first-order condition tells us that the sensitivity of demand to expected profit differs from that of the other large investors because of the minimum bidding penalty \(\chi\):

\[
M_d \equiv [\nu_I\hat{\tau}_1 + (1 - \chi)dp/dq_d]^{-1}. \tag{52}
\]

The price sensitivity to demand \(dp/dq_d\) depends on whether this is dealer 1 (the one who lost the business of 1 large investor) or another dealer. Differentiating the market-clearing condition, we find that the price impact \(dp/dq_{d1}\) of dealer 1 is

\[
\left(\frac{dp}{dq_{d1}}\right)^{-1} = \left(\frac{dp}{dq_L}\right)^{-1} - M_L \left(\frac{\beta_{pL} - \tilde{B}}{B}\right) + M_{d=1} \left(\chi + \left(\frac{\beta_{p,d1} - \tilde{B}}{B}\right)\right), \tag{53}
\]

and the sensitivity of the price to the demand of any other dealer \(d \neq 1\) is:

\[
\left(\frac{dp}{dq_{d\neq1}}\right)^{-1} = \left(\frac{dp}{dq_L}\right)^{-1} - M_L \left(\frac{\beta_{pL} - \tilde{B}}{B}\right) + M_{d\neq1} \left(\chi + \left(\frac{\beta_{p,d\neq1} - \tilde{B}}{B}\right)\right). \tag{54}
\]

Finally, we substitute the demands of all agents back into the market clearing condition,
use our price guess to write the price signal as \( s(p) = (p - A - C \bar{x})/\bar{B} \), to obtain

\[
1 = x + (\nu_1 \hat{\tau}_1 \rho^{-1} + (\nu_l - 1) M_{L(d=1)} + M_{d=1}) \left( (1 - \beta_{s, d1} - \beta_{p, d1}) \mu + \beta_{s, d1} \bar{s}_1 + \beta_{p, d1} \frac{p - A - C \bar{x}}{B} - p \right)
\]

\[
+ \sum_{d=2}^{N_D} (\nu_1 \hat{\tau}_2 \rho^{-1} + \nu_l M_{L(d \neq 1)} + M_{d \neq 1}) \left( (1 - \beta_{s, d \neq 1} - \beta_{p, d \neq 1}) \mu + \beta_{s, d \neq 1} \bar{s}_d + \beta_{p, d \neq 1} \frac{p - A - C \bar{x}}{B} - p \right)
\]

\[
+ \chi M_{d=1} + \chi (N_D - 1) M_{d \neq 1} + M \left( (1 - \beta_{s, L} - \beta_{p, L}) \mu + \beta_{s, L} \bar{s}_{N_1 + 1} + \beta_{p, L} \frac{p - A - C \bar{x}}{B} - p \right).
\]

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

\[
A = C \left( 1 + (\nu_1 \hat{\tau}_1 \rho^{-1} + (\nu_l - 1) M_{L(d=1)} + M_{d=1}) \left( (1 - \beta_{s, d1} - \beta_{p, d1}) \mu - \beta_{p, d1} \frac{A + C \bar{x}}{B} \right) \right)
\]

\[
+ C (N_D - 1) \left( (\nu_1 \hat{\tau}_2 \rho^{-1} + \nu_l M_{L(d \neq 1)} + M_{d \neq 1}) \left( (1 - \beta_{s, d \neq 1} - \beta_{p, d \neq 1}) \mu - \beta_{p, d \neq 1} \frac{A + C \bar{x}}{B} \right) \right)
\]

\[
+ CM \left( (1 - \beta_{s, L} - \beta_{p, L}) \mu - \beta_{p, L} \frac{A + C \bar{x}}{B} \right)
\]

\[
B_1 = C \left( (\nu_1 \hat{\tau}_1 \rho^{-1} + (\nu_l - 1) M_{L(d=1)} + M_{d=1}) \beta_{s, d1} \right)
\]

\[
B_2 = C \left( (N_D - 1) \left( (\nu_1 \hat{\tau}_2 \rho^{-1} + \nu_l M_{L(d \neq 1)} + M_{d \neq 1}) \beta_{s, d \neq 1} \right) \right)
\]

\[
F = CM \beta_{s, L}(1)
\]

\[
C^{-1} = -M \left( \beta_{p, L} - \frac{\bar{B}}{\bar{B}} \right) \left( (\nu_1 \hat{\tau}_1 \rho^{-1} + (\nu_l - 1) M_{L(d=1)} + M_{d=1}) \left( \beta_{p, d1} - \frac{\bar{B}}{\bar{B}} \right) - \chi M_{d=1} \right)
\]

\[
- (N_D - 1) \left( (\nu_1 \hat{\tau}_2 \rho^{-1} + \nu_l M_{L(d \neq 1)} + M_{d \neq 1}) \left( \beta_{p, d \neq 1} - \frac{\bar{B}}{\bar{B}} \right) - (N_D - 1) \chi M_{d \neq 1}. \right)
\]

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

The solution to this model is a joint solution to the price coefficient equations (55) - (59), the M equation (23), in conjunction with the price sensitivities (49) - (54) and the \( \beta \)'s ('), which in turn depend on the \( \Pi \)'s (44), (45) and (46) and the variances of the orthogonal shocks (43). Setting \( \chi = 0 \) yields the solution to the model without the minimum bidding requirement.

A.4 Indirect Bidding with Perfect Information Sharing and Bidding Penalties.

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 his dealer’s beliefs. We begin with the preliminaries of belief updating.

Define the weight the dealer’s signal puts on each small investor’s signal as \( \omega_s \equiv \tau_{e, s} / (\nu_s \tau_{e, s} + (1 + \nu_l) \tau_{e, l}) \), where \( \tau \)'s are exogenous signal precisions. \( \nu_s \equiv N_L/N_D \) is small investors per dealer and \( \nu_l \equiv N_{l}/N_D \) is large investors per dealer (both exogenous). Then \( s_{d} \) is a signal-precision-weighted average of all of the dealer’s customer signals as well as the dealer’s own
signal:
\[ \bar{s}_d = \frac{\tau_{s,s} \sum_{k \in \mathcal{I}_d} s_k + \tau_{s,l} (\sum_{j \in \mathcal{I}_d} s_j + s_d)}{\nu_s \tau_{s,s} + (1 + \nu_l) \tau_{s,l}} \equiv \omega_s \sum_{k \in \mathcal{I}_d} s_k + \omega_l \sum_{k \in \mathcal{I}_d} s_k + \omega_l s_d. \]

For a given agent who trades through dealer \( d \), there are only 2 relevant signals: the dealer's signal \( s_d = f + \bar{\epsilon}_d \) where \( \bar{\epsilon}_d \sim N(0, \tau_d^{-1}) \) and \( \tau_d = \nu_s \tau_s + (\nu_l + 1) \tau_l \). The \( \nu \)'s, \( \tau_s \) and \( \tau_l \) are all primitives. We can express each dealer's signals and the price signal as the following Π loading matrix times a vector \( Z \) of orthogonal shocks:
\[ S_d = f + \bar{\Pi}_d \bar{Z} : \]
\[ \bar{\Pi}_d = \begin{bmatrix} \nu_d & 0 \\ \frac{1}{N_d} & \frac{1}{N_d} \end{bmatrix} C/B \] (60)
\[ \bar{Z} = [ \bar{\epsilon}_1 \ldots \bar{\epsilon}_{N_d} \ x ]' \quad \text{s.t.} \]
\[ \nu [\bar{Z}] = \text{diag}([\tau_d^{-1} 1_{N_d}, \tau_x^{-1}]). \] (61)
where \( \nu_d \) is a vector of zeros with a 1 in entry \( d \) and \( B \) and \( C \) are equilibrium pricing coefficients.

Then, for all agents \( j \) who are clients of dealer \( d \), signal variance-covariance is
\[ \nu(S_j) = \tau_f^{-1} + \bar{\Pi}_d \nu[\bar{Z}] \bar{\Pi}_d' \quad \forall j \in \mathcal{I}_d. \] (62)

Then, the Bayesian updating weights and posterior precision are the same formulas as before (21) and (??). Using (2) and (4), we find that the posterior belief is
\[ \mathbb{E}[f | X_d S] = (1 - \beta_s - \beta_p) \mu + \beta_s \bar{s}_d + \beta_p \frac{p - A - C \bar{x}}{B}, \] (63)
with conditional variance \( \nu [f | X_d S] = \tau_f^{-1} - \tau_f^{-1} (\beta_s + \beta_p) \equiv \tau^{-1} \).

**Proof of Result 5: 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 (8) and the large investors’ bids by equation (22). Dealers’ bids are (16). Substituting these bids into the market clearing condition \( x + \sum_{i=1}^{N} q_i = 1 \) yields
\[ 1 = x + \sum_{d=1}^{N_D} \left( \nu_s \tau \rho^{-1} + (\nu_l + 1) M_L \right) \left( \mathbb{E}[f | X_d S] - p \right) + \chi N_D M_D p. \]
Substituting in (63) for the conditional expectation and then taking the derivative with respect to \( q_L \), we obtain the price impact of a large investor
\[ 0 = 1 + \left( N_L \tau \rho^{-1} + (N_L + N_D - 1) M_L \right) \left( \frac{\beta(2) - B}{B} + \chi N_D M_D \right) \frac{dp}{dq_L}. \]
so that $M_L$ (defined as in (23)) is given implicitly by

$$M_L^{-1} = \rho L \hat{\tau}^{-1} - B \left( (N_I \hat{\tau} \rho^{-1} + (N_L + N_D - 1) M_L) (\beta(2) - B) \right)^{-1}. \quad (64)$$

Similarly, taking the derivative with respect to $q_d$, we obtain the price impact of a dealer $dp/dq_d$. Substituting into (23) yields the dealer demand coefficient

$$M_D^{-1} = \rho L \hat{\tau}^{-1} - (1 - \chi) \left[ (N_D (\nu_s \hat{\tau} \rho^{-1} + \nu_l M_L + M_D) - M_D) \frac{\beta(2) - B}{B} + \chi (N_D - 1) M_D \right]^{-1}. \quad (65)$$

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

$$A = C \left( -1 + N_D (\nu_s \hat{\tau} \rho^{-1} + \nu_l M_L + M_D) \left( (1 - \beta (1) - \beta (2)) \mu - \frac{\beta(2) B}{\beta(1)} (A + C \bar{x}) \right) \right) \quad (66)$$

$$B = C N_D (\nu_s \hat{\tau} \rho^{-1} + \nu_l M_L + M_D) \beta(1) \quad (67)$$

$$C = - \left[ N_D \left( (\nu_s \hat{\tau} \rho^{-1} + \nu_l M_L + M_D) \left( \frac{\beta(2) - B}{B} \right) + \chi M_D \right) \right]^{-1}. \quad (68)$$

The solution to this model is a solution to (21), (64)-(65) and (66)-(68).

**B Computational Details**

**Initial Guesses: Guess for M’s** The object $M_I$ is the effect of higher expected profit per share on demand for the asset. A natural initial guess for $M_I$ is the $M_I$ that would arise with no monopoly profit. That is $M_I^{\text{guess}} = 1/\rho I * V[F | X_I S]$. This is also written as $M_I^{\text{guess}} = \hat{\tau}_l/\rho l$. This should be the guess for every agent $l$ in each model. What differs across agents and models is the posterior precision $\hat{\tau}_l$. I will work through that next. Risk aversion $\rho_l$ differs across agents, but is exogenous.

The problem is that $\hat{\tau}$ depends on $\beta$’s, which depends on the signal variances, which depend on the pricing coefficients in non-linear ways. Suppose we ignore the signal covariances. That would be like a competitive economy limit. But then the variance of the price signal still depends on price coefficients. But in a competitive economy, we get $B/C = -N_J \sigma_{s/\rho} + N_D \sigma_{d/\rho} + N_L \sigma_{L/\rho_l}/\rho_l$.

Then, we can express the posterior precision as

$$\hat{\tau} = \tau_f + \frac{N_s \tau_s/\rho_s + N_d \tau_d/\rho_d + N_L \tau_L/\rho_l}{N_s \rho_s + N_d \rho_d + N_L \rho_l} + (B/C)^2 \tau_x$$

Substitute $B/C$ into the equation for $\hat{\tau}$ and then divide $\hat{\tau}$ by $\rho$ to get $M^{\text{guess}}$. All of the objects in the resulting expression are exogenous parameters.
C Extension: Learning from the intermediation decision of the large investor

In this appendix, we consider an extension of the model where dealer 1 incorporates the information that can be learned from the intermediation choice of the large investor in the average signal shared with the rest of his clients. When the large investor chooses to bid through the dealer, the dealer observes the investor’s signal directly and the equilibrium outcomes are the same as before. We solve this model generally with a minimum bidding penalty. Then we can set \( \chi = 0 \) to get the results for the no-penalty model.

Consider now the case when the large investor bids directly in the auction. Recall that the large investor chooses to bid directly whenever \( \mathbb{E} U(L_d) \geq \mathbb{E} U(L_i) \). Using (10), we can rewrite this as

\[
(1 + 2\theta_{Ld}\Delta V_{Ld})^{-\frac{1}{2}} \exp \left( -\frac{\theta_{Ld}\mu_{r,Ld}^2}{1 + 2\theta_{Ld}\Delta V_{Ld}} \right) \leq (1 + 2\theta_{Li}\Delta V_{Li})^{-\frac{1}{2}} \exp \left( -\frac{\theta_{Li}\mu_{r,Li}^2}{1 + 2\theta_{Li}\Delta V_{Li}} \right).
\]

Equivalently, the large investor chooses to bid directly when

\[
\exp \left( -\frac{\theta_{Ld}\mu_{r,Ld}^2}{1 + 2\theta_{Ld}\Delta V_{Ld}} + \frac{\theta_{Li}\mu_{r,Li}^2}{1 + 2\theta_{Li}\Delta V_{Li}} \right) \leq \sqrt{\frac{1 + 2\theta_{Ld}\Delta V_{Ld}}{1 + 2\theta_{Li}\Delta V_{Li}}},
\]

or

\[
\frac{\theta_{Ld}\mu_{r,Ld}^2}{1 + 2\theta_{Ld}\Delta V_{Ld}} - \frac{\theta_{Li}\mu_{r,Li}^2}{1 + 2\theta_{Li}\Delta V_{Li}} \geq -\frac{1}{2} \log (1 + 2\theta_{Ld}\Delta V_{Ld}) + \frac{1}{2} \log (1 + 2\theta_{Li}\Delta V_{Li}).
\]

The left hand side of the above is a quadratic function of the signal of the large investor, \( s_{N_{j+1}} \). Let \( \Upsilon_{l,u} \) be the two solutions to the quadratic equation for the boundaries of the indirect bidding region, so that the large investor bids indirectly whenever \( \Upsilon_l \leq s_{N_{j+1}} \leq \Upsilon_u \); equivalently, the large investor bids directly if \( s_{N_{j+1}} \geq \Upsilon_u \) or \( s_{N_{j+1}} \leq \Upsilon_l \).

To solve the model where all agents make rational inferences from the intermediation decision, we first solve our model without this information and determine the cutoffs \( \nu_l \) and \( \nu_h \). We simulate the model to determine the probability of direct bidding \( Pr(v) \). Then, we also compute, conditional on choosing to bid directly, what the direct bidder’s average signal is \( (s_v) \) and the variance of that signal \( (\tau_v^{-1}) \). Next, we construct a normally-distributed, conditionally independent approximating signal with the same mean and variance: \( s_v = f + \epsilon_v \) where \( \epsilon_v \sim N(0, \tau_v^{-1}) \). In cases where the large investor bids directly, we allow dealer 1 to observe \( s_v \) and incorporate it in his advice to his clients. When the large investor bids indirectly, his dealer observes his signal exactly, as in the indirect bidding model we solved before.

Variance-covariance matrix of signals When the large investor bids directly, the first dealer’s signal is the average of the first \( \nu_l \) investors’, the first \( \nu_l - 1 \) large investors’, the first dealer’s signal and the new signal \( s_v \) that arises from observing the direct bidding decision.
By Bayes’ law, the posterior expectation, conditional on all these signals is:

\[
\bar{s}_1 = \frac{\tau_{\varepsilon,s} \sum_{k \in I_1} s_k + \tau_{\varepsilon,l} \left( \sum_{j \in I_1} s_j + s_d \right) + \tau_\nu s_\nu}{\nu_s \tau_{\varepsilon,s} + \nu_l \tau_{\varepsilon,l} + \tau_\nu}
\]

Bayes law also tells us that the precision is \( \hat{\tau}_{d1} = \nu_s \tau_{\varepsilon,s} + \nu_l \tau_{\varepsilon,l} + \tau_\nu \). The signals of all other dealers are the same as before because those dealers are not aware of the large investor’s intermediation decision.

As in the previous model, investor \( i \) who bids through intermediary \( d \) observes signals \( \tilde{S}_i = [s_i, \bar{s}_d, s(p)] \). Since dealer share information perfectly here, the investor’s own signal is redundant. Thus, for notational convenience, we continue as if \( \tilde{S}_i = [\bar{s}_d(i), s(p)] \), \( \forall i \neq L \).

The signals of the large investor bidding directly observes only his own signal and the price information: \( \tilde{S}_L = [s_L, s(p)] \).

We can construct the signals as the true payoff \( f \) plus weights \( \Pi \) on orthogonal shocks \( Z \). Let

\[
\tilde{Z}_v = \begin{bmatrix} \epsilon_L & \bar{\epsilon}_1 & \bar{\epsilon}_2 & \ldots & \bar{\epsilon}_{Nd} & (x - \bar{x}) \end{bmatrix}^T
\]

where \( \epsilon_i \equiv S_L - f \) and \( \bar{\epsilon}_d \equiv \bar{s}_d - f \) are the noise the the signals of the large investors and the dealer and \( x \) is the non-competitive bids. Note that the information from the intermediation decision of the direct bidder in incorporated in the first dealer’s signal, and thus in \( \bar{\epsilon}_1 \). The variances of these shocks are:

\[
\mathbb{V} \left[ \tilde{Z}^{\text{direct}} \right] = \text{diag}(\tau_L^{-1}, \hat{\tau}_{d1}^{-1}, \tau_{Nd}^{-1}, \tau_x^{-1})).
\]

The signals of the large investor \( L \), dealer \( 1 \) and his clients \( D1 \) and all others \( d \neq 1 \) have loadings on the orthogonal shocks given by (44), (45) and (46). Then, for all agents \( j \) who are clients of dealer \( d(j) \), or for the large investor \( d(j) = L \), the signal variance-covariance is \( \mathbb{V}(S_j) = \tau_f^{-1} + \Pi_d(j) \mathbb{V}(Z_v) \Pi_d(j)^T \) \( \forall j \).

**Bayesian updating weights \( \beta \) and posterior precisions \( \tau \)** The posterior expectation of the asset value is a linear combination of the prior and two signals: (1) the signal \( s_L \) or \( s_d(i) \) provided by the dealer; and (2) the signal conveyed by the price \((p - A - C\bar{x})/B\). \( \beta_s \) and \( \beta_p \) are the weights the agent places on the two signals. The weights are given by (21), the resulting conditional expectation is given by (2), with precision (48).

**Equilibrium Price** Once we’ve adjusted the Bayesian updating weights \( \beta \), the rest of the solution of the direct bidding model follows just as before. The equilibrium price coefficients are a joint solution to the coefficient equations (55) - (59), the \( M \) equation (23), in conjunction with the price sensitivities (49) - (54) and the \( \beta \)’s (21), which in turn depend on the \( \Pi \)’s (44), (45) and (46) and the variances of the orthogonal shocks (69). Setting \( \chi = 0 \) yields the solution to the model without the minimum bidding requirement.
Table 4: 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**

<table>
<thead>
<tr>
<th></th>
<th>Revenue</th>
<th>Dealer allocation</th>
<th>Direct allocation</th>
<th>Indirect allocation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>37.5957</td>
<td>42.0538</td>
<td>0.5868</td>
<td>50.8466</td>
</tr>
<tr>
<td><strong>Std. Dev.</strong></td>
<td>69.8128</td>
<td>20.1859</td>
<td>0.7675</td>
<td>57.0472</td>
</tr>
<tr>
<td><strong>Skew</strong></td>
<td>-0.0150</td>
<td>0.0072</td>
<td>0.9225</td>
<td>0.1887</td>
</tr>
<tr>
<td><strong>Kurtosis</strong></td>
<td>3.0101</td>
<td>3.0185</td>
<td>2.5776</td>
<td>4.6919</td>
</tr>
</tbody>
</table>

**Panel B: Direct Bidding**

<table>
<thead>
<tr>
<th></th>
<th>Revenue</th>
<th>Dealer allocation</th>
<th>Direct allocation</th>
<th>Indirect allocation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>52.7388</td>
<td>41.3514</td>
<td>1.3302</td>
<td>45.0985</td>
</tr>
<tr>
<td><strong>Std. Dev.</strong></td>
<td>66.7869</td>
<td>20.3264</td>
<td>0.5887</td>
<td>21.0972</td>
</tr>
<tr>
<td><strong>Skew</strong></td>
<td>0.0019</td>
<td>0.0033</td>
<td>0.0870</td>
<td>0.0011</td>
</tr>
<tr>
<td><strong>Kurtosis</strong></td>
<td>3.0634</td>
<td>3.0590</td>
<td>3.0824</td>
<td>3.0663</td>
</tr>
</tbody>
</table>

**Panel C: Indirect Bidding**

<table>
<thead>
<tr>
<th></th>
<th>Revenue</th>
<th>Dealer allocation</th>
<th>Direct allocation</th>
<th>Indirect allocation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>25.6443</td>
<td>42.6081</td>
<td>0.0000</td>
<td>55.3831</td>
</tr>
<tr>
<td><strong>Std. Dev.</strong></td>
<td>69.8299</td>
<td>20.0570</td>
<td>0.0000</td>
<td>73.6539</td>
</tr>
<tr>
<td><strong>Skew</strong></td>
<td>0.0126</td>
<td>0.0126</td>
<td>-</td>
<td>-0.0126</td>
</tr>
<tr>
<td><strong>Kurtosis</strong></td>
<td>2.9827</td>
<td>2.9827</td>
<td>-</td>
<td>2.9827</td>
</tr>
</tbody>
</table>

**Panel D: High signal**

<table>
<thead>
<tr>
<th></th>
<th>Revenue</th>
<th>Dealer allocation</th>
<th>Direct allocation</th>
<th>Indirect allocation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>56.6878</td>
<td>42.5339</td>
<td>1.4342</td>
<td>43.8340</td>
</tr>
<tr>
<td><strong>Std. Dev.</strong></td>
<td>66.6329</td>
<td>20.2831</td>
<td>0.5781</td>
<td>21.0393</td>
</tr>
<tr>
<td><strong>Skew</strong></td>
<td>0.0091</td>
<td>0.0089</td>
<td>0.0739</td>
<td>-0.0081</td>
</tr>
<tr>
<td><strong>Kurtosis</strong></td>
<td>3.0774</td>
<td>3.0716</td>
<td>3.1070</td>
<td>3.0837</td>
</tr>
</tbody>
</table>

**Panel E: Low signal**

<table>
<thead>
<tr>
<th></th>
<th>Revenue</th>
<th>Dealer allocation</th>
<th>Direct allocation</th>
<th>Indirect allocation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>17.3851</td>
<td>40.2359</td>
<td>0.0000</td>
<td>64.0946</td>
</tr>
<tr>
<td><strong>Std. Dev.</strong></td>
<td>69.5897</td>
<td>19.9880</td>
<td>0.0000</td>
<td>73.4005</td>
</tr>
<tr>
<td><strong>Skew</strong></td>
<td>0.0250</td>
<td>0.0250</td>
<td>-</td>
<td>-0.0250</td>
</tr>
<tr>
<td><strong>Kurtosis</strong></td>
<td>2.9751</td>
<td>2.9751</td>
<td>-</td>
<td>2.9751</td>
</tr>
</tbody>
</table>
Figure 6: 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