Preliminary Version - Please do not Quote

Understanding Limit Order Book Depth: Conditioning on Trade Informativeness

Héléna Beltran[a], Joachim Grammig[b], Albert J. Menkveld[c]

Abstract

We study how a limit order book reacts to informed trades and adverse selection. We estimate Sandas'(2001) version of the classical Glosten (1994) order book model and accept it, but only for the first two prices displayed on each side of the book. We then relax one of the assumption and allow the level of private information in market orders to be stochastic. By conditioning on the information level, we find support for deeper order books, larger market orders and shorter inter-trade durations at times of relatively uninformative market orders, which is consistent with liquidity traders concentrating their orders at uninformative times.

[a]Université Catholique de Louvain - CORE
[b]Eberhard Karls Universität Tübingen
[c]Vrije Universiteit Amsterdam

Contact details authors. Héléna Beltran, `beltran@core.ucl.ac.be`, FNRS Research Fellow at the Center for operations research and econometrics (CORE), Voie du Roman Pays 34, 1348 Louvain-la-Neuve, Belgium, phone: +32 10 478160, fax: +32 10 474301. Joachim Grammig, `joachim.grammig@uni-tuebingen.de`, Eberhard-Karls-Universität-Tübingen, Wirtschaftswissenschaftliche Fakultät, Abteilung Statistik und Ökonometrie, Mohlstr. 36, 72074 Tübingen, Germany, +49 7071 29 78159, fax: +49 7071 29 5546. Albert J. Menkveld, `ajmenkveld@feweb.vu.nl`, Vrije Universiteit Amsterdam, FEWEB, De Boelelaan 1105, 1081 HV Amsterdam, Netherlands, phone: +31 20 4446130, fax: +31 20 4446020. The first author gratefully acknowledges financial support from the Belgian "Fonds National de la Recherche Scientifique" (FNRS). The usual disclaimer applies.

An increasing number of financial assets trade in limit order markets around the world. In such markets, traders can submit market and limit orders. A market order trades at the best available price posted by previously submitted limit orders that make up the limit order book. A limit order is a conditional buy or sell at a pre-specified price and is stored in the book until cancelled or hit by an incoming market order.[1]

A limit order trades at a better price than a market order. However, there are two types of costs to submitting a limit order. First, the order may fail to execute. Second, the order may execute against a market order that is based on private information (adverse selection). In this case, limit orders trade at the wrong side of the market. This is referred to as picking-off risk. Various models have been proposed that formalize this trade-off between prices, execution probability, and picking-off risk. They essentially take either of two fundamentally different approaches. The first treats the arrival of traders at the market exogenously and models their decision to submit a limit or market order and, potentially, conditions the order's choice on the state of the book (see, e.g., Cohen et al. (1981), Foucault (1999), Harris (1998), Parlour (1998), Hollifield, Miller, and Sandas (2004), and Goettler, Parlour, and Rajan (2003)). The second approach considers endogenous limit order traders arrivals, as it retrieves the state of the book from a zero-profit condition on the marginal limit order. In other words, at each point in time profit opportunities in the book can only be short-lived as they attract traders who exploit them immediately. As in the first stream of the literature, market orders arrive exogenously in time, consume liquidity, and potentially leave profit opportunities in the book that are, again, immediately exploited (see, e.g., Glosten (1994), Biais, Martimort, and Rochet (2000), Seppi (1997), and Parlour and Seppi (2003)).

The distinction in the theoretical literature carries over to the empirical literature. Many recent studies focus on whether an arriving trader chooses a market or limit order

[1] If a limit order is submitted at a price such that it executes immediately (a *marketable* limit order), we consider it a market order.

and how market conditions affect this choice (see, e.g., Biais, Hillion, and Spatt (1995a), Griffiths et al. (2000), and Ranaldo (2004)). Most of them condition on the best bid and ask quotes and depth offered at these quotes. Relatively few studies, however, focus on explaining the state of the order book beyond these quotes. This is surprising as the best quotes only partially reflect overall liquidity. A nice illustration is the sequential tick size reduction in the U.S. that led to a 20% reduction in quoted spreads, but, at the same time, reduced aggregate depth in the order book (see, e.g., NYSE (2000), Goldstein and Kavajecz (2000)). Sandas (2001) tests the celebrated Glosten (1994) model and rejects it for a sample of Swedish stocks. Coppejans, Domowitz, and Madhavan (2003) study order book resiliency to shocks in volatility and returns for the Swedish stock index future market.

In this paper, we aim at understanding how order book depth is affected by the picking-off risk and trade informativeness, by testing the Glosten (1994) model directly (as in Sandas (2001)) and indirectly through a dynamic setting that allows for a stochastic level of private information in order flow. Contrary to the direct test, the dynamic setting does not rely on structural (model-based) equations and is thus a more general test of any model that assumes that informed trades decreases the order book depth. In the indirect test, we model the level of private information by stochastic price impacts coefficients, i.e. Glosten-α's, in a regression of midquote changes on order imbalance. Model estimates show that the α-coefficients can drift away from the unconditional mean, but are mean reverting. This is consistent with the notion that trading can at times experience either highly informative order flow (e.g. around news releases) or uninformative order flow. As the level of private information is a key determinant of the limit order book equilibrium, we study, at each point in time, the state of the book conditional on traders' expectations about next trade informativeness $\hat{\alpha}_t$. These predictions are done based on the information set containing all trades and quotes until the next market order arrives.

We perform our tests on limit order book data in continuous time from one of the largest limit order markets: the Xetra system maintained by the Deutsche Börse Group. We select the three most liquid securities (DaimlerChrysler, Deutsche Telekom, and SAP), which make up almost one-third of the main German index ("DAX") during the sample period. The sample runs from August 2, 1999, through October 29, 1999. We consider this an ideal laboratory for a test of the Glosten model, as high liquidity arguably attracts (implicit) market makers who keep the order book in equilibrium by promptly stepping in on any profit opportunity. Furthermore, the use of hidden (or iceberg) orders was not possible during the sample period. This is highly unusual as, generally, investors are allowed to hide limit orders in the book, like for example in the CAC system operated by the Paris Bourse, and in the Stockholm Stock Exchange's trading system which generated the data analysed in Sandas (2001).

We contribute to the literature in three important ways. First, we consider the state of limit order books in a dynamic setting with time-varying trade informativeness. This, essentially, allows us to test the importance of picking off risk, a key ingredient to limit order book models. If the market *expects* very informative order flow, the book should be shallow. Second, we use a new dataset with fully transparent, hidden-order-free order books of three *highly* liquid shares, for which books are most likely in equilibrium. Hidden orders hamper a test on the Glosten model as it assumes full transparency of the book, so that investors can condition their decision on whether or not to submit a limit order on the state of the book. With hidden orders, they cannot. Last, we do not impose any restriction on the shape of the relation between the order book depth and the adverse selection component of the trade. More precisely, we do not assume that the depth in the book is linear in the inverse of trade informativeness.

Contrary to Sandas (2001), we accept the static Glosten (1994) model applied to the first two prices and depths of the book. Thus, the rejection in Sandas (2001) cannot be extended to active hidden order-free markets where traders compete to provide liquidity. Nevertheless, the Glosten's model is rejected when we go further in the book. Turning to a dynamic setting with stochastic and potentially non-linear price impact we find evidence that picking-off risk is important for order book depth. When the market expects (conditional on the trade history) informative market orders, depth in the book is significantly lower. These effects are also economically significant, as depth decreases up to 25%. Moreover, the drop in liquidity is far from being proportional to the increase in trade informativeness. This supports the intuition that the rejection of the Sandas setting is partially due to the assumption of a linear price impact. In addition to the book, we study order size and duration conditional on order informativeness. We find evidence of larger orders and shorter inter-trade durations at times when upcoming markets orders are expected to be less informative. This supports the notion that (uninformed) liquidity traders optimally concentrate their orders in relatively uninformative times.

The rest of the paper is structured as follows. Section 1 elaborates on the trading platform and the data. Section 2 briefly describes the Glosten/Sandas modelling framework, and presents GMM parameter estimates and test results. Section 3 introduces the dynamic setting. In section 4, we present model estimates and discuss how depth, market order size and inter-trade durations are affected by estimated trade informativeness. Last section concludes.

1 The trading platform and the data

1.1 A pure limit order market: the Xetra platform

The Xetra platform was developed and is maintained by the German Stock Exchange and has operated since 1997 as the main trading platform for German blue chip stocks.

Similar to the Paris Bourse's CAC and the Toronto Stock Exchange's CATS trading systems, a computerized trading protocol keeps track of entry, cancellation, revision, execution and expiration of market and limit orders. Except for the daily open, close and midday auctions and extraordinary batch auctions, trading in the Xetra system is continuous during the opening hours and is based on the order book trading mechanism (also called continuous double auction). The matching of orders is automatically performed by the computer based on the usual rules of price and time priority.[2]

Until September 17, 1999, Xetra trading hours extended from 8.30 a.m to 5.00 p.m. CET. Beginning with September 20, 1999 trading hours were shifted to 9.00 a.m to 5.30 p.m. CET, in order to stick to the main European exchanges trading hours. A call auction takes place at midday, before and after the trading hours. No trading is allowed outside the regular trading hours, although it is possible to enter, revise and cancel orders.[3]

Until October 2000, Xetra screens displayed not only best bid and ask prices, but the whole content of the order book to the market participants. This implies that liquidity supply and the potential price markup of a market order (or marketable limit order) are exactly known to the trader. This was a great difference compared to e.g. Paris Bourse's CAC system where "hidden" orders (also called "iceberg" orders) may be present in the order book. As the name suggests, a hidden limit order is not visible in the order book. This implies that if a market order is executed against a hidden order, the trader submitting the market order may receive an unexpected price improvement. Iceberg orders were allowed in Xetra from October 2000 onwards, heeding the request of investors who were reluctant to see their (potentially large) limit orders, i.e. their investment decisions, disclosed in the open order book. The transparency of the Xetra

[2]See Biais, Hillion, and Spatt (1995b) or Bauwens and Giot (2001) for a complete description of an order book market.

[3]Further information about the organization of the Xetra trading process is provided in Deutsche Börse AG (1999).

order book does not extend to revealing the identity of the traders submitting market or limit orders. Instead, Xetra trading is completely anonymous and dual capacity trading, i.e. trading on behalf of customers and principal trading by the same institution is not forbidden. Finally, prices are decimal, leading to a tiny minimum tick size (0.01 euros).

In the remaining of the paper we focus on three highly liquid stocks: DaimlerChrysler (hereinafter "DCX"), Deutsche Telekom (hereinafter "DTE"), and SAP. All three belong to the main German stock index, the DAX, and jointly amounted (during the sample period) to roughly a third of its market capitalization. Large-caps, like the stocks studied here, are under world-wide scrutiny by a large number of financial analysts. Every single piece of new information (or analysis) is rapidly disseminated via Reuters or Bloomberg and reach traders in a tiny amount of time. This enhances market transparency and the incorporation of new information to stock prices. Moreover, numerous financial institutions hold those stocks in their portfolio and trade them on a daily basis. As such, they are traded around the world in large foreign trading venues (e.g. the New York Stock Exchange). Table 1 reports some daily statistics on the level of trading and book activity. Indeed, the stocks are actively traded as more than 600 trades are executed every day. 2,000 to 3,000 limit orders are submitted every trading day; among them, around 60% are cancelled. The figures confirm that numerous traders compete every day to provide liquidity to the book. Thus, it is quite unlikely that profit opportunities remain.

Overall, the trading rules that produced the order books used in this study are as close as you can get to what a Glosten type limit order book model assumes: the book is fully displayed to market participants, no hidden orders blur the picture, and no dedicated market maker participates in the extremely active trading of the stocks.

1.2 The data

The German Stock Exchange granted access to a database containing all order book events (entries, cancellations, revisions, expirations, partial-fills and full-fills of market and limit orders) that occurred between August 2, 1999 and October 29, 1999. Based on the event histories we perform a real-time reconstruction of the order book sequences. Starting from an initial state of the order book, we track each change in the order book implied by entry, partial or full fill, cancellation and expiration of market and limit orders. This is done by implementing the rules of the Xetra trading protocol outlined in Deutsche Börse AG (1999) in the reconstruction program. We obtain, at each point in time during the 3-month period, the exact limit order book available to market participants.

The main objective of this paper is to investigate to which extent limit order books react to changes in expected trade informativeness. Therefore, we are interested in the shape of the order book just before the trade takes place. It may seem straightforward then to take a snapshot of the book each time we observe a trade. However, for very frequently traded stocks, like the ones in our sample, many transactions occur within short time periods; even buy and sell trades within one second are not impossible.[4] To be able to defend the assumption that the order book snapshots represent equilibrium states, we have to ensure that the snapshots are taken reasonably far apart from the last trade event in order to grant limit order traders enough time to assimilate the information from the trade and subsequently to modify their orders appropriately. Limit order traders need time to react to news and trades, by revising their positions. We pointed in the previous section that a) new public information is quickly transmitted to the market; b) the whole limit order book and all trades are available in real time to market participants; and c) a plethora of limit order traders compete to provide

[4] E.g. for DTE, 2% of the trades occur within one second.

liquidity. Furthermore, the routing of orders from the traders' terminals to the Xetra order book is nearly immediate. This points towards a relatively short amount of time that needs to be granted to allow limit order traders to react to the occurrence of a trade event. The 30-second consolidation scheme described below is a reasonable choice. It is also in line with previous work and thus facilitates comparisons. The procedure works as follows. When a trade is initiated by a market order, we take the snapshot of the limit order book just before the submission and execution of the market order. We then aggregate all *signed* market orders that are submitted within the following 29 seconds, but refrain from taking new order book snapshots within this interval. The market order volume associated with the trade event that triggered the order book snapshot is then the sum of the signed market orders aggregated within the 29 second interval. As an example, assume that we observe three consecutive market order arrivals, a market sell order of size $X_t = -1,000$ at time t, another market sell order of size $X_{t+15} = -100$ at $t + 15$s and a market buy order of size $X_{t+45} = +100$ at $t + 45$s. No further trade occurs within the next 30 seconds. According to our consolidation scheme, the first snapshot of the order book is taken at time t, disregarding the effect that the order exerts on the book. The market order volume associated with the time t trade is then the aggregated order volume of the first and the second market order, i.e. $X_t = -1,100$. The second snapshot is taken at $t + 45$s (again without taking into account the effect of the "triggering" market order on the book). The market order volume associated with this snapshot is $X_{t+45} = +100$, the size of the market order that triggered the snapshot. No aggregation is performed as no other trade occurs within the next 30 seconds. We operate as if the book did not come back to its equilibrium between the first and the second trade, and we implicitly assume that the aggregated signed order (of trades at time t and $t + 15$) conveys information about the underlying value of the stock.[5]

[5]Typically, when market orders from both sides of the book arrive within the interval, a large order is typically added to a small one.

The upper part of table 2 reports the activity of limit order traders between two order book snapshots taken according to the 30-second aggregation scheme. On average, traders submit three limit orders that improve the current best prices, and they cancel six orders at the best prices. The book updating activity between two order book snapshots is thus quite significant. It seems then plausible to assume that, within the 30-second interval, the book has reached its equilibrium state and that all trade related information has been incorporated.

Another practical problem is induced by the fact that Xetra features a fine tick size of one euro cent. If traders quote at all the possible prices, we should observe that, on average, the spread between two consecutive prices (e.g. the best ask price and the second best ask price) is of one tick. In reality, we observe an average spread of 3 cents (20 cents for SAP), which is three times the official tick size (respectively 20 times the official tick size for SAP).[6] Thus, typically not all ticks are actually used by limit order traders. We therefore have to choose an appropriate "virtual" tick size by aggregating the limit order volume at adjacent ticks. One has to be careful, however, since if the virtual tick size is too large then one would aggregate too many different prices and potentially relevant information about the shape of the price-volume relation will be lost. The average spread seems then the more reasonable choice for the "virtual" tick size.

2 A direct test of the Glosten (1994) model

Sandas (2001) presents a variant of Glosten's (1994) limit order book model with discrete price ticks which ideally lends itself to GMM estimation and testing and which is particularly suited for the Xetra data generating process. The two key ingredients of the approach are as follows. First, the fundamental asset value v_t is described by a

[6] Tables of the book spreads are available upon request.

random walk with innovations depending on a constant adverse selection parameter α (or picking-off risk), which gives the informativeness of a signed market order of size X_t:

$$v_t = c_v + v_{t-1} + \alpha X_t + \eta_{v,t} \tag{1}$$

Negative values of X_t denote sell orders, positive values buy orders and $E(X_t) = 0$. $\eta_{v,t}$ is an innovation orthogonal to v_{t-1} and X_t, and c_v gives the expected change in the fundamental value. Equation 1 is quite standard in microstructure models. It is worth stressing that the formulation assumes that the price impact of a trade is linear.

To obtain a model ready to be estimated, a second assumption is introduced: market buy and sell orders arrive with equal probability and with a two-sided exponential density describing the distribution of the order size:[7]

$$f(X_t) = \begin{cases} \frac{1}{2\lambda} e^{\frac{-X_t}{\lambda}} & \text{if } X_t > 0 \text{ (market buy)} \\ \frac{1}{2\lambda} e^{\frac{X_t}{\lambda}} & \text{if } X_t < 0 \text{ (market sell).} \end{cases} \tag{2}$$

The remaining assumptions are as outlined in Glosten (1994). Risk neutral, limit order submitting traders face a fixed order submission cost, γ. They have knowledge about the distribution of market orders (2) and the adverse selection component α and choose their limit order price and quantity such that their expected profit is maximized. If the last unit, asked and bid, at any discrete price tick exactly breaks even, i.e. has expected profit equal to zero, the order book is in equilibrium. Denote the ordered discrete price ticks on the ask (bid) side by p_{+k} (p_{-k}) with $k = 1, 2, \ldots$ and the limit order volumes at these prices by q_{+k} (q_{-k}). With $q_{0,t} \equiv 0$, the equilibrium order book at time t can then recursively be constructed as follows:

$$q_{+k,t} = \frac{p_{+k,t} - v_t - \gamma}{\alpha} - q_{+k-1,t} - \lambda \quad k = 1, 2, \ldots, N \quad \text{(ask side)}$$

[7]We deviate from Sandas (2001) in that we assume that the expected market order size can be described by the same parameter, λ. If we allow this parameter to differ between buy and sell side we would face methodological problems in the GMM estimation procedure. We come back to this point below.

$$q_{-k,t} \quad = \quad \frac{v_t - p_{-k,t} - \gamma}{\alpha} - q_{-k+1,t} - \lambda \quad k = 1, 2, \ldots, N \quad \text{(bid side)}. \tag{3}$$

Equation (3) highlights that order book depth and adverse selection component are inversely and proportionally related. The proportionality follows from the assumption of a linear price impact in equation 1, which is not part of the Glosten (1994) model.

Allowing for mean zero random deviations from the order book equilibrium at each price tick the following unconditional moment restrictions, referred to as break-even conditions, can be used for GMM estimation:

$$E\left(p_{+k,t} - p_{-k,t} - 2\gamma - \alpha\left(\sum_{i=+1}^{+k} q_{i,t} + 2\lambda \sum_{i=-1}^{-k} q_{i,t}\right)\right) = 0 \quad k = 1, 2, \ldots, N \tag{4}$$

A second set of unconditional moment conditions, referred to as updating restrictions, relate the expected changes in the order book to the market order flow:

$$\begin{aligned} E\left(\Delta p_{+k,t} - \alpha\left(\sum_{i=+1}^{+k} q_{i,t+1} - \sum_{i=+1}^{+k} q_{i,t}\right) - c_v - \alpha X_t\right) &= 0 \qquad k = 1, 2, \ldots, N \\ E\left(\Delta p_{-k,t} + \alpha\left(\sum_{i=-1}^{-k} q_{i,t+1} - \sum_{i=-1}^{-k} q_{i,t}\right) - c_v - \alpha X_t\right) &= 0 \qquad k = 1, 2, \ldots, N \end{aligned} \tag{5}$$

where $\Delta p_{j,t} = p_{j,t} - p_{j,t-1}$. Finally, an obvious moment condition to identify the expected market order size is given by[8]

$$E(|X_t| - \lambda) = 0. \tag{6}$$

We estimate the model parameters exploiting the break even (4) and the updating conditions (5) along with (6). For $N = 2$, seven moment conditions (two break even

[8]If different expected market order sizes on the buy and sell side were possible then we would face the problem that the number of observations used to compute the sample means that replace the expectations in (6) would be different from those used for the other moment conditions (4) and (5). It is not clear how proper GMM inference can be conducted in this situation. The problem is avoided by assuming identical expected buy and sell volumes which is a reasonable assumption from an economic point of view anyway.

conditions, four updating restrictions plus one for the expected market order size) are employed to estimate the four model parameters.

We estimate the model for 5 different numbers of quote-depth pairs ($N = \{2, 3, 4, 6, 8\}$). Table 3 contains estimation results both of first stage GMM as well as iterated GMM, for the two best quotes and the four best quotes ($N = 2$ and 4). For other values of N, we reproduce only the estimate for the α and the $J-$statistic resulting from the iterated GMM (see table 4). Additionally, the two tables reports the ratio of the average spread (avg.spr.) and the price impact at the estimated mean market order volume λ computed as $\frac{\alpha\lambda}{\text{avg.spr.}}$. This is a useful measure to compare the adverse selection component across stocks and different studies.

For all values of N, we obtain two sets of results: in the second one we use only order book observations after 10 a.m. One could argue that the first trading hours are supposed to be quite different from the remaining hours in that incorporation of information accumulated during the non-trading night hours is taking place (see Chung, Van Ness, and Van Ness (1999) or Madhavan, Richardson, and Roomans (1997)). Moreover, depth is consistently lower every morning than during the rest of the day (see Beltran, Giot, and Grammig (2003) for a description of intraday patterns). For the rest of the day we do not expect significant seasonal patterns. As GMM estimation requires stationarity, we leave out the first hour of trading. We will follow the same approach when we estimate stochastic price impacts by means of state-space models.

The results can be summarized as follows. First, the estimates do not differ much between first stage and iterated GMM, both regarding the estimates and the parameter standard errors. The estimates have small standard errors which is of course also due to the large sample size. For all stocks we observe significant estimates for the α. Second, estimates based on all observations are very close to estimates based on observations after 10 a.m. Indeed, the relative adverse selection estimates ($\frac{\hat{\alpha}\hat{\lambda}}{\text{avg.spr.}}$) are only marginally smaller when observations before 10 a.m. are excluded. The estimates, which range from about .42 to about .55, are remarkably close to the figures which Sandas (2001) reports. Finally, based on the J-test the model, unlike in Sandas (2001), the model for the best two quotes ($N = 2$) cannot be rejected at any reasonable significance level. We stressed in section 1 that our dataset suits perfectly to a test of a Glosten-type order book. Indeed, we find that the estimation results are quite robust across different intraday sample periods, and the model, unlike in Sandas (2001), cannot be rejected. The Swedish stocks studied by Sandas are not traded actively; thus, the lack of competition between traders may explain the rejection of theoretical models like Glosten (1994) that assume a large number of limit order traders. Another important explanation for non-rejection is the absence of hidden orders in our data, a feature that renders the Xetra data generating process much closer to Glosten's (1994) theoretical framework.

When we go beyond the best two prices ($N > 2$), we always reject the Sandas model (see table 4). What causes the rejection of the model? Two assumptions made by Sandas may be too strong. First, he assumes that the price impact α is constant, while empirical studies underline the existence of periods with highly-informative trades followed by non-informative periods. Sandas partially tests that hypothesis by allowing the price impact to depend on the stock-specific volatility, the market volatility, and trading volume. But he still rejects the Glosten's model with the time-varying α. It is thus likely that we would obtain similar results in our dataset. Another crucial assumption in Sandas (2001) is that the depth at consecutive ticks is linear in the inverse of the adverse selection component. Whatever the size of the trade, the impact on the fundamental value is constant and linear. The linearity is carried over the book, and depths and prices $\{p_{k,t}, q_{k,t}\}$ are linked by a fixed constant $\frac{1}{\alpha}$, whatever the price level k we consider. However, the estimate for α reported in table 4 are far from being constant for different N. Thus, the rejection of the Glosten (1994) model may originate in the linear impact on depth that Sandas implicitly assumes.

3 A simple model for trade informativeness

We opt for a more flexible framework that not only allows the price impact to vary through the trading day, but also that does not impose any restriction on the shape of the relation between depth and trade informativeness. To do so, we separate the analysis in two parts. First, in the current section, we model price dynamics by allowing time-varying adverse selection α_t. Then, in the following section, we condition the state of the book on next trade expected informativeness.

Typically, microstructure models assume that the book midquote at time t M_t (i.e. the average of the best bid and ask prices) is equal to the current (unobserved) value of the stock v_t plus some (transitory) noise ε_t due to microstructure effects. t denotes the time of the (aggregated) trade in our setting, and successive t are not regularly spaced.[9] As pointed out in most theoretical works on order books (see e.g. Glosten (1994), Seppi (1997) or Foucault (1999)), sell (buy) limit orders are posted at a price larger (smaller) than or equal to the expected value of the stock conditionally on the limit order being hit. Thus, in an order-driven market, the midquote is set at the underlying value as expected before the trade takes place, conditionally on the next trade being of signed size X_t. We write:

$$M_t = v_t + \varepsilon_t, \tag{7}$$

where v_t denotes the value of the stock after the trade takes place. New private information revealed by a signed trade at time t enables to update the unobserved value:

$$v_t = c_v + v_{t-1} + \alpha_t X_t + \eta_{v,t}. \tag{8}$$

X_t is the aggregated signed order defined in section 1, and α_t is the (unobserved) stochastic price impact coefficient of the trade in t. It measures how informative the trade at

[9] M_t is the midquote as displayed in the book just before the submission of the market order.

time t is. Thus, the fundamental value v_t follows a random walk with a linear price impact, in line with most models for the efficient price, and as in Sandas (2001). The main contribution here is that we do not assume that the linear setting is carried over the book depth, that is there is no assumption of a linear relation between depth and the inverse of the adverse selection component of the trade. Finally, we assume that the price impact coefficient follows a mean-reverting, first-order autoregressive process:

$$\alpha_t - \bar{\alpha} = 1_{[t \text{ and } t-1 \text{ on same day}]} \phi(\alpha_{t-1} - \bar{\alpha}) + \eta_{\alpha,t}. \tag{9}$$

$\bar{\alpha}$ is the unconditional mean of the price impact coefficient. The price impact is reset to zero every day by means of the Dirac function 1 that takes value zero if trade in $t-1$ belongs to the previous trading day. This rules out any link between today and past information.

In the following section, we present estimates for the three stocks of the model defined by equations 7 to 9, and we test if the book is indeed inversely related to the stochastic price impact coefficient. To avoid any seasonal pattern in the efficient price volatility or in the adverse selection component, we consider only observations after 10 a.m. The Kalman filter is used for inference and signal extraction (see Durbin and Koopman (2001) for a more detailed description of the algorithms). The parameters are estimated using maximum likelihood. The estimation was done in Ox using SsfPack software (see Doornik (2001) and Koopman, Shephard, and Doornik (1999)). In a second step, the α_t are estimated with a filtered algorithm, which means that at each point in time t we use *only past information* (observations) to estimate the price impact coefficient at t.

4 Conditioning on expected trade informativeness: the results

Table 5 presents the estimated parameters for the model presented in section 3. We first analyze the stochastic price impact dynamics (equation 9). The estimated autoregressive coefficient is 0.22 for DCX, 0.37 for DTE and 0.91 for SAP, and are all significant at the 1% confidence level. This indicates that for two of the three stocks, there is little clustering in trade informativeness. In fact, there are probably only two or three consecutive liquidity-motivated trades (when α_t is lower than its unconditional mean); the same conclusion holds for informed trades (when α_t is above its unconditional mean). This result is quite surprising, as it means that for DCX and DTE, informative periods last only a couple of minutes; the same holds for trading opportunities for liquidity traders. This feature is also clear when we graph the evolution of forecasted α_t (see figure 2).

A crucial difference with Sandas (2001) is that in the following we will not condition the state of the book on the price impact (supposedly known in Sandas (2001)) at the time of the trade, but on the *forecasted* price impact, denoted by $\hat{\alpha}_t$. Traders submit limit orders at a price and for a volume which depend on *future* trades and the amount of information they convey, that is on expected adverse changes in the fundamental value of the stock (picking-off risk).[10]

We further define ten levels of expected trade informativeness, from (expected) purely liquidity-motivated trades to (expected) highly informed trades. First, we consider all $\hat{\alpha}_t$ that belong to a two-standard deviation interval around the unconditional $\bar{\alpha}$, i.e. we take $\hat{\alpha}_t \in [\bar{\alpha} - 2\sigma_{\hat{\alpha}}, \bar{\alpha} + 2\sigma_{\hat{\alpha}}]$. We cut the interval in ten buckets, with a step size of $0.4\sigma_{\hat{\alpha}}$. Then, we classify all observations as belonging to one of the 10 buckets. For instance, the

[10]In our setting we assume that traders know the size of the next trade. We plan to relax this assumption in the future.

first bucket corresponds to periods where the forecasted α_t was between 2 to 1.6 times its standard deviation lower than its unconditional mean ($\hat{\alpha_t} \in [\bar{\alpha} - 2\sigma_{\hat{\alpha}}, \bar{\alpha} - 1.6\sigma_{\hat{\alpha}}]$), the second bucket between $1.6\sigma_{\hat{\alpha}}$ and $1.2\sigma_{\hat{\alpha}}$, and so on. The fifth and sixth buckets correspond to periods where the forecasted α_t was around its unconditional mean. At that time, market participants expect "normally informed" trades. Lower buckets define periods where few information was expected to spread to the market, while periods that correspond to expected large flows of information are in the last buckets. This scheme enables to classify all our observations as belonging to periods with expected liquidity-motivated trades (buckets 1 to 4) or expected informed trades (buckets 7 to 10).

Then, we assign each observed order book depth, market order size and inter-trade duration to the level of trade informativeness that market participants expect. Each observation is allocated to one of the ten levels of expected trade informativeness. We then compute the averages of the variables. Figures 3 to 5 plot the mean cumulated depth for the four best ask prices for each level of expected trade informativeness. Figures 6 to 8 report the same averages for the bid side. For the four price levels and both sides of the book, we do find a negative relationship between trade informativeness and depths: when limit order traders expect informed trades, the book is shallow. Nevertheless, when market participants expect liquidity-motivated trades (first four buckets), the cumulated depth offered at any of the four ask prices is remarkably constant. Thus, it seems that adverse selection has a non linear impact on the depth offered at a certain level k: higher expected trade informativeness leads to drops in the depth offered at the level k when trade informativeness is around or above its average value. When trades are not very informative, variations in the portion of the trade that is informed have virtually no impact on the provision of liquidity by the order book. Additionally, for a similar value of $\hat{\alpha_t}$, the drop in depths at levels further in the book seem more pronounced than for the depth at the best price. The two non-linear features, across different values of (expected) trade informativeness and across different levels for the book depths, confirm the intuition that the rejection of the Sandas (2001)' framework when we analyze depths beyond the best two prices originates in a non-linear impact of trade informativeness.

Figure 9 indicates that on average large market orders are submitted either when the next trade is expected to be mainly liquidity-motivated (and then the book is full) or when limit orders submitters expect new information from an informed trade (and the book is shallow). Large market orders walk up in the book and pay a markup for the provision of liquidity beyond the volumes at the best prices. That large orders arise also in periods without any new information confirm findings in Beltran, Giot, and Grammig (2003) or Coppejans, Domowitz, and Madhavan (2003): market order traders step in when liquidity is cheap (small markups) and thus takes advantage of a kind of “liquidity opportunity”. On the other hand, markups are quite large when the period is expected to be informative and thus only investors with a private valuation steadily different from the best prices will be willing to pay it.

Finally, we compute the average durations between two consecutive (aggregated) trades for each level of expected trade informativeness. Results are reported in figure 10. The pattern confirms the result on market order size: trade durations are smaller for liquidity-motivated trades and informed trades. This is consistent with traders quickly stepping in and trading when short-lived liquidity or profit opportunities arise.

It is worth noting that for SAP the results are somewhat different. We still find that adverse selection and depths are inversely related, and that this fashion is non-linear. But, we do not find that smaller market orders (and longer durations) are submitted during “normally informative” periods. Indeed, market order size and duration are inversely related to expected trade informativeness. The fact that the largest market orders are not submitted during highly-informative periods is surprising. Nevertheless, the drop in liquidity when we move further in the book is more pronounced for SAP than for the other two stocks (see figure 5). The mark-up for large orders is quite large and likely deter informed investors to trade against the market and to take advantage of their information.

We replicate the dynamic analysis in two different directions to check that the results were not dependant on some potentially arbitrary assumptions.[11] First, we perform the same temporal aggregation, but at 5 seconds; book activity statistics are reported in the bottom part of table 2. We estimated the same model with the new snapshots and aggregated order sizes and obtain similar results.

Second, there is typically a peak in trading activity around the opening of the New York Stock Exchange (NYSE). As some information, at least on the orientation of the whole market, leaks to Europeans market places, we could find a spurious correlation between the alphas on one hand, and the state of the book and the market order size on the other hand. Thus, we redid the analysis for each hour in the day. Unconditional $\bar{\alpha}$ estimates did not vary significantly throughout the day, except in the period following the opening of the exchange. This comforts our choice to exclude observations before 10 a.m.

5 Conclusion

In the paper we show that there is a non-linear relation between the depths displayed by the book and expected trade informativeness. Indeed, when market participants expect large informed trades, they withdraw from the book, and depths are low. Nevertheless, the relation is far from being linear as prices and depths further in the book are more affected by adverse selection than best quotes and depths. This is consistent with the rejection of Sandas (2001) model when we move further in the book. Besides, variations in the level of informativeness of the trades have an asymmetric impact on the book: the

[11]The results presented in this section are available from the authors upon request.

depth offered at the, say, third best price decreases when trades are expected to be more informed than on average, but liquidity trades have virtually no impact on the depth provided by traders. Moreover, traders quickly step in to take advantage of short-lived profit opportunities.

This framework is, up to our knowledge, the first clear confirmation that indeed the book is strongly dependent on the expected informativeness of the trade. This opens room for numerous research paths. A very interesting issue is to explain the non linear patterns. Do traders posting orders far in the book ask for an additional reward for longer time to executions (and thus larger picking-off risk)? Are limit orders traders risk-averse? Another interesting aspect is modelling the arrival process of market orders. We will address those questions in future research works.

References

Bauwens, L., and P. Giot. 2001. *Econometric modelling of stock market intraday activity.* Kluwer Academic Publishers.

Beltran, H., P. Giot, and J. Grammig. 2003. "Liquidity, volatility and trading activity in the XETRA automated auction market." Mimeo, University of Namur.

Biais, B., P. Hillion, and C. Spatt. 1995a. "An Empirical Analysis of the Limit Order Book and the Order Flow in the Paris Bourse." *Journal of Finance*, no. 50:1655–1689.

———. 1995b. "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(4):799–837.

Chung, K.H., B.F. Van Ness, and R.A. Van Ness. 1999. "Limit orders and the bid-ask spread." *Journal of Financial Economics* 53:255–287.

Cohen, K., S. Maier, R. Schwartz, and D. Whitcomb. 1981. "Transaction Costs, Order Placement Strategy and Existence of the Bid-Ask Spread." *Journal of Political Economy* 89:287–305.

Coppejans, M., I. Domowitz, and A. Madhavan. 2003. "Dynamics of Liquidity in an Electronic Limit Order Book Market." Technical Report, Department of Economics, Duke University.

Deutsche Börse AG. 1999. "Xetra Market Model Release 3 Stock Trading." Technical Report.

Doornik, J.A. 2001. *Object-Oriented Matrix Programming using Ox 3.0.* London: Timberlake Consultants Press.

Durbin, J., and S.J. Koopman. 2001. *Time Series Analysis by State Space Models.* Oxford: Oxford University Press.

Foucault, T. 1999. "Order Flow Composition and Trading Costs in a Dynamic Limit Order Market." *Journal of Financial Markets* 2:99–134.

Glosten, L. 1994. "Is the Electronic Limit Order Book Inevitable?" *Journal of Finance* 49:1127–1161.

Goettler, R.L., C.A. Parlour, and U. Rajan. 2003. "Equilibrium in a Dynamic Limit Order Market." Technical Report, EFA 2003 Annual Conference Paper `www.ssrn.com/abstract=410021`.

Goldstein, M.A., and K.A. Kavajecz. 2000. "Eights, Sixteenths, and Market Depth: Changes in Tick Size and Liquidity Provision on the NYSE." *Journal of Financial Economics* 1 (56): 125–149.

Griffiths, M., B. Smith, D.A. Turnbull, and R. White. 2000. "The Costs and Determinants of Order Aggressiveness." *Journal of Financial Economics* 56:65–88.

Harris, L. 1998. "Optimal Order Submission Strategies in Some Stylized Trading Problems." *Financial Markets, Institutions, and Instruments*, vol. 7(2).

Hollifield, B., R. Miller, and P. Sandas. 2004. "Empirical Analysis of Limit Order Markets." Technical Report, Review of Economic Studies (Forthcoming).

Koopman, S.J., N. Shephard, and J.A. Doornik. 1999. "Statistical Algorithms for Models in State Space using Ssfpack 2.2." *Econometrics Journal* 2:113–166.

Madhavan, A., M. Richardson, and M. Roomans. 1997. "Why do Security Prices Change? A Transaction-Level Analysis of NYSE Stocks." *Review of Financial Studies*, no. 10:1035–1064.

NYSE. 2000. "Presentation at SEC Roundtable discussion on Decimalization, December 11, 2000." Technical Report.

Parlour, C. 1998. "Price Dynamics in Limit Order Markets." *Review of Financial Studies*, no. 11:789–816.

Parlour, C.A., and D.J. Seppi. 2003. "Liquidity-Based Competition for Order Flow." *Review of Financial Studies* 16(2):301–343.

Ranaldo, A. 2004. "Order Aggressiveness in Limit Order Book Markets." *Journal of Financial Markets* 7(1):53–74.

Sandas, P. 2001. "Adverse Selection and Competitive Market Making: Empirical Evidence from a Limit Order Market." *Review of Financial Studies* 14 (3): 705–734.

Seppi, D. 1997. "Liquidity Provision with Limit Orders and a Strategic Specialist." *Review of Financial Studies* 10:103–150.

Table 1: Summary statistics

The table reports summary statistics for the three stocks. Trades is the number of trades, price is the transaction price. Traded volumes are in millions of euros. LO means Limit Order (marketable limit orders are excluded). All values are daily averages, except the transaction price (sample average) and the return. The sample return is computed with the first and last transaction prices of the sample.

Stock	Trades	Price	Volume	Return	LO Submitted	LO Cancelled
DCX	1296.6	69.9	189.5	5.9	3139.1	1967.7
DTE	922.3	40.7	133.2	13.5	2311.7	1394.5
SAP	661.3	402.5	124.1	17.4	3038.5	2346.3

Table 2: Submissions and cancellations between aggregated trades

The table reports the number of limit orders submitted and cancelled in the open limit order book in between two market orders, for a duration between aggregated market orders of 30 seconds. All figures are averages. The first three rows correspond to the number of limit orders submitted at a price improving the current best price (Better), at the best price (Best), and beyond the best price (Worse). The next two rows correspond to the number of orders standing at the best price (Best) or beyond (Worse) which were cancelled in between the 30-second duration. The last row sums all submissions and cancellations taking place within 30 seconds.

		DCX	DTE	SAP
Submission	Better	3.7	3.3	3.9
	Best	0.5	0.5	0.4
	Worse	2.7	2.1	3.6
Cancellation	Best	7.1	6.7	6.4
	Worse	3.1	2.7	4.2
Total		17.2	15.2	18.5

The following table reports the same figures for market orders aggregated over a 5-second duration.

		DCX	DTE	SAP
Submission	Better	1.2	1.1	1.2
	Best	0.1	0.1	0.1
	Worse	0.6	0.4	0.7
Cancellation	Best	3.0	2.8	2.8
	Worse	0.8	0.7	1.0
Total		5.8	5.1	5.7

Table 3: A GMM test of the Glosten (1994) model

This table reports GMM estimation results for the Glosten (1994) model with the additional assumptions in Sandas (2001); see equations 4 to 6 for the moment conditions. The avg.spr column reports the average spread. The expected buy and sell volume is assumed to be identical (λ). α is the constant price impact of a trade, γ is the order processing cost, and c_v is the expected change in the fundamental value v_t. For each stock , we estimate the model 4 times. We take the first N best prices $\{p_{k,t}\}_{k=-N}^{k=+N}$ and the associated cumulative volumes $\{q_{k,t}\}_{k=-N}^{k=+N}$ from the order book, for N equals to 2 (rows "best 2") and 4 (rows "best 4"). We further check that the opening period is not more informative than the rest of the day (and thus may bias the estimates) by re-estimating the same model only with observations after 10 a.m. (rows "> 10 a.m.").

			First Stage GMM						Iterated GMM					
	avg.spr.	$\frac{\hat{\alpha}\hat{\lambda}}{\text{avg.spr.}}$	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\lambda}$	$\hat{c_v}$	$J-$stat	$p-$val.	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\lambda}$	$\hat{c_v}$	$J-$stat	$p-$val.
DCX														
best 2, all	0.0609	0.5504	0.0079	-0.0252	4.2382	-0.0024	0.0003	1.0000	0.0079	-0.0252	4.2382	-0.0024	0.0003	1.0000
			(0.00004)	(0.00042)	(0.03261)	(0.00042)			(0.00003)	(0.00037)	(0.03253)	(0.00040)		
best 4, all	0.0609	0.4989	0.0072	-0.0192	4.2382	-0.0022	2007.8	0.0000	0.0075	-0.0212	4.1811	-0.0019	1862.4	0.0000
			(0.00002)	(0.00035)	(0.03261)	(0.00038)			(0.00002)	(0.00032)	(0.03242)	(0.00038)		
best 2, > 10 a.m.	0.0583	0.5550	0.0076	-0.0245	4.2680	-0.0021	0.0000	1.0000	0.0076	-0.0245	4.2680	-0.0021	0.0000	1.0000
			(0.00004)	(0.00044)	(0.03495)	(0.00046)			(0.00003)	(0.00039)	(0.03488)	(0.00044)		
best 4, > 10 a.m.	0.0583	0.5054	0.0069	-0.0190	4.2680	-0.0019	1607.7	0.0000	0.0072	-0.0207	4.2067	-0.0015	1490.6	0.0000
			(0.00002)	(0.00036)	(0.03495)	(0.00042)			(0.00002)	(0.00033)	(0.03475)	(0.00042)		
DTE														
best 2, all	0.0528	0.4949	0.0040	-0.0197	6.5242	-0.0007	0.0000	1.0000	0.0040	-0.0197	6.5242	-0.0007	0.0000	1.0000
			(0.00002)	(0.00036)	(0.05036)	(0.00034)			(0.00002)	(0.00031)	(0.05017)	(0.00034)		
best 4, all	0.0528	0.4635	0.0038	-0.0163	6.5242	-0.0006	1307.6	0.0000	0.0038	-0.0175	6.5359	-0.0005	1273.5	0.0000
			(0.00001)	(0.00029)	(0.05036)	(0.00032)			(0.00001)	(0.00026)	(0.05000)	(0.00032)		
best 2, > 10 a.m.	0.0506	0.5018	0.0039	-0.0193	6.5748	-0.0009	0.0005	1.0000	0.0039	-0.0193	6.5749	-0.0009	0.0005	1.0000
			(0.00002)	(0.00037)	(0.05410)	(0.00038)			(0.00002)	(0.00032)	(0.05391)	(0.00038)		
best 4, > 10 a.m.	0.0506	0.4701	0.0036	-0.0160	6.5748	-0.0008	1141.7	0.0000	0.0037	-0.0173	6.5868	-0.0008	1109.6	0.0000
			(0.00001)	(0.00030)	(0.05410)	(0.00035)			(0.00001)	(0.00026)	(0.05372)	(0.00036)		
SAP														
best 2, all	0.5438	0.4460	0.3059	-0.1237	0.7927	-0.0187	0.0022	1.0000	0.3059	-0.1237	0.7927	-0.0187	0.0022	1.0000
			(0.00234)	(0.00388)	(0.00769)	(0.00444)			(0.00199)	(0.00337)	(0.00766)	(0.00443)		
best 4, all	0.5438	0.4370	0.2998	-0.1080	0.7927	-0.0182	1578.8	0.0000	0.2966	-0.1154	0.8005	-0.0165	1609.6	0.0000
			(0.00154)	(0.00325)	(0.00769)	(0.00434)			(0.00140)	(0.00282)	(0.00764)	(0.00430)		
best 2, > 10 a.m.	0.5185	0.4353	0.2932	-0.1134	0.7699	-0.0200	0.0002	1.0000	0.2932	-0.1134	0.7699	-0.0200	0.0002	1.0000
			(0.00238)	(0.00386)	(0.00779)	(0.00494)			(0.00203)	(0.00335)	(0.00775)	(0.00493)		
best 4, > 10 a.m.	0.5185	0.4296	0.2893	-0.1019	0.7699	-0.0196	1193.0	0.0000	0.2869	-0.1086	0.7776	-0.0186	1212.2	0.0000
			(0.00157)	(0.00326)	(0.00779)	(0.00484)			(0.00144)	(0.00281)	(0.00773)	(0.00481)		

Table 4: Estimated adverse selection component with additional quotes

The table reports changes in the GMM estimates for the adverse selection component when we consider quotes posted further in the book. The estimation is described in table 3; see equations 4 to 6 for the moment conditions. α is the constant price impact of a trade, λ is the expected trade volume, and avg. spread is the average spread over the sample. We repeat the estimation 5 times and we use each time an increasing number of quote-volume pairs (column "Ticks"). More precisely, we estimate the model with the series $\{p_{k,t}, q_{k,t}\}_{k=-N}^{k=+N}$ for $N = \{2, 3, 4, 6, 8\}$ successively. Only observations after 10 a.m. were kept.

	Ticks	$\frac{\hat{\alpha}\hat{\lambda}}{avg.spread}$	$\hat{\alpha}$	$J-$stat	$p-$value
DCX	2	0.5550	0.0076	0.0	1.00000
			(0.00003)		
	3	0.4912	0.0070	1225.7	0.00000
			(0.00002)		
	4	0.5054	0.0072	1490.6	0.00000
			(0.00002)		
	6	0.4781	0.0067	3705.1	0.00000
			(0.00002)		
	8	0.4904	0.0069	4539.9	0.00000
			(0.00002)		
DTE	2	0.5018	0.0039	0.0	1.00000
			(0.00002)		
	3	0.4582	0.0035	722.8	0.00000
			(0.00002)		
	4	0.4701	0.0037	1109.6	0.00000
			(0.00001)		
	6	0.4652	0.0036	1318.4	0.00000
			(0.00001)		
	8	0.4691	0.0036	1628.9	0.00000
			(0.00001)		
SAP	2	0.4353	0.2932	0.0	1.00000
			(0.00203)		
	3	0.4646	0.3124	246.8	0.00000
			(0.00179)		
	4	0.4296	0.2869	1212.2	0.00000
			(0.00144)		
	6	0.4277	0.2855	1532.2	0.00000
			(0.00127)		
	8	0.4395	0.2967	2012.9	0.00000
			(0.00124)		

Table 5: Stochastic Price Impact Coefficients

This table extends the standard microstructure model to include stochastic price impact coefficients and temporary noise. The model is:

$$
\begin{aligned}
v_t &= c_v + v_{t-1} + \alpha_t X_t && + \eta_{v,t} \\
\alpha_t - \bar{\alpha} &= 1_{[t \text{ and } t-1 \text{ on same day}]} \phi(\alpha_{t-1} - \bar{\alpha}) && + \eta_{\alpha,t} \\
M_t &= v_t && + \varepsilon_t \\
\eta_{v,t} &\sim \mathrm{N}(0, \sigma_v^2) \text{ for intraday innovations} \\
\eta_{v,t} &\sim \mathrm{N}(0, \kappa_v^2) \text{ for day-over-day innovations} \\
\alpha_{v,t} &\sim \mathrm{N}(0, \sigma_\alpha^2) \text{ for intraday innovations} \\
\alpha_{v,t} &\sim \mathrm{N}(0, \kappa_\alpha^2) \text{ for day-over-day innovations} \\
\varepsilon_t &\sim \mathrm{N}(0, \sigma_\varepsilon^2)
\end{aligned}
$$

where v_t is the unobserved efficient price, X_t is the signed (aggregate) market order, α_t is the (stochastic) price impact coefficient, and M_t is the observed midquote. Standard deviations are in parentheses.

	DCX	DTE	SAP
σ_v (in ticks)	2.42	2.17	23.27
	(0.05)	(0.05)	(0.59)
κ_v (in ticks)	64.31	44.54	717.97
	(11.45)	(7.94)	(127.68)
σ_α	0.29	0.16	4.95
	(0.01)	(0.01)	(0.92)
κ_α	0.48	0.24	18.40
	(0.21)	(0.11)	(6.54)
σ_ε (in ticks)	1.06	1.01	10.83
	(0.06)	(0.06)	(0.62)
c_v	-0.13	0.01	-0.37
	(0.03)	(0.03)	(0.39)
$\bar{\alpha}$	0.44	0.24	21.17
	(0.01)	(0.01)	(0.98)
ϕ	0.22	0.37	0.91
	(0.06)	(0.07)	(0.03)
N	26,776	22,149	16,674

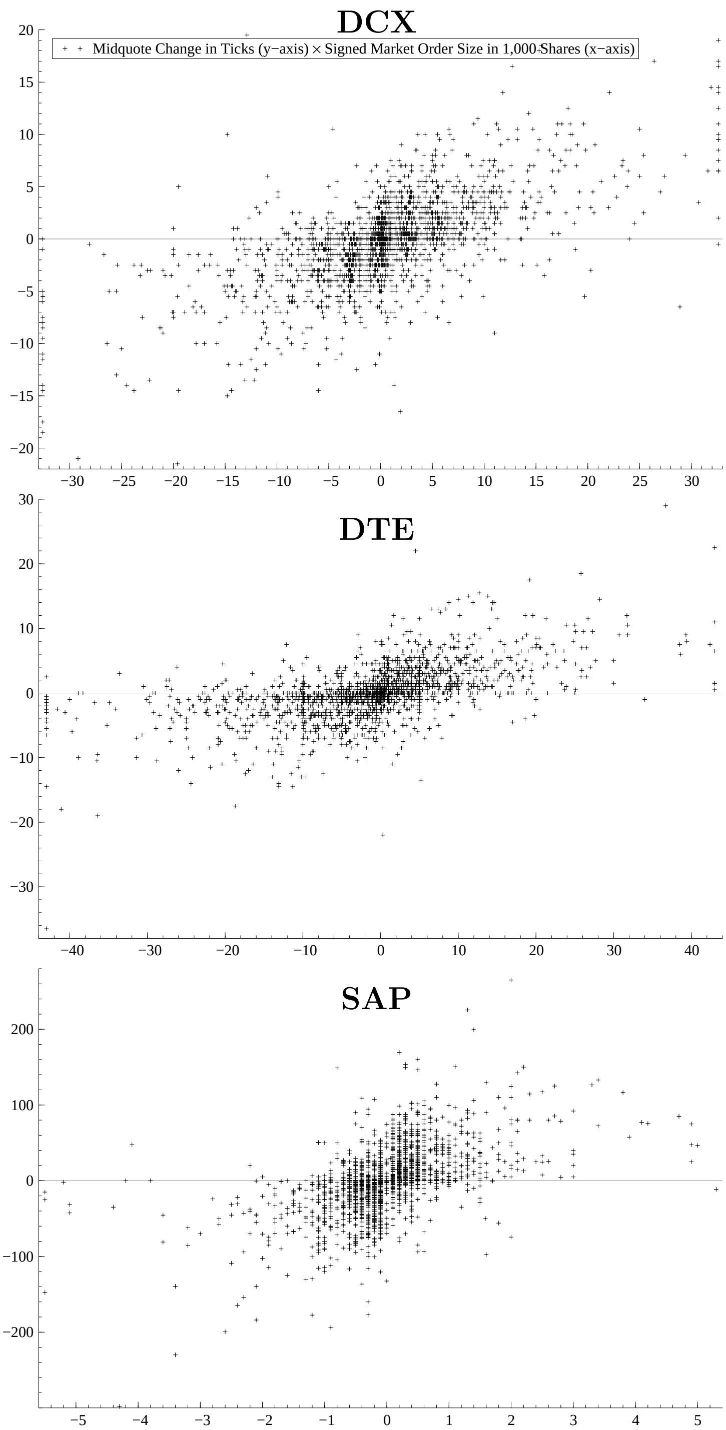

Figure 1: **Scatterplot of Signed Market Order Size against Midquote Change in Ticks (First 2000 Observations).**

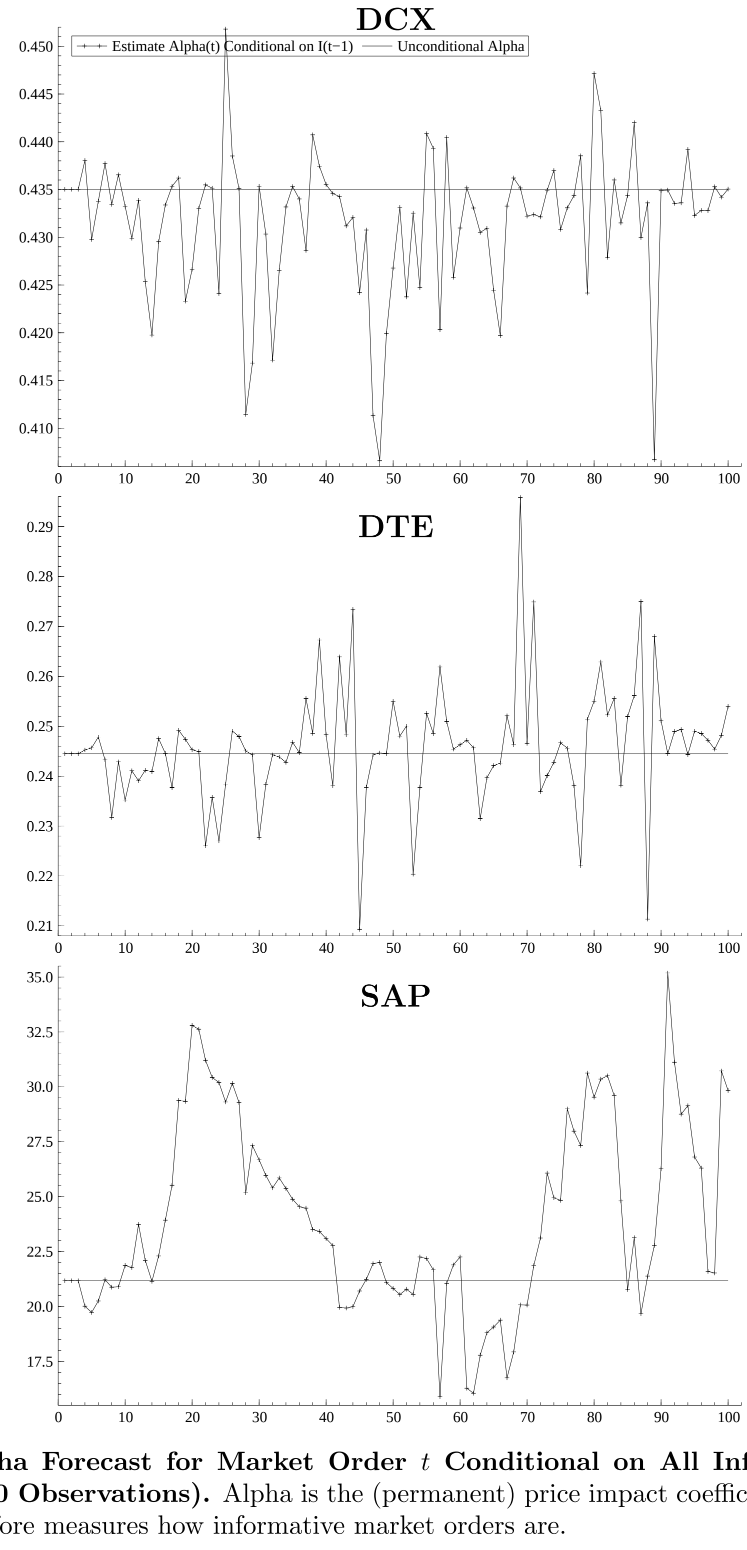

Figure 2: **Alpha Forecast for Market Order t Conditional on All Information through $t-1$ (First 2000 Observations).** Alpha is the (permanent) price impact coefficient of signed market orders and therefore measures how informative market orders are.

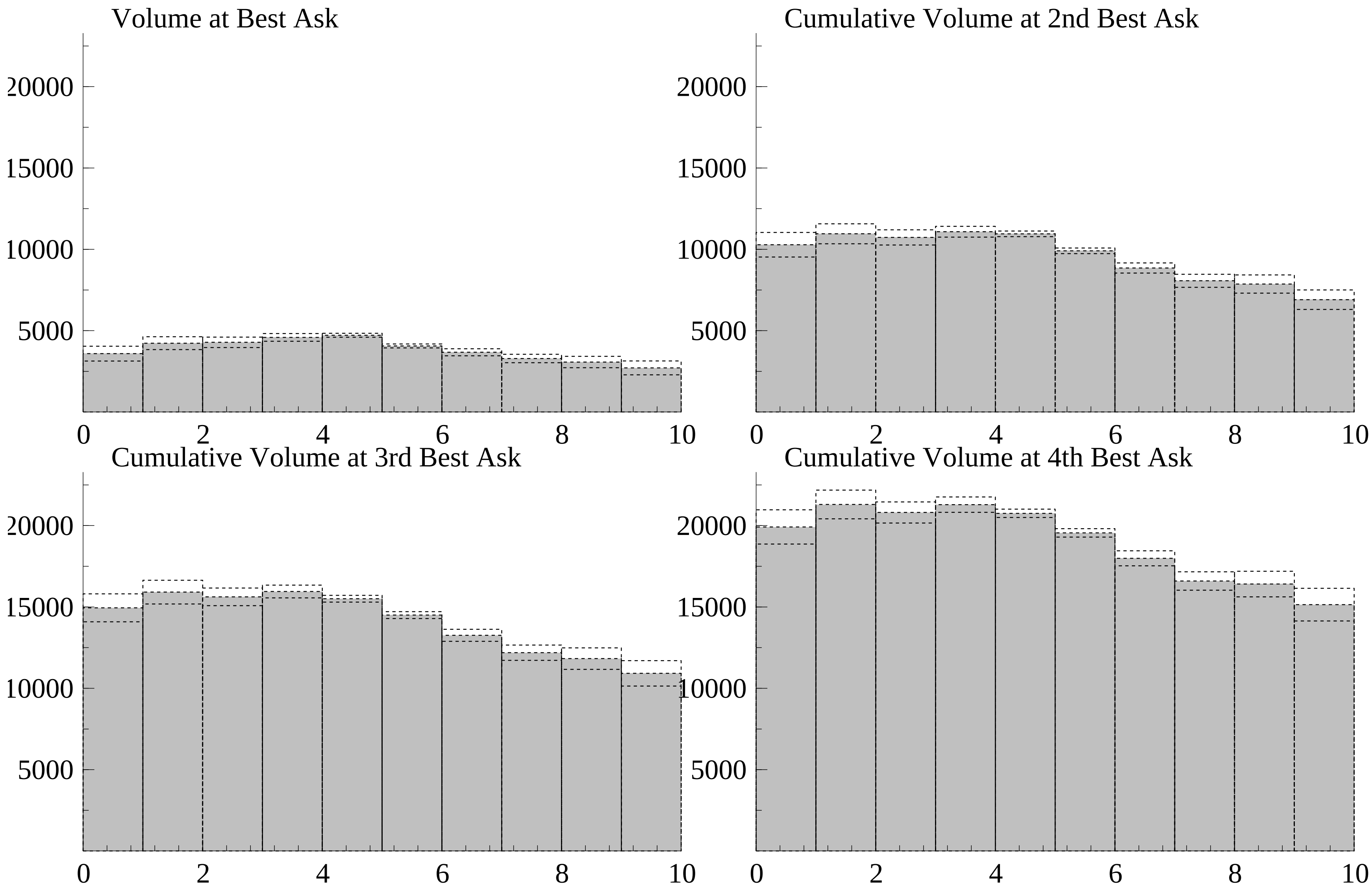

Figure 3: **DCX: Order Book Depth on Ask Side vs. Alpha**. These four graphs illustrate the cumulative depth available at the four best ask price buckets. Along the horizontal axis is the alpha forecast for market order t. The buckets for the alpha forecast are centered around the unconditional alpha ($\bar{\alpha}$) stretching from minus twice the standard deviation of the alpha forecast through plus twice the standard deviation. The step size for the bucket is 0.4 times the standard deviation of the alpha forecast.

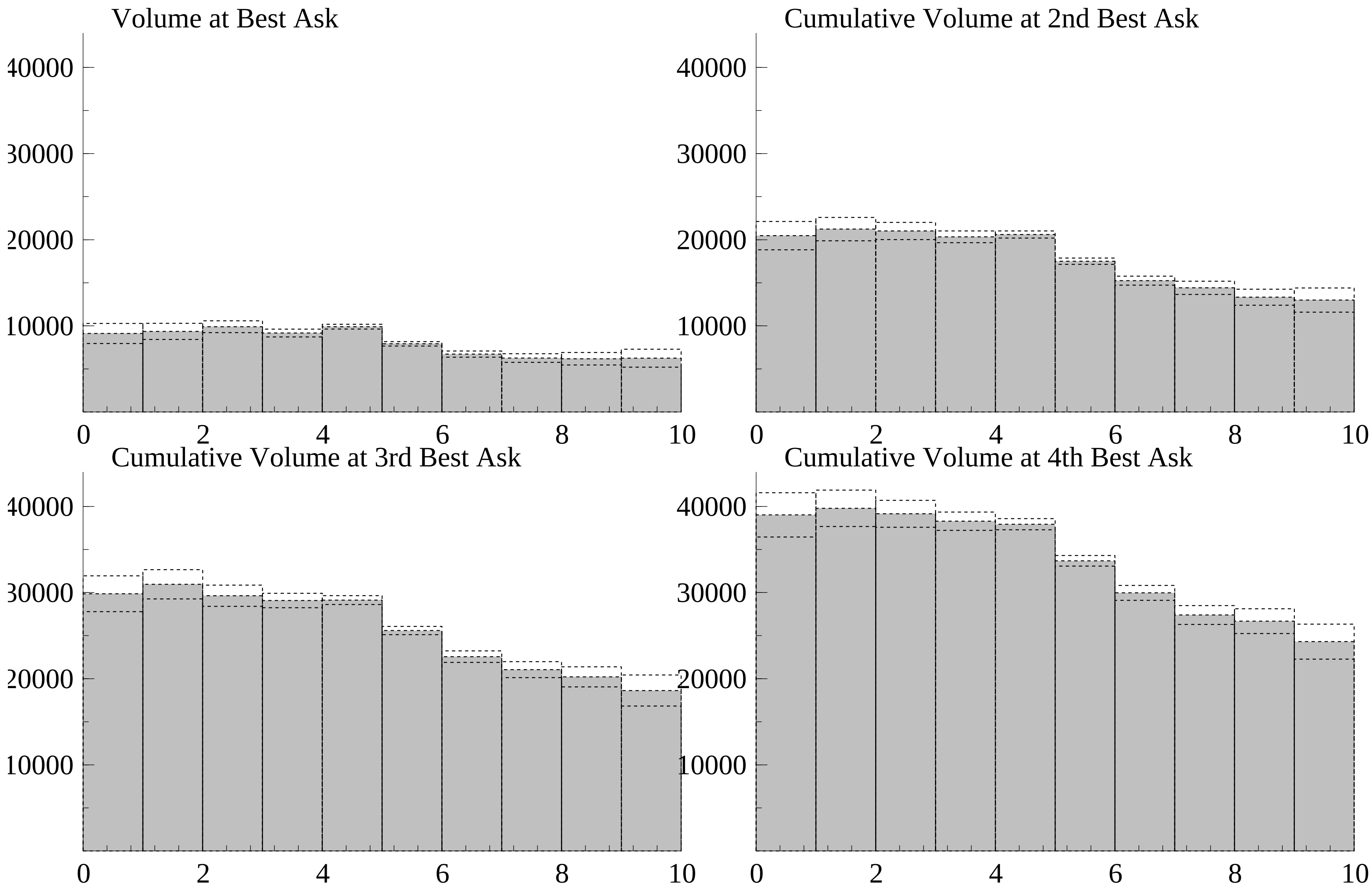

Figure 4: **DTE: Order Book Depth on Ask Side vs. Alpha**. These four graphs illustrate the cumulative depth available at the four best ask price buckets. Along the horizontal axis is the alpha forecast for market order t. The buckets for the alpha forecast are centered around the unconditional alpha ($\bar{\alpha}$) stretching from minus twice the standard deviation of the alpha forecast through plus twice the standard deviation. The step size for the bucket is 0.4 times the standard deviation of the alpha forecast.

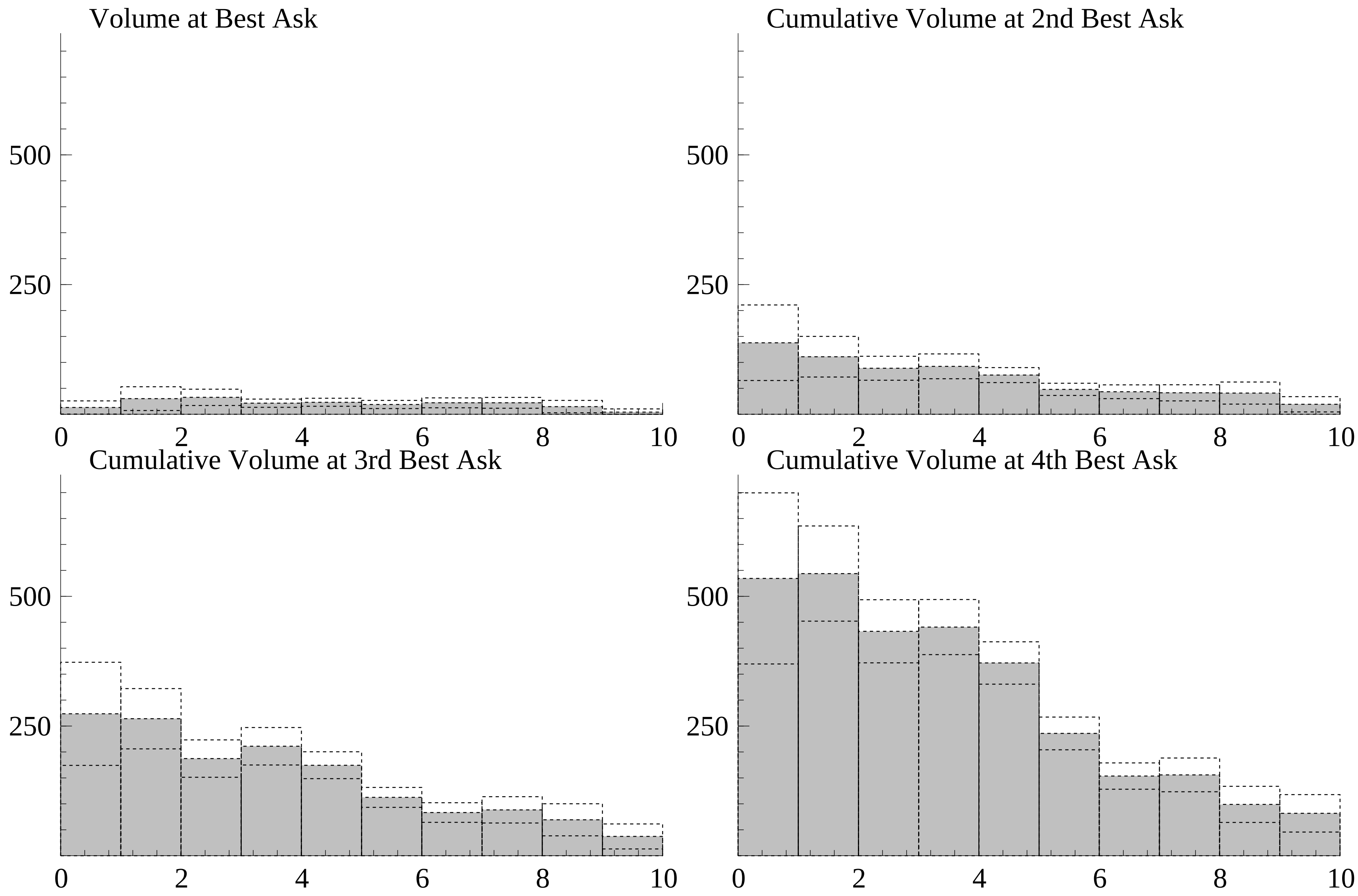

Figure 5: **SAP: Order Book Depth on Ask Side vs. Alpha**. These four graphs illustrate the cumulative depth available at the four best ask price buckets. Along the horizontal axis is the alpha forecast for market order t. The buckets for the alpha forecast are centered around the unconditional alpha ($\bar{\alpha}$) stretching from minus twice the standard deviation of the alpha forecast through plus twice the standard deviation. The step size for the bucket is 0.4 times the standard deviation of the alpha forecast.

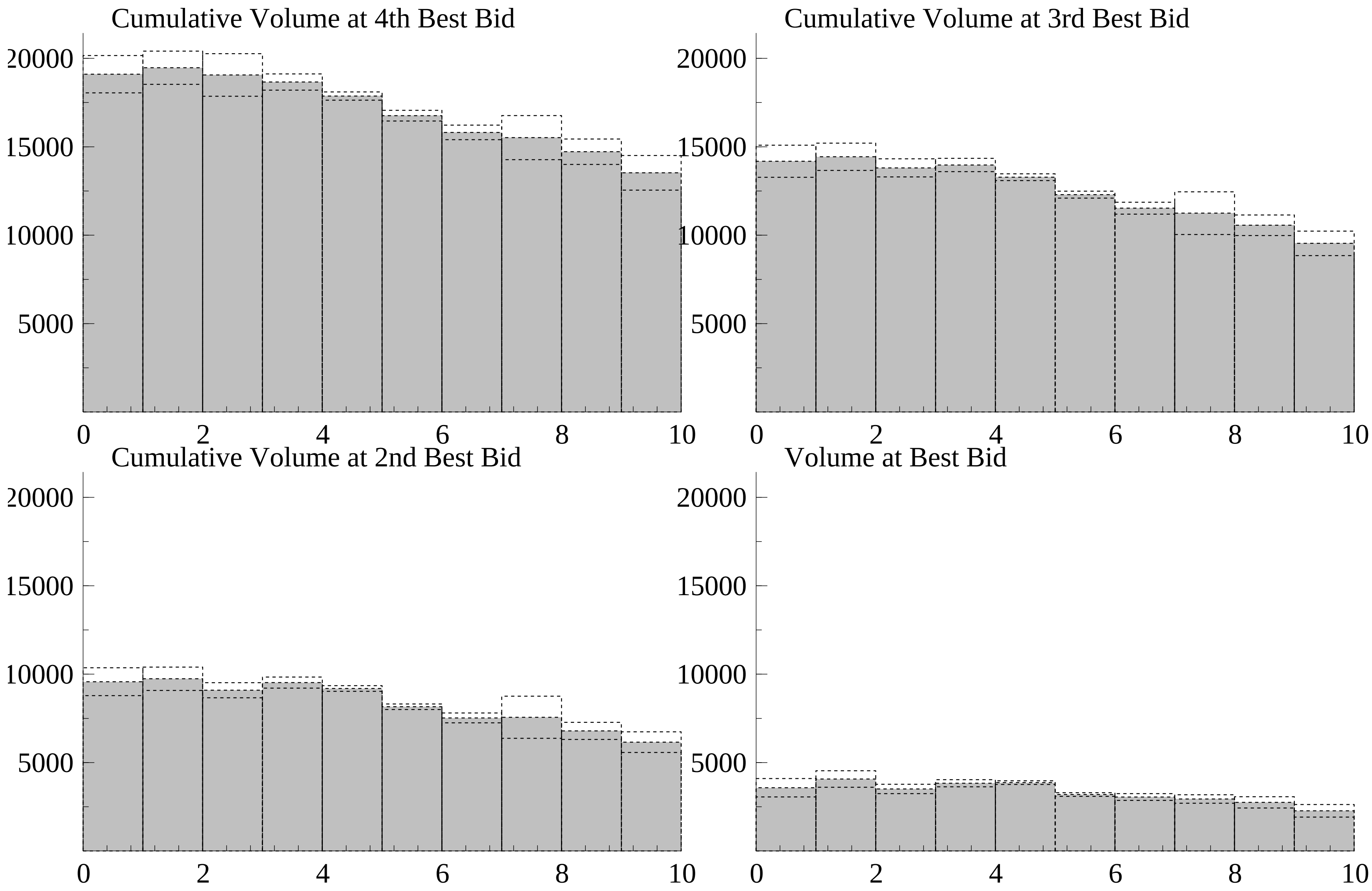

Figure 6: **DCX: Order Book Depth on Bid Side vs. Alpha**. These four graphs illustrate the cumulative depth available at the four best ask price buckets. Along the horizontal axis is the alpha forecast for market order t. The buckets for the alpha forecast are centered around the unconditional alpha ($\bar{\alpha}$) stretching from minus twice the standard deviation of the alpha forecast through plus twice the standard deviation. The step size for the bucket is 0.4 times the standard deviation of the alpha forecast.

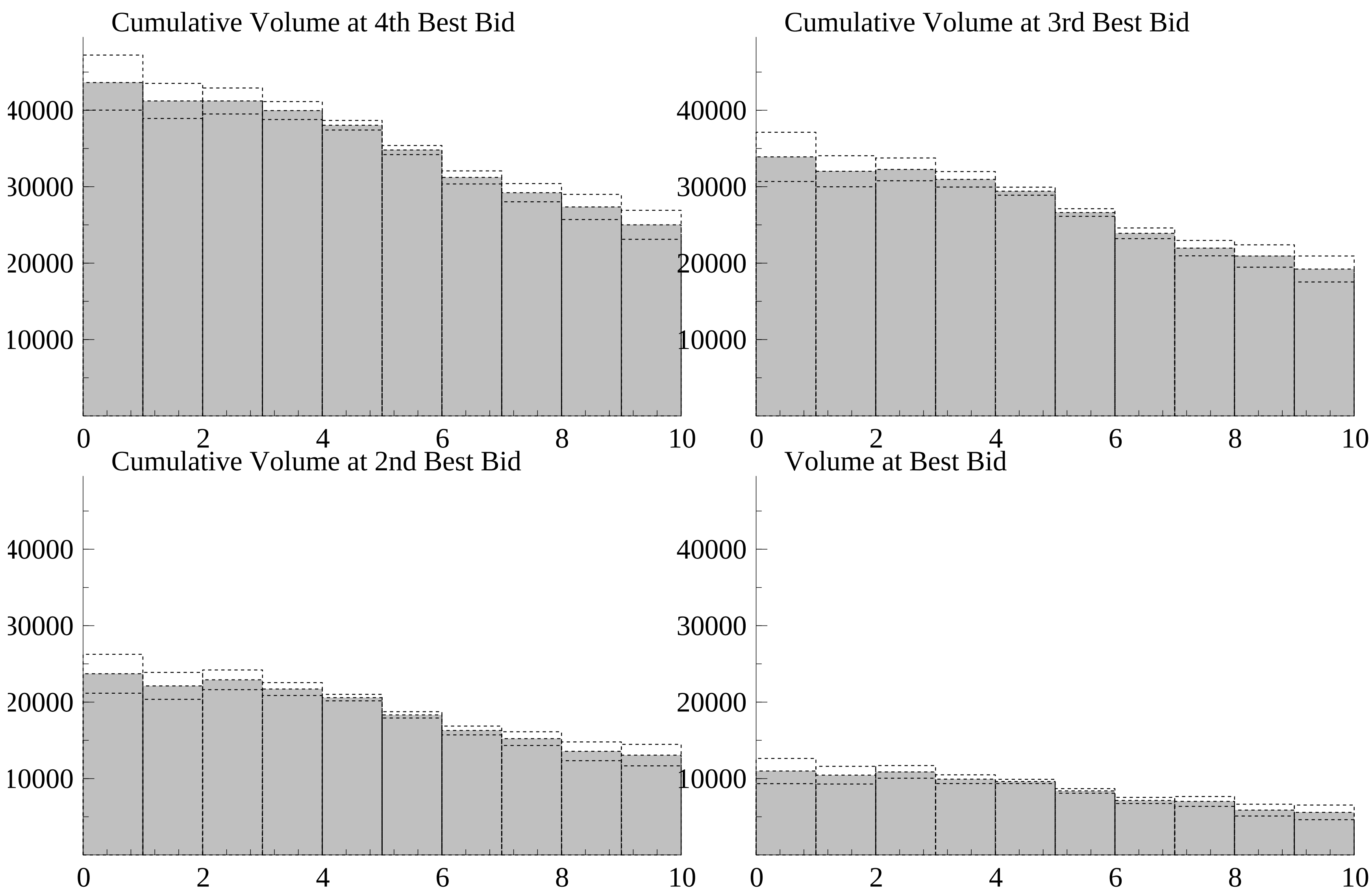

Figure 7: **DTE: Order Book Depth on Bid Side vs. Alpha**. These four graphs illustrate the cumulative depth available at the four best ask price buckets. Along the horizontal axis is the alpha forecast for market order t. The buckets for the alpha forecast are centered around the unconditional alpha ($\bar{\alpha}$) stretching from minus twice the standard deviation of the alpha forecast through plus twice the standard deviation. The step size for the bucket is 0.4 times the standard deviation of the alpha forecast.

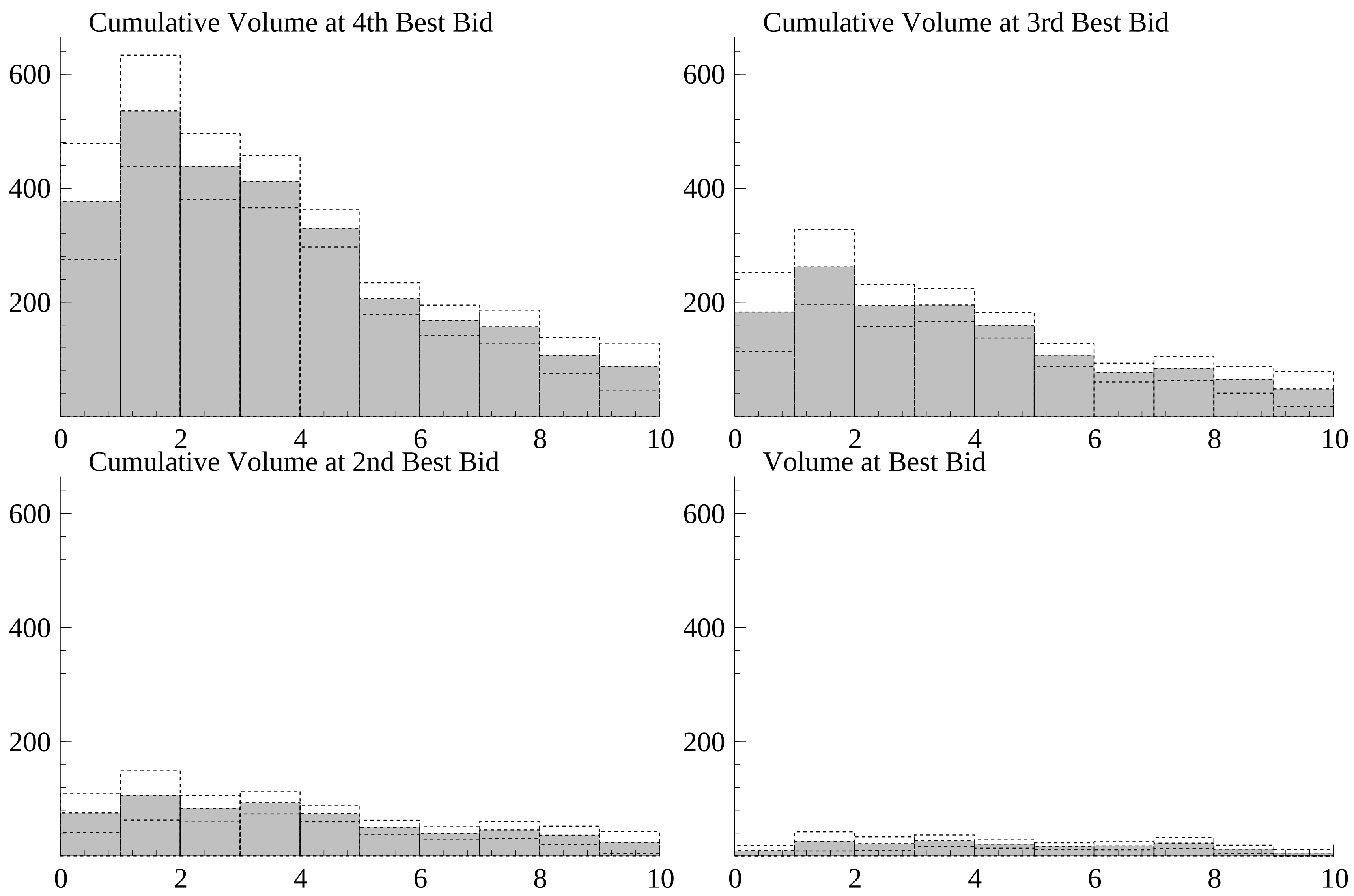

Figure 8: **SAP: Order Book Depth on Bid Side vs. Alpha**. These four graphs illustrate the cumulative depth available at the four best ask price buckets. Along the horizontal axis is the alpha forecast for market order t. The buckets for the alpha forecast are centered around the unconditional alpha ($\bar{\alpha}$) stretching from minus twice the standard deviation of the alpha forecast through plus twice the standard deviation. The step size for the bucket is 0.4 times the standard deviation of the alpha forecast.

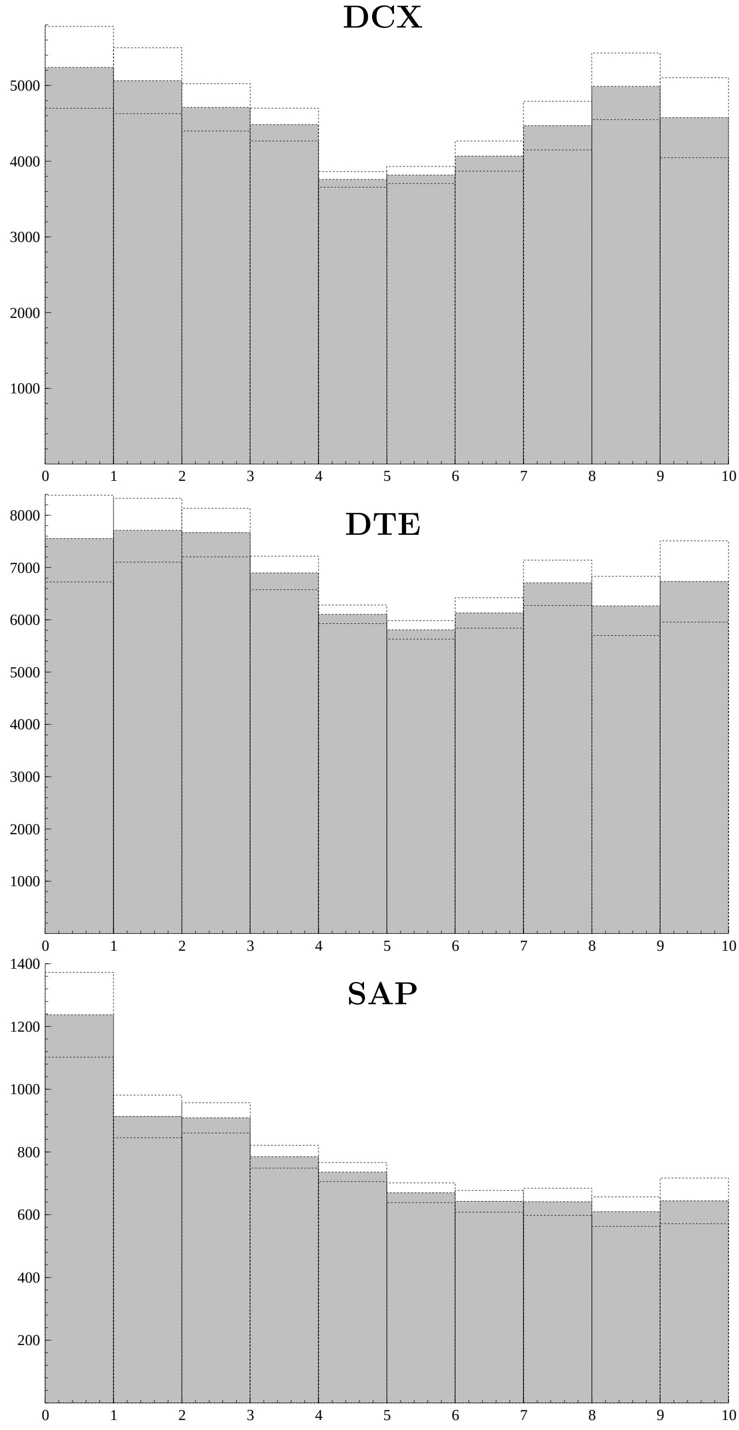

Figure 9: **Mean Market Order Size (y-axis) vs. Alpha (x-axis)**. This graph depicts the mean size of the (aggregated) market order for different values of alpha. Along the horizontal axis is the alpha forecast for market order t. The buckets for the alpha forecast are centered around the unconditional alpha ($\bar{\alpha}$) stretching from minus twice the standard deviation of the alpha forecast through plus twice the standard deviation. The step size for the bucket is 0.4 times the standard deviation of the alpha forecast.

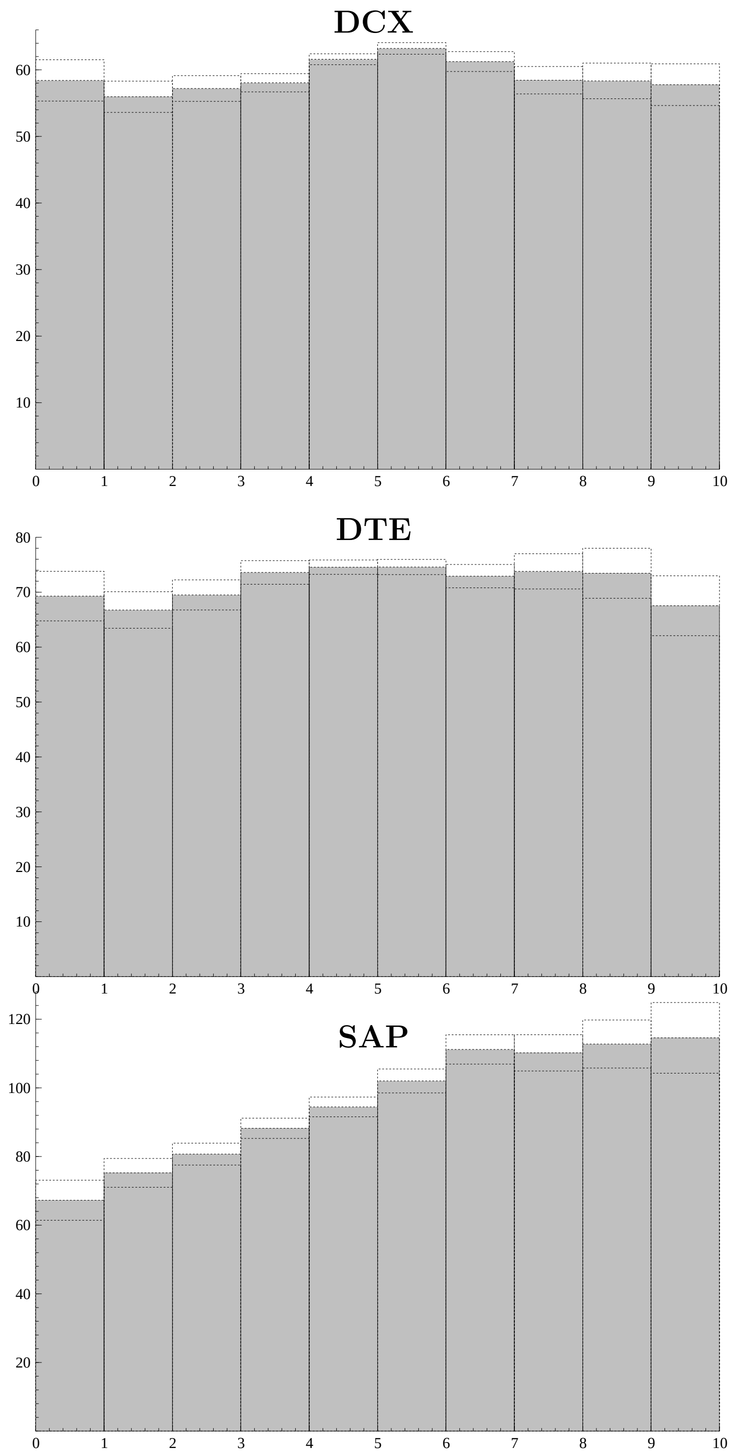

Figure 10: **Mean Duration in Seconds (y-axis) vs. Alpha (x-axis)**. This graph depicts the mean duration for different values of alpha. Along the horizontal axis is the alpha forecast for market order t. The buckets for the alpha forecast are centered around the unconditional alpha ($\bar{\alpha}$) stretching from minus twice the standard deviation of the alpha forecast through plus twice the standard deviation. The step size for the bucket is 0.4 times the standard deviation of the alpha forecast.