

Dispersion and Skewness of Bid Prices¹

Boyan Jovanovic and Albert J. Menkveld

First version: July 7, 2014

This version: February 2, 2016

¹Boyan Jovanovic, Department of Economics, New York University, 19 West 4th Street, New York, NY, USA, tel +1 212 998 8900, fax +1 212 995 4186, bj2@nyu.edu. Albert J. Menkveld, VU University Amsterdam, FEWEB, De Boelelaan 1105, 1081 HV, Amsterdam, Netherlands, tel +31 20 598 6130, fax +31 20 598 6020, albertjmenkveld@gmail.com. We thank the NSF and NWO for support, and Sean Flynn, Wenqian Huang, Sai Ma, and Gaston Navarro for assistance.

Dispersion and Skewness of Bid Prices

Abstract

Competitive bidding by homogeneous agents in a first-price auction can yield a non-degenerate bid price distribution. This price dispersion is the unique equilibrium in a setting where bidders “pay to play.” *Ex ante*, bidders decide simultaneously on whether to play or not. *Ex post*, those who play submit their bid simultaneously not knowing who else is in the market. The price-dispersion result is applied to high-frequency bidding in limit-order markets. The parsimonious model fits the bid-price dispersion for S&P 500 stocks remarkably well.

1 Introduction

This paper develops an auction model that is applied to a limit-order market. This type of market is relevant as many securities now trade in electronic markets with a limit-order book.

Model. The model is a first-price common-value auction with N potential bidders who can decide to pay a participation cost to submit a bid. These players are identical *ex ante* and the value of the object is known to them. The game has two stages. First, players decide simultaneously whether to pay the cost to submit a bid, or to stay out. Second, bids are simultaneous and the highest bidder wins; he receives the object and pays his bid. Bidders do not observe how many others are bidding. The unique equilibrium is symmetric: Each player chooses the same participation probability and, if he decides to participate, draws a bid from the same non-degenerate (endogenously determined) bid-price distribution. The model has no *ex-ante* heterogeneity, no exogenous randomness, and always yields negative skewness in the distribution of bids.

We stress that this unique equilibrium has inertia. Having more middleman around *ex ante* might imply having fewer around *ex post*, in expectation. It always raises the probability of the event that no middleman shows up *ex post*, even in the case where the expected number of middlemen to show up *ex post* is *higher*. Such event is socially costly as it implies that gains from trade are not realized. It constitutes a deadweight loss.

Application. We apply the model to a limit-order market for a financial security. The bidders are interpreted to be “middlemen” (i.e., high-frequency market makers).¹ Their bids are interpreted as limit orders placed in the limit-order book. We then estimate the model based on realized bid-price dispersions for all S&P 500 stocks, using data from the Flash-Crash episode of May 6, 2010. Although the model only has two free parameters, it fits the realized price dispersions surprisingly

¹In the past three decades, human-intermediated equity markets around the world were gradually replaced by limit-order markets, essentially continuous double-sided auctions (Jain, 2005). In this new electronic environment, the role of human market makers (e.g., NYSE specialists) was taken over mostly by high-frequency traders submitting price quotes (SEC, 2010a).

well. In particular, it captures one of its salient and highly robust features: a negative skew in the bid distribution. Model estimates further suggest that, relative to total gains from trade, the participation cost for bidders increased substantially in the course of the Flash Crash, while the seller's outside option decreased in value. This combination (increased cost and less valuable outside option) makes that, again relative to total gains to trade, the inefficiency of having too many middlemen around is highest at the time of the crash.

The model has no bidder heterogeneity in the form of object values or private signals. Although adding these features may be critically important to fit price distributions in other markets, it seems Occam's razor would remove them for bid prices of S&P 500 stocks.

A closely related model is Baruch and Glosten (2013) who also have bidder homogeneity and a mixed-strategies equilibrium in limit-order markets. They find that "fleeting orders" are a natural outcome as bidders rebid in each round to avoid undercutting risk. Their model however does not rule out a pure-strategy equilibrium. The non-zero bidding cost in our setting rules out such equilibrium. Mixed strategies can explain the high quote-to-trade ratio (Angel, Harris, and Spatt, 2015, Fig. 2.16) and the extreme quote volatility (Hasbrouck, 2015) that are characteristic of modern securities markets.²

Contribution to the auction literature. In the literature on auctions there are many models with a random number of bidders (see Klemperer, 1999, Section 8.4, for a discussion). In some of them entry is endogenous and in this group, the closest models to ours are Hausch and Li (1993) and Cao and Shi (2001). Indeed, our model is a special case of these two models when the signal is uninformative or prohibitively costly. These two papers analyze only symmetric equilibria and our contribution is to show that asymmetric equilibria do not exist. We show that the equilibrium is unique and that it is symmetric. However, related arguments are in papers on Bertrand competition

²Another important difference is that some equilibria in their model yield positive middleman rents for finite N that tend to zero when N is taken to infinity. In our model, middlemen never earn any positive rents, i.e., for all N their rents are zero.

among firms when customers have switching costs, e.g., Shilony (1977) and, in particular, Rosenthal (1980) who proves the non-existence of asymmetric equilibria using reasoning analogous to ours which is that players can be taken advantage of if they act deterministically. Finally, in the search literature the closest model is Butters (1977) where firms quote prices to customers.

We further add to this literature by stressing inertia. We show that more middlemen *ex ante* could result in each middleman reducing his participation probability to such a low level that fewer middlemen are expected to participate *ex post*. We prove that more middlemen *ex ante* always results in a higher probability of the event that *no* middleman shows up *ex post* (see Proposition 2). As, in this event, the asset does not transfer from seller to buyer, gains from trade are not realized. The upside is that in this event no middleman pays the participation cost. We prove that *net* (of participation cost) gains from trade decrease with any middleman added beyond the first two of them (see Proposition 3).³ In some sense, this result could be interpreted as generalized Bertrand competition; only two middlemen are needed for investors to reap the full benefit of middlemen competition.

We argue that the common value auction of this type is a good description of the process generating the limit order book. The two assumptions that are key to delivering the observed left skewness in the distribution of bids are (i) *homogeneity* of potential bidders in a *common-value* auction and (ii) an *inability to coordinate* entry decisions which rules out asymmetric pure strategy equilibria. The paper's main contribution is showing that with just two free parameters the model achieves a remarkable fit of bid prices in the stock market.

We cannot claim to have ruled out other explanations for the variance and skewness in stock-price bids. Heterogeneous bidder valuations or differential speeds of market access could be producing such an outcome provided that the distribution among the bidders is of an appropriate

³It is worth pointing out that this result differs from Biais, Martimort, and Rochet (2000) who have Cournot competition. In their model each additional middleman lowers their overall rent, and thus benefits investors. We have Bertrand price competition; the winner takes all. Like us, Dennert (1993) also finds that investor transaction cost increases when more middleman participate in the bidding game. His model is more elaborate as investors are either informed or uninformed. We get the result without assuming such heterogeneity.

form — bidder heterogeneity may indeed be key for explaining bid distributions in auctions more generally. All we claim here is that one can completely dispense with it and yet end up with a coherent explanation for why the distribution of bids in a limit-order market takes the form that it does.

Other evidence on bid-price distributions. In their structural estimation of a limit-order book equilibrium with heterogeneous agents, Hollifield et al. (2006, Table 3) document that bids are left skewed for Swedish stocks. The number of shares available at the second-best bid is 242,900. This bid is on average 1.1 price ticks below the best bid. The *additional* amount available at double that price distance is 167,900, i.e., 31% less. This is consistent with left skewness in bids, but only partial evidence.⁴ We add to this evidence by estimating our model for the *entire* bid price distribution of all S&P 500 stocks. The data sample comprises a single day: May 6, 2010. This day enables us to compare parameter estimates for two market conditions: “normal” for the start of the day and “extreme” for the half-hour from 2:30 p.m. until 3:00 p.m., the period of the “Flash Crash.” The estimation illustrates how the model’s primitive parameters changed in the course of the day. It turns out that investors need their intermediaries more in extreme market conditions (the value of their outside option is reduced) but these intermediaries suffer a higher cost of participating.

2 Model

Consider a common value, sealed bid, first-price auction; $N \geq 2$ players can bid for an object that each values at v . The seller’s reservation price for the object is u , and it is exogenous. To bid, a player must pay c , where $c < v - u$. The parameters of the game are thus N , c , u and v , and they are common knowledge. The notation used throughout the manuscript is summarized in Appendix A.

⁴Goldstein and Kavajecz (2004, Figure 1), Naes and Skjeltorp (2006, Table 2), Degryse, de Jong, and van Kervel (2015, Table 2) document left skewness for the entire bid-price distribution for U.S., Norwegian, and Dutch stocks, respectively. Biais, Hillion, and Spatt (1995, Figure 1) find linearity for the best bid prices for French stocks.

Actions.—A player has two actions: an entry decision and, in the event of entry, a bid p . A player acts with no access to signals about the actions of others. In particular, although the entry decision is made before p is chosen, when bidding, a player does not know how many others are bidding.

Payoffs.—Entering and bidding p yields a payoff of $v - p - c$ if the bid wins and $-c$ otherwise. Not entering yields a payoff of zero.

2.1 Equilibrium

Equilibrium is unique, symmetric, and in mixed strategies. Both actions — entry and bid price — have non-degenerate distributions. Uniqueness arises because the number of competing bids is uncertain at the time of bidding.

Proposition 1 *A unique Nash equilibrium exists, and it is symmetric and in mixed strategies. A player enters with probability*

$$\lambda = 1 - \left(\frac{c}{v - u} \right)^{1/(N-1)}, \quad (1)$$

and draws a bid from the CDF

$$H(p) = \frac{1}{\lambda} \left(\frac{c}{v - p} \right)^{1/(N-1)} - \frac{1 - \lambda}{\lambda}, \quad (2)$$

where H has support $[u, v - c]$.

Proof. Eq.s (1) and (2) arise in the model of Hausch and Li (1993) when their parameter $\alpha = 0$ and when, instead of being zero as they assume, the outside option for the seller is u . Then the distribution M is the same as H . It only remains to show that no asymmetric equilibria exist in either pure or in mixed strategies. This is done in Appendix D. ■

Bid aggressiveness as function of N .— $H(p)$ is increasing in N . Bid aggressiveness refers to how closely the bid of a given bidder approaches his true valuation. Here the true valuation is

v , and since the price-bid distribution for N first-order stochastically dominates the distribution for $N + 1$, aggressiveness declines with N . That is, as *ex-ante* competition intensifies, the *ex-post* distribution of bids shifts to the left.

Winning-bid distribution as function of N .—The distribution of the highest bid is

$$F(p) = \sum_{k=0}^N H^k(p) \binom{N}{k} \lambda^k (1 - \lambda)^{N-k} = (1 - \lambda + \lambda H(p))^N = \left(\frac{c}{v - p} \right)^{N/(N-1)}. \quad (3)$$

The second equality in (3) uses the binomial formula, and the third equality follows from (2). The first equality in (3) is based on the technical assumption that if $k = 0$, the outcome is a maximum bid to equal u , which yields the same utility for the seller as not receiving any bids. For $p \in [u, v - c]$, $F(p)$ is strictly *increasing* in N .⁵ F therefore shifts to the left when N increases just like H (the individual player's bid-price distribution) in spite of there being more players around *ex ante*. We will revisit this result and illustrate it with an example later in this section.

Skewness of the bid-price distribution.—The density of bids,

$$h(p) = \frac{1}{\lambda(N-1)} c^{\frac{1}{N-1}} (v - p)^{-\frac{N}{N-1}}, \quad (4)$$

is increasing in p for all values of N . The distribution is skewed to the left, i.e., “negatively skewed.” We compute the level of left-skewness as the slope of the density of p averaged over the range of bids. This is the ratio of the density at the largest possible bid relative to the smallest:

$$r \equiv \frac{h(v - c)}{h(u)} = \left(\frac{v - u}{c} \right)^{N/(N-1)}. \quad (5)$$

r is decreasing in N . This mirrors the earlier result that bid aggressiveness is increasing in N . As one would expect, r is increasing in $(v - u)$ and is decreasing in c .

The probability of winning the game.—The probability that a player wins the game is not simply $1/N$ as no one wins in the event that no one made a bid. The following proposition establishes

⁵Rosenthal (1980) proves an analogous property for the distribution of asking prices in the symmetric mixed-strategy equilibrium in his model, where the distribution of asking prices shifts to the right as N rises.

the probability of winning.

Proposition 2 *The probability that no one shows up is*

$$(1 - \lambda_N)^N,$$

and it increases in N . The probability that a player wins the game is

$$p_N = \frac{1}{N} \left(1 - (1 - \lambda_N)^N \right)$$

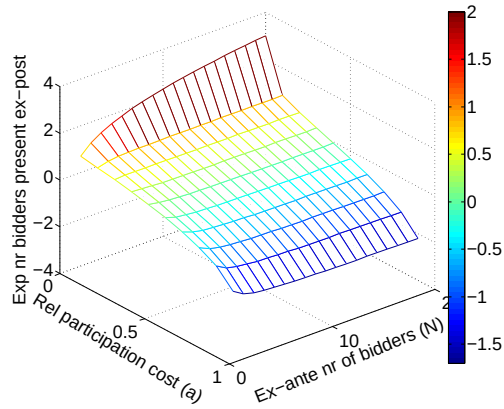
and p_N decreases in N .

Proof is in Appendix D.

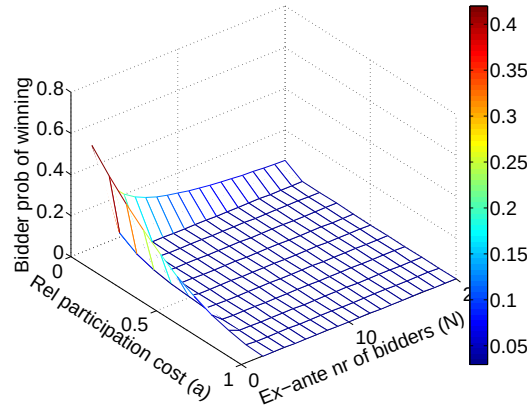
Figure 1: Expected number of bidders, probability of winning, and net gains from trade

This figure plots the expected number of bidders present *ex post* (a), a bidder's probability of winning the game (b), and the expected gains from trade net of participation cost (c). They are all determined by the model's two primitive parameters. N is the number of bidders who consider entering the game *ex ante*. a is the relative cost of participating in the game, i.e., $a = c/(v - u)$.

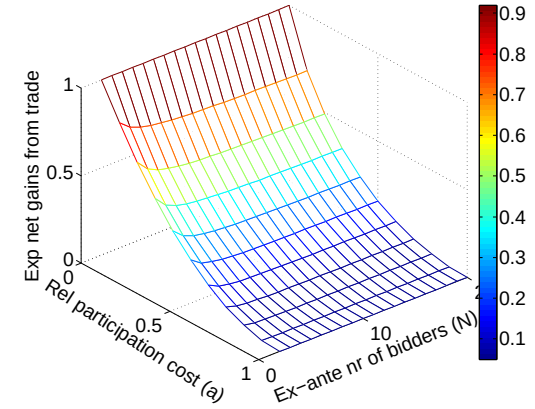
(a) Expected number of bidders *ex post*



(b) Bidder's probability of winning



(c) Net gains from trade



An illustrative example. We illustrate these results by analyzing some comparative statics. To that end, we define a as the cost of participating, divided by the size of the profit opportunity for middlemen:

$$a = c / (v - u) .$$

The expected number of middlemen participating *ex post*, the probability of winning, and the net gains from trade can all three be expressed in a and N . Figure 1 illustrates the results with three-dimensional plots. A somewhat surprising pattern emerges:

1. The expected number of bidders present *ex post* ($N\lambda$) is non-monotonic in the number of bidders present *ex ante* (N). This pattern arises for high values of a . The expected number of bidders declines initially but rises eventually.
2. The probability of winning the game declines monotonically in N as predicted by proposition 2.

The surprising result is that there is a parameter region — high a and low N — where a bidder's probability of winning *declines* in N while fewer bidders show up for larger N . This is counter-intuitive given bidder homogeneity. How can these two findings be reconciled? The force that runs counter to the fewer-bidders-higher-likelihood-of-winning is that the event of no one showing up increases in likelihood for larger N (the proof of Proposition 2 has the result that $(1 - \lambda_N)^N$ increases in N). If no bidder shows up, then there is no winner. This counterforce is strongest for high a (relative cost of bidding) and low N , which explains why the surprising result pops up in that region.

The bidder game analysis reveals inertia. There is a higher probability that the potential gains from trade are not realized when there are *more* bidders around *ex ante*. This is due to a lower probability of participating for each of them in equilibrium, so much so that the event of no one showing up becomes more likely. This result seems to only hold, however, in a particular parameter region (high a and low N). Everywhere else there are more middlemen around *ex post* when N

increases. Here the gross gains from trade increase in N . This observations are summarized in the following remark.

Remark 1 *The expected number of middlemen showing up ex post is neither uniformly increasing, nor uniformly decreasing in N .*

More middlemen around *ex post* also implies that, collectively, they pay a higher participation cost. Panel (c) of Figure 1 illustrates that the *net* gains from trade seem to always decrease in N . This turns out to be true a general property and is therefore stated as a proposition.

Proposition 3 *Net gains from trade (welfare) are decreasing in the ex-ante number of middlemen (N).*

Proof. Let W denote the net gains from trade, then

$$W = \frac{c}{a} \left(1 - a^{N/(N-1)}\right) - c \left(1 - a^{1/(N-1)}\right) N. \quad (6)$$

If there are infinitely many middlemen available *ex ante*, we have

$$\lim_{N \rightarrow \infty} W_N = \frac{c}{a} (1 - a) + c \ln(a).$$

Let us consider the more general case of $N \in \mathbb{R}$ then

$$\frac{\partial}{\partial N} \left(\frac{W}{c} \right) = - \left(1 - a^{1/(N-1)}\right) + (N-1) \frac{\partial}{\partial N} a^{1/(N-1)} = a^n (1 - n \ln a) - 1$$

for $n = 1/(N-1) > 0$. To show that it is negative, we need to show that

$$a^{1/(N-1)} \left(1 - \frac{1}{N-1} \ln a\right) < 1.$$

Denote

$$F(a) = a^{1/(N-1)} \left(1 - \frac{1}{N-1} \ln a\right).$$

We have

$$F'(a) = na^{n-1} (1 - n \ln a) - a^n \frac{n}{a} = -n^2 a^{n-1} \ln a.$$

$a \in (0, 1)$ implies $\ln a < 0$ and therefore

$$F'(a) > 0 \text{ for } a \in [0, 1].$$

We further have $F(1) = 1$ therefore

$$F(a) < 1 \text{ for } a \in [0, 1].$$

■

Proposition 3 could be read as “more competition” is bad for welfare. Having more middlemen around *ex ante* increases social cost in either of two ways. First, the state of no one bidding becomes more likely. Second, if more middlemen are expected to show up *ex post* then (aggregate) participation costs are higher.

This claim is true for $N \geq 2$. When $N = 1$, however, adding the middleman cannot raise welfare. The only equilibrium then is for the seller to post an ask himself. To wait for a bid from the middleman would mean foregoing his option to post, and would expose the seller to a monopsony bid below the outside option that he has foregone. The equilibrium outcome is then equivalent to the situation in which $N = 0$, i.e., in which there are no middlemen.

The Poisson limit as $N \rightarrow \infty$. The expected number of entrants is λN . Appendix D shows that

$$\lambda N \xrightarrow{N \rightarrow \infty} \ln \frac{v - u}{c} \equiv m. \quad (7)$$

Therefore, as N grows, the distribution of k , defined as the number of middlemen who show up *ex*

$post$, approaches the Poisson distribution with mean m , i.e.,

$$\Pr(\tilde{k} = k) = \frac{m^k e^{-m}}{k!}.$$

And, from (4), the bid-price density and distribution become

$$h(p) \xrightarrow{N \rightarrow \infty} \frac{1}{m} \frac{1}{v - p} \quad \text{and} \quad H(p) \xrightarrow{N \rightarrow \infty} 1 + \frac{1}{\ln(1 - u) - \ln(c)} \ln \frac{c}{1 - p} \quad (8)$$

for $p \in [u, v - c]$.

For $N \in \{2, 3, 12, \infty\}$, Panel (a) in Figure 2 plots $h(p)$ for the case where $u = 0$, $v = 1$, and $c = 0.1$. H therefore has support $[0, 0.9]$. The green curve ($N = 2$) is steeper than the blue curve ($N = 3$), which is again steeper than the red curve ($N = \infty$). In the estimation of the model we will use results for $N = \infty$ as these expressions are clean and easy to work with. We do not know how many middlemen were present *ex ante* in the data, but we do know that there are at least 12 of them (to be discussed in detail in Section 5.2). We therefore add the $N = 12$ red curve here to show that it is not far off from the $N = \infty$ curve, at least for these parameters. This makes us more comfortable with using the $N = \infty$ expressions in the estimation.

As N rises, h rotates clockwise reflecting decreased aggressiveness of bidding. Moreover, convergence in N is very fast. Panel (b) plots the winning-bid distribution F for the same parameter values. The mass points at zero coincide with the value $(1 - \lambda)^N$.⁶ The panel shows that also the winning-bid is less aggressive when N increases.

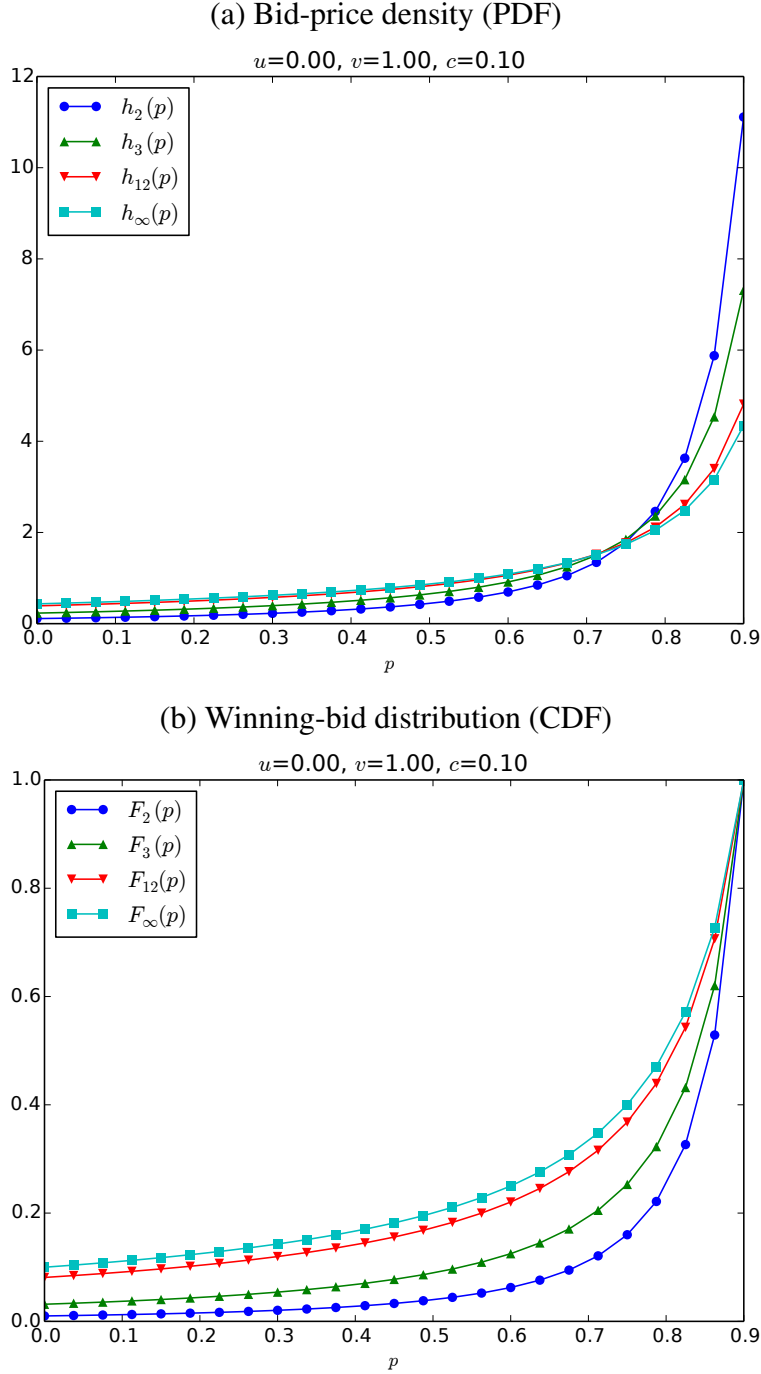
3 The high-frequency trader game

High-frequency traders (HFTs) are natural intermediaries in a game between an early-arriving seller and a late-arriving buyer, where the seller leaves a price quote (limit order) for the buyer to consider. This is in essence how modern limit-order markets work. An important friction is that the

⁶The mass point at zero is fully the result of the simplification that $k = 0$ was added as discussed on page 6.

Figure 2: Bid-price and winning-bid distribution

Panel (a) graphs the density functions that middlemen use in equilibrium in case there are two, three, twelve, or infinitely many of them, i.e., $N \in \{2, 3, 12, \infty\}$. Panel (b) graphs the winning-bid distribution for these cases. All results are based on setting the seller's reservation value to zero ($u = 0$), the value of the object to the bidder to one ($v = 1$), and the cost to be paid for participating to a tenth ($c = 0.1$).



buyer might have witnessed common-value changes that occurred after the seller left the market. He will adversely select the seller's price quote based on it. The seller anticipates such behavior and, if this adverse-selection cost is enough, he might forego leaving a price quote altogether.

HFTs can remove this “trade deadlock” as they have negligible cost of staying in the market and effectively add to it a capacity of quickly refreshing quotes based on information arrival. Their intermediation could remove the “stale quote” friction. Instead of posting a price himself, a seller would pass the security off to an HFT who would maintain a price quote with less or no adverse-selection risk vis-à-vis the buyer. This is the spirit of the application we develop in the remainder of this section.

Players.—The game has $N + 2$ players. We shall use the acronyms **S** for the seller, **B** for the buyer, and **M** for a high-frequency traders (HFTs) or “middlemen.” Two of the players, **S** and **B**, enter the market regardless of any other event that may occur; **S** arrives first and **B** arrives later. Their entry and timing decisions are exogenous and all this is common knowledge. The remaining N players are middlemen (**M**). $k \leq N$ of these will enter the market after paying the entry cost c , and the remaining $N - k$ will stay out.

Preferences.—All players are risk neutral. An investor's valuation of an asset is the sum of a common value z and a private value. **S** has private value $x > 0$ and **B** has private value $y > x$. **M** has a private value of zero. **S**'s utility of ending up with the asset is therefore $x + z$, **B**'s is $y + z$, and **M**'s is just z . The values x and y are common knowledge and fixed. Only z is random, with mean zero, and CDF $\Phi(z/\sigma)$. The model has only 4 parameters, namely (c, x, y, σ, Φ) . The first four are scalars, the fourth is a distribution.

Because we assumed that there is just one unit for sale an HFT would not place more than one order. If he were to make two bids, the lower bid would have a zero chance of winning. And because bids are real-valued random variables, the probability of landing on the same price as another bidder would be zero.

Timing of news and actions.—There are five stages:

1. $k \leq N$ middlemen pay c and enter with no securities. **S** also enters with one unit,
2. entering **M** post bids $p^{\mathbf{M}, \text{bid}}$ *without* seeing k ,
3. **S** either accepts the highest $p^{\mathbf{M}, \text{bid}}$, or he posts $p^{\mathbf{S}, \text{ask}}$ and leaves,
4. z is realized, and **M** and **B** see it. If **M** has the asset, he then posts $p^{\mathbf{M}, \text{ask}} = y + z$,
5. **B** arrives, sees $p^{\mathbf{S}, \text{ask}}$ or $p^{\mathbf{M}, \text{ask}}$, accepts or rejects it, and the game ends.

The homogeneity of middlemen makes that, in the bidding stage, the value of acquiring the security is worth v to them, i.e., the same for all **M**. This v is not to be confused with the “common value” z . v is determined before z is realized, and is common knowledge, as opposed to z which is realized only at stage 4.

Strategies.—

(i) **B** is the last to move and his action is binary, “accept or reject” $p^{\mathbf{S}, \text{ask}}$ or $p^{\mathbf{M}, \text{ask}}$, whichever is on the table when **B** arrives. If indifferent, **B** accepts, and therefore the optimal strategy is to accept if $p \leq y + z$.

(ii) **S** has two actions: The first is “accept or reject” the highest $p^{\mathbf{M}, \text{bid}}$. If indifferent, **S** accepts. If he rejects, he then chooses $p^{\mathbf{S}, \text{ask}}$ and leaves.

(iii) Each **M** enters with probability λ and if he enters, he draws $p^{\mathbf{M}, \text{bid}}$ from the CDF $H(\cdot)$. Thus **M**’s strategy is the pair (λ, H) .

Payoff functions.—From the preferences stated at the outset and from the structure of the game,

(i) **B**’s outside option at stage 5 is zero, and his payoff is $\max(0, y + z - p)$, where p is either $p^{\mathbf{S}, \text{ask}}$ or $p^{\mathbf{M}, \text{ask}}$.

(ii) **S**’s outside option at stage 3 is U (to be defined presently) and his payoff is $\max(U, p^{\mathbf{M}, \text{bid}})$. If **S** rejects **M**’s bids, he can post $p^{\mathbf{S}, \text{ask}} = p$ which **B** will accept iff $y + z \geq p$. Therefore at stage 3, **S**’s outside option is

$$U = \max_p \left\{ p \left[1 - \Phi \left(\frac{p - y}{\sigma} \right) \right] + \int_{-\infty}^{p-y} (x + z) d\Phi \left(\frac{z}{\sigma} \right) \right\} \quad (9)$$

and $p^{\mathbf{S}, \text{ask}}$ is the argmax of the RHS of (9).

(iii) $\mathbf{M}$'s outside option at stage 1 is zero, and so his expected payoff must be non-negative. If an $\mathbf{M}$ acquires the asset, he then sets $p^{\mathbf{M}, \text{ask}} = y + z$ and sells for sure, extracting all the rent from $\mathbf{B}$. So, if he pays c and enters and subsequently wins the bidding at $p^{\mathbf{M}, \text{bid}} = p$, his expected payoff (since $E(z) = 0$) is $y - p - c$ whereas, if he is not the highest bidder, his payoff is $-c$.

The key simplification is that at the time of bidding, $\mathbf{M}$ knows that whoever of them manages to get the security, he will be able to extract $y + z$ from $\mathbf{B}$, so that at the bidding stage (which occurs before z is revealed), the value to $\mathbf{M}$ of the object is y . For $\mathbf{M}$ to participate it is necessary and sufficient that

$$c < y - U. \quad (10)$$

Now U is not a parameter but, rather, is given by (9) and so, for (10) to hold it is necessary (but not sufficient) that $c < y - x$. Rather than provide conditions on the primitives now, we shall verify (10) *ex post*, i.e., show that it holds in equilibrium. This part will, however, be easy because U is increasing in x , and values of x can always be found that will deliver (10) which guarantee a positive entry probability for $\mathbf{M}$. We thus have

Proposition 4 *If (10) holds, the game between the N middlemen is equivalent to the auction game if*

$$u = U \quad \text{and} \quad v = y. \quad (11)$$

With c as the participation cost, the results of section 2 all apply, particularly (1) and (2).

4 Some final results needed to fit aggregate realized price data

We shall fit the HFT game to data in which the uncertain number of bidders is appropriate, namely, the stock market. Stocks are traded in limit-order markets where (i) bidders compete in a first-price auction because market sell orders are matched with the highest bid and (ii) a bidder does not instantaneously observe how many others are bidding.

In this section we develop two final sets of results that allow us to fit aggregated realized price dispersions. First, we show under what condition realized price dispersions for different securities can be “aggregated.” The model’s primitive parameters for an S&P 500 stock can then be estimated by matching the model-implied density with the empirical density that is available to us only as an aggregate across all S&P 500 stocks. We refer to this as the scaling condition. Second, we derive the distribution of bid prices relative to the (realized) best bid price. Under the scaling condition the resulting distribution is stock invariant. This allows for cross-sectional aggregation.

4.1 Scaling condition

The S&P 500 index comprises stocks of varying price levels or “sizes,” and each stock can be considered as having its own market. So, there are 500 distributions characterized by the stock-specific parameters which in our HFT game are (c, x, y, σ, Φ) .

The following proposition states the scalability result.

Proposition 5 *For $\alpha > 0$, if $(c, x, y, \sigma, \Phi(z/\sigma))$ are scaled by α to $(\alpha c, \alpha x, \alpha y, \alpha \sigma, \Phi(z/(\sigma \alpha)))$, then λ remains unchanged and $H(p)$ scales to $H(p/\alpha)$.*

Proof. Evidently, λ and H are homogeneous of degree zero in the vector (c, u, y, σ, p) . Now, suppose that when $(c, u, y, \sigma,) = (c_0, u_0, y_0)$, the solution to (2) is $H_0(p)$ for $p \in [u_0, y_0 - c_0]$. Then when $(c, u, y, \sigma,) = (\alpha c_0, \alpha u_0, \alpha y_0)$, the solution to (2) is:

$$H(p) = H_0(p/\alpha) \quad \text{for } p \in [\alpha u_0, \alpha(y_0 - c_0)]. \quad (12)$$

Now, the variable u is not a primitive of the model. Rather, as shown in (9), u depends on x, y, σ and $\Phi(\cdot)$. Scalability requires that if $(x_0, y_0, \sigma_0, \Phi_0(z/\sigma_0))$ in (9) leads to $u = u_0$, then $(\alpha x_0, \alpha y_0, \alpha \sigma_0, \Phi_0(z/(\sigma_0 \alpha)))$ in (9) leads to $u = \alpha u_0$. This turns out to be true since at $(\alpha x_0, \alpha y_0, \alpha \sigma_0, \Phi_0(z/(\sigma_0 \alpha)))$, the RHS of (9) reads

$$\begin{aligned} & \max_p \left\{ p \left[1 - \Phi_0 \left(\frac{p - \alpha y_0}{\sigma_0 \alpha} \right) \right] + \int_{-\infty}^{p - \alpha y_0} (\alpha x_0 + z) d\Phi_0 \left(\frac{z}{\sigma_0 \alpha} \right) \right\} \\ &= \max_{p'} \left\{ \alpha p' \left[1 - \Phi_0 \left(\frac{p' - y_0}{\sigma_0} \right) \right] + \alpha \int_{-\infty}^{p' - y_0} (x + z') d\Phi_0 \left(\frac{z'}{\sigma_0} \right) \right\} \\ &= \alpha u_0. \end{aligned}$$

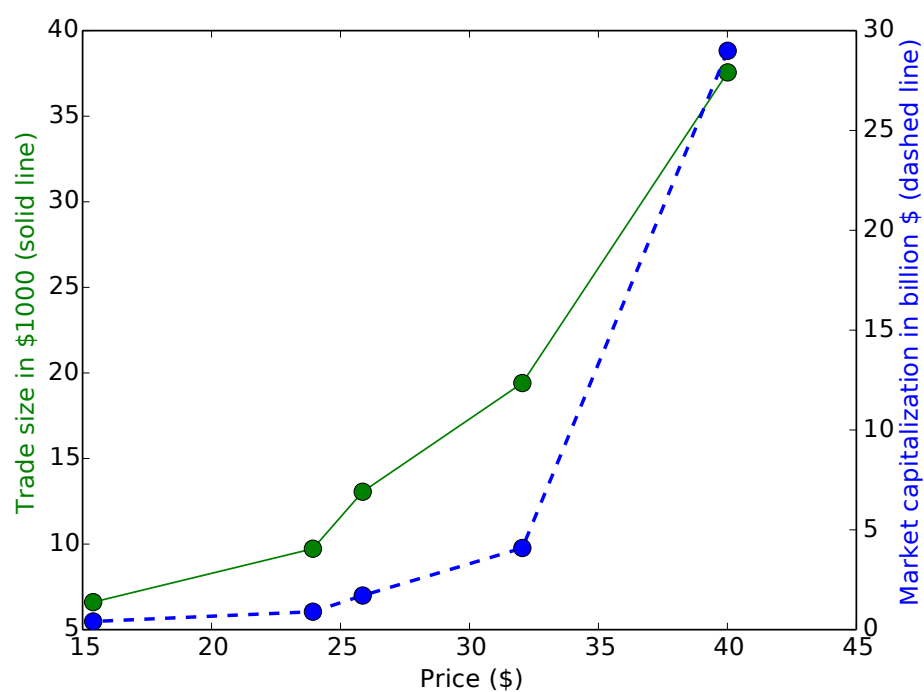
The first equality is based on the following logic. In the first integral the variable of integration is z . After a change of variable from z to $z' = z/\alpha$, $z = p - \alpha y_0$ is equivalent to $z' = p' - y_0$, where and $p' = p/\alpha$. ■

Is the scaling condition stated in Proposition 5 a reasonable assumption for the cross-section of stocks? We consider the condition that the private value and participation cost scale with the common value of a security a reasonable first-order approximation. It is standard practice in financial economics to model those who trade to lock in a private value, as “noise traders.” Their desire for trade is expressed by adding a constant volatility noise term to a model of log price changes (e.g., Kyle, 1985). The distribution therefore scales with price level, i.e., the “size of the security.” This is a reasonable assumption to make as agents who are in the market for some fixed level of exposure to a security’s value, say \$1500, will submit an order of size 100 (securities) if the security’s (common) value is \$15 but only 10 if its value is \$150.

The participation cost scales with price as higher priced stocks are associated both with larger transaction sizes and larger firms (see Figure 3). Middlemen should be inclined to devote more attention and effort to larger transactions. But, as transaction size increases with firm size, they would have to process more information for the simple reason that the flow of information is larger

Figure 3: The cross-section: trade size and market cap vs. price

This figure shows how trade size and market capitalization scale with the price of a stock. They are based on the average of these variables in the five size quintiles of U.S. stocks (Hendershott, Jones, and Menkveld, 2011, Table I).



for these firms.⁷ It is therefore natural that the model's assumption that the cost c scales with price bears out in the cross-section; c is likely to increase with firm size, and therefore with transaction size, and therefore with price.

4.2 The distribution of bids relative to the highest bid

Let p denote a bid drawn from $H(p)$, let p^* be the highest bid, and let $P = p/p^*$ denote a bid relative to the highest bid. Assume $N \geq 2$.

The conditional distribution of P .—For $p^* \in [u, y - c]$ and

$$P \in \left[\frac{u}{p^*}, 1 \right] \Leftrightarrow p^* \in \left[\frac{u}{P}, y - c \right]. \quad (13)$$

Now, p^* is distributed with CDF $H^k(p^*)$. Conditional on k and p^* , the second-highest bid is the largest of $k - 1$ draws that are all lower than p^* . Since u and, hence, p^* are positive, $\tilde{P} \leq P \Leftrightarrow p \leq p^*P$. Therefore,

$$\Pr(\tilde{P} \leq P \mid k, p^*) = \Pr(p \leq p^*P \mid k, p^*) = \frac{H(p^*P)}{H(p^*)}, \quad (14)$$

which is one at $P = 1$.

Conditioning just on k , the distribution of P is (dropping the superscript $*$ from p)

$$\Psi(P \mid k) = \int_{u/P}^{y-c} \frac{H(pP)}{H(p)} dH^k(p) = k \int_{u/P}^{y-c} H(pP) H^{k-2}(p) h(p) dp \quad (15)$$

for $P \in \left[\frac{u}{y-c}, 1 \right]$, because $dH^k(p) = kH^{k-1}(p)h(p)$. When $k = 1$, the distribution of P is concentrated on $P = 1$, which is how we define $\Psi(P \mid 1)$.

The unconditional distribution of P .—When $k = 0$ there is no bid, and this event occurs with probability $(1 - \lambda)^N$. The unconditional distribution of P is therefore equal to:

⁷This might also explain the empirical fact that large firms get more analyst coverage (see, e.g., Barth, Kasznik, and McNichols, 2001).

$$\Psi(P) = \frac{1}{1 - (1 - \lambda)^N} \sum_{k=1}^N \Psi(P | k) \binom{N}{k} \lambda^k (1 - \lambda)^{N-k}. \quad (16)$$

The scalability of P .—We now show that $\Psi(P)$ is invariant to scaling.

Proposition 6 *For $\alpha > 0$, if $(c, x, y, \Phi(z))$ are scaled by α to $(\alpha c, \alpha x, \alpha y, \Phi(z/\alpha))$, the solution for Ψ is invariant to α .*

Proof. From (1), λ does not change with α , so that by (16) it suffices to prove that $\Psi(P | k)$ does not depend on α . The proof of Proposition 5 revealed that H and, hence, h are of the form $H(p/\alpha)$ and $h(p/\alpha)$. Then choosing an arbitrary α , the RHS of (15) becomes (after noting that $dH(p/\alpha) = h(p/\alpha) \frac{dp}{\alpha}$)

$$\int_{\alpha u/P}^{\alpha(y-c)} H^k\left(\frac{p}{\alpha} P\right) \frac{h(p/\alpha)}{H(p/\alpha)} \frac{dp}{\alpha} = \int_{u/P}^{y-c} \frac{H(wP)}{H(w)} h(w) dw$$

after a change of variable to $w = p/\alpha$. Therefore $\Psi(P | k)$ does not depend on α , which proves the claim. ■

The density of P .—The density does not have mass points as it is based on H which does not have any mass points either. We can therefore derive it by considering the full domain except for $P = 1$ so as to avoid technical difficulties. Now, the derivative w.r.t. P in the lower limit of the integral in (15), $(u/P)^2 \frac{H^k(u)}{H(u/P)} h(u/P)$ is zero, because $H(u) = 0$ since it cannot have mass points, which implies $\frac{H^k(u)}{H(u/P)} = H^{k-1}(u) = 0$. Therefore the density of P is

$$\psi(P) = \frac{1}{1 - (1 - \lambda)^N} \sum_{k=1}^N \psi(P | k) \binom{N}{k} \lambda^k (1 - \lambda)^{N-k}, \quad (17)$$

where

$$\psi(P | k) = k \int_{u/P}^{y-c} \frac{\partial}{\partial P} H(pP) H^{k-2}(p) h(p) dp = k \int_{u/P}^{y-c} p h(pP) H^{k-2}(p) h(p) dp. \quad (18)$$

At the lowest point in the support of P , namely $\frac{u}{y-c}$, the density $\psi(P | k) = 0$ for all k because the upper and lower limits of the integral on the RHS of (18) coincide. On the other hand, at $P = 1$, ψ is positive. That is, for all k ,

$$\psi\left(\frac{u}{y-c} | k\right) = 0 \quad \text{and} \quad \psi(1 | k) = \int_u^{y-c} H^{k-1}(p) \frac{p[h(p)]^2}{H(p)} dp. \quad (19)$$

The Poisson case ($N = \infty$). In the Poisson case

$$\Psi(P) = \frac{m}{1 - e^{-m}} \int_{u/P}^{y-c} e^{-m(1-H(p))} \frac{H(pP)}{H(p)} h(p) dp \quad (20)$$

which is one at $P = 1$. This result is derived in Appendix D. The density is

$$\psi(P) = \frac{m}{1 - e^{-m}} \int_{u/P}^{y-c} e^{-m(1-H(p))} \frac{1}{H(p)} p h(pP) h(p) dp. \quad (21)$$

5 Estimation of the model

We estimate the model using bid-price distributions of financial securities that are actively traded through centralized limit-order books. This environment is particularly attractive for estimating our model for a couple of reasons. First, financial securities are largely common-value “goods” as they represent claims to future cash flows.

Second, high-frequency traders bidding in a centralized market could reasonably be thought of as competition among homogeneous middlemen. HFTs are likely to have the same information sets as argued in section 6. Boehmer, Li, and Saar (2015) use Canadian data to show that HFT market-making strategies are highly correlated, thus lending some support to this assumption of homogeneous middlemen. The centralized market further creates homogeneity as middlemen cannot, for example, benefit from a better position on a network that is often used to characterize a decentralized, over-the-counter (OTC) setting.

Third, the limit-order trading protocol ensures that incoming market orders are matched with the best price quote. It therefore is a first-price auction.

Fourth, submitting a limit order to the exchange is costly. HFTs have extremely large, but not infinite computer capacity and (colocation) bandwidth to submit (and maintain) price quotes for a large set of securities. Now that exchanges clock at a microsecond (one millionth of a second) frequency, HFT capacity becomes a binding constraint and submitting a price quote therefore entails the shadow cost of not being able to do something else at that instant of time.

Finally, even if one believes HFTs revisit the market an order of magnitude more often than investors do, a non-degenerate distribution still emerges. For example, if **S** showed up with probability α whereas **B** always showed up, with risk-neutral bidders this amounts to raising the participation cost by a factor of $1/\alpha$. If the arrival probability was constant, the same game is repeated and the equilibrium bid distribution would not change since it is the outcome of a unique equilibrium play.

5.1 Data

The data pertain to one of the most discussed days in recent trade history: May 6, 2010. In just a couple of minutes U.S. stock index securities (e.g., E-mini index futures and ETFs) along with index constituent stocks experienced steep price declines. Prices recovered in about the same time span. The event came to be known as the Flash Crash. It created investor anxiety, media attention, and substantial follow-up by the SEC who published a detailed report later that year (SEC, 2010b). The report zeroed in on massive selling by a single trader in the E-mini market as a key contributing factor.

Another important reason to pick this day is that data on the full order-book distribution of all S&P 500 member stocks is publicly available. SEC (2010b) reveals such information for a snapshot taken at every full minute of the day. Figure 4 depicts minute by minute snapshots of the bid price distribution in the combined order books of NYSE, NASDAQ, and BATS in the half-hour interval

Figure 4: Aggregate order book of S&P 500 stocks on May 6, 2010

This figure graphs the order book evolution in the half hour of the Flash Crash on May 6, 2010. The color bands are used to represent order book liquidity supply. For example, the lightest blue band reveals how many shares were bid for at prices between the midquote and the midquote minus 10 basis points (the midquote is the average of the best bid and ask price). “Minimum Executed Price” refers to the minimum trade price in each minute and “Net Aggressive Buy Volume” sums across the size of all trades in a minute where buyer-initiated trades (execution at the ask quote) get a positive sign and seller-initiated trades get a negative sign. The graph was obtained from SEC (2010b, p. 34). It combines information from NYSE Openbook, Arcabook, NASDAQ ModelView, and BATS.

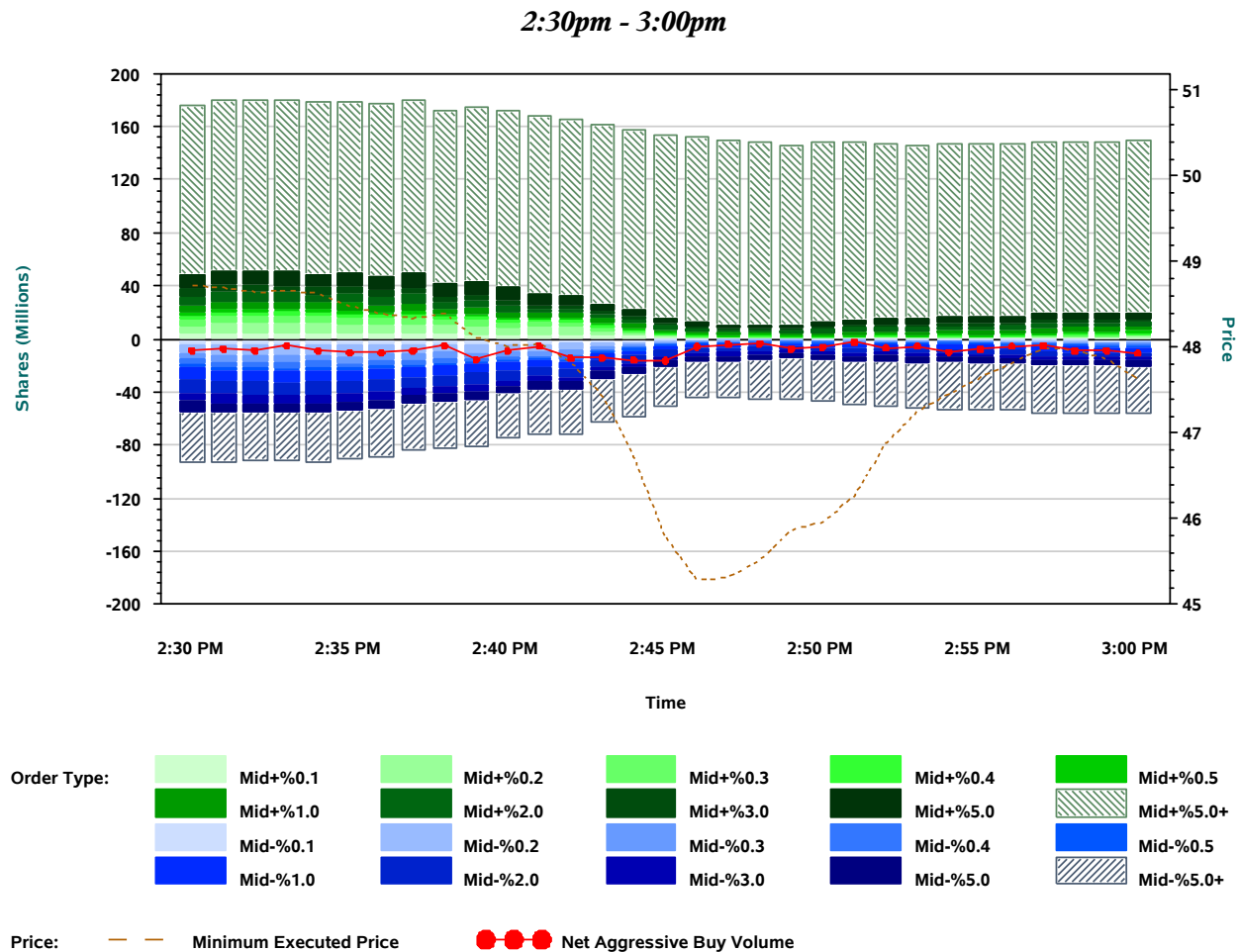
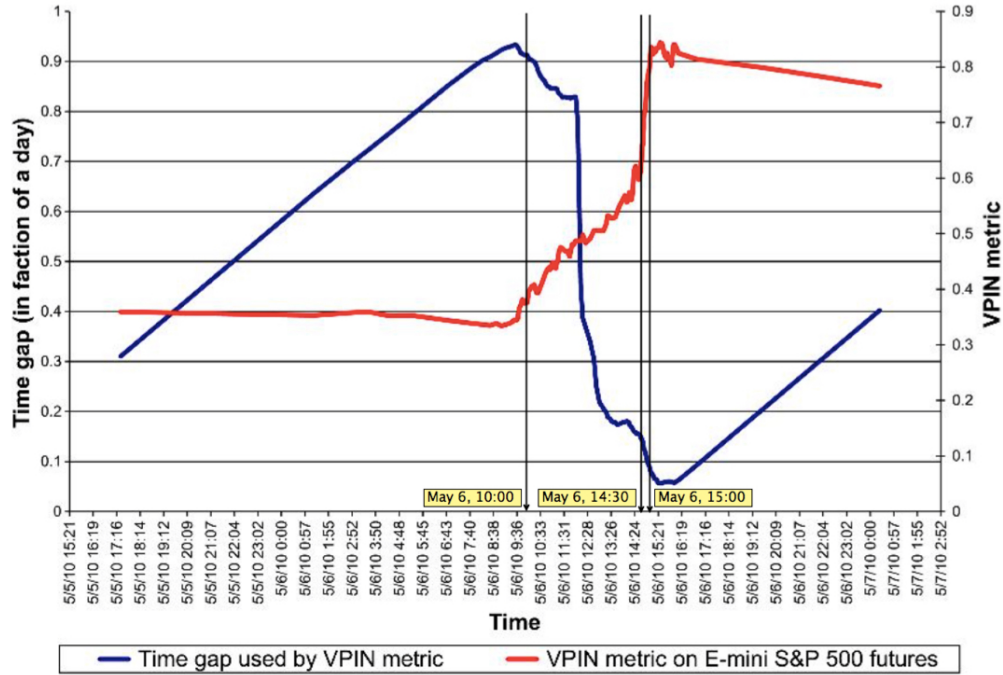


Figure 5: Order flow toxicity on May 6, 2010

This figure is taken from Easley, López de Prado, and O'Hara (2011, Fig. 2). The authors plot the “toxicity” of order flow in the period leading up to the Flash Crash. Three vertical lines have been added to emphasize the time points we focus on in our analysis: 10:00 a.m., 2:30 p.m., and 3:00 p.m.



of the Flash Crash (SEC, 2010b, Chart 1B, p. 34).⁸ The color bands correspond to the aggregate amount of shares available at various bid price ranges. The graph shows that at 2:30 p.m., about 30 million shares were available to a seller at a price range from the bid-ask midpoint to 1% below that midpoint. Another 30 million were available in the range from -1% to -5%. In total there were about 100 million shares available for sale. The snapshot pattern suggests a monotonically increasing probability density function for bids, which is consistent with the theoretical result that h increases in p , see “skewness” paragraph on page 6 and Figure 2. This is encouraging first evidence in favor of the model.

We pick four time points to estimate the model’s primitive parameters.

⁸The report has a similar picture for the full day but not for other days.

- 10:00 a.m.: The 10:00 a.m. order book snapshot captures a “normal day” bid-price distribution. Easley, López de Prado, and O’Hara (2011) plot E-mini order flow “toxicity” for the period leading up to the Flash Crash. The plot reveals that toxicity was at a normal level at the start of May 6, but then steadily rose in the course of the day (see Figure 5 which was taken from their paper). The equity markets open at 9:30 a.m. We picked 10:00 a.m. as representative of a normal market to avoid any contamination by heavy trading in lieu of the opening auction.
- 2:30 p.m.: The start of the Flash-Crash half-hour. Menkveld and Yueshen (2015) report that, by that time, the large seller had not initiated selling yet. Two minutes later he started.
- 2:46 p.m.: The market reached its deepest point. It is the snapshot just after a five-second trading halt in the E-mini market. 2:30 p.m. is a somewhat arbitrarily chosen time point. We could also have taken the snapshot just before the halt. It turns out to be irrelevant as the price dispersion is very similar in the minutes before the halt (see Figure 4).
- 3:00 p.m.: The last snapshot in the Flash-Crash half-hour.

5.2 Estimation

The bid-price dispersions depicted in Figure 4 are quite useful for a number of reasons. First, the aggregation across the 500 index stocks smooths out the noise in stock-specific price dispersions (note that Proposition 5 allows us to meaningfully aggregate relative price distributions across stocks in the model). Second, middlemen are particularly active in equity as an SEC report earlier that year states: “... estimates of HFT volume in equity markets vary widely, though they typically are 50% of total volume or higher (SEC, 2010a, p. 45).” Third, the data are comprehensive in that they not only have the supply of shares at a few best price levels as is often the case, but they also include total supply. One needs the latter to characterize skewness or “aggressiveness” at the top of the book.

Model to be taken to the data.—We decide to preset two model parameters and estimate the other two. First, the *ex-ante* number of HFTs is set to infinity. The number of middlemen must exceed 12 as, in their Flash-Crash report, the SEC identified 12 HFTs who are active in these markets (SEC, 2010b, p. 45). We believe that $N = \infty$ is a relatively innocent choice as the bidding functions seem to quickly converge when N gets large (see Figure 2 which includes both $N = 12$ and $N = \infty$).

Second, since the distribution of bids relative to the best bid, $\Psi(P)$, is homogeneous of degree zero in (c, u, y) , the buyer’s value, y , is set to one. The estimates for c and u should therefore be interpreted as *relative values*, i.e., they are measured as fractions of y . The parameters that will be estimated are c and u . We will further analyze time variation in u as variation in adverse-selection risk for the seller. The “extended” version of the baseline model, the HFT application presented in Section 3, allows us to do so by setting x to zero, and allowing σ to vary, see (9); the maximum gains-from-trade are normalized to one. Note that this indeed allows us to interpret the parameter σ as essentially the size of the adverse-selection friction between the seller and the buyer.

Estimation routine.—The two model parameters are estimated by matching the model-implied CDF with the empirical CDF. The procedure involves minimization of a sum of squared errors.⁹ The summation is across the following price levels: the bid-ask midpoint -0.2%, -0.3%, -0.4%, -0.5%, -1%, -2%, -3%, and -5%. A detailed description of the estimation procedure is in Appendix B.

Estimates.—Table 1 reports the parameter estimates for the four time snapshots. Before discussing their values, we first analyze the fit as the model is extremely parsimonious. The homogeneity assumption restricts the set of parameters to only two: u and c .

One standard procedure to verify equality of distributions is the Kolmogorov-Smirnov test. It

⁹It is essentially a standard moment-matching exercise as a CDF evaluated at a particular value X is a moment, i.e., the expectation of a dummy that is one for a value less than X and zero otherwise. Standard GMM estimation is not feasible as the individual data points are not available to us, only the empirical moments are known.

Table 1: Parameter estimates

This table presents the parameter estimates for the bid-price distribution of S&P 500 stocks at four time points on May 6, 2010. All values are to be interpreted as values *relative to* the maximum gains-from-trade (as buyer private value y is set to one and seller private value x is set to zero). The two estimated parameters are middleman participation cost c and the seller's outside option u . We assume there are infinitely many middlemen around *ex ante*, i.e., $N = \infty$. The Kolmogorov-Smirnov test statistic is reported to test whether the empirical distribution is significantly different from the distribution implied by the fitted model. This is true if it exceeds the critical level. The table further presents the values the parameter estimates imply for various other model variables: the number of middlemen that are expected to show up *ex post* m , middleman bid aggressiveness r , defined in (5), and common-value volatility σ . The relative social cost of having infinitely many middlemen around is gauged by the welfare differential between having infinitely many of them around (W_∞) and having only two around (W_2). The latter is the best outcome from a planner's perspective. The welfare values are computed based on (6).

Time snapshot	10:00 a.m.	2:30 p.m.	2:46 p.m.	3:00 p.m.
<i>Parameter estimates and fit</i>				
c	0.0017	0.0007	0.0043	0.0050
u	0.53	0.66	0.15	0.23
Kolmogorov-Smirnov test statistic (95% critical value)	0.051 (0.072)	0.031 (0.087)	0.010 (0.053)	0.037 (0.051)
<i>Other values implied by parameter estimates</i>				
m	3.32	6.17	5.28	5.04
r	28	486	198	154
σ	0.28	0.18	0.88	0.68
W_2	0.464	0.339	0.841	0.760
W_∞	0.456	0.335	0.823	0.740
$W_2 - W_\infty$	0.008	0.004	0.019	0.020
$\frac{W_2 - W_\infty}{W_\infty}$	0.017	0.011	0.022	0.027

involves computing the maximum distance between the fitted and the empirical distribution. In our case, the test statistic is:

$$\max_{i \in \{1, \dots, 8\}} |\Psi(P_i) - \hat{\Psi}(P_i)|$$

where Ψ and $\hat{\Psi}$ are the fitted and empirical CDF respectively. In our application we only have eight values at which we can evaluate the distribution (instead of all values in the support) but, for each of them, we have that it is based on 500 stocks. We therefore cannot use the standard distribution of the test statistic, but use simulations to establish the distribution of this modified Kolmogorov-Smirnov statistic (see Appendix C for details). The table illustrates that we cannot reject the null of equality for each of the four time points at a 5% significance level.

Figure 6: Empirical and fitted bid-price distribution

This figure illustrates the estimation result by plotting both the realized and the fitted price dispersion across all S&P 500 stocks. The estimation is done separately for four time points on May 6, 2010, the day of the Flash Crash: 10:00 a.m., 2:30 p.m., 2:46 p.m., and 3:00 p.m. The top graphs depict the empirical and the fitted CDFs. The bottom graphs depict the corresponding PDFs. The empirical CDFs correspond to the color bands in Figure 4. Relative prices were obtained by dividing each bid price quote by the bid-ask midquote. The estimated parameter values are added on top of each graph. The estimated parameters are u and c ; y was set to one and N was set to infinity.

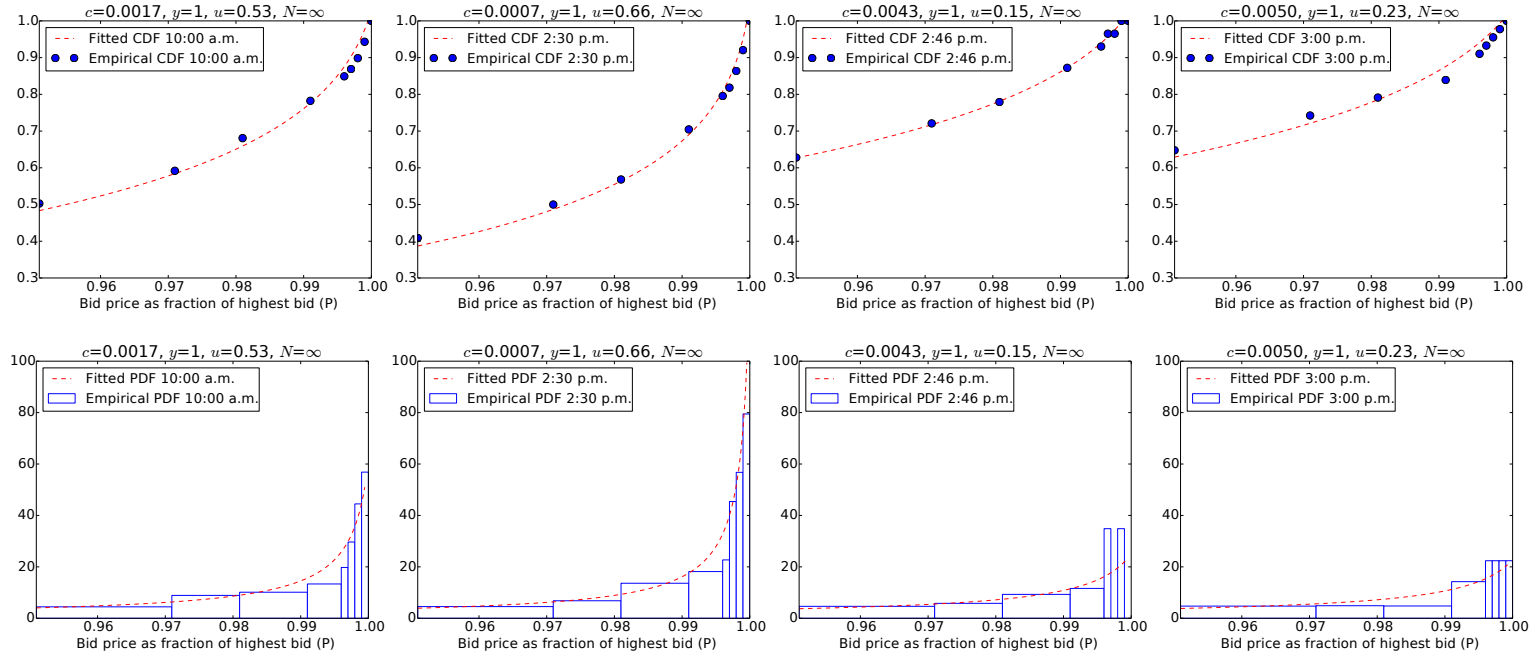


Figure 6 illustrates the results of the estimation. The top graphs illustrate the empirical and model-implied CDFs corresponding to 10:00 a.m., 2:30 p.m., 2:46 p.m., and 3:00 p.m., respectively. The model-implied CDFs denoted by the dashed line in these graphs are close the eight points that were used to fit the empirical CDF. This is remarkable as the model has only two parameters to create the fit, c and u , the cost of participation for middlemen and the value of the outside option for the seller respectively. Their estimated values are reported at the top of the graphs. Finally, the bottom graphs illustrate the implied PDFs for the four time points.

In the next couple of paragraphs we will discuss the time series pattern in parameter estimates and what they imply for other variables in the model. We caution that, throughout, we interpret time series patterns *relative to* maximum gains-from-trade. The reason is that all model variables are calculated for a model where maximum gains-from-trade are normalized to one. If one assumes these gains-from-trade remain constant throughout the sample period, then these interpretations can be taken literally.

The estimated parameter values suggest that it became very costly for middlemen to participate, yet sellers needed them as their outside option declined in value. The participation cost c declined somewhat leading up to the crash, from 0.0017 at 10:00 a.m. to 0.0007 at 2:30 p.m. It then shot up to 0.0043 in the middle of the crash (2:46 p.m.) and increased somewhat further to 0.0050 right after the market recovered (3:00 p.m.).

One possible reason for the extreme cost of participating in the crash and its aftermath is that middlemen needed to exert more effort to process the data as data feeds became unreliable. In their Flash Crash report, based on data analysis and interviews with market participants, the SEC emphasized that the publicly distributed data experienced integrity issues. This automated data stream, for example, reports the “national best bid offer” (NBBO) in real time. This NBBO is the highest bid across all exchanges, along with the lowest ask across these exchanges. If some exchange experiences delay in reporting their best bid offer to the consolidator, then the national best bid might be above the lowest national offer. This is unlikely to occur other than for a fleeting

moment in normal market conditions. Arbitrageurs will be quick to lock in a profit by selling to the highest bid and buying from the lowest ask (at the two exchanges where these quotes originate from).¹⁰

Alternatively, middlemen themselves might have become capacity constrained when trying to process all data so that, for each security, the shadow cost of putting together a price quote increased. The SEC report (SEC, 2010b) states:

“Some firms experienced their own internal system capacity issues due to the significant increase in orders and executions they were initiating that afternoon, and were not able to properly monitor and verify their trading in a timely fashion.” (p. 36)

The outside option for the seller (u) was 0.53, 0.66, 0.15, and 0.23 at the four consecutive time points. It increased somewhat leading up to the crash, but then suddenly dropped to very low levels in the middle of the crash, only to show partial recovery after prices rebounded. Apparently, in the context of the model, the adverse-selection cost was substantially higher for sellers in the crash period, and stayed at elevated levels right after. Note that the implied common-value volatility pattern in Table 1 corroborates this interpretation; it is 0.28, 0.18, 0.88, and 0.68, respectively. This is arguably due to the rare nature of the event.

It would be costly for anyone to make “free options” available for others to consider, i.e., price quotes. This however is particularly true for end-users, i.e., sellers who are further removed from the inner circle of the market. The SEC report notes that the participants most exposed to the data integrity issues described above are, indeed, the buyer and seller type in our model:

“We note, however, that while these types of firms are not generally market makers or liquidity providers, they can be significant fundamental buyers and sellers.” (p. 36)

Finally, we note that the best time for middlemen to operate was at 2:30 p.m., just before the

¹⁰Note that such arbitrage trade is costly as traders typically pay a fee on market orders. The price discrepancies at the time of the crash were of such magnitudes that they dwarfed exchange fees. A detailed discussion of these data integrity issues is in (SEC, 2010b, Section III.3).

Crash. The number of middlemen expected to show up *ex post* is highest, $m = 6.17$, and they bid most aggressively, $r = 486$. Notice that this is a non-trivial result. Middleman participation cost c is at its lowest level ($c = 0.0007$) but the seller's outside option is at its *highest* level (i.e., the seller's option of posting himself is attractive, $u = 0.66$). The net effect of more aggressive middlemen must be due to the order of magnitude of these extreme values. That is, the ratio of the two, c/u , is at its lowest level at 2:30 p.m.: 0.001. This ratio at the other time points is at least 20 times higher. The relative decline in participation cost must therefore dominate the relative increase in the seller's outside option, and this could explain why we find the middlemen to be most aggressive at 2:30 p.m.

Welfare comparisons. Next, we use the parameter estimates to study the time series pattern of the social cost of having (infinitely) many middlemen around. It reaches its highest levels in the post-crash period. In the model, one only needs two middlemen to reap the full benefit of competition. Each additional middleman only adds cost to the economy. To judge when the cost of having too many middlemen around is highest we compute the welfare differential between having two middlemen around versus infinitely many: $W_2 - W_\infty$. It is no surprise that this cost is lowest right before the crash as middleman cost is at its lowest level, and the seller's outside option is at its highest level.

It is interesting that the post-crash welfare differential is much higher than the morning differential in spite of a much lower middleman cost in the afternoon. For example, at 3:00 p.m. the relative differential is 2.7% and middleman cost is 0.0050 whereas the 10:00 a.m. differential is only 1.7% while middleman cost is 0.0170. This illustrates that the value of the seller's outside option is just as important for welfare. The value of this option is much lower in the afternoon so, in some sense, the seller is driven into the arms of middlemen out of pure necessity. Middlemen collectively respond by increasing their participation probability: m decreases in u , see (7). The estimate for m is 5.04 at 3:00 p.m. relative to 3.32 at 10:00 a.m., an increase of 52%. The aggregate

participation cost therefore increases, which adds cost to the economy.

Parameter estimates relative to other literature. How do our parameter estimates compare to earlier work? Let us first focus on the participation cost c . Most closely related are Sandås (2001) and Hollifield, Miller, and Sandås (2004) who both estimate a structural model of a limit-order market. Sandås (2001) estimates Glosten (1994) and finds a puzzling negative “order-processing” cost. Hollifield, Miller, and Sandås (2004) propose a model with participation cost c but set it to zero in the estimation. Other studies report the explicit part of participation cost, i.e., the fee that one pays when submitting an order (e.g., an exchange fee). Bodurtha and Courtadon (1986), for example, report a 12 basis-point fee for trading in foreign currency options in the mid-eighties. Colliard and Foucault (2012, Figure 1) documents equity fees in the late zeros that range from 1 to 12 basis points.¹¹ The estimation of our model suggests that (total) participation cost is between 7 and 17 basis points ahead of the crash, and between 43 to 50 points at and after the crash. These estimates are therefore of the same order of magnitude.

The outside option value for the seller, u , is at low levels before the crash, and at extremely low levels during and after the crash. In their limit-order model, Hollifield, Miller, and Sandås (2004, Table VII) estimate that the standard deviation of private values is 21% (relative to common value). This implies that, under normality, the expected gains from trade (GFT) are $\sqrt{\frac{2}{\pi}} \times \sqrt{2} \times 21\% \approx 24\%$.¹² The standard deviation of GFT is $\sqrt{\sqrt{2} \times 0.21^2 - 0.24^2} \approx 7\%$. The u in our model is expressed in terms of buyer valuation. Its 10:00 a.m. value of 0.53 therefore implies that only if the gains-from-trade are larger than 47% will the seller consider this option. Chronologically, our u estimates imply an outside-option cost for the seller that is 4.1, 1.4, 8.7, and 7.6 GFT standard-deviations above average GFT, as implied by the Hollifield et al. study. The level of u is therefore low for the two pre-crash snapshots, and extremely low for the crash and the post-crash snapshot.

¹¹Some equity exchanges applied a maker/taker model where a limit-order submitter receives a “maker” rebate upon execution, and the market order that executes against it is charged a “taker” fee that is slightly higher.

¹²Note that the expected value of $|X|$ with $X \sim N(0, \sigma)$ is $\sqrt{\frac{2}{\pi}}\sigma$; private values are orthogonal by definition.

In a recent study, Yueshen (2015) identifies price dispersion for U.S. stocks based on the time series dynamics the midquote price (the average of the bid and ask quote) and signed order flow. His empirical identification of dispersion is essentially in the extent to which midquote returns “excessively” respond to the information in order flow. Interestingly, his results suggest that the price dispersion is an order magnitude larger than the long-term price impact of order flow. For 2010, he estimates it to be four to five times larger, and there is strong upward trend as of 2005. Although not directly comparable to our estimates, both studies suggest that price dispersion is sizeable, characteristic of modern markets, and worthy of understanding.

Finally, we want to reiterate that our estimates come from a homogeneous-agent model. Other structural models, including the two limit-order models referred to above, have some heterogeneity in their primitives (e.g., a dispersion in informativeness of market orders or agents’ private values). Such models might fit the bid price distribution equally well, but come at the cost of additional free parameters (that characterize the dispersion). Other papers that fit price-distribution data using models with homogeneous price-setting agents include Head et al. (2012) and Kaplan and Menzio (2014).

6 Literature survey

The paper relates to three bodies of literature: Auctions, Bertrand pricing, and Search. Most papers focus on ask instead of bid prices and find positive instead of negative skewness. Like us, however, they also have that, for one reason or another, the number of “bidders” is random.

In the auctions literature, Hausch and Li (1993), Piccione and Tan (1996) and Cao and Shi (2001) allow bidders to choose whether to bid and to acquire a costly signal about the common value. When signals are coarse, several experts will generally have seen the same level of signal, and if the number of such bidders is not common knowledge, it is an equilibrium for them to then use a mixed strategy. Moreover, Cao and Shi (2001) and Silva, Jeitschko, and Kosmopoulou (2009)

find that the *ex-post* rent (excluding *ex-ante* “participation” cost) of bidders rises as their number grows. We find that this is not necessarily the case, in particular when a is high and N is low.¹³ But that the number of bidders endogenous, random, and not common knowledge among bidders, and that having more potential bidders is not necessarily better are issues that the auction literature has discussed, beginning with Harstad (1990).

The above papers derive price distributions from an uncertainty that a player has over how much competition he faces when setting his price — ask or bid as the case may be. By contrast, in papers on Bertrand style competition price dispersions arise when one adds fixed costs and monopoly power. Absent such monopoly power a firm sets the price equal to marginal cost. In Shilony (1977), for example, monopoly power arises as sellers are spatially dispersed and buyers pay a transportation cost to travel between locations. In Rosenthal (1980) there is no such exogenous fixed cost, but a firm can participate in market-wide price competition by giving up profit in some captive segment of the market. This foregone profit is the counterpart of entry cost in our model.

In the search literature, Butters (1977) has multiple agents on both sides of the market and the process by which price quotes reach customers is random. It is an auction for customers in which a bidder does not know how many actual other bids he is competing with. Butters takes limits as the number of firms and the number of customers get large; he does not calculate the equilibrium when N is finite.¹⁴ Burdett and Judd (1983) is also closely related, they derive non-degenerate ask-price distributions in a similar context and for the same reason — the firm is not sure whether or not its customers will see other bids; the distribution of information over customers they take to be exogenous whereas we endogenize it.

The above models all assume, as we do, that the price setters are identical *ex ante*, but choose

¹³The expected number of bidders present *ex post* declines in N (see panel (a) in Figure 1). Therefore, per bidder, the expected participation cost has to decline in N and since a bidder is on a zero profit condition, his expected rent has to decline as well.

¹⁴In Butters’ model the outcome of skewness of asking prices is confounded by the ability of firms to choose their advertising intensity; firms that charge higher asking prices advertise more intensively since a sale yields a higher profit. This makes high asking prices more profitable than they would be if firms were not able to advertise.

different prices *ex post*. We believe that in our application to bidding by HFTs this is approximately correct. HFTs are programs run on computers. Their information sets are arguably very similar (they will parse anything available in digital form, e.g., recent trades or quotes in correlated securities, press releases, etc.). Moreover, as order book information (i.e., outstanding quotes) is revealed almost instantaneously (in microseconds) to HFTs, there is little room for information heterogeneity to persist among them. Information heterogeneity alone is therefore hard to reconcile with, for example, Hasbrouck (2015, Fig. 1) which shows that the best bid for a U.S. stock shows bursts of extreme volatility that persists for minutes.

7 Conclusion

In a model with homogeneous bidders, we have solved for a unique distribution of bids for a homogeneous good. We distinguished the forces determining the dispersion of bids from the forces determining the degree of negative skewness in the bid distribution, and we found that only the latter depends on the number of potential bidders. We then fitted the model to the bid price distribution as conveyed by the limit-order book of S&P 500 stocks.

With just two free parameters in the estimation, the simple analytic solutions fit the data surprisingly well. This suggests that two salient features of the model capture trade frictions quite well: Middlemen incur non-zero participation costs and are unable to coordinate participation decisions. This inability is what ultimately delivers the non-degeneracy of the price distribution that the unique equilibrium entails.

Appendix

A Notation summary

The following table summarizes the notation used throughout the manuscript.

a	Relative cost of bidding, i.e., $a = c/(v - u)$
B	Buyer in the HFT game
c	The price a bidder must pay to participate in the bidding game
F	Distribution of the winning bid
Φ	Distribution of common value in HFT game
h	The PDF of the bid price a bidder submits, a choice variable for a bidder
H	The CDF of the bid price a bidder submits, a choice variable for a bidder
k	The number of middlemen who show up <i>ex post</i> , i.e., those who (probabilistically) decided to participate
λ	The probability of play, a choice variable for a bidder
m	The expected number of bidders for $N \rightarrow \infty$, see (7), note $m = -\ln(a)$
M	Middlemen in the HFT game
N	Number of candidate bidders
p	A bid price
P	A bid price divided by the highest bid
Ψ	CDF of bid prices divided by the highest bid
ψ	PDF of bid prices divided by the highest bid
r	Bid aggressiveness as defined in (5)
S	Seller in the HFT game
u	The seller's reservation value for the object
U	The seller's outside option in the HFT game
v	The value of the object to a bidder
W	Expected net gains from trade (welfare)
x	Private value of object to seller in HFT game
y	Private value of object to buyer in HFT game
z	Common value of object value in HFT game, added to private value for each participant

B Description of the estimation procedure

This appendix describes how the model is estimated to fit the realized bid price dispersion for stocks in the S&P 500 index. In short, the empirical strategy involves two steps. First, instead of

actual price dispersion, the price dispersion relative to the best bid is calculated. The “scalability” property for the latter distribution is invoked to aggregate across stocks (Proposition 6). Second, the sum of squared differences between the model-implied CDF of prices relative to the best bid and the “empirical CDF” for these relative bid prices is minimized to identify the primitive parameters. We shall set $N = \infty$ and $y = 1$ for reasons discussed in the section 5.2.

Model-implied CDF of bid-to-best-bid ratio. The model-implied CDF of P , the ratio of a bid to the best bid, expressed in the primitive parameters is

$$\Psi(P) = \frac{\ln(1-u) - \ln(c)}{1 - c/(1-u)} \int_{u/P}^{1-c} \left(\frac{c}{1-u} \right)^{1-H(p)} \left(\frac{H(pP)}{H(p)} \right) h(p) dp \quad \text{for } P \in \left[\frac{u}{1-c}, 1 \right]$$

in which

$$H(p) = 1 + \frac{1}{\ln(1-u) - \ln(c)} \ln \frac{c}{1-p} \quad \text{for } p \in [u, 1-c]$$

and

$$h(p) = \frac{1}{\ln((1-u)/c)} \frac{1}{1-p}.$$

This CDF is derived combining (8) and (20).

Empirical CDF of bid-to-best-bid ratio. The empirical CDF for all stocks in the S&P 500 is taken for four snapshots of the aggregate bid price distribution for May 6, 2010, the day of the Flash Crash: 10:00 a.m., 2:30 p.m., 2:46 p.m., and 3:00 p.m. The data are taken from SEC (2010b, Chart 1B, partially shown in our Figure 4). The graph allows us to derive the empirical CDF by observing the supply of shares at eight relative bid price levels: the bid-ask midpoint -0.2%, -0.3%, -0.4%, -0.5%, -1%, -2%, -3%, and -5%.¹⁵ To compute the empirical CDF at these price levels we

¹⁵The estimation implicitly assumes that the size of the bid-ask spread is negligible in the following sense. The relative prices available are all measured relative to the bid-ask midpoint as opposed to the best bid for which the model-implied CDF that was derived in (20). The distance between the midpoint and the best bid is equal to half the bid-ask spread. This spread is most likely smaller than 0.1% for the S&P 500 stocks (see, e.g., summary statistics in Hendershott, Jones, and Menkveld, 2011). The bin that contains all bids from the midpoint to -0.1% was left out as

also need to observe the total number of shares offered at the bid. This information is retrieved by adding what is available at price levels below 5%, i.e., the cross-hatched areas in Figure 4. The empirical CDF is then computed as the number of shares supplied in a bin up to a price level, divided by the total number of shares supplied.

Approach. We minimize the following objective function with respect to (c, u)

$$L(c, u) \equiv \sum_{i=1}^8 \left(\Psi(P_i) - \hat{\Psi}(P_i) \right)^2,$$

where $\hat{\Psi}$ is the empirical CDF, and P_i correspond to the bid-to-best-bid ratio associated with the bid-ask midpoint -0.2%, -0.3%, -0.4%, -0.5%, -1%, -2%, -3%, and -5%, respectively.

C Distribution of the modified Kolmogorov-Smirnov statistic

The critical value of the modified Kolmogorov-Smirnov (MKS) statistic is obtained by simulation. The procedure involves the following steps:

1. Sample 500 times from the estimated distribution (as there are 500 stocks in the S&P500 index).
2. Compute the empirical distribution value, say $\tilde{\Psi}(P_i)$, that it implies for the eight relative bid price levels that we have data for, i.e., $P_i = -0.2\%, -0.3\%, \dots, -5\%$ (this now corresponds to one observation from the data-generating process that we assume generated our sample);
3. Calculate the simulated value of the MKS statistic

$$MKS = \max_{i \in \{1, \dots, 8\}} |\tilde{\Psi}(P_i) - \hat{\Psi}(P_i)|.$$

that observation is most affected by the assumption.

Repeat to obtain a distribution for the MKS statistic. The 95% critical value correspond to the (empirical) quantiles.

D Proofs

Proof that no asymmetric equilibria exist, missing part in proof of Proposition 1. This part of the proof was moved here because of its straightforward nature.

No pure strategy equilibria exist.—Suppose, on the contrary, that the equilibrium number of players that enter the bidding was k . Then $k = 0$ is not an equilibrium for then a sole entrant would bid $p = u$, obtain the object, and earn $v - c > 0$. Also, $k \geq 2$ is not an equilibrium for after the entry cost was sunk, firms would Bertrand compete on bids and all set $p = v$ and earn zero rents *ex post*, and would therefore be unable to cover the entry cost c . Finally, $k = 1$ is not an equilibrium, for then the sole bidder would bid $p = u$ and would collect a positive profit. But this would invite a second entrant who could bid $\varepsilon < v - u - c$ and for ε small enough would win the object and make a positive profit.

No asymmetric mixed strategy equilibria exist. Denote the mixed strategies by $(\lambda_j, H_j)_{j=1}^N$. First we take the case $N = 2$. Player i 's indifference about entering implies that for p in the support of H_i

$$c = (v - p) \left[1 - \lambda_j + \lambda_j H_j(p) \right], \quad (\text{A1})$$

which means that if for some p we had $H_i(p) \neq H_j(p)$, it would follow that $\lambda_i \neq \lambda_j$. Let us say $\lambda_i > \lambda_j$. But then at the lowest bid of $p \equiv p_{\min}$ (assuming for the moment that this minimum is the same across both distributions), the expected profit $(v - p_{\min})(1 - \lambda_j)$ for i would exceed those for j . But then player j 's expected profit at $p_{\min}$ must be less than c , and therefore $p_{\min}$ would be outside of the support of H_j . Hence the supports of H_i and H_j must differ, but that cannot be optimal (for the same reason as that for why in Proposition 1, H cannot have holes). Next, the case

$N > 2$. Again there must be a pair of players for whom $\lambda_i \neq \lambda_j$. But for player i we must have $\frac{c}{v-u} = \prod_{s \neq i} (1 - \lambda_s)$, and for player j we must have $\frac{c}{v-u} = \prod_{s \neq j} (1 - \lambda_s)$. Taking the ratio of these conditions we obtain

$$\frac{1 - \lambda_i}{1 - \lambda_j} = 1 \Rightarrow \lambda_i = \lambda_j.$$

And the argument then proceeds as it did for the case $N = 2$.

Proof of more-middleman results of Proposition 2. More middlemen *ex ante* implies a higher probability of no-middleman present *ex post* and a lower probability of winning for single player.

Since $(1 - \lambda)^N$ is the probability of there being no winner and since each player is equally likely to win, p_N is the resulting expression. Formally,

$$\begin{aligned} p_N &= \lambda \sum_{k=0}^{N-1} \frac{1}{1+k} \binom{N-1}{k} \lambda^k (1-\lambda)^{N-1-k} = \lambda \sum_{l=1}^N \frac{1}{l} \binom{N-1}{l-1} \lambda^{l-1} (1-\lambda)^{N-l} = \\ &= \lambda \sum_{l=1}^N \binom{N}{l} \frac{1}{\lambda N} \lambda^l (1-\lambda)^{N-l} \end{aligned}$$

because $l(N-1-(l-1))!(l-1)! = (N-l)!l!$. To show that p_N declines in N it suffices to show that $(1 - \lambda_N)^N$ is increasing in N . Now because $a \in (0, 1)$:

$$(1 - \lambda_{N+1})^{N+1} = a^{\frac{N+1}{N}} = a^{1+\frac{1}{N}} > a^{1+\frac{1}{N-1}} = a^{\frac{N}{N-1}} = (1 - \lambda_N)^N.$$

The Poisson case ($N = \infty$). In the Poisson case

$$\begin{aligned}
\Psi(P) &= \frac{1}{1 - e^{-m}} \sum_{k=1}^{\infty} \int_{u/P}^{y-c} H(pP) H^{k-2}(p) h(p) dp \frac{m^k e^{-m}}{(k-1)!} \\
&= \frac{m}{1 - e^{-m}} \int_{u/P}^{y-c} \left(\sum_{k=1}^{\infty} H^{k-1}(p) \frac{m^{k-1} e^{-m}}{(k-1)!} \right) \frac{H(pP)}{H(p)} h(p) dp \\
&= \frac{m}{1 - e^{-m}} \int_{u/P}^{y-c} e^{-m(1-H(p))} \left(\sum_{k=1}^{\infty} H^{k-1}(p) \frac{m^{k-1} e^{-mH(p)}}{(k-1)!} \right) \frac{H(pP)}{H(p)} h(p) dp \\
&= \frac{m}{1 - e^{-m}} \int_{u/P}^{y-c} e^{-m(1-H(p))} \frac{H(pP)}{H(p)} h(p) dp \tag{A2}
\end{aligned}$$

which is one at $P = 1$. To show this, note that when $P = 1$, the integral above becomes

$$\int_u^{y-c} e^{-m(1-H(p))} h(p) dp = - \frac{1}{m} e^{-m(1-H(p))} \Big|_u^{y-c} = \frac{1 - e^{-m}}{m}.$$

Limit of λN as $N \rightarrow \infty$. Since $\lambda \rightarrow 0$, $\lim_{N \rightarrow \infty} \lambda N = \lim_{N \rightarrow \infty} \lambda(N-1)$. Therefore let $S \equiv N-1$ and let $a = c/(v-u) < 1$. Then $\lim_{N \rightarrow \infty} \lambda N = \lim_{S \rightarrow \infty} S(1 - a^{1/S})$. Letting $s = 1/S$, $\lim_{S \rightarrow \infty} S(1 - a^{1/S}) = \lim_{s \rightarrow 0} \frac{1-a^s}{s} = -\ln a$, which implies (7).

References

- Angel, James J., Lawrence E. Harris, and Chester S. Spatt. 2015. “Equity Trading in the 21st Century: An Update.” *Quarterly Journal of Finance* 5.
- Barth, Mary E., Ron Kasznik, and Maureen F. McNichols. 2001. “Analyst Coverage and Intangible Assets.” *Journal of Accounting Research* 39:1–34.
- Baruch, Shmuel and Lawrence R. Glosten. 2013. “Fleeting Orders.” Manuscript, Columbia University.
- Biais, B., P. Hillion, and C. Spatt. 1995. “An Empirical Analysis of the Limit Order Book and the Order Flow in the Paris Bourse.” *Journal of Finance* 50:1655–1689.
- Biais, B., D. Martimort, and J.C. Rochet. 2000. “Competing Mechanisms in a Common Value Environment.” *Econometrica* 68:799–837.
- Bodurtha, James N. and Georges R. Courtadon. 1986. “Efficiency Tests of the Foreign Currency Options Market.” *Journal of Finance* 41:151–162.

- Boehmer, Ekkehart, Dan Li, and Gideon Saar. 2015. "Correlated High-Frequency Trading." Manuscript, Cornell University.
- Burdett, Kenneth and Kenneth L. Judd. 1983. "Equilibrium Price Dispersion." *Econometrica* 51:955–969.
- Butters, Gerard. 1977. "Equilibrium Distributions of Sales and Advertising Prices." *Review of Economic Studies* 44:465–491.
- Cao, Melanie and Shouyong Shi. 2001. "Screening, Bidding, and the Loan Market Tightness." *European Finance Review* 5:21–61.
- Colliard, Jean-Edouard and Thierry Foucault. 2012. "Trading Fees and Efficiency in Limit Order Markets." *Review of Financial Studies* 25:3389–3421.
- Degryse, Hans, Frank de Jong, and Vincent van Kervel. 2015. "The Impact of Dark Trading and Visible Fragmentation on Market Quality." *Review of Finance (forthcoming)*.
- Dennert, Jürgen. 1993. "Price Competition between Market Makers." *Review of Economic Studies* 60:735–751.
- Easley, David, Marcos M. López de Prado, and Maureen O'Hara. 2011. "Flow Toxicity and Liquidity in a High Frequency World." Manuscript, Cornell University.
- Glosten, L. 1994. "Is the Electronic Limit Order Book Inevitable?" *Journal of Finance* 49:1127–1161.
- Goldstein, M.A. and K.A. Kavajecz. 2004. "Trading Strategies during Circuit Breakers and Extreme Market Movements." *Journal of Financial Markets* 7:301–333.
- Harstad, Ronald. 1990. "Alternative Common-Value Auction Procedures: Revenue Comparisons with Free Entry." *Journal of Political Economy* 98:421–429.
- Hasbrouck, Joel. 2015. "High Frequency Quoting: Short-Term Volatility in Bids and Offers." Manuscript, NYU.
- Hausch, Donald B. and Lode Li. 1993. "A Common Value Auction Model with Endogenous Entry and Information Acquisition." *Economic Theory* 3:315–334.
- Head, Allen, Lucy Qian Liu, Guido Menzio, and Randall Wright. 2012. "Sticky Prices: A New Monetarist Approach." *Journal of the European Economic Association* 10:939–973.
- Hendershott, Terrence, Charles M. Jones, and Albert J. Menkveld. 2011. "Does Algorithmic Trading Improve Liquidity?" *Journal of Finance* 66:1–33.
- Hollifield, B., R. Miller, and P. Sandås. 2004. "Empirical Analysis of Limit Order Markets." *Review of Economic Studies* 71:1027–1063.

- Hollifield, Burton, Robert A. Miller, Patrik Sandås, and Joshua Slive. 2006. “Estimating the Gains from Trade in Limit Order Markets.” *Journal of Finance* 61:2753–2804.
- Jain, P.K. 2005. “Financial Market Design and the Equity Premium: Electronic versus Floor Trading.” *Journal of Finance* 60:2955–2985.
- Kaplan, Greg and Guido Menzio. 2014. “Shopping Externalities and Self-Fulfilling Unemployment Fluctuations.” Manuscript, Princeton University.
- Klemperer, Paul. 1999. “Auction Theory: a Guide to the Literature.” *Journal of Economic Surveys* 13:227–286.
- Kyle, A. 1985. “Continuous Auctions and Insider Trading.” *Econometrica* 53:1315–1335.
- Menkveld, Albert J. and Bart Z. Yueshen. 2015. “The Flash Crash: A Cautionary Tale about Highly Fragmented Markets.” Manuscript, VU University Amsterdam.
- Naes, Randi and Johannes A. Skjeltorp. 2006. “Order Book Characteristics and the Volume-Volatility Relation: Empirical Evidence from a Limit Order Market.” *Journal of Financial Markets* 9:408–432.
- Piccione, Michele and Guofu Tan. 1996. “A Simple Model of Expert and Non Expert Bidding in First Price Auctions.” *Journal of Economic Theory* 70:501–515.
- Rosenthal, Robert W. 1980. “A Model in Which an Increase in the Number of Sellers Leads to a Higher Price.” *Econometrica* 48:1575–1579.
- Sandås, P. 2001. “Adverse Selection and Competitive Market Making: Empirical Evidence from a Limit Order Market.” *Review of Financial Studies* 14:705–734.
- SEC. 2010a. “Concept Release on Equity Market Structure.” Release No. 34–61358; File No. S7–02-10.
- . 2010b. “Findings Regarding the Market Events of May 6, 2010.” Report of the staffs of the CFTC and SEC to the joint advisory committee on emerging regulatory issues.
- Shilony, Yufal. 1977. “Mixed Pricing in Oligopoly.” *Journal of Economic Theory* 14:373–388.
- Silva, Dakshina De, Thomas Jeitschko, and Georgia Kosmopoulou. 2009. “Entry and Bidding in Common and Private Value Auctions with an Unknown Number of Rivals.” *Review of Industrial Organization* 35:73–93.
- Yueshen, Bart Z. 2015. “Market Making Uncertainty.” Manuscript, INSEAD.