CENTER FOR FINANCIAL STUDIES

CFS Working Paper Series

No. 468

Gökhan Cebiroglu, Nikolaus Hautsch, Ulrich Horst

Order Exposure and Liquidity Coordination: Does Hidden Liquidity Harm Price Efficiency?

Center for Financial Studies | Goethe University Frankfurt | Grüneburgplatz 1 | 60323 Frankfurt am Main | Germany
Phone +49 69 798 30050 | Fax +49 69 798 30077 | info@ifk-cfs.de | www.ifk-cfs.de

The CFS Working Paper Series

presents ongoing research on selected topics in the fields of money, banking and finance. The papers are circulated to encourage discussion and comment. Any opinions expressed in CFS Working Papers are those of the author(s) and not of the CFS.

The Center for Financial Studies, located in Goethe University Frankfurt's House of Finance, conducts independent and internationally oriented research in important areas of Finance. It serves as a forum for dialogue between academia, policy-making institutions and the financial industry. It offers a platform for top-level fundamental research as well as applied research relevant for the financial sector in Europe. CFS is funded by the non-profit-organization Gesellschaft für Kapitalmarktforschung e.V. (GfK). Established in 1967 and closely affiliated with the University of Frankfurt, it provides a strong link between the financial community and academia. GfK members comprise major players in Germany's financial industry. The funding institutions do not give prior review to CFS publications, nor do they necessarily share the views expressed therein.

Order Exposure and Liquidity Coordination: Does Hidden Liquidity Harm Price Efficiency?*

Gökhan Cebiroğlu[†]
University of Vienna

Nikolaus Hautsch[‡]
University of Vienna
Center for Financial Studies

Ulrich Horst[§]
Humboldt Universität zu Berlin

September 18, 2014

Abstract

We develop a model of an order-driven exchange competing for order flow with off-exchange trading mechanisms. Liquidity suppliers face a trade-off between benefits and costs of order exposure. If they display trading intentions, they attract additional trade demand. We show, in equilibrium, hiding trade intentions can induce mis-coordination between liquidity supply and demand, generate excess price fluctuations and harm price efficiency. Econometric high-frequency analysis based on unique data on hidden orders from NASDAQ reveals strong empirical support for these predictions: We find abnormal reactions in prices and order flow after periods of high excess-supply of hidden liquidity.

JEL classification: G02, G10, G23

Keywords: liquidity externalities, order flow, trade signaling, limit order book

*The paper formerly circulated under the title ”Does Hidden Liquidity Harm Price Efficiency?”. We thank Torben Andersen, Damir Filipovic, Tony Hall, Joel Hasbrouck, Albert Kyle, Bruce Lehmann, Helmut Lütkepohl, Albert Menkveld and Mathieu Rosenbaum for their helpful comments and suggestions. We would also like to thank the participants at the Workshop "Market Microstructure Confronting Many Viewpoints", Paris, December 2012, the sixth "Erasmus Liquidity Conference" at the Rotterdam School of Management, Erasmus University, August 2013, the Workshop "Market Microstructure and Nonlinear Dynamics", Evry, June 2013, the Bachelier World Congress, Brussels, June 2014, the Banff "Workshop on New Directions in Financial Mathematics and Mathematical Economics", Banff, Canada, July 2014, the Workshop "Measuring and Modeling Financial Risk with High Frequency Data", European University Institute, Florence, June 2014, and the finance seminars at the HEC Lausanne and University of Technology, Sydney, for their helpful discussions and comments. Financial support and data provision from Deutsche Bank and the Collaborative Research Center 649 “Economic Risk” are gratefully acknowledged. Hautsch also acknowledges financial support from the Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF).

[†]Faculty of Business, Economics and Statistics, Oskar-Morgenstern-Platz 1, A-1090 Vienna, *email*: goekhan.cebiroglu@univie.ac.at.

[‡]Faculty of Business, Economics and Statistics, Oskar-Morgenstern-Platz 1, A-1090 Vienna, *email*: nikolaus.hautsch@univie.ac.at.

[§]Department of Mathematics, Unter den Linden 6, D-10099 Berlin, *email*: horst@mathematik.hu-berlin.de.

1 Introduction

Today's electronic markets organize trading by matching liquidity suppliers with liquidity demanders. Liquidity suppliers expose their trade interest by submitting limit orders to the electronic limit order book and provide trading opportunities to counterparties. Exposure of trade intention is thus central to the process of trading. Over the most recent decade, markets have nevertheless increased the proliferation of *hidden liquidity*, enabling investors to conceal their trading intentions. In fact, hidden liquidity has become a sizable proportion of the overall market liquidity. Bessembinder et al. (2009) report that 44% of volume in Euronext-Paris is hidden and Frey and Sandas (2009) report a proportion of 16% of hidden liquidity in the German Xetra market.

Although these numbers suggest that a growing trading population perceive an individual benefit in trading hidden orders, there is, however, an ongoing debate about the impact of hidden liquidity for the overall market. The controversy has increasingly moved to the center of regulatory debates.[1] Critics claim that intransparent markets can damage the role of stock exchanges as a venue for investors to establish fair and efficient prices (see, e.g., Hendershott and Jones (2005)). On the other hand, proponents argue that hidden liquidity attracts market depth that would otherwise not partake in trading (see Aitken et al. (2001)). The ongoing controversy, particularly the question of how hidden liquidity affects market quality, is arguably relevant for both stock exchanges and regulators.

A second fundamental question concerns the origination of hidden liquidity: What is the impact of trade design on the supply of hidden liquidity? The answer has important implications for stock exchange operators. Imposing adequate trading rules might help reducing downside effects of hidden liquidity while preserving some of its benefits. In this case, exchange operators are challenged to broaden the range of market instruments to ensure adequate transparency levels in markets.

To address these questions, we analyze the role of hidden liquidity when public exchanges compete for order flow with off-exchange trading mechanisms. As markets become increasingly fragmented, alternative trading venues, such as dark pools, crossing networks and private broker dealer networks, have significantly increased their share of market liquidity at the ex-

[1] In view of new rules for the markets in financial instruments (MIFID II), European Commissioner Barnier echoed the growing concerns: "... Strict transparency rules will ensure that [hidden] trading of shares and other equity instruments which undermine efficient and fair price formation will no longer be allowed."

pense of the public exchange. To understand the effects and determinants of hidden liquidity in a given market, it is therefore necessary to account for the interplay between inter-market and intra-market liquidity competition.

We theoretically and empirically show that large hidden orders on public exchanges can intensify market fragmentation, increase transaction costs, and induce excess returns and excess volatility. We illustrate that these price effects are *not* related to information but arise from *mis-coordination* between trading counterparties: when liquidity suppliers do not (or incompletely) display their trade intentions, counterparties (i.e., liquidity demanders) cannot effectively coordinate the pricing and timing of mutually beneficent trades. A key result is that the induced counterparty mismatch not only impairs overall welfare but also generates artificial price pressure, causes liquidity migration to off-exchange trading venues and thus harms liquidity provision in the public exchange.

Our analysis builds on a model with two markets: a primary exchange and an alternative trading venue. The primary market is organized as an electronic limit order book. In our model, three traders arrive sequentially and act strategically. The *liquidity supplier* in the primary market has discretion over the displayed/hidden size of his order and faces liquidity competition from a *liquidity competitor*. The latter has discretion over order aggressiveness and may front-run (i.e., overbid) the liquidity supplier's order. The third trader is a *large investor* and has discretion over the trading venue. He can decide whether to trade in the primary market or in the alternative trading venue. Grossmann (1992) argues that it is more cost-efficient for large investors to trade *reactively* and to monitor the market for trading opportunities instead of actively expressing their trading intentions in the public exchange.[2] We assume that our large trader is such a *latent* trader. He has an incentive to trade in the public market when the market discloses a critical mass of liquidity.

Thus, liquidity suppliers can elicit order flow from the latent investor if they actively expose their trading intentions. The downside of revealing trading intentions, however, is to be confronted with liquidity competition. The first trader therefore faces a trade-off between losses incurred from being overbid and the risk of non-execution as counterparty order flow is not attracted. On the other hand, the liquidity competitor decides whether to overbid the hidden order. Not overbidding the hidden order results in loss of execution priority which aggravates with the liquidity supplier's exposure (display) size. Hence, the liquidity competitor's decision ultimately involves a trade-off between the loss of priority and the costs of trading aggressively

[2] See Hasbrouck and Saar (2009), Bessembinder et al. (2009), and Harris (1997) for a similar argument.

(i.e., overbidding). Our key results are therefore driven by the interaction of both mechanisms, liquidity competition in the primary exchange and order flow attraction from latent investors.

A central finding of our study is that due to *mis-coordination* (large) hidden orders can significantly harm price discovery and induce excess price fluctuations. When traders conceal their trade intentions by using hidden orders, they are less likely to attract latent counterparties, leading to increased execution risk.[3] When outstanding orders are not executed and when traders are pre-committed to trading, remaining (hidden) orders eventually have to be canceled and re-submitted as aggressive market orders to ensure liquidation. Ultimately, this switch to aggressive trading, however, induces price pressures (i.e., price fluctuations) as the market is confronted with excess trade demand. These price pressures do not arise when large orders are fully revealed. Large *displayed* orders attract counterparty order flow and thus get likely executed right-away, absorbing any price pressure arising from potentially non-executed orders. In this sense, large hidden orders can impair price efficiency by causing mismatches between liquidity suppliers and (latent) liquidity demanders. The implication that *large* hidden orders can harm market quality is relevant as the empirical literature suggests that in fact most large orders are hidden.[4]

Our work distinguishes from the traditional information-based literature (see, e.g., Kyle (1985)): In our framework, price fluctuations are due to *non-informational* frictions and arise from impaired trade coordination (see Admanti and Pfleiderer (1991)). The latter, however, is affected by liquidity competition and the level of liquidity externalities to which a market is exposed. These driving forces have been recently identified as key mechanisms of modern trading (see, e.g., Foucault et al. (2005) and Hendershott and Mendelson (2002)). Our findings complement this line of research and demonstrate that fundamental empirical market microstructure relationships in limit order book markets (see, e.g., Biais (1993), Ranaldo (2004), and Hall and Hautsch (2006)) can be explained by the interplay between liquidity competition and the disclosure of trading intentions. For instance, we show that liquidity suppliers hide a larger fraction of their orders when the spread is wide and when the opposite side of the book is thin. In this case, liquidity suppliers have a greater incentive to bypass (i.e., overbid) orders, forcing liquidity suppliers to limit the display size. Second, we find that public exchanges exhibit a

[3] Bessembinder et al. (2009) provide evidence that hidden orders have a low execution probability.

[4] For instance, Bessembinder et al. (2009) reports that on average 76% of large orders contain a hidden size and that 75% of the volume of these orders is hidden. Frey and Sandas (2009) find that (partially) hidden orders are on average 12 to 20 times larger than exposed limit orders.

higher level of hidden depth when the costs of trading in off-exchange venues are low. Consequently, latent traders are less likely to trade in the primary market and liquidity suppliers tend to hide more to avoid liquidity competition. Third, the role of the tick size is ambiguous, as two counteracting mechanisms exist. In line with Harris (1997), we find that a larger tick size reduces the likelihood of overbidding and therefore increases the incentives for order exposure. However, we also illustrate that a larger tick size reduces the incentives of latent traders to trade in the primary exchange, thus increasing execution risk.

Finally, our model provides a novel perspective on the causal relation between bid-ask spreads and market volatility. We show that the spread and market volatility are linked through hidden liquidity. The logic is as follows: Wider spreads increase liquidity competition as overbidding existing orders by a single price increment is less costly. Consequently and in line with Harris (1997) and Buti and Rindi (2013), liquidity suppliers hide an increasing fraction of their order to compensate for stronger liquidity competition. This reduces the presence of latent traders and thus increases liquidity suppliers' risk of non-execution. Eventually, as liquidity suppliers have to trade a larger fraction of their non-executed orders via aggressive market orders, stronger price reactions emerge. Thus, in our model, wider spreads cause larger price fluctuations as a result of liquidity competition, whereas the traditional models based on information asymmetry suggest the exact opposite causation.[5]

Our theoretical hypotheses are empirically tested using two unique data sets. To infer the presence of hidden depth, we use NASDAQ ModelView data, which provides minute-by-minute snapshots of the entire hidden and displayed depth in the order book of S&P 500 stocks. Moreover, order flow information is reconstructed from ITCH order-message data contained in NASDAQ's TotalView data set, as processed via the limit order book platform LOBSTER.[6] This data set contains the reconstructed limit order book at each instant, including all order messages, i.e., order cancelations, modifications, submissions, and executions.

The empirical analysis confirms the main theoretical predictions. First, on the basis of cross-sectional simultaneous equations regressions, we find that the fraction of hidden liquidity is higher when (on average) spreads are wider and (relative) tick sizes are larger. Similarly, we show that stocks with a larger proportion of hidden liquidity are more volatile. Second, to

[5]For instance, Glosten and Milgrom (1985) and Copeland and Galai (1983) argue that wider spreads compensate market makers for the risk of adverse selection in volatile markets, implying that higher volatility increases spreads.

[6]See https://lobster.wiwi.hu-berlin.de/

test whether price reactions are triggered by hidden order submissions, we estimate the short-run and long-run impact of (one-sided) hidden liquidity using impulse response functions. We model the order book and order flow dynamics using a high-frequency vector autoregressive process with the vector of endogenous variables capturing order flow and order book characteristics. In line with our theoretical predictions, we find that (buy) hidden order submissions cause significant subsequent (buy) market order submission activities, which, in turn, generate price pressures resulting in significant mid-quote reactions. In contrast, as correctly predicted by our model, excess supply in *displayed* liquidity elicitates counterparty order flow and therefore reduces the price reactions emanating from non-executed, outstanding orders.

Our results are important for both market regulators and exchange operators. Public markets compete for order flow in an increasingly fragmented market. If they lose this "battle for liquidity", the public price formation process may be harmed. Extant literature suggests that these liquidity externalities are closely related to market transparency. In this work, we show that transparency on primary exchanges can enhance market quality and price efficiency, as displayed liquidity attracts additional order flow from latent investors. To increase market share and improve price formation of public exchanges, our analysis suggests that market operators should broaden their network with other liquidity pools, enhance their order routing infrastructure and provide large institutional investors direct market access and real-time monitoring capabilities such that liquidity opportunities can be seized instantly as they arise.

The remainder of this paper is structured as follows: In Section 2, we introduce the theoretical model. Equilibrium results follow in Section 3, yielding main theoretical implications and testable predictions: We first analyse a baseline version of the model without latent trader and derive conclusions on different aspects of liquidity competition. In a second step, we extend our analysis to allow for the existence of alternative trading venues. Employing a partial equilibrium analysis, Section 4 derives key insights regarding the impact of hidden order submissions on several dimensions of the market. In Sections 5 and 6, we empirically verify the key predictions based on cross-sectional and time series analysis. Finally, Section 7 concludes.

2 A Model of Liquidity Competition and Trade Signaling

We introduce a sequential trading model with discrete timing in two markets: a primary exchange that operates as an electronic limit order book and an alternative trading platform that is specialized in trading large blocks. The trading population consists of sequentially arriving liquidity suppliers, a large (block) trader, and a noise trader. While liquidity suppliers and the large block trader interact strategically, the noise trader trades for exogenous reasons. In the next section, we explain the details of the institutional framework, the timing of the events, and the trading population.

2.1 Institutional Framework

The primary exchange is *order driven*. Thus, investors can openly quote prices and trade against the public order book. In this market, liquidity suppliers have the possibility of hiding their orders. Hence, not all liquidity is fully disclosed to other market participants. Most generally, the alternative trading venue is a brokerage mechanism that locates counterparties to investors, similar to upstairs markets, crossing networks, or broker-dealer networks. The alternative trading venue provides a second source of liquidity, which is particularly important when liquidity in the primary exchange is thin. This trading structure is particularly tailored to large institutional investors and block traders (see, e.g., Madhavan and Cheng (1997) and Keim and Madhavan (1996)).

The Electronic Exchange

Orders are submitted on a discrete price grid with a minimum tick size Δ. Prices at time t are given relative to the prevailing best ask, A_t, and bid, B_t. To keep the model tractable, we assume a liquid market, where the *spread resiliency*, i.e., the speed of reversion of spreads to their equilibrium level (see, e.g., Foucault et al. (2005)), is higher than the time scale underlying our model.[7] This assumption is in line with the assumption of a *competitive* spread as discussed by Kyle (1985) and Glosten and Milgrom (1985).

Assumption 1. *Market makers are competitive such that the spread instantly reverts back to the competitive level S, once a change occurred, i.e.,*

$$S_t = S \qquad \forall\, t > 0. \tag{2.1}$$

This assumption implies that order submissions, cancelations, and executions on one side of the market are instantly corrected by price revisions on the opposite side, such that the spread

[7] The speed of resiliency is known to be linked to the liquidity of a market; see Biais et al. (1995), Degryse et al. (2005), or Domowitz and Madhavan (2003).

always remains at its equilibrium value. We assume that limit orders can be submitted either on the best bid (ask) quote or with a price improvement of one tick (into the spread). Limit orders are executed against incoming market orders in a discriminatory fashion by using a hierarchy of (i) price priority, (ii) display priority, and (iii) time priority. In line with Harris and Hasbrouck (1996), we assume that limit orders pose a *pre-commitment* to trade. Hence, suppose that a buy (hidden) limit order trader has a trading horizon until time T and that the trader enters the market at time t, aiming at buying N shares at the (submission) price p_t^l. Consequently, at terminal time T, non-executed shares are canceled and turned into market orders to guarantee trade execution. The trader's implementation shortfall according to Perold (1988) is given by

$$\Pi_T := \underbrace{\left(p_t^l - B_t\right) X_T}_{\text{executed limit order}} + \underbrace{\left(p_T^m - B_t\right)\left(N - X_T\right)}_{\text{non-executed limit order}}, \tag{2.2}$$

where $B_t N$ is the benchmark price at arrival time t, p_T^m denotes the marginal price of the market order at time T and X_T is the volume executed at the initial submission price level p_t^l until time T. While the trader's limit order submission price p_t^l is a matter of choice, the market order price p_T^m at T is determined by the order's price impact and thus depends on the execution volume $(N - X_T)$ and the prevailing opposite-side (i.e., sell-side) depth. In order to keep the analysis tractable, we assume that the price impact is linear in the remaining shares, i.e.,

$$p_T^m := B_T + \underbrace{S + \frac{1}{2}\beta(N - X_T)}_{\text{effective spread}}. \tag{2.3}$$

Note that β represents the marginal *price impact* of the market (buy) order and corresponds to the inverse of the buy-side depth density. The *effective* or *realized* spread gives the cost (i.e., price impact) of trading aggressive market orders.

The Alternative Trading Venue

In the alternative off-exchange trading venue, investors employ the services of brokers or brokerage mechanisms to locate counterparties. Consistent with Keim and Madhavan (1996), Booth et al. (2002), and Harris (2003), we assume that the settlement price in this market, p_t^γ, equals a (reference-) price from the public exchange plus a commission fee γ that compensates the broker's counterparty search costs:

$$p_t^\gamma = \begin{cases} A_t + \gamma & \text{in case of a buy,} \\ B_t - \gamma & \text{in case of a sell.} \end{cases} \tag{2.4}$$

We implicitly assume that the brokerage market is sufficiently liquid to assure trade execution.

2.2 Market Participants and Timing

We consider four market participants that arrive in sequential order: a *hidden trader* (H) arriving at t_0 with probability q, a *liquidity competitor* (C) arriving at t_1, a *latent* block trader (L) arriving at t_2 with probability μ, and a *noise trader* arriving at T with $t_0 < t_1 < t_2 < T$. The noise trader represents exogenous and random liquidity demand, reflected by a market order with size x. For simplicity, we assume that x is exponentially distributed with mean λ. The remaining three traders interact strategically and are risk-neutral. We denote the respective strategies of each strategic investor by σ_H, σ_C, and σ_L.

Traders in the Primary Exchange

The (buy) hidden trader H has N_H shares to trade until time T. He has discretion over the display size D_H. The remaining $N - D_H$ are kept hidden from the public. Hence, his strategy σ_H consists of deciding on the magnitude of D_H. In line with the trading mechanism of the primary exchange (see Section 2.1), at arrival time t_0, he submits his order at the best prevailing bid price B_{t_0}. As the trading horizon T is fixed, we henceforth omit the time subscript and index variables that are associated with the hidden trader or the liquidity competitor by H and C, respectively. According to (2.2) and (2.3), the hidden trader's payoff is then given by

$$\Pi_H = \left(S + \frac{1}{2}\beta(N_H - X_H(\sigma_H, \sigma_C, \sigma_L))\right)\left(N_H - X_H(\sigma_H, \sigma_C, \sigma_L)\right). \tag{2.5}$$

The liquidity competitor C arrives at t_1 with size N_C. His strategy σ_C consists of deciding on the optimal submission price level: submission at the prevailing bid price B_{t_0} ("$stay$") or overbidding at $B_{t_0} + \Delta$ ("$step$"). Accordingly, his payoff derived from (2.2) is

$$\begin{aligned}\Pi_C = {} & \Delta \cdot 1_{\{\sigma_C = step\}} X_C(\sigma_H, \sigma_C, \sigma_L) \\ & + \left(\Delta \cdot 1_{\{\sigma_C = step\}} + S + \frac{1}{2}\beta(N_C - X_C(\sigma_H, \sigma_C, \sigma_L))\right)\left(N_C - X_C(\sigma_H, \sigma_C, \sigma_L)\right).\end{aligned} \tag{2.6}$$

The execution volumes X_H and X_C (and thereby the payoffs) depend on the trading strategies of all traders, σ_C, σ_H, and σ_L and are derived in Lemma A1 and Lemma A2 (see Appendix).

2.2.1 The Latent (Block) Trader

At t_2, a (sell) block trader arrives with a total trade demand of N_L shares. We assume that he monitors the primary exchange with probability μ and that he has the strategic choice σ_L

between trading on-exchange ($\sigma_L = on$) or off-exchange, i.e., in the alternative market ($\sigma_L = off$).[8] If he does not monitor the primary exchange, he only trades on the alternative market, i.e., $\sigma_L = off$. As the alternative market provides an infinite liquidity reservoir at price p_t^γ (2.4), there is no execution risk. In contrast, the liquidity supply in the primary exchange is limited, and the total trade demand of block traders may not be executed. Therefore, the block trader uses the alternative market as a market of *last resort* to enforce the execution of non-executed shares.[9]

We assume that the latent trader is large in the sense that his trading demand exceeds the combined liquidity supply of both liquidity suppliers:

Assumption 2 (Latent Trader Demand)**.** *The latent trader is large in the sense of*

$$N_L > N_H + N_C. \tag{2.7}$$

For large investors, continuous expression of trade demand (in terms of limit orders) in public markets is costly (see, e.g., Grossmann (1992)); hence, we assume that the latent trader uses market orders only. Thus, using (2.4) and assuming that the latent trader only trades the amount of shares that is openly displayed, i.e., at most $D_H + N_C$ shares, the payoff at time T in excess of the arrival price B_{t_0} is given by

$$\Pi_L = \begin{cases} (\Delta \cdot 1_{\{\sigma_C = step\}} - \gamma)N_L & \text{if} \quad \sigma_L = off, \\ \Delta N_C \cdot 1_{\{\sigma_C = step\}} - (\Delta + \gamma)(N_L - D_H - N_C) & \text{if} \quad \sigma_L = on. \end{cases} \tag{2.8}$$

Hence, in the case of off-exchange trading, the latent trader has to pay the off-exchange fee γ but benefits from an increase in the off-exchange settlement price if the liquidity competitor decides to overbid the prevailing quote by one tick Δ. If the latent trader enters the primary exchange, he "saves" the fee γ but can maximally trade $D_H + N_C$ shares. By completely executing the (displayed) liquidity supply $D_H + N_C$, he removes the order book depth of the price levels $B_{t_0} + \Delta$ (in case of overbidding by the liquidity competitor) and B_{t_0} and thus pushes the bid quote down to the level $B_{t_0} - \Delta$. Because of (2.4), this also lowers the corresponding settlement (sell) price and makes it more costly to execute the remaining $N_L - D_H - N_C$ shares on the off-exchange venue. The mechanism in which the latent block trader moves quotes on

[8] Traditionally, the off-exchange market for large block traders is called the upstairs market, whereas the exchange market is referred to as the downstairs market.

[9] Conrad et al. (2003) shows that 60% of large block trades use off-exchange trading mechanisms as a market of last resort.

the primary exchange, which in turn affects off-exchange settlement prices, is crucial to our model. As discussed in detail below, this mechanism induces the latent trader to stay away from the primary market if (displayed) liquidity supply is too low and off-exchange executions of remaining volume thus become too costly.

The Trading Game

Denote the action spaces of the hidden trader, liquidity competitor, and latent trader by Σ_H, Σ_C, and Σ_L, respectively:

$$\Sigma_H = [0, N_H], \qquad \Sigma_C = \{stay, step\}, \qquad \Sigma_L = \{on, off\}. \tag{2.9}$$

As both H and C are assumed to be buyers, a *better* strategy *reduces* the respective payoff, i.e., reduces transaction costs. Accordingly, the Nash equilibrium is characterized as follows.

Definition 1 (Nash Equilibrium Trading Strategies)**.** *The triple* $\sigma^* = (\sigma_C^*, \sigma_H^*, \sigma_L^*)$ *constitutes a Nash equilibrium if for any other set of strategies* $(\sigma_H, \sigma_C, \sigma_L)$ *the following relations hold:*

$$\mathbb{E}_{t_0}\left[\Pi^H(\sigma_H^*, \sigma_C^*, \sigma_L^*)|\right] \leq \mathbb{E}_{t_0}\left[\Pi^H(\sigma_H, \sigma_C^*, \sigma_L^*)|\right] \tag{2.10}$$

$$\mathbb{E}_{t_1}\left[\Pi^C(\sigma_H^*, \sigma_C^*, \sigma_L^*)|\right] \leq \mathbb{E}_{t_1}\left[\Pi^C(\sigma_H^*, \sigma_C, \sigma_L^*)|\right] \tag{2.11}$$

$$\mathbb{E}_{t_2}\left[\Pi^L(\sigma_H^*, \sigma_C^*, \sigma_L^*)|\right] \geq \mathbb{E}_{t_2}\left[\Pi^L(\sigma_H^*, \sigma_C^*, \sigma_L)|\right] \quad . \tag{2.12}$$

The *existence* of the Nash equilibrium in such a finite dynamic game with complete information is guaranteed by *Zermelo's Theorem*, and the equilibrium strategies of each player can be derived by applying the principle of sequential rationality or backward induction (see Mas-Colell et al. (1995)). The uniqueness of the equilibrium will be shown below.

3 Equilibrium Analysis

Solving for equilibrium in the general case $\mu \in [0, 1]$ is possible but rather tedious. Therefore, we confine our analysis to two benchmark cases that help to disentangle and highlight the key mechanisms. We therefore focus on a baseline model without latent demand ($\mu = 0$), and in a second step, we include a latent block trader with $\mu = 1$. The proofs for the lemmata and propositions in this section are provided in the Appendix. Corollaries follow straightforwardly from previous propositions.

3.1 Equilibrium without a Latent Block Trader

Under the absence of a latent block trader, the model is reduced to a game of pure liquidity competition between trader H and C. We first derive the competitor's best response strategy at t_1 and, subsequently, solve for the hidden trader's equilibrium strategy.

Lemma 1 (Liquidity Competitor's Best Response)**.** *Given the hidden trader's display size D_H, the competitor's best response σ_C^* obeys*

$$\sigma_C^* = \begin{cases} stay & \text{if} \quad D_H \leq \Phi_C, \\ step & \text{else}, \end{cases} \qquad \Phi_C := \begin{cases} \lambda \log\left(\frac{1}{1-g}\right) & \text{if} \quad g < 1, \\ \infty & \text{else}, \end{cases} \tag{3.1}$$

with

$$g := \frac{N_C}{\lambda} \frac{\Delta}{S\left(1 - e^{-\frac{N_C}{\lambda}}\right) + \beta\left(N_C - \lambda\left(1 - e^{-\frac{N_C}{\lambda}}\right)\right)}. \tag{3.2}$$

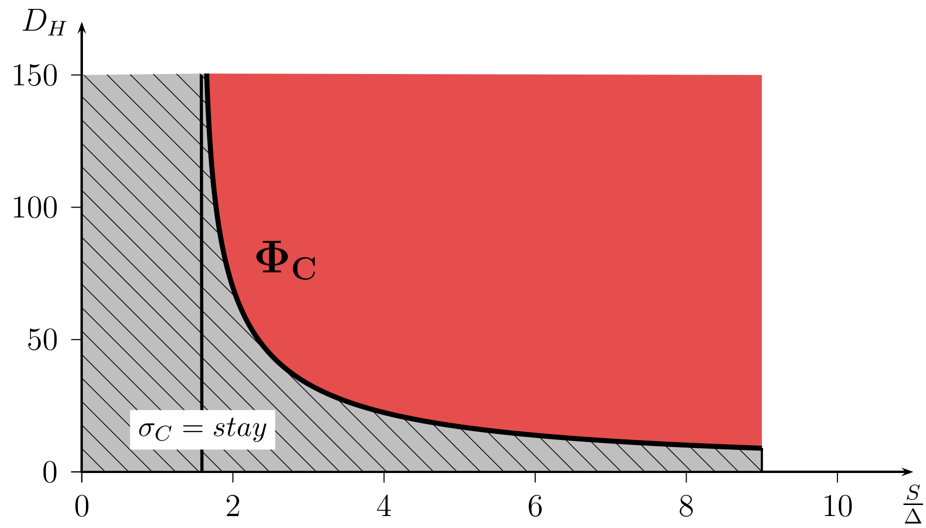
Hence, the liquidity competitor's order aggressiveness is governed by a (display) threshold Φ_C. The reason for this result is intuitive: for large display sizes D_H, the competitor faces a reduced execution probability owing to an increasing loss in time priority. To counter-balance this effect, he needs to overbid the hidden trader by a single tick Δ. This strategy is particularly beneficial if the spread S is wide, as in this case, it becomes costly to cross the spread and to trade the non-executed shares as market orders. On the other hand, the incentive for overbidding the current bid quote declines with increasing minimum tick size Δ as it represents the costs for overbidding (i.e., order aggressiveness).

Figure 1 illustrates these relationships by depicting the threshold Φ_C in dependence of the display size D_H and the ratio between the spread and the minimum tick size. The figure shows that for large displayed order volumes, it is beneficial for the competitor to increase liquidity competition by ”stepping ahead” even if the (relative) spread is comparably small. Nevertheless, a lower bound for S/Δ exists for which the competitor never overbids, regardless of D_H. [10] As of Figure 1, liquidity competitors never overbid in markets that have a spread less than 2 ticks Δ.

In equilibrium, the hidden trader limits his exposure to the critical size, which marginally prevents the competitor from overbidding ($\sigma_C = step$) and thus from gaining price priority. As consequence, he displays at most Φ_C shares:

[10]Note that the fraction S/Δ represents the spread in multiples of the tick size and can never be below 1.

Figure 1: Overbidding region triggered by the threshold Φ_C. The graph illustrates the scenario for $\lambda = N_C = 100$ shares and $\beta = 0$ (i.e., "thick" opposite-side depth). When the hidden trader's display size is large, i.e., $D_H > \Phi_C$, the liquidity competitor will improve the price, i.e., $\sigma_C = step$. When the spread (tick) is wide (small), his incentive to "step ahead" increases. Φ_C diverges at $\frac{S}{\Delta} = (1 - e^{-1}) \approx 1.582$.



Proposition 1 (Equilibrium)**.** *The hidden trader's and liquidity competitor's equilibrium strategies σ^*_H ($= D^*_H$) and σ^*_C obey*

$$\sigma^*_C = stay, \qquad D^*_H = \begin{cases} \Phi_C & \text{if} \quad N_H > \Phi_C, \\ N_H & \text{else.} \end{cases} \tag{3.3}$$

Hence, only large (hidden) traders need to limit their exposure because of competition in liquidity supply. Particularly, according to the next corollary, large traders hide an equally large portion of their order:

Corollary 1 (Large orders are hidden)**.** *In equilibrium,*

$$\lim_{N_H \to \infty} \frac{D^*_H}{N_H} = 0 \quad \textit{holds.} \tag{3.4}$$

The fact that particularly large investors use hidden orders to reduce their exposure costs is well documented in the empirical literature. For instance, Frey and Sandas (2009) report that iceberg or hidden orders are on average 12-20 times larger than ordinary limit orders. Bessembinder et al. (2009) show that 75% of "large" orders with a notional value exceeding 50,000 EUR are at least partially hidden. They also find that 87% of the volume of large orders is hidden. Summarizing the insights from Lemma 1 and Proposition 1 yields the following corollary:

Corollary 2 (Determinants of Hidden Liquidity without Latent Block Traders)**.** *Liquidity competition and thus the provision of hidden liquidity*

(i) rises with larger same-side depth D_H,

(ii) rises with wider bid-ask spread S,

(iii) declines with opposite-side depth,

(iv) declines with tick size Δ,

(v) rises with larger demand N_C.

These results are in line with extant empirical findings. For instance, Bessembinder et al. (2009) report that the decision to hide and the magnitude of the hidden size are positively affected by the size of the spread and negatively affected by opposite-side depth. Likewise, Biais et al. (1995), Ranaldo (2004), Cao et al. (2009) and Hall and Hautsch (2006) show that liquidity competition increases if spreads widen. Ranaldo (2004) and Cao et al. (2009) report that liquidity competition increases when same-side depth is large. Similarly, Harris (1994, 1996) and De Winne and D'Hondt (2009) report that hidden liquidity provision is low in stocks with large tick sizes. Finally, Harris (1994, 1996) shows that the presence of liquidity competition forces traders to hide their orders.

3.2 Equilibrium with Latent Block Traders

We extend the baseline model by allowing for an additional strategic trader: the latent block trader. This agent is actively monitoring the order book of the primary exchange for liquidity opportunities and has discretion over the trading place. Because he strategically chooses between both market places, he effectively acts as a source for liquidity externalities, as liquidity suppliers in the primary exchange have incentives to expose sufficient liquidity to elicit order flow from the latent trader. In this section, we derive the equilibrium by first computing the latent investor's optimal strategy. Then, the liquidity competitor's and hidden trader's best responses are derived recursively.

Lemma 2 (Latent Trader's Best Response)**.** *Given, $\sigma_H = D_H$ and σ_C, the latent trader's* *optimal strategy σ_L^* obeys*

$$\sigma_L^* = \begin{cases} off & if \quad 0 \le D_H < \Phi_{L_a}, \\ on & if \quad \Phi_{L_a} \le D_H < \Phi_{L_b} \quad and \quad \sigma_C = stay, \\ off & if \quad \Phi_{L_a} \le D_H < \Phi_{L_b} \quad and \quad \sigma_C = step, \\ on & if \quad \Phi_{L_b} \le D_H, \end{cases} \tag{3.5}$$

with $\Phi_{L_a} \le \Phi_{L_b}$ and

$$\Phi_{L_a} := \frac{(N_L - N_C)\Delta - N_C\gamma}{\gamma + \Delta}, \qquad \Phi_{L_b} := \frac{2(N_L - N_C)\Delta - N_C\gamma}{\gamma + \Delta}. \tag{3.6}$$

The latent trader chooses a strategy of optimally benefiting from the liquidity supply on the primary exchange without making the off-exchange execution of remaining shares too costly. Recalling that quote shifts in the primary exchange also shift off-exchange settlement prices (see (2.4)), the latent trader only trades in the primary exchange if his payoff (because of saving the commission fee γ) over-compensates for the increased off-exchange execution costs. If $D_H < \Phi_{L_a}$, the liquidity supply is too low to outweigh the costs of executing the remaining shares at a higher settlement price. Accordingly, if $D_H \ge \Phi_{L_a}$ holds, he only enters the primary market if the liquidity competitor decides not to overbid. If the competitor revises the bid quote ($\sigma_C = step$), however, the latent trader takes advantage of an improved settlement price on the off-exchange market and stays away from the primary exchange. This situation arises unless the displayed liquidity supply exceeds a second threshold, Φ_{L_b}. Then, it becomes beneficial for the latent trader to trade in the public market, as he can execute a sufficiently large portion of his order volume.

Hence, we can conclude that the structure of the public exchange affects market fragmentation. Precisely, latent order flow is attracted when the commission fee γ is high, tick sizes Δ are small, and the total displayed liquidity is high. These predictions resonate well with the findings of the empirical literature. For instance, Griffiths et al. (1998), Griffiths et al. (2001), and Hendershott and Jones (2005) report that the proportion of off-exchange trading increases when the displayed depth is low and tick sizes are wide.

Since the competitor knows that the latent trader is only attracted if a minimum level of liquidity is displayed ($D_H \ge L_a$) and if the current bid quote is not overbid ($\sigma_C = stay$), it is optimal for the competitor to "stay" at the initial best bid quote. The advantage of this strategy is twofold: first, he attracts the latent block trader, and second, he ”saves” the extra tick Δ in transaction costs. If the displayed liquidity is too low ($D_H < \Phi_{L_a}$), however, the block trader will never enter the primary market, regardless of the competitor's action σ_C. In this case, the competitor's optimal strategy σ_C^* reduces to the optimal strategy of the baseline game without the latent block trader (see Lemma 1). Formally summarizing this strategy yields Lemma 3:

Lemma 3 (Liquidity Competitor's Best Response)**.** *Given σ_L^* and D_H, the liquidity competitor's best response obeys*

$$\sigma_C^* = \begin{cases} \sigma_C^{*0}(D_H) & \text{if} \quad 0 \leq D_H < \Phi_{L_a}, \\ stay & \text{if} \quad \Phi_{L_a} \leq D_H, \end{cases} \tag{3.7}$$

with σ_C^{*0} denoting the competitor's optimal strategy in the case without the latent investor according to Proposition 1.

Finally, the equilibrium is obtained by deriving the hidden trader's best response, given the competitor's best strategy.

Proposition 2 (Equilibrium Strategies)**.** *The equilibrium strategies σ_H^* ($\equiv D_H^*$), σ_C^* and σ_L^* obey*

$$D_H^* = \begin{cases} D_H^{*0} & \text{if} \quad N_H \leq \Phi_{L_a} \\ N_H & \text{else} \end{cases}, \quad \sigma_C^* = stay, \quad \sigma_L^* = \begin{cases} on & \text{if} \quad N_H \geq \Phi_{L_a} \\ off & \text{else} \end{cases}, \tag{3.8}$$

with D_H^{*0} denoting the hidden trader's optimal strategy in the case without latent investors according to Proposition 1.

Hence, in equilibrium, it is optimal for large liquidity suppliers, i.e., $N_H > \Phi_{L_a}$, to fully reveal their trade intentions and thus to display the entire volume, i.e., $D_H^* = N_H$. In this case, they benefit from counterparty attraction while avoiding or outweighing the downside of liquidity competition. Conversely, small traders with $N_H < \Phi_{L_a}$ cannot attract latent demand and thus choose a display size (i.e., $D_H^* \leq \Phi_C$) that merely prevents overbidding by the liquidity competitor. In this case, the optimal display size is limited by D_H^{*0} and is exclusively driven by liquidity competition, as derived in Proposition 1 in Section 3.1.

The critical size Φ_{L_a} can be interpreted as a liquidity premium demanded by the latent block trader to offset the costs of trading in the off-exchange trading venue. Thus, all factors that increase Φ_{L_a} make attracting latent counterparties less likely and implicitly increase – though not necessarily monotonously – the size of the hidden volume. The following corollary summarizes the determinants of hidden liquidity in the presence of latent trade demand:

Corollary 3 (Determinants of Hidden Liquidity with Latent Block Traders)**.** *The provision of hidden liquidity increases*

(i) with larger spreads S,

(ii) lower opposite-side depth,

(iii) the size of the latent block trader N_L,

(iv) lower upstairs commission fee γ.

The tick size Δ and the extent of liquidity competition N_C affect the provision of hidden liquidity in the primary exchange through opposite channels, liquidity competition and the attraction of latent order flow. For instance, while a smaller tick size increases liquidity competition, it increases the incentives for latent block traders to enter the public market. Similarly, a higher N_C increases competition but also attracts more latent demand. Consequently, the effects of Δ and N_C on hidden liquidity provision are not straightforward.

4 Impact of Hidden Orders: A Partial Equilibrium Analysis

How market transparency affects market quality is a key issue in market microstructure research. This question, however, cannot be solely addressed in the full equilibrium analysis in Section 3.2, where traders act rationally and have complete information.[11] In such a setting, it is optimal for a (sufficiently large) liquidity supplier *not* to use hidden orders at all. Consequently, the impact of the use of hidden orders can only be assessed by deviating from the full equilibrium. Therefore, we use a partial equilibrium setting to analyze the case, where liquidity suppliers hide too much. Then, the choice variable D_H is no longer fixed to its equilibrium value but is a free parameter with $D_H \neq D_H^*$. This approach is further justified by the vast empirical evidence that the majority of large orders are in fact kept hidden and not exposed.[12]

In this section, we assume that the hidden trader is not small, i.e., $N_H \geq \Phi_{L_a}$, and that the latent trader is present with certainty, i.e., $\mu = 1$. Let σ^* denote the Nash equilibrium triple as

[11] In particular, assuming that liquidity suppliers in the primary exchange are aware of the presence of the latent investor (i.e., either $\mu = 0$ or $\mu = 1$) is a strong assumption.

[12] Deviating from the full-equilibrium concept does not necessarily weaken the normative implications of our model. Traders who deviate from the equilibrium are 'irrational' in light of our setting but are not necessarily under risk aversion or uncertainty regarding the presence of a latent counterparty in an extended setting.

in Proposition 2, and let $\tilde{\sigma}$ denote the partial equilibrium triple with

$$\tilde{\sigma}(D_H) := (\sigma_H = D_H, \sigma_C^*, \sigma_L^*) \tag{4.1}$$

where σ_C^* and σ_L^* denote the Nash-equilibrium strategies of the competitor and latent trader as in Proposition 2 and D_H is fixed with $D_H \neq D_H^*$ (or $\sigma_H \neq \sigma_H^*$).

In limit order book markets, price changes materialize through the impact of incoming order flows on the (initial) state of the order book. Analyzing the impact of hidden orders on the order flow composition is thus important to quantify the effect of hidden liquidity supply on prices.

4.1 Impact on the Order Flow Composition

To assess the impact of the hidden order submission at time t_0 on the *future* incoming order flow, we consider order decisions taken after t_0 until T. Denote by $EXB_{t_0}^T$ and $CAB_{t_0}^T$ the expected execution volumes and cancelation volumes of (limit) buy orders until time T, respectively. Similarly, $EXS_{t_0}^T$ denotes the expected execution volume of (limit) sell orders until time T.

The expected execution volume of buy limit orders, $EXB_{t_0}^T$, stems from market sell orders x_L, submitted by the latent investor and market sell orders x submitted by the noise trader. Both market orders execute against the liquidity provided by the hidden trader and the liquidity competitor. Recalling that $x \sim Exp(\lambda)$ and that x_L obeys

$$x_L = \begin{cases} D_H + N_C & if \qquad \sigma_L = on, \\ 0 & else. \end{cases} \tag{4.2}$$

Then, $EXB_{t_0}^T$ is given by

$$EXB_{t_0}^T(\tilde{\sigma}(D_H)) := \mathbb{E}[X_C + X_H | \sigma = \tilde{\sigma}(D_H)] = \mathbb{E}[\min(x + x_L, N_C + N_H) | \sigma = \tilde{\sigma}(D_H)], \tag{4.3}$$

where X_C and X_H denote the execution volumes of the hidden trader and the liquidity supplier until time T, respectively, corresponding to the minimum of liquidity supply $N_C + N_H$ and demand $x + x_L$. To emphasize the dependence of the partial equilibrium triple $\tilde{\sigma}$ on the display size choice D_H, we use the notation $\tilde{\sigma} = \tilde{\sigma}(D_H)$.

One central assumption of our model is that liquidity suppliers trade under liquidation constraints. Therefore, non-executed shares at T have to be canceled and (re-) submitted as market sell orders to ensure liquidation. Consequently, the expected amount of buy-side cancelations and market sell order (re-) submissions until time T have to equal the expected amount of non-executed buy-side shares, i.e.,

$$CAB_{t_0}^{T}(\tilde{\sigma}(D_H)) = EXS_{t_0}^{T}(\tilde{\sigma}(D_H)) = N_C + N_H - EXB_{t_0}^{T}(\tilde{\sigma}(D_H)). \tag{4.4}$$

The next corollary describes the relation between the display size D_H and the expected buy executions, sell executions, and cancelations.

Corollary 4 (Expected Execution and Cancelation Volumes)**.** *Assume a large (buy) hidden trader, i.e., $N_H > \Phi_{L_a}$ and $D_H \neq D_H^*$. Then, the following relations hold:*

$$EXB_{t_0}^{T}(\sigma^*) > EXB_{t_0}^{T}(\tilde{\sigma}(D_H)), \tag{4.5}$$

$$CAB_{t_0}^{T}(\sigma^*) < CAB_{t_0}^{T}(\tilde{\sigma}(D_H)), \tag{4.6}$$

$$EXS_{t_0}^{T}(\sigma^*) < EXS_{t_0}^{T}(\tilde{\sigma}(D_H)). \tag{4.7}$$

In particular, $EXB_{t_0}^{T}(\tilde{\sigma})$ ($CAB_{t_0}^{T}(\sigma^*)$ and $EXS_{t_0}^{T}(\sigma^*(D_H))$) is monotonically increasing (decreasing) in the display size D_H.

Hence, the lower the display size is, i.e., the larger the proportion of hidden liquidity is, the lower the number of buy-side limit orders executed by the latent investor will be. Consequently, a larger volume of buy-side limit orders remain non-executed, need to be canceled, and need to be (re-)submitted as the number of market sell orders increase. Therefore, we can summarize that the decision to hide orders affects the composition of the future order flow and increases the aggressiveness on the opposite side of the market. In contrast, displayed (buy) orders do not generate sell market order executions, as they find a trading counterpart and get fully executed.

4.2 Impact on Prices

In this section, we show that the market orders resulting from non-executed (and canceled) hidden orders generate price fluctuations. By defining $M_t = (A_t + B_t)/2$ as the mid-quote, we denote $R_{t_0}^{T} = M_T - M_{t_0}$ as the midquote return between time t_0 and T. We define excess returns and excess volatility in a way providing sensible measures of how the deviation from the socially desirable Nash equilibrium materializes in terms of price fluctuations:

Definition 2 (Excess Return)**.** *The excess return is defined as the difference between the expected midquote return under the strategy-triple $\tilde{\sigma}$ with fixed D_H and the expected midquote return under the Nash equilibrium-triple σ^*, i.e.,*

$$\tilde{R}_{t_0}^{T}(D_H) := \mathbb{E}[R_{t_0}^{T}|\tilde{\sigma}(D_H)] - \mathbb{E}[R_{t_0}^{T}|\sigma^*]. \tag{4.8}$$

Definition 3 (Excess Volatility)**.** *Excess volatility is defined as the difference between the variance of the midquote return under the strategy-triple $\tilde{\sigma}$ with fixed D_H and the variance of the midquote return under the Nash equilibrium-triple σ^*, i.e.,*

$$\tilde{V}_{t_0}^T(D_H) := Var[R_{t_0}^T|\tilde{\sigma}(D_H)] - Var[R_{t_0}^T|\sigma^*]. \tag{4.9}$$

The following propositions provide the excess returns and excess volatility conditional on the hidden size $H_H = N_H - D_H$, the slope of the price impact function β, and the hidden trader's arrival probability q.

Proposition 3 (Excess Returns)**.** *Assume a large buy hidden trader, i.e., $N_H \geq \Phi_{L_a}$. Then, excess returns obey*

$$\tilde{R}_{t_0}^T(H_H) = \begin{cases} q\beta H_H & \textit{for} \quad 0 < \Phi_{L_a} \leq D_H, \\ q\beta(N_H + N_C) & \textit{for} \quad 0 \leq D_H < \Phi_{L_a}, \\ q\beta(H_H + N_C) - \beta N_C & \textit{for} \quad \Phi_{L_a} \leq 0. \end{cases} \tag{4.10}$$

The proposition shows that hiding trading intentions can cause price changes. It is important to note that these price effects are driven not by information but by insufficient signaling of trading intentions. Not signaling results in a mis-coordination between trading counterparties and a mis-match between liquidity supply and demand. Because of the underlying liquidation time constraint, this liquidity mis-match induces price pressure as liquidity suppliers are ultimately forced to increase their order aggressiveness and to post market orders. These market orders confront the market with additional buy demand, which has not been previously visible. Consequently, prices rise with the hidden size H_H. The following proposition shows that these effects also translate into higher volatility:

Proposition 4 (Excess Volatility)**.** *Assume a large buy hidden trader, i.e., $N_H \geq \Phi_{L_a}$. Then, excess volatility obeys*

$$\tilde{V}_{t_0}^T(H_H) = \beta^2(1-q)q \begin{cases} H_H^2 - 2H_H N_C, & 0 < \Phi_{L_a} \leq D_H, \\ N_H^2 - N_C^2, & 0 \leq D_H < \Phi_{L_a}, \\ H_H^2 - N_C^2, & \Phi_{L_a} \leq 0. \end{cases} \tag{4.11}$$

Hence, the decision to hide induces excess volatility, which is increasing in the hidden size. In contrst, excess-volatility vanishes for large display sizes $D_H \approx N_H$ (i.e., $H_H \approx 0$) that are close to the rational Nash strategy $\sigma_H^* = D_H^* = N_H$. The empirical literature has provided extensive evidence of the negative relation between observable depth ("open interest") and market volatility.[13]

Recall from Lemma 2 that latent trading counterparties in the public order book market are more likely to be attracted when off-exchange brokerage fees γ are large and when the (relative) tick size Δ is small. Therefore, liquidity suppliers are more likely to reveal their trading intentions and to execute their orders by directly avoiding the use of market orders. This mitigates price pressures from non-executed shares and reduces excess volatility. Consequently, Δ and γ affect price volatility indirectly through Φ_{L_a}. The resulting relationships are summarized in Corollary 5 and follow directly from Proposition 4.

Corollary 5 (Hidden Liquidity and Volatility)**.** *Assume a large hidden trader, i.e.,* $N_H > \max(N_C, \Phi_{L_a})$*. Then, ceteris paribus (excess) volatility increases with*

(i) larger hidden size $N_H - D_H$*,*

(ii) smaller display size D_H*,*

(iii) lower off-exchange-brokerage fees γ*,*

(iv) larger tick size Δ*.*

An interesting finding is that the spread does not play a role in the case $\mu = 1$, i.e., when the presence of the latent investor is known with certainty by all market participants. Although the spread is a major determinant of the liquidity competitor's decision to overbid the hidden trader in the case *without* a hidden trader (see Lemma 1), this mechanism changes as soon as the liquidity competitor knows that the hidden trader will attract the latent counterparty. Then, there is no incentive for him to overbid the hidden trader and to post more aggressive prices. Consequently, liquidity competition and bid-ask spreads do not affect the supply of hidden liquidity and thereby will not affect prices.

However, as soon as there is uncertainty (i.e., $0 < \mu < 1$) or incomplete information about the presence of the latent trader, the bid-ask spreads will affect volatility through liquidity competition and the hidden trader's willingness to expose his shares. This follows because with probability $1 - \mu$, the liquidity competitor will face the benchmark case without the latent

[13]For instance, Bessembinder and Seguin (1993) find that the relation between open interest and market volatility is negative for a wide range of markets, including agricultural, financial, metal, and currency trading. More evidence is reported in Ahn et al. (2001), Watanabe (2001), Ragunathan and Peker (1997) and Fung and Patterson (1999).

trader. Thus, if at least μ is small enough, by continuity, the spread will affect the competitor's order aggressiveness in the same way as of Proposition 2 and in turn will affect the hidden traders exposure decision likewise.

Corollary 6 (Volatility and Spread)**.** *Assume a large hidden trader, i.e., $N_H > \max(N_C, \Phi_{L_a})$ and that $0 \geq \mu < 1$ holds. Then, for sufficiently small μ, volatility increases with wider spreads.*

For sake of brevity, we limit our arguments to the previous discussion and do not show a rigorous proof. Nevertheless, substantial empirical evidence for this relation has been documented for various markets.[14]

5 Testing Cross-Sectional Implications

5.1 Testable Hypotheses

The equilibrium model in Sections 3 and 4 establishes a set of cross-sectional predictions on the origination of hidden liquidity and its effects on different characteristics of the order book market. Key cross-sectional hypotheses are generated from Corollaries 3, 5, and 6 as follows:

Hypothesis 1 (Cross-sectional Predictions: Spread, Tick and Volatility)**.**

(i) Markets with smaller spreads have a higher proportion of hidden liquidity.

(ii) Markets with wider (relative) tick sizes have a higher proportion of hidden liquidity.

(iii) Markets with higher proportion of hidden liquidity have a higher return volatility.

5.2 Data

Our empirical analysis uses a combination of two data sets based on NASDAQ trading. Information on consolidated hidden and displayed depth for each price level on a minute-by-minute basis for all NASDAQ traded stocks originates from the NASDAQ ModelView data set. The initial sample covers the constituents of the S&P500 universe through the period from November to December 2008. To reduce the impact of very illiquid stocks, we restrict the analysis to

[14] See, for instance Harris (1996), Aitken et al. (2001), De Winne and D'Hondt (2009), Bollerslev and Melvin (1994), Bollerslev and Domowitz (2012), Hasbrouck and Saar (2001), Plerou et al. (2005), Wang and Yau (2000) or Kalimipalli and Warga (2002).

stocks that have a quoted spread below 25 cents on average. This leaves us with a sample of $N = 468$ stocks.

Moreover, to utilize information on order flow *between* the minute-by-minute snapshots, we augment the snapshots by using NASDAQ TotalView message-level data, which is processed via the data service "LOBSTER"[15]. The data contains information on any (visible) order activity and the corresponding fully reconstructed (displayed) NASDAQ limit order book at each instant. We aggregate order executions, cancelations, and submissions stemming from NASDAQ TotalView for each minute and merge this information with the minute-by-minute snapshots on hidden depth from NASDAQ ModelView. The merged data set then consists of 390 daily minute-by-minute information observed over 40 trading days, resulting into $15,600$ observations per stock.

We limit our analysis to the ten best price levels in the order book. These levels represent the most active parts of the limit order book and are thus most suitable to test the predictions of our theoretical model. Table 1 reports averages across stock groups and time for mid-quote levels, spreads, visible and hidden depth, and limit order activities. We group the stocks into quintiles based on the average daily trading volumes (ADV). We observe distinct variations in trading activities, as reflected by inter-trade durations ranging from 2.65 seconds for the least actively traded stocks to 0.35 seconds for stocks in the largest liquidity quintile. Similar monotonic relationships across the liquidity quintiles are found for trade sizes (increasing in ADV), bid-ask spreads (decreasing in ADV), price levels (decreasing in ADV), first-level order book depth (increasing in ADV), and daily volatility, measured based on the daily high-low range relative to the daily average mid-quote (increasing in ADV). Hence, the highest trading activity (in terms of both number of transactions and size of shares) is observed for stocks with small spreads, low price levels, and high depth.

Most interestingly, we observe that the proportion of hidden shares in total posted shares is declining for less liquid stocks, amounting to approximately 17% on average. The relative amount of shares executed against standing hidden orders, however, is decreasing with the overall underlying daily trading volume. While 26% of trading volume on average is executed against hidden orders in the smallest liquidity quintile, this number declines to 7% on average for the most actively traded stocks. Hence, hidden liquidity is more prevalent for less liquid stocks with wider spreads and lower displayed depth.

[15] See http://lobster.wiwi.hu-berlin.de/.

Table 1: Averages across stocks and time for daily trading volume (ADV), inter-trade durations (DUR), daily high-low ranges standardized by corresponding daily average mid-quotes (HL), and trade sizes (TS), and averages of minute-by-minute snapshots of bid-ask spreads (SPR), mid-quotes (MQ), visible depth on top (first level) of the book ($D1$), and total hidden depth on the first 10 levels ($HD10$). Moreover, we report the average ratios of hidden to total depth on the first 10 levels (evaluated based on minute-by-minute snapshots) ($RHD10$), the average number of shares traded against hidden volume (THD), and the corresponding ratio of executed hidden shares to average trading volumes ($RTHD := THD/ADV$). The amount of traded hidden volume, THD, is computed as the average daily trade volume executed on the best quotes. The averages are computed within liquidity quintiles based on ADV. The stock universe consists of all S&P500 constituents that are traded on NASDAQ, excluding stocks with an average spread below 25 cents. The sample ultimately includes 468 stocks for the period between November and December 2008.

Liquidity Quintile	Observable Stock Properties							Hidden Liquidity posted		Hidden Liquidity traded	
	ADV $(10^6 sh.)$	DUR $(sec.)$	HL $(ratio)$	TS $(sh.)$	SPR $(ticks)$	MQ $(\$)$	$D1$ $(sh.)$	$HD10$ $(sh.)$	$RHD10$ $(ratio)$	THD $(10^6 sh.)$	$RTHD$ $(ratio)$
q_{20}	1.39	2.65	0.07	147	4.91	36.46	308	656	0.19	0.37	0.26
q_{40}	2.72	1.38	0.08	158	3.39	32.84	576	1318	0.20	0.57	0.20
q_{60}	4.23	0.94	0.09	165	2.40	27.41	800	1671	0.17	0.69	0.15
q_{80}	7.13	0.61	0.10	178	1.87	24.59	1278	2292	0.16	0.83	0.11
q_{100}	16.98	0.35	0.11	219	1.38	23.32	3490	6202	0.13	1.10	0.07
Total	6.57	1.19	0.09	174	2.79	28.91	1305	2440	0.17	0.71	0.16

5.3 Econometric Analysis

Hypothesis 1 posits predictions regarding the relationships between the proportion of hidden liquidity ($RHD10$), the relative bid-ask spread ($RSPR$), the relative tick size ($RTCK$), and volatility (RV). The relative bid-ask spread $RSPR$ is defined as the ratio between the spread SPR and the mid-quote MQ, while the relative tick size $RTCK$ is defined as the ratio between the tick size TCK and the mid-quote price MQ. Volatility is estimated as the daily realized volatility (computed as the sum of squared 10-min returns).

According to our theory, hidden liquidity triggers market volatility but not vice versa, suggesting the following reduced model:

$$
\begin{aligned}
RHD10_i &= \alpha_h + \beta_{h,2} RTCK_i + \beta_{h,3} RSPR_i + \varepsilon_{hi}, &(5.1)\\
RV_i &= \alpha_v + \beta_{v,1} RHD10_i + \beta_{v,2} RTCK_i + \beta_{v,3} RSPR_i + \varepsilon_{vi}, &(5.2)
\end{aligned}
$$

for $i = 1, \ldots, N$ and white noise error terms ε_{hi} and ε_{vi}. Although not predicted by our theory, we also include the relative tick size $RTCK_i$ and $RSPR_i$ as additional control variables in equation (5.2).[16] All variables are entered in logarithmic form of time averages across days and (in the case of $RHD10_i$ and $RSPR_i$) one-minute snapshots within a day. This leaves us with $N = 468$ cross-sectional observations.[17]

However, although it is not captured by our framework, the causality between market volatility and hidden liquidity might be reversed. For instance, Harris (1996) suggests that liquidity suppliers use hidden orders to reduce the risk of being picked-off. Since the risk of being picked-off is particularly high in volatile markets, causality may run from volatility to hidden liquidity. Moreover, simultaneity can simply arise because of the use of time averages of volatility and hidden liquidity. To account for this effect, we consider a second specification in which we explicitly include RV_i in the first equation, resulting into a bivariate simultaneous equations system:

$$
\begin{aligned}
RHD10_i &= \alpha_h^{IV} + \beta_{h,1}^{IV} RV_i + \beta_{h,2}^{IV} RTCK_i + \beta_{h,3}^{IV} RSPR_i + \varepsilon_{hi}^{IV}, &(5.3)\\
RV_i &= \alpha_v^{IV} + \beta_{v,1}^{IV} RHD10_i + \beta_{v,2}^{IV} RTCK_i + \beta_{v,3}^{IV} RSPR_i + \varepsilon_{vi}^{IV}. &(5.4)
\end{aligned}
$$

As soon as both $\beta_{h,1}^{IV}$ and $\beta_{v,1}^{IV}$ are truly non-zero, $RHD10_i$ and RV_i are simultaneous, and the parameters cannot be consistently estimated by OLS. We therefore use two-stage least squares (2SLS) to estimate the system equation by equation. We use the squared daily mid-quote return, $RET2_i$, as an obvious instrument for RV_i. As a second instrument, we utilize the displayed depth $D10_i$. Both $RET2_i$ and $D10_i$ are correlated with the endogenous variables. While the uncorrelatedness of $RET2_i$ and ε_{hi}^{IV} is easily justified (given that RV_i serves as a regressor

[16] For instance, the extant information-based literature suggests that volatility and spreads are linked through information asymmetry (e.g., Kyle (1985)).

[17] An alternative to using averaged variables would be to estimate the model in a panel setting, which would allow us to exploit not only cross-sectional variation but also time variation. Properly capturing the strong serial (cross-)dependencies of most variables (see Section 6) in a panel setting would be very challenging, however, and would require panel VAR approaches, which would be cumbersome, or even impossible, to estimate given the amount of underlying observations.

Table 2: Estimation results of cross-sectional regressions of (5.1), (5.2), (5.3) and (5.4). The first two columns give the OLS estimates of (5.1) and (5.2). The next columns give the 2SLS estimates of (5.3) and (5.4) with instruments $RET2_i$ and $D10_i$. Standard errors are shown in brackets. Below, we report the F-statistics based on the first-stage regressions as tests for weak instruments and the Sargan test for over-identification.

	Structural Model		*Simultaneous Equations Model*			
	$RHD10_i$	RV_i	$RHD10_i$	RV_i	$RHD10_i$	RV_i
	(5.1)	(5.2)	(5.3)	(5.4)	(5.3)	(5.4)
$RHD10_i$		3.356***		3.356***		1.560***
		(0.427)		(0.427)		(0.163)
RV_i			0.298***		0.294***	
			(0.038)		(0.038)	
$RTCK_i$	−0.464***	1.877***	−0.559***	1.877***	−0.558***	1.043***
	(0.027)	(0.217)	(0.029)	(0.217)	(0.029)	(0.090)
$RSPR_i$	0.172***	−0.234*	0.070**	−0.234*	0.071**	0.075
	(0.031)	(0.126)	(0.033)	(0.126)	(0.033)	(0.063)
Const.	−1.663***	−5.438***	1.620***	−5.438***	1.581***	−8.426***
	(0.106)	(0.791)	(0.431)	(0.791)	(0.430)	(0.332)
Instruments	–	–	$RET2_i$		$RET2_i, D_i$	
N	468	468	468	468	468	468
Weak-Instr.	–	–	1067.78***	68.08***	532.46***	86.52***
Sargan	–	–	–	–	71.75***	109.30***

Note: *p<0.1; **p<0.05; ***p<0.01

in (5.3) and thus captures most volatility-associated variation in ε_{hi}^{IV}), the uncorrelatedness of $RET2_i$ and ε_{vi}^{IV} is more critical and relies on the ability of the regressors that are included in (5.4) to sufficiently capture variations in RV_i. We conjecture, however, that cross-sectional variation in RV_i in particular is captured by the included regressors rather than by squared daily returns, diminishing the remaining explanatory power of $RET2_i$ for RV_i and making correlations between $RET2_i$ and ε_{vi}^{IV} unlikely. The uncorrelatedness between $D10_i$ and both ε_{hi}^{IV} and ε_{vi}^{IV} can be similarly justified, as $D10_i$ and $RTCK_i$ are strongly correlated[18], and thus, we expect the explanatory power of $RTCK_i$ to capture most of the variation in both equations (5.3) and (5.4), making correlations between $D10_i$ and both ε_{v_i} and ε_{h_i} less likely.

Table 2 presents the equation-by-equation OLS estimates of (5.1) and (5.2) and 2SLS estimates of (5.3) and (5.4). Without exception, the coefficient estimates for all the model specifications confirm the predictions of Hypothesis 1 and are significant at the 5% level. In particular, hidden liquidity provision ($RHD10_i$) is higher for stocks that trade at wider spreads ($RSPR_i$) and smaller tick sizes ($RTCK_i$). Furthermore, markets that exhibit a higher proportion of hidden liquidity supply ($RHD10_i$) are more volatile (RV_i). Moreover, as expected, we find evidence for simultaneity between volatility and hidden liquidity provision. Although the test of over-identification does not fully support our choice of over-identifying moment conditions, the results are nevertheless qualitatively similar across the different specifications. The results are also qualitatively similar if additional or other instruments are employed (not shown here). We therefore conclude that simultaneity effects do not fundamentally influence the coefficient estimates of our variables of interest.

6 Testing the Dynamic Implications

6.1 Testable Hypotheses

Beyond cross-sectional relationships among hidden liquidity supply, minimum tick sizes, and volatility, our theory implies that order (non-)display has causal effects on order executions, order cancelations, and market returns. In this section, we employ a multivariate time-series approach to test these time series implications while controlling for dynamic interdependencies. This setting allows us to empirically assess whether and to which extent hidden liquidity causes price shifts and volatility.

[18] The estimated correlation coefficient between $D10_i$ and $RTCK_i$ equals 0.81.

The following hypotheses originate from Corollary 4. Since the postulated relationships are symmetric for buy and sell orders, it is sufficient to formulate the hypotheses for one side of the market only. Accordingly, the hypotheses are associated with a hidden/displayed order submission on the buy side.

Hypothesis 2 (Oder Executions)**.**

(i) Displayed limit buy order submissions do (do not) increase limit buy (sell) order executions.

(ii) Hidden limit buy order submissions do (do not) increase limit sell (buy) order executions.

Hypothesis 3 (Order Cancelations)**.**

(i) Displayed limit buy order submissions do not (do not) increase subsequent buy (sell) side cancelation rates.

(ii) Hidden limit buy order submissions do (do not) increase subsequent buy (sell) side cancelation rates.

Hypothesis 4 (Returns)**.**

(i) Displayed limit buy order submissions do not cause positive excess returns.

(ii) Hidden limit buy order submissions cause positive excess returns.

6.2 Descriptive Statistics

Testing the dynamic implications of hidden liquidity for order executions, cancelations, and returns requires a time-series setting that accounts for high-frequency order book dynamics. To limit the computational burden induced by data processing and dynamic order book modeling, we conduct the following analysis based on 10 randomly selected stocks from the S&P500 universe. The tickers are APC, AZO, CAH, GAS, GOOG, EMR, LEG, PAYX, STJ, and TDC, reflecting an arbitrary cross-section of differently liquid S&P500 constituents.

Time-series averages of mid-quotes, bid-ask spreads, visible and hidden depth, and order activities based on one-minute aggregates for the 10 stocks are provided in Table WA1 (see web appendix).[19] The statistics indicate that order submission behavior and market dynamics

[19]For the sake of brevity, some of the descriptive statistics are provided in a web appendix available at http://homepage.univie.ac.at/nikolaus.hautsch/media/files/webappendix_CHH14.pdf.

are obviously strongly driven by liquidity competition and non-trade order activity. In line with Hautsch and Huang (2012), we observe that most order activity derives from order submission and order cancelation activity: on average, approximately 47% of the order flow volume originates from order submissions, 49%, from cancelations, and only approximately 4%, from trades/executions.

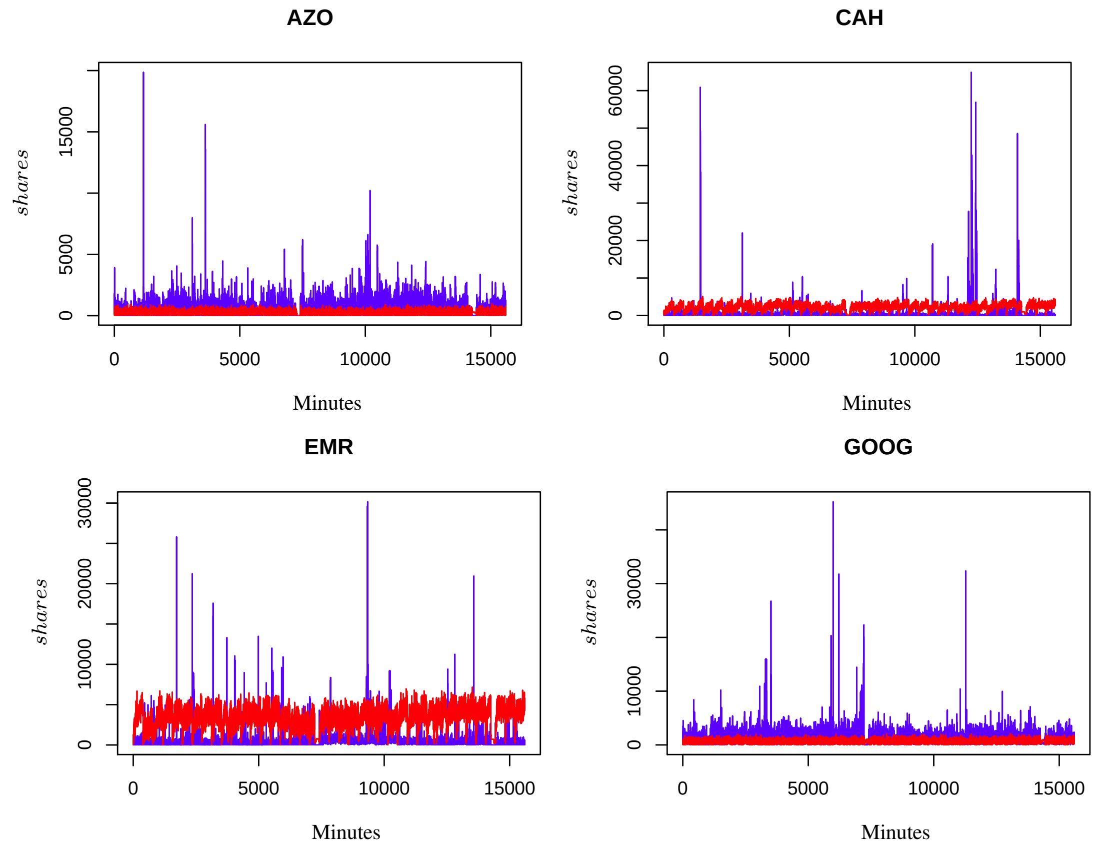
Table WA2 (see web appendix) provides evidence on the time variability of minute-by-minute activities for one exemplary stock in our sample that belongs to the automotive industry, AOZ (AutoZone, Inc.). The table reports mid-quotes, bid-ask spreads, visible and hidden depth, order imbalances, and order activities based on one-minute aggregates. We observe distinct time-series variations in bid-ask spreads and the number of limit order submissions and cancelations. Interestingly, the overall liquidity supply (through the first 10 price levels) and the number of limit order executions are clearly more stable over time. These patterns are quite representative for other S&P500 stocks and show that recent NASDAQ market activities are strongly driven by significant order submission and cancelation activities that mostly occur on the first level but that obviously not substantially affect the liquidity supply deeper in the book. A second noticeable observation is that hidden liquidity, in terms of both total order supply and order imbalances, exhibits high variation and has larger extreme values than displayed liquidity.

The time variability of depth is illustrated in Figure 2, which shows minute-by-minute time series of total hidden and displayed depth up to the first 10 levels for four randomly selected stocks. We observe that the liquidity supply tends to be clustered over time and that it exhibits substantial time variation. The latter finding is particularly true for hidden liquidity, as significant spikes and thus large quantities of hidden volume tend to occur on an irregular but relatively frequent basis.

Figure WA3 (see web appendix) shows cross-sectional averages of autocorrelation functions (ACFs) of one-minute returns, 10-min volatilities, and one-minute snapshots of depth and displayed depth imbalances, defined as standing buy volume in excess of sell volume. Moreover, we report one-minute aggregates of limit order submissions (SUB and SUS), cancelations (CAB and CAS), and executions (EXB and EXS). The 10-min volatility is computed as a 10-min realized volatility, corresponding to the sum of squared one-minute mid-quote returns, computed over subsequent 10-min intervals.

All variables (except mid-quote returns) are strongly autocorrelated over time and show very persistent (i.e., slowly decaying) autocorrelation patterns in most cases. Interestingly, the liquidity supply (reflected by submissions and cancelations) is more persistent than the liquidity

Figure 2: Minute-by-minute time series of total hidden and displayed depth $HD10$ and $D10$ for the stocks AZO (AutoZone, Inc.), CAH (Cardinal Health, Inc.), EMR (Emerson Electric Co.), and GOOG (Google) traded on NASDAQ, November to December 2008. The total period consists of 40 days with 390 trading minutes each, corresponding to 15600 minutes in total. The total hidden (displayed) depth on the first 10 levels is shown in blue (red).



demand (reflected by order executions). In line with the finding that liquidity competition is a substantial driver of market dynamics, this finding suggests that traders actively micro-manage, modify, and cancel passive orders when they feel that orders are mis-priced or have a low chance of execution. The presence of strong serial dependence in execution volumes (i.e., market orders) is in line with the fact that traders do not execute their position by using a single market order, but rather slice larger orders into smaller orders and feed them sequentially into the market. This is in line with the literature on optimal liquidation (e.g., Obizhaeva and Wang (2013)).

Moreover, we find that depth imbalances, representing excess demand for trading on one side of the book, are less persistent. Hence, excess trade demand does not persist over longer periods, as traders naturally have to liquidate excess positions over a reasonable time period. We observe that hidden order imbalances are more persistent than displayed imbalances. This finding supports our theory predicting that displayed order imbalances provide a signal for counterparties to trade larger volumes at lower costs. Thus, displayed imbalances are eventually absorbed by counterparties' market order flows. Conversely, hidden order imbalances "survive" longer, as their presence cannot easily be detected by counterparties.

6.3 Econometric Modeling

To test Hypotheses 2, 3, and 4 and to isolate the dynamic implications of hidden order submissions on subsequent returns and order activities, we must account for the underlying serial dependence of the processes. Moreover, an evaluation of the cross-autocorrelations (not reported here) reveals significant dynamic *cross*-dependences between the individual variables.

In line with, e.g., Hasbrouck (1991) and Hautsch and Huang (2012), we capture these multivariate dynamics by using a vector autoregressive model for the underlying order book process. In particular, we suggest modeling the state of the market in terms of mid-quote returns, volatility, and variables representing the state of the order book as well as the incoming order flow. In particular, we model high-frequency market dynamics based on one-minute mid-quote returns (RET), bid-ask spreads (SPR), minute-by-minute rolling window estimates of 10-min realized volatilities (RV), minute-by-minute snapshots of hidden and displayed order imbalances ($HI10$ and $DI10$, respectively), total depth (sum of hidden and displayed depth on the first 10 levels; $TD10$), and per-minute numbers of submitted, executed, and canceled buy and sell limit orders (SUB, SUS, EXB, EXS, CAB, and CAS).[20] Accordingly, the state of the order book at t is represented by the 12-dimensional vector y_t, consisting of the variables RET, SPR, $VOLA$, $HI10$, $DI10$, TD, SUB, SUS, EXB, EXS, CAB, and CAS. We propose modeling y_t in terms of a vector autoregressive model of the order p (VAR(p)) of the form

$$y_t = \sum_{j=1}^{p} A_j y_{t-j} + u_t, \tag{6.1}$$

with A_j denoting (12×12) coefficient matrices for $j = 1, ..., p$ and u_t denoting the vector of zero mean white noise error terms with $E\left[u_t u_t'\right] = \Sigma_u$.

[20]We log-transform total depth $TD10$, realized volatility RV, and spread SPR. Residual diagnostics widely confirm this choice.

Using order imbalances allows us to test market-side-specific effects of hidden and displayed volume. Biais et al. (1995), Ranaldo (2004), Chordia et al. (2002), Hall and Hautsch (2006), and Cao et al. (2009) show that order imbalances carry more information than individual order depth levels about the state of the market. Moreover, directly modeling buy-sell depth imbalances simplifies the analysis, as these variables are less persistent than the underlying depth levels. A similar argument provides justification for modeling mid-quote returns and bid-ask spreads instead of separate quote levels. While quote processes are non-stationary, first differences of quote processes and spreads are covariance stationary, see, e.g., Engle and Patton (2004) and Hautsch and Huang (2012). Moreover, the two fundamental mechanisms in our theoretical model – liquidity competition and counterparty attraction – are driven by the magnitude of bid-ask spreads and returns but not quote *levels*. Therefore, in line with the analysis of Hautsch and Huang (2012), the resulting vector autoregressive model can be considered as a restricted stationary version of a more general co-integrated VAR model for levels of quotes and depth. Indeed, the stationarity of the underlying variables, including order flows and standing order depth, is confirmed by underlying unit root (Augmented Dickey-Fuller) tests (not reported here).

A natural way to test Hypotheses 2, 3, and 4 while controlling for dynamic order book interdependencies is to evaluate the impulse response of hidden and displayed order imbalances. In particular, in our framework, incoming buy (sell) hidden or displayed orders are identified as positive (negative) shocks to the corresponding order imbalances $HI10$ or $DI10$. Impulse responses in a VAR system are straightforwardly derived based on the moving-average representation,

$$y_t = \Phi_0 u_t + \Phi_1 u_{t-1} + \Phi_2 u_{t-2} + \Phi_3 u_{t-3} + \dots, \tag{6.2}$$

with $\Phi_0 = I_K$ and $\Phi_s = \sum_{j=1}^{p} \Phi_{s-j} A_j$ for $s > 0$. Following Pesaran and Shin (1998), we consider *generalized* impulse response functions that are obtained by shocking element j by δ_j while integrating out the effects on other elements, i.e.,

$$\Theta_j(n) := E\left[y_{t+n} \mid u_{jt} = \delta_j, \Omega_{t-1}\right] - E\left[y_{t+n}, \Omega_{t-1}\right], \tag{6.3}$$

with Ω_t denoting the information set up to time t. Assuming multivariate normality for u_t, the conditional expectation given a scaled shock $\delta_j := \sqrt{\sigma_{jj}}$ in one variable yields $E\left[u_t \mid u_{jt} = \delta_j\right] = \Sigma_u e_j \sigma_{jj}^{-1} \delta_j$, with e_j denoting the unit vector. The generalized impulse , i.e., $\Theta(n) := (\Theta_1(n), \Theta_2(n), \dots, \Theta_j(n)$ with the j-th response obeying

$$\Theta_j(n) = \frac{\Phi_n \Sigma_u e_j}{\sqrt{\sigma_{jj}}}, \quad j = 1, \dots, K, \tag{6.4}$$

measures the effect of a one-standard-error shock to the j-th equation at time t on the conditional expectation of y_{t+n}. Accordingly, the *cumulative (generalized) impulse response* $\Xi(n) := (\Xi_1(n), \Xi_2(n), \ldots, \Xi_j(n), \ldots \Xi_K(n))$ is defined by

$$\Xi_j(n) := \sum_{k=1}^{n} \Theta_j(k) = \sum_{k=1}^{n} \Phi_k \frac{\Sigma_u e_j}{\sqrt{\sigma_{jj}}}, \quad j = 1, \ldots, K \tag{6.5}$$

and is consistently estimated by

$$\hat{\Xi}_j(n) = \sum_{k=1}^{n} \hat{\Phi}_k \frac{\hat{\Sigma}_u e_j}{\sqrt{\hat{\sigma}_{jj}}}. \tag{6.6}$$

The main advantage of this approach is that the generalized impulse response functions are invariant to the re-ordering of the endogenous variables. As shown by Pesaran and Shin (1998), orthogonalized impulse responses coincide with orthogonalized impulse responses (based on a Cholesky decomposition of Σ_u) if the respective variable is the first one in the ordering. Pesaran and Shin (1998) derive the asymptotic properties of the impulse response functions based on a co-integrated VAR model. In the web appendix, we adapt these derivations and provide the asymptotic distributions of the generalized impulse response functions.

6.4 Estimation Results

The high persistence of the underlying order book process requires using a VAR process with high lag order, which could be more parsimoniously captured by using a vector autoregressive moving average (VARMA) specification. The latter specification, however, is more cumbersome to estimate, particularly in case of a 12-dimensional process. Therefore, we use a VAR approximation of the underlying process, which is consistently estimated with OLS equation by equation. Information criteria suggest a lag order of 30. Portmanteau tests for the presence of serial correlation in the resulting residuals (not shown here) widely confirm this choice. To check the robustness of this choice with respect to the resulting impulse response functions, we also estimate alternative specifications that are parameterized more parsimoniously, particularly a VAR(5) and VAR(15) specification. In line with the results of Jorda (2005) showing that impulse-response estimates are relatively stable regarding the choice of the underlying lag order (given that a dominant part of the serial dependence is sufficiently captured), we find that our results are not qualitatively affected and that they are remarkably quantitatively stable with respect to the model choice. We refrain from reporting individual VAR estimates, which are hardly interpretable for such a highly parameterized process and which are not the focus of our analysis.

Figures 3 to 5 show the estimates of the cumulative impulse responses of one-minute buy and sell limit order execution volumes ($\hat{\Xi}_{EXB}$ and $\hat{\Xi}_{EXS}$), buy and sell limit order cancelation volumes ($\hat{\Xi}_{CAB}$ and $\hat{\Xi}_{CAS}$), and mid-quote returns ($\hat{\Xi}_{RET}$) triggered by a positive one-standard-error shock in hidden and displayed order imbalances ($HI10$ and $DI10$). The reported impulse response functions are cross-sectional averages (across the analyzed $M = 10$ stocks). Thus, given a fixed time interval $t = n$ after the shock and variable j, the corresponding asymptotic variance of the averaged impulse response function is approximated by $M^{-2} \sum_{l=1}^{M} \Lambda_{jn}^{l}$, with Λ_{jn}^{l} denoting the asymptotic covariance of the generalized impulse response (see web appendix for details).[21] Note that a positive shock in order imbalances is associated with an increase in (hidden or displayed) liquidity on the buy side. Since the effects are symmetric in the sign of the shock, we restrict our analysis to positive shocks only and refrain from showing the opposite case.

According to Figure 3 below, the cumulative impulse response estimates with respect to order execution volumes confirm Hypothesis 2: In line with prediction (ii), an increase in buy-side hidden orders does not generate an increase in buy order executions (EXB) but does generate an increase in sell order executions (EXS). Second, and in line with prediction (i), we observe that an increase in buy-side displayed orders generates an increase in buy order executions (EXB) but does not generate an increase in sell order executions (EXS). Our predictions are significant at the 5% level.

[21]This approximation obviously ignores potential cross-equation correlations between the estimated asset-specific impulse response functions of stock l. Given the high parameterization, the latter is not straightforwardly taken into account. We therefore use this approximation as a convenient but still sufficiently precise way to assess and compactly illustrate the overall significance of our estimates. The latter – and thus – our conclusions regarding the empirical validity of our hypotheses are not affected by this approximation and is confirmed by the individual (asset-specific) estimates, which are not reported here.

Figure 3: Estimates of cross-sectional averages of the cumulative generalized impulse response of one-minute aggregated limit order buy and sell **execution volumes** ($\hat{\Xi}_{EXB}$ and $\hat{\Xi}_{EXS}$) due to positive one-standard-deviation shocks in hidden (blue) and displayed (red) order imbalances. The dashed lines show the approximate 95% confidence intervals of the averaged impulse response functions. Based on one-minute aggregates of NASDAQ ITCH data and one-minute snapshots of NASDAQ ModelView data for the stocks APC, AZO, CAH, GAS, GOOG, EMR, LEG, PAYX, STJ, and TDC, November to December 2008.

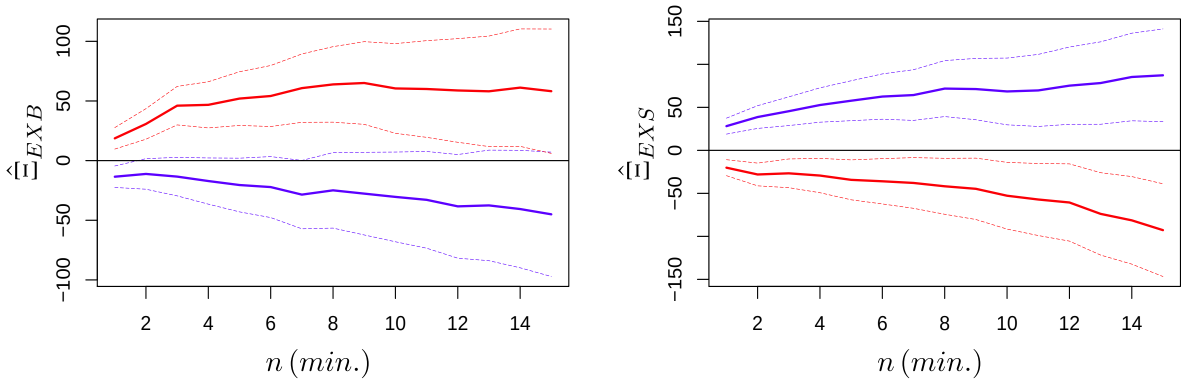
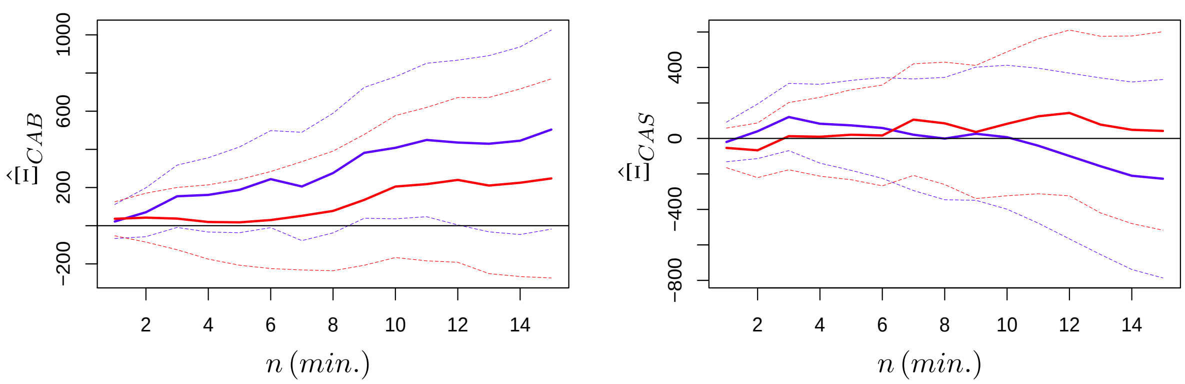


Figure 4: Estimates of cross-sectional averages of the cumulative generalized impulse response of one-minute aggregated limit order buy and sell **cancelation volumes** ($\hat{\Xi}_{CAB}$ and $\hat{\Xi}_{CAS}$) due to positive one-standard-deviation shocks in hidden (blue) and displayed (red) order imbalances. The dashed lines show the approximate 95% confidence intervals of the averaged impulse response functions. Based on one-minute aggregates of NASDAQ ITCH data and one-minute snapshots of NASDAQ ModelView data for the stocks APC, AZO, CAH, GAS, GOOG, EMR, LEG, PAYX, STJ, and TDC, November to December 2008.



As depicted by Figure 4 above, shocks in hidden or displayed order imbalances do not cause any significant reactions in cancelation volumes. While this result confirms Hypothesis 3(i), it does not support Hypothesis 3(ii) at first sight. This lack of support can be easily explained by data limitations as NASDAQ ModelView data do not contain the cancelation of hidden orders. Consequently, Hypothesis 3(ii) cannot be directly tested. Recorded cancelations of displayed orders, however, can (at least partly) correspond to displayed parts of larger (partially) hidden orders. In this case, the cancelation of partly hidden orders might also trigger a fraction of displayed cancelations. This reasoning might explain the borderline significance (though non-significance throughout the entire period) of buy impulse responses triggered by hidden liquidity.

Figure 5: Estimates of cross-sectional averages of the cumulative generalized impulse response of one-minute aggregated **mid-quote returns** ($\hat{\Xi}_{RET}$) due to positive one-standard-deviation shocks in hidden (blue) and displayed (red) order imbalances. The dashed lines show the approximate 95% confidence intervals of the averaged impulse response functions. Based on one-minute aggregates of NASDAQ ITCH data and one-minute snapshots of NASDAQ ModelView data for the stocks APC, AZO, CAH, GAS, GOOG, EMR, LEG, PAYX, STJ, and TDC, November to December 2008.

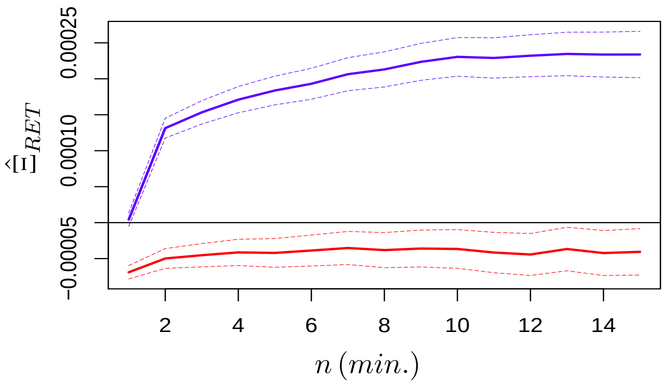
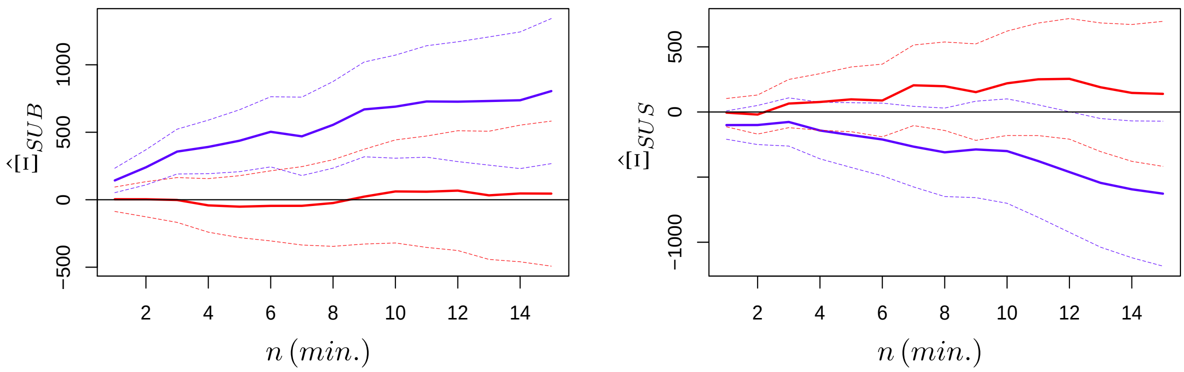


Figure 5 above displays the cumulative impulse response of mid-quote returns. As predicted by Hypothesis 4, we observe significantly positive price reactions after submissions of hidden orders. Conversely, the reaction of displayed orders is negative with a significantly smaller magnitude. Although Hypothesis 4 does not predict negative return reactions owing to shocks in displayed orders, there is a straightforward explanation for this effect: displayed orders are executed immediately, i.e., market sell orders consume liquidity on the buy side of the book and thus push prices downward. This is also in line with the observation that price reactions occur almost instantly, and prices do not change thereafter. Hence, displayed buy orders are executed once they are submitted and observed by other market participants. Hence, this finding empirically supports one of the major predictions of our theory: excessively hiding order volume is not optimal, causes a mis-coordination between liquidity supply and demand, and ultimately causes price movements.

Although it is not a subject of our predictions, for completeness, we also show estimated cumulative impulse responses for submission volumes (see Figure 6). We observe that (buy side) hidden order submissions increase the rate of subsequent (buy side) displayed order submissions. Even if this effect is not captured by our theory, we interpret this effect to arise from hidden traders who are not able to execute their position during the intended trading horizon and to re-submit their volume in terms of *limit orders* (instead of market orders). This behavior might be considered as an alternative – though less aggressive – strategy to enforce execution.

Figure 6: Estimates of cross-sectional averages of the cumulative generalized impulse response of one-minute aggregated limit buy and sell **order submission volumes** ($\hat{\Xi}_{SUB}$ and $\hat{\Xi}_{SUS}$) due to positive one-standard-deviation shocks in hidden (blue) and displayed (red) order imbalances. The dashed lines show the approximate 95% confidence intervals of the averaged impulse response functions. Based on one-minute aggregates of NASDAQ ITCH data and one-minute snapshots of NASDAQ ModelView data for the stocks APC, AZO, CAH, GAS, GOOG, EMR, LEG, PAYX, STJ, and TDC, November to December 2008.



7 Conclusion

We model a liquidity suppliers' trade-off between the costs and benefits of order exposure. While order exposure fosters liquidity competition and increases the risk of being front-run by liquidity competitors, it has the advantage of possibly attracting latent trade demand. Latent trade demand can stem from large investors who generally employ off-exchange trading mechanisms and who only enter the public primary market because of trading (liquidity) opportunities. In a (partial) equilibrium model, we show that excessively hiding trading interests can be harmful if traders are pre-committed to trade. The key mechanism is that hidden orders cannot elicit latent trade demand; thus, they are less likely executed, as they are not visible to latent traders. Consequently, to enforce execution, (hidden) liquidity suppliers need to cancel their orders and post them as market orders. This activity causes excess trade demand in the market, pushes prices, and ultimately increases volatility. These price reactions are caused not by information but by mis-coordination between liquidity supply and demand.

Our model does not assume the presence of asymmetric information but builds on the effects of order exposure and liquidity externalities across markets. While omitting the effects of asymmetric information is arguably a simplification, our model nevertheless allows us to demonstrate that several fundamental market microstructure relationships can be explained by pure mechanisms of order setting and liquidity competition. Indeed, empirical evidence based on unique data on non-displayed orders at NASDAQ strongly confirms our predictions. The key theoretical and empirical result is that greater use of hidden liquidity ultimately induces a higher mix of (costly) more aggressive orders and thereby generates excess price fluctuations. In this sense, hidden orders can impair price efficiency.

An important fact is that the elicitation of latent order flow depends on a critical mass of displayed orders. As there are several sources of risks and costs to trade in the public exchange, latent investors demand a minimum liquidity premium (i.e., a minimum display size) to trade publicly. Thus, hidden liquidity might be beneficial when exposure does not exceed the liquidity premium or when the likelihood of latent investors monitoring the market is low. Ultimately, whether hidden liquidity is beneficial depends on the heterogeneity of traders. We predict that the downside effect of hidden orders increases with the fraction of large investors compared to small and medium-sized traders.

What are the practical consequences of our findings for exchange operators? Certainly, banning hidden orders is not a viable option, as large investors would increasingly switch to off-exchange mechanisms, such as dark pools and brokerage networks, ultimately reducing liquidity in the public market. We nevertheless show that exchange operators can control pre-trade transparency through key market variables, such as the spread, tick size, and fees. First, the relative tick size can be controlled through optimal stock split rules. Tick sizes that are too large (relatively) can prevent latent investors from trading publicly and thus reduce the benefits of order exposure. On the other hand, tick sizes that are too small (relatively) increase liquidity suppliers' risk of being undercut, ultimately increasing the incentives for dark liquidity. Second, exchange operators can enhance market makers and liquidity suppliers' incentives to keep spreads narrow, reducing the costs of liquidity competition and thus increasing the incentives for exposure. Finally, our results suggest that a market's fee-and-rebate policy should provide greater incentives for large investors and large trades. Since market monitoring and exposure are both expensive for large investors, providing higher rebates and increasing monitoring as well as *direct market access* capabilities can increase the chances of exposure but also increase the rate at which latent investors observe the public market to seize those liquidity opportunities.

Ultimately, the issue at hand stems from the fact that continuous double-auction markets are less efficient than other markets in matching and coordinating large trades, as gaining exposure and monitoring large positions are more expensive in continuous markets. Providing incentives for large trades can help attract greater order flow from off-exchange trading mechanisms to public exchanges, improve public market liquidity, and enhance the quality of the price discovery process.

Appendix

To proof the propositions and lemmata of Section 3 and 4, we first establish preparatory results in Section *Preparatory Lemmata*. The proofs of the key results are then provided in the succeeding sections. Lemmata, propositions and remarks that are established in this appendix are numbered with an "A" prefix, while corresponding equations are numbered with an "a" prefix.

A.1 Preparatory Lemmata

Traders' Execution Volumes

Lemma A1 (Liquidity Competitor's Execution Volume)**.** *Let x denote the market sell order size at time T, N_C the liquidity supplier's total order size and D_H the hidden trader's display size. Then, the liquidity competitor's execution volume X_C at time T obeys*

$$X_C = \begin{cases} \min(x, N_C) & \textit{if} \quad (\sigma_C, \sigma_L) = (step, off), \\ \min((x - D_H)^+, N_C) & \textit{if} \quad (\sigma_C, \sigma_L) = (stay, off), \\ N_C & \textit{if} \quad \sigma_L = on. \end{cases} \tag{a1}$$

Proof. Consider the first case, i.e., $(\sigma_C, \sigma_L) = (step, off)$. Then, because the competitor undercuts the hidden trader and submits at $B_{t_0} + \Delta$, he has price priority over the hidden trader. Hence, incoming market order shares x get first matched against the N_C of the competitor, i.e., $X_C = \min(x, N_C)$. Now assume $(\sigma_C, \sigma_L) = (stay, off)$. This time price priority between the competitor and the hidden trader is equal. However, the displayed part - having arrived at t_0 - has time priority over the competitor's order. Thus, the liquidity suppliers order only executes against the remaining $(x - D_H)^+$ shares of the initial market sell order of size x. Hence, the execution volume reads $X_C = \min((x - D_H)^+, N_C)$. Finally assume $\sigma_L = on$. Because of the block-trader's large demand, i.e. $N_L > N_H + N_C$, he will trade all shares from the competitor, i.e., $X_C = N_C$. □

Lemma A2 (Hidden Trader's Execution Volume)**.** *Let x denote the market sell order size at time T, N_C the liquidity supplier's total order size and D_H the hidden trader's display size. Then, the hidden trader's execution volume X_H obeys*

$$X_H = \begin{cases} \min\left((x - N_C)^+, N_H\right) & \text{if} \quad (\sigma_C, \sigma_L) = (step, off), \\ \min(x, D_H) + \min((x - D_H - N_C)^+, N_H - D_H) & \text{if} \quad (\sigma_C, \sigma_L) = (stay, off), \\ D_H + \min(x, N_H - D_H) & \text{if} \quad \sigma_L = on. \end{cases} \tag{a2}$$

Proof. We can essentially recycle the arguments of the proof in Lemma A1. In case $(\sigma_C, \sigma_L) = (step, off)$, the hidden trader has lower price priority than the competitor, thus his order gets executed only after a market order of size x has executed the competitor's N_C shares, i.e. $X_H = \min((x - N_C)^+, N_H)$. In case $(\sigma_C, \sigma_L) = (stay, off)$, the displayed part of the hidden trader gets served first (i.e., D_H), then the competitor (i.e., N_C shares) and at last the hidden trader's hidden part of the order (i.e., $N_H - D_H$). Thus,

$$X_H = \min(x, D_H) + \min((x - D_H - N_C)^+, N_H - D_H).$$

Finally assume $\sigma_L = on$. Because his demand is large, i.e., $N_L > N_H + N_C$, the latent trader will trade all displayed D_H shares. The remaining $N_H - D_H$ will be traded against the noise trader. Hence, the hidden trader's execution volume reads $X_H = D_H + \min(x, N_H - D_H)$. □

Traders' Payoffs

Lemma A3 (The Block-Trader's Payoff)**.** *Given the strategies σ_L, σ_C and $\sigma_H \equiv D_H$, the block-trader's payoff Π_L obeys*

$$\Pi^L = \begin{cases} -(\Delta+\gamma)\,(N_L - D_H - N_C)^+ & \textit{if} \quad (\sigma_L, \sigma_C) = (on, stay), \\ N_C\Delta - (\Delta+\gamma)\,(N_L - D_H - N_C)^+ & \textit{if} \quad (\sigma_L, \sigma_C) = (on, step), \\ -\gamma N_L & \textit{if} \quad (\sigma_L, \sigma_C) = (off, stay), \\ (\Delta-\gamma)\,N_L & \textit{if} \quad (\sigma_L, \sigma_C) = (off, step). \end{cases} \tag{a3}$$

Proof. Consider the first case, i.e., $(\sigma_L, \sigma_C) = (on, stay)$. The block-trader trades all displayed depth, i.e., $D_H + N_C$ shares, at the price $B_{t_0} = 0$. Consequently, downstairs market price shifts to $B_{t_0} - \Delta$. Thus, the remaining $N_L - N_C - D_H$ shares will get executed at the upstairs price $B_{t_0} - \Delta + \gamma$ and the (relative) payoff reads $\Pi_L = -(\Delta+\gamma)\,(N_L - D_H - N_C)^+$. Now, consider the second case, i.e., $(\sigma_L, \sigma_C) = (on, step)$. In this case, everything remains the same, except the block-trader executes N_C shares at one-Δ better price. Therefore, the payoff obeys $\Pi_L = N_C\Delta - (\Delta+\gamma)\,(N_L - D_H - N_C)^+$. Consider the case $(\sigma_L, \sigma_C) = (off, stay)$. The block-trader trades all N_L shares in the upstairs market by paying a fee γ for each of the shares, thus $\Pi_L = -\gamma N_L$. Finally, assuming $(\sigma_L, \sigma_C) = (off, step)$, i.e., the block-trader again trades all N_L in the upstairs market. As the liquidity competitor improves the public best bid price, the upstairs prices shifts as well according to (2.4). Therefore, payoff reads $\Pi_L = (\Delta - \gamma)N_L$. ☐

Lemma A4 (Liquidity Competitor's Payoff)**.** *Given the strategies σ_L, σ_C and $\sigma_H \equiv D_H$, the liquidity competitor's payoff Π_C obeys*

$$\Pi_C = \begin{cases} \Delta \min(x, N_C) + (S + \Delta + \tfrac{1}{2}\beta(N_C - x)^+)(N_C - x)^+ & if \quad (\sigma_C, \sigma_L) = (step, off), \\ \Delta N_C & if \quad (\sigma_C, \sigma_L) = (step, on), \\ \left(S + \tfrac{1}{2}\beta\,(N_C - (x - D_H)^+)^+\right)(N_C - (x - D_H)^+)^+ & if \quad (\sigma_C, \sigma_L) = (stay, off), \\ 0 & if \quad (\sigma_C, \sigma_L) = (stay, on). \end{cases} \tag{a4}$$

Proof. Follows directly from equation (2.2) and Lemma A1. For instance, assume $\sigma_L = off$ and $\sigma_C = stay$, the execution volume according to Lemma A1 equals $\min((x - D_H)^+, N_C)$ shares. Thus, the payoff according to (2.2) reads

$$\left(S + \frac{1}{2}\beta\left(N_C - (x - D_H)^+\right)^+\right)\left(N_C - (x - D_H)^+\right)^+.$$

On the other hand, when the liquidity competitor "steps ahead", i.e., $\sigma_C = step$, then the opportunity costs associated with executing the order increases marginally by one tick Δ, i.e., the payoff reads

$$\Delta \min(x, N_C) + (S + \Delta + \frac{1}{2}\beta(N_C - x)^+)(N_C - x)^+.$$

Now assume the case when the latent trader trades downstairs, i.e., $\sigma_L = on$. Then when the liquidity competitor improves the best bid (i.e., $\sigma_C = step$), his total payoff reads ΔN_C. If the competitor submits his limit order at the benchmark price $B_{t_0} = 0$ however, his execution costs are zero. □

Lemma A5 (Hidden Trader's Payoff). *Given the strategies σ_L, σ_C and $\sigma_H \equiv D_H$, the hidden trader's payoff Π_H obeys*

$$\Pi_H(\sigma_C, \sigma_L, \sigma_H) = \begin{cases} (S + \Delta + \frac{1}{2}\beta(N_H - (x - N_C)^+)^+)(N_H - (x - N_C)^+)^+ & if \quad (\sigma_C, \sigma_L) = (step, off), \\ (S + \frac{1}{2}\beta X_H)\,(N_H - X_H) & if \quad (\sigma_C, \sigma_L) = (stay, off), \\ (S - \Delta + \frac{1}{2}\beta(N_H - D_H - x)^+)(N_H - D_H - x)^+ & if \quad \sigma_L = on, \end{cases} \tag{a5}$$

with $X_H := \min(x, D_H) + \min((x - D_H - N_C)^+, N_H - D_H)$.

Proof. Follows directly from (2.2) and the execution volume X_H derived from Lemma A2. We proceed in the same fashion as before. Therefore, consider first $\sigma_L = off$ and assume $\sigma_C = step$. Because the competitor's order has priority over the hidden trader's order, in total $(x - N_C)^+$ standing (iceberg) order shares get executed at the benchmark price $B_{t_0} = 0$. Thus remaining $(N_H - (x - N_C)^+)^+$ shares have to get executed via markets orders at the (relative) price $(S + \Delta + \frac{1}{2}\beta(N_H - (x - N_C)^+)^+)$. Consider now $\sigma_C = stay$. In this case, the execution volume reads

$$X_H = \min(x, D_H) + \min((x - D_H - N_C)^+, N_H - D_H).$$

Together with (2.2), one obtains the result. Finally, in the case $\sigma_L = on$, the execution volume according to Lemma A2 reads $D_H + \min(x, N_H - D_H)$. As all visible liquidity has been replenished at B_{t_0}, the price shifts by a tick Δ downwards. Therefore, the remaining $N_H - D_H - \min(x, N_H - D_H) = (N_H - D_H - x)^+$ shares are executed as market orders at the price $(S - \Delta + \frac{1}{2}\beta(N_H - D_H - x)^+)$. □

A.2 Equilibrium Results of Section 3

Proof of Lemma 1 (Liquidity Competitor's Best Response). We use Lemma A4 and the fact that $\mu = 0$ or equivalently $\sigma_L = off$ holds. Hence, the competitor's payoff as in (a4) reduces to

$$\Pi^C(\sigma_C, D_H) = \begin{cases} S\left(N_C - (x - D_H)^+\right)^+ & if \quad \sigma_C = stay, \\ \Delta \min(x, N_C) + (S + \Delta)(N_C - x)^+ & if \quad \sigma_C = step. \end{cases} \tag{a6}$$

We want to find the strategy σ_C^* that minimizes the competitor's expected payoff given the hidden trader choses to display D_H shares, i.e.,

$$\sigma_C^* \equiv \arg\min_{\sigma_C \in \Sigma_C} \mathbb{E}[\Pi_C(\sigma_C, D_H)].$$

From (a6), we infer that $\mathbb{E}[\Pi_C(\sigma_C = stay, D_H = 0)] < \mathbb{E}[\Pi_C(\sigma_C = step, D_H = 0)]$ holds. Thus because of continuity, for sufficiently small display sizes D_H, $\sigma_C = stay$ is the optimal strategy for the competitor. On the other hand, $\mathbb{E}[\Pi_C(\sigma_C = stay, D_H)]$ is monotonously increasing in the display size D_H, whereas it is constant for $\sigma_C = step$. Let us denote Φ_C the critical threshold when both strategies exactly trade-off (if no such finite threshold exists, we symbolically write $\Phi_C = \infty$). Then, the optimal strategy can be expressed in the following way:

$$\sigma_C^* = \begin{cases} stay & \text{if} \quad D_H \leq \Phi_C, \\ step & \text{else.} \end{cases}$$

We obtain Φ_C by simply equating both payoffs $\mathbb{E}[\Pi_C(\sigma_C = stay, D_H)]$ and $\mathbb{E}[\Pi_C(\sigma_C = step, D_H)]$ and solving for $D_H \equiv \Phi_C$. That is

$$\begin{aligned} 0 &= \mathbb{E}[\Pi_C(\sigma_C = stay, \Phi_C)] - \mathbb{E}[\Pi_C(\sigma_C = step, \Phi_C)] \\ &= \left(\left(1 - e^{-\frac{N_C}{\lambda}}\right)\left(1 - e^{-\frac{\Phi_C}{\lambda}}\right)\lambda(\beta\lambda - S) + N_C e^{-\frac{\Phi_C}{\lambda}}\left(\beta\lambda + e^{\frac{\Phi_C}{\lambda}}(\Delta - \beta\lambda)\right)\right). \end{aligned}$$

Solving for Φ_C, we can finally rewrite the latter expression as

$$\Phi_C = \begin{cases} \lambda \log\left(\frac{1}{1-g}\right) & \text{if} \quad g < 1, \\ \infty & \text{else.} \end{cases}$$

with

$$g := \frac{N_C}{\lambda} \frac{\Delta}{S\left(1 - e^{-\frac{N_C}{\lambda}}\right) + \beta\left(N_C - \lambda\left(1 - e^{-\frac{N_C}{\lambda}}\right)\right)}.$$

□

Proof of Proposition 1 (Equilibrium). Because of Lemma 1 and the fact that $\lambda, N_C, \Delta > 0$ holds, the display threshold is positive, i.e., $\Phi_C > 0$. Let us therefore consider the first case, i.e., $N_H < \Phi_C$. Then, because of $D_H \leq N_H$ and by Lemma 1, the liquidity competitor *stays* at the same price level as the hidden trader, i.e., $\sigma_C^* = stay$. Thus according to Lemma A5, the hidden trader's (expected) payoff reads

$$\mathbb{E}\Big[\Pi_H(\sigma_C = stay, \sigma_H = D_H)\Big] = \mathbb{E}\Big[(S + \frac{1}{2}\beta(N_H - X_H))(N_H - X_H)\Big]$$
$$= SN_H + \frac{1}{2}\beta N_H^2 - \mathbb{E}[X_H](S + \beta N_H) + \frac{1}{2}\beta\mathbb{E}[(X_H)^2].$$

By Lemma A5, the hidden trader's payoff is monotonously decreasing in the display size D_H. Hence, $D_H^* = N_H$ and $\sigma_C^* = stay$.

Now assume the opposite case, i.e., $N_H \geq \Phi_C$ holds. Following the same reasoning, in case $D_H \leq \Phi_C$, the competitor chooses the $stay$-strategy and therefore we have

$$\mathbb{E}[\Pi_H(\sigma_H = D_H)] \geq \mathbb{E}[\Pi_H(\sigma_H = \Phi_C)], \quad D_H \leq \Phi_C.$$

Hence, $D_H^* \geq \Phi_C$. It remains to be shown that $D_H^* \leq \Phi_C$ holds. Therefore, consider the following expression

$$\begin{aligned}
&\mathbb{E}[\Pi_H | D_H \leq \Phi_C] - \mathbb{E}[\Pi_H | D_H > \Phi_C] = \\
&= \mathbb{E}\Big[(S + \frac{1}{2}\beta(N_H - X_H))(N_H - X_H)\Big| D_H \leq \Phi_C\Big] \\
&\qquad - \mathbb{E}\Big[(S + \Delta + \frac{1}{2}\beta(N_H - X_H))(N_H - X_H)\Big| D_H > \Phi_C\Big] \\
&\overset{(*)}{=} \underbrace{-\Delta\left(N_H - \mathbb{E}[X_H \big| D_H > \Phi_C]\right)}_{\leq 0} + (S + \beta N_H)\underbrace{\left(\mathbb{E}[X_H \big| D_H > \Phi_C] - \mathbb{E}[X_H \big| D_H \leq \Phi_C]\right)}_{<0} \\
&\qquad + \frac{1}{2}\beta\underbrace{\left(\mathbb{E}[(X_H)^2 \big| D_H \leq \Phi_C] - \mathbb{E}[(X_H)^2 \big| D_H > \Phi_C]\right)}_{<0} < 0.
\end{aligned}$$

The negativity of the first term in $(*)$ follows because $N_H \geq X_H$ by definition. The signs of the second and third terms follow directly from Lemma A2 and the fact that in equilibrium the competitor choses $\sigma_C = step$ in case $\sigma_H > \Phi_C$ and $\sigma_C = stay$ otherwise. Thus finally, $D_H^* \leq \Phi_C$ and therefore $D_H^* = \Phi_C$. □

Proof of Lemma 2 (Laten Trader's Best Response). First, assume $\sigma_C = stay$ and $0 \leq D_H \leq$

Φ_{L_a}. Then, according to the block investor's payoff (a3) and the definition of Φ_{L_a}, we have

$$\begin{aligned}\Pi_L(\sigma_L = off, \sigma_C = stay) - \Pi_L(\sigma_L = on, \sigma_C = stay) = \\ &= -\gamma N_L + (\Delta+\gamma)(N_L - D_H - N_C)^+ \\ &\geq -\gamma N_L + (\Delta+\gamma)(N_L - \Phi_{L_a} - N_C)^+ = 0.\end{aligned}$$

Thus, for $D_H \leq \Phi_{L_a}$ and $\sigma_C = stay$, the block-trader's optimal strategy obeys $\sigma_L^* = on$. In an analogous way, assuming $\sigma_C = stay$ and $\Phi_{L_a} < D_H$, we find $\sigma_L^* = off$. Thus, we can summarize

$$\sigma_L^* = \begin{cases} on & \text{for} \quad D_H \leq \Phi_{L_a} \\ off & \text{else} \end{cases} \qquad \text{for} \quad \sigma_C = stay. \tag{a7}$$

Now we consider the case $\sigma_C = step$. We proceed in the same fashion, according to (a3) and for $0 \leq D_H \leq \Phi_{L_b}$, we have

$$\begin{aligned}\Pi_L(\sigma_L = off, \sigma_C = step) - \Pi_L(\sigma_L = on, \sigma_C = step) = \\ &= (\Delta - \gamma)N_L - N_C\Delta + (\Delta+\gamma)(N_L - D_H - N_C)^+ \\ &\geq (\Delta - \gamma)N_L - N_C\Delta + (\Delta+\gamma)(N_L - \Phi_{L_b} - N_C)^+ = 0.\end{aligned}$$

In other words, for $D_H \leq \Phi_{L_b}$ and $\sigma_C = step$, the block-trader's optimal strategy is $\sigma_L^* = on$. Analogously we obtain for $\Phi_{L_b} < D_H$, that the block-trader's optimal strategy is $\sigma_L^* = off$. We thus have

$$\sigma_L^* = \begin{cases} on & \text{for} \quad D_H \leq \Phi_{L_b} \\ off & \text{else} \end{cases} \qquad \text{for} \quad \sigma_C = step. \tag{a8}$$

Because of $N_L > N_C$, we have $\Phi_{L_a} < \Phi_{L_b}$ and we can finally sum up both results (a7) and (a8)

$$\sigma_L^* = \begin{cases} off & \text{if} \quad 0 \leq D_H < \Phi_{L_a}, \\ on & \text{if} \quad \Phi_{L_a} \leq D_H < \Phi_{L_b} \quad \text{and} \quad \sigma_C = stay, \\ off & \text{if} \quad \Phi_{L_a} \leq D_H < \Phi_{L_b} \quad \text{and} \quad \sigma_C = step, \\ on & \text{if} \quad \Phi_{L_b} \leq D_H. \end{cases}$$

□

Proof of Lemma 3 (Liquidity Competitor's Best Response with Latent Investor). First, assume that $D_H \leq \Phi_{L_a}$ holds. Then, because of Lemma 2, the latent trader will never trade in the public exchange, i.e., $\sigma_L = off$. For the liquidity competitor and the hidden trader, this problem effectively reduces to the case without latent trader. We can thus recycle the results of Proposition 1, i.e., $\sigma_C^* = \sigma_C^{*0}$ with $0 \leq D_H < \Phi_{L_a}$, where σ_C^{*0} is referring to the competitor's equilibrium strategy without latent demand (i.e., $\mu = 0$).

Now assume $\Phi_{L_a} < D_H \leq \Phi_{L_b}$. According to Lemma 2, the latent trader trades in the primary exchange if (and only if) the competitor does not improve the best bid price, i.e., if $\sigma_C = stay$ holds. However, according to Lemma A3,

$$\Pi_C(\sigma_L = off, \sigma_C = step) - \Pi_C(\sigma_L = on, \sigma_C = stay) > 0$$

holds for any $x \geq 0$. Thus, $\sigma_C^* = stay$ for $\Phi_{L_a} < D_H \leq \Phi_{L_b}$.

Finally, consider the case $\Phi_{L_b} < D_H$. Again using Lemma 2, the latent trader will trade on the primary exchange, i.e., $\sigma_L = on$ and the payoff according to Lemma A3 obey

$$\Pi_C(\sigma_L = on, \sigma_C = stay) - \Pi_C(\sigma_L = on, \sigma_C = step) = -\Delta N_C < 0, \qquad x \geq 0.$$

If the latter inequality holds for all x, so it also holds in expectation. Thus, in this case, the liquidity competitor's optimal strategy is $\sigma_C^* = stay$ when $D_H > \Phi_{L_a}$ holds. □

Proof of Proposition 2 (Equilibrium with Block Investors). The best response strategies of the liquidity competitor and the latent trader have been shown in Lemma 2 and 3. To derive the equilibrium, the hidden trader's optimal strategy remains to be shown. For now, assume $N_H \leq \Phi_{L_a}$. Because $D_H \leq N_H \leq \Phi_{L_a}$ and because of Lemma 2, the latent trader will never trade in the public exchange, i.e., $\sigma_L = off$. Hence, the hidden trader's game reduces to the baseline model without the latent investor and therefore $\sigma_H^* = \sigma_H^{*0}$ for $N_H \leq \Phi_{L_a}$, where σ_H^{*0} denotes the hidden trader's equilibrium strategy without the latent investor as of Proposition 1.

Now assume the opposite case, i.e., $N_H > \Phi_{L_a}$. Because of Lemma 2 and Lemma 3, in case $D_H > \Phi_{L_a}$, the competitor will chose $\sigma_C = stay$ and the latent investor will chose $\sigma_L = on$. Hence in this case, the payoff according to Lemma A5 reads

$$\begin{aligned}
\mathbb{E}[\Pi_H \big| D_H > \Phi_{L_a}] =& \mathbb{E}[\Pi_H \big| \sigma_L = on, \sigma_C = stay, D_H > \Phi_{L_a}] \\
=& \mathbb{E}[(S + \frac{1}{2}\beta(N_H - X_H))(N_H - X_H) \big| \sigma_L = on, \sigma_C = stay, D_H > \Phi_{L_a}] \\
\geq& \mathbb{E}[(S + \frac{1}{2}\beta(N_H - X_H)) \underbrace{(N_H - X_H)}_{=0} \big| \sigma_L = on, \sigma_C = stay, D_H = N_H] \\
=& \mathbb{E}[\Pi_H \big| D_H = N_H] = 0.
\end{aligned}$$

On the other hand,

$$
\begin{aligned}
\mathbb{E}[\Pi_H \Big| D_H \leq \Phi_{L_a}] \geq & \mathbb{E}[\Pi_H \Big| D_H \leq \Phi_{L_a}, \sigma_C = stay] \\
= & \mathbb{E}[(S + \frac{1}{2}\beta(N_H - X_H)) (N_H - X_H) \Big| D_H \leq \Phi_{L_a}, \sigma_C = stay] \\
\geq & S\mathbb{E}[(N_H - X_H) \Big| D_H \leq \Phi_{L_a}, \sigma_C = stay] \\
\geq & S\mathbb{E}[(N_H - X_H) \Big| D_H \leq \Phi_{L_a}, \sigma_C = stay, N_C = 0] \\
= & S\left(N_H - \lambda(1 - e^{-\frac{N_H}{\lambda}})\right) > 0
\end{aligned}
$$

holds for finite $\lambda > 0$. Thus, we have $\mathbb{E}[\Pi_H] \geq \mathbb{E}[\Pi_H \Big| D_H = N_H]$, i.e., $D_H^* = N_H$. □

A.3 Partial Equilibrium Results of Section 4

Proof of Corollary 4 (Expected Execution and Cancelation Volumes). It suffices to show the assertion for $EXB_{t_0}^T$. The rest follows by definition. Therefore, first observe that the latent trader's market order size is governed by

$$
x_L = \begin{cases} D_H + N_C & if \qquad \sigma_L = on, \\ 0 & else. \end{cases} \tag{a9}
$$

But according to Lemma 2, in equilibrium (i.e., $\Sigma_C = stay$), if the hidden trader displays more, the latent trader will adopt the strategy $\sigma_L = on$. Thus x_L is a monotonically increasing function of D_H and so is $x + x_L$ as the noise trader's market order does not depend on D_H. In particular, $\min(x + x_L, N_C + N_H)$ is a monotonously increasing function of D_H and so must the expectation

$$
EXB_{t_0}^T = \mathbb{E}[\min(x + x_L, N_C + N_H)] \tag{a10}
$$

be a monotonously increasing function of D_H. □

Proofs of Proposition 3 and 4

To derive Proposition 3 and 4, we first establish some basic properties related to the binomial arrival of the hidden trader. Subsequently, we provide a representation of the midpoint returns in our model as established in Lemma A6. The results of excess returns and excess volatility will be based on this representation. Finally, Proposition A1 and A2 are established to prove Proposition 3 and 4 of Section 4.

Preparatory Remarks and Results

Remark A1 (Hidden Trader's Binomial Order Arrival). *Assume the hidden trader arrives with probability q at t_1. When he arrives, then $N_H = n > 0$, with $0 \leq D_H = d \leq N_H = n$, else $N_H = D_H = 0$. Thus, N_H and D_H can be described as binomial random variables with mean and variance obeying*

$$\begin{aligned} \mathbb{E}[\bar{N}_H] &= qN_H, & \mathbb{E}[\bar{D}_H] &= qD_H, \\ Var[\bar{N}_H] &= q(1-q)N_H^2, & Var[\bar{D}_H] &= q(1-q)D_H^2. \end{aligned} \tag{a11}$$

Both, $\bar{N_H}$ and $\bar{\Delta}$, are correlated with covariance

$$\begin{aligned} Cov[\bar{N}_H, \bar{D}_H] &= \mathbb{E}[\bar{N}_H \bar{D}_H] - \mathbb{E}[\bar{N}_H]\mathbb{E}[\bar{D}_H] \\ &= q(1-q)N_H D_H. \end{aligned} \tag{a12}$$

Moreover, denote the total sell market order volume at T by $\bar{X}$. The sell market order volume consists of both, the sell orders issued by the noise trader, i.e., x, and the market orders issued from the latent block trader, i.e., $x_L \leq N_L$. Thus, we have $\bar{X} = x + x_L$. Remember that

$$\bar{x_L} = \begin{cases} \bar{D}_H + N_C & \textit{if} \quad \sigma_L = on, \\ 0 & \textit{else}. \end{cases} \tag{a13}$$

Also by assumption arrival of the noise trader is independent of the actions of the hidden trader and the latent trader, i.e.,

$$Cov[x, x_L] = 0, \qquad Cov[x, \bar{D}_H] = 0, \qquad Cov[x, \bar{N}_H] = 0. \tag{a14}$$

Lemma A6 (Midpoint Returns). *The midpoint-return between time $t = 0$ and T, i.e., R_T obeys*

$$R_T = p_T^{mid} - p_0^{mid} = \frac{\beta}{2}(\bar{N}_H + N_C - \bar{X}), \tag{a15}$$

with $p_t^{mid} = \frac{A_t + B_t}{2}$.

Proof. When liquidity supply is higher than liquidity demand, i.e., $\bar{N}_H + N_C > \bar{X}$, then the (buy) liquidity suppliers issue $\bar{N}_H + N_C - \bar{X}$ shares of market buy orders to complete their trades at time T. Since we assume linear price impact of market orders as of (2.3) and a constant spread, i.e., $A_t = S + B_t \quad \forall t > 0$. This results in a price shift of the best ask price A_T and bid price B_T as follows

$$A_T = A_t + \frac{\beta}{2}(\bar{N}_H + C - \bar{X}), \qquad B_T = B_t + \frac{\beta}{2}(\bar{N}_H + C - \bar{X}).$$

The same holds in the presence if excess liquidity demand, i.e., $\bar{N}_H + N_C \leq \bar{X}$. Thus, ultimately, the midpoint return reads

$$R_{t_0}^T := m_T - m_{t_0} = \frac{\beta}{2}(\bar{N}_H + C - \bar{X}),$$

with $m_t = \frac{A_t + B_t}{2}$. □

To prove Proposition 3, we first establish the following proposition.

Proposition A1 (Expected Equilibrium Returns)**.**

(i) ***Partial Equilibrium****: Assume that $\sigma_L = \sigma_L^*$ and $\sigma_C = stay$ as of Proposition 2 and that the hidden trader chooses any strategy, i.e., $\sigma_H = D_H \equiv D_H^*$. Then the expected market return obeys*

$$\mathbb{E}[R_{t_0}^T | D_H] = \frac{\beta}{2} \begin{cases} qN_H + N_C - \lambda, & 0 \leq D_H < \Phi_{L_a}, \\ q\big(N_H - N_C - D_H\big) + N_C - \lambda, & 0 < \Phi_{L_a} \leq D_H, \\ q\big(N_H - D_H\big) - \lambda, & \Phi_{L_a} \leq 0. \end{cases} \tag{a16}$$

(ii) ***Full Equilibrium****: Assume all traders employ their equilibrium strategies, i.e., $\sigma_L = \sigma_L^*$, $\sigma_C = stay$ and $\sigma_H = \sigma_H^*$ as of Proposition 2, then the expected market return obeys*

$$\mathbb{E}[R_{t_0}^T | D_H = D_H^*] = \frac{\beta}{2} \begin{cases} qN_H + N_C - \lambda, & 0 \leq N_H < \Phi_{L_a}, \\ N_C(1-q) - \lambda, & 0 < \Phi_{L_a} \leq N_H, \\ -\lambda, & \Phi_{L_a} \leq 0. \end{cases} \tag{a17}$$

Proof. We first compute the expected return for the case $\Phi_{L_a} > D_H \geq 0$. According to Proposition 2, in this case the latent trader does not trade in the order book market, i.e., $\sigma_L \neq on$ and because of (a13), $\bar{x_L} = 0$ and $\bar{X} = x$ follow. Thus, together with (a15), we get

$$\begin{aligned} \mathbb{E}[R_{t_0}^T] &= \mathbb{E}[\beta(\bar{N}_H + N_C - \bar{X})] \\ &= \mathbb{E}[\beta(\bar{N}_H + N_C - x)] \\ &= \beta\big(qN_H + N_C - \lambda\big). \end{aligned} \tag{a18}$$

The conclusion for the case $0 < \Phi_L \leq D_H$ is found in total analogy. Thus, according to Proposition 2, $\sigma_L = on$ (i.e., $\bar{x}_L = D_H + N_C$) holds, when the hidden trader arrives and

$\sigma_L = off$ (i.e., $\bar{x}_L = 0$) otherwise. In particular, $\mathbb{E}[\bar{x_L}] = q(D_H + N_C)$ holds. Together with (a15), we obtain

$$\begin{aligned}\mathbb{E}[R^T_{t_0}] &= \mathbb{E}[\beta(\bar{N}_H + N_C - \bar{X})] \\ &= \beta \underbrace{\mathbb{E}[\bar{N}_H]}_{=qN_H} + \beta N_C - \beta \underbrace{\mathbb{E}[x]}_{=\lambda} - \beta \underbrace{\mathbb{E}[\bar{x}_L]}_{=q(D_H+N_C)} \\ &= \beta\Big(q\big(N_H - N_C - D_H\big) + N_C - \lambda\Big),\end{aligned} \tag{a19}$$

where we used (a11) and (a13). Least, we consider the case $\Phi_L \leq 0 \leq D_H$. Since $\Phi_L \leq 0$ holds, the latent trader trades always all visible shares, independent of the hidden trader's arrival. In particular, $\mathbb{E}[\bar{x}_L] = qD_H + N_C$ holds. Thus,

$$\begin{aligned}\mathbb{E}[R^T_{t_0}] &= \mathbb{E}[\beta(\bar{N}_H + N_C - \bar{X})] \\ &= \beta\mathbb{E}[(\bar{N}_H + N_C - x - \underbrace{\bar{x}_L}_{=\bar{D}_H+N_C})] \\ &= \beta\Big(qN_H - qD_H - \lambda\Big).\end{aligned}$$

(ii) Follows directly from (i) and Proposition 2. The cases $\Phi_{L_a} < 0$ and $N_H \geq \Phi_{L_a}$ follow immediately from $\sigma^*_H = D^*_H = N_H$. While the third case is independent of D_H. □

Proof of Proposition 3. Follows directly from Proposition A1 and the definition of excess returns (see Definition 2 on page 19). □

Similarly, to prove Proposition 4, we first establish the following proposition.

Proposition A2 (Equilibrium Volatility)**.**

(i) ***Partial Equilibrium****: Assume that $\sigma_L = \sigma^*_L$ and $\sigma_C = stay$ as of Proposition 2 and that the hidden trader chooses any strategy, i.e., $\sigma_H = D_H \equiv D^*_H$. Then, the midpoint return variance obeys*

$$Var[R^T_{t_0}] = \frac{1}{4}\beta^2\lambda^2 + \frac{1}{4}\beta^2(1-q)q\begin{cases} N_H^2, & 0 \leq D_H < \Phi_{L_a}, \\ (N_H - D_H - N_C)^2, & 0 < \Phi_{L_a} \leq D_H, \\ (N_H - D_H)^2, & \Phi_{L_a} \leq 0. \end{cases} \tag{a20}$$

(ii) ***Full Equilibrium****: Assume all traders employ their equilibrium strategies, i.e.,* $\sigma_L = \sigma_L^*$*,* $\sigma_C = stay$ *and* $\sigma_H = \sigma_H^*$ *as of Proposition 2, then volatility (i.e., return variance between time* t_0 *and* T*) obeys*

$$Var[R_{t_0}^T] = \frac{1}{4}\beta^2\lambda^2 + \frac{1}{4}\beta^2(1-q)q \begin{cases} N_H^2, & 0 \leq N_H < \Phi_{L_a}, \\ N_C^2, & 0 < \Phi_{L_a} \leq N_H, \\ 0, & \Phi_{L_a} \leq 0. \end{cases} \tag{a21}$$

Proof. Consider the case $0 \leq D_H < \Phi_{L_a}$. Due to Proposition 2 and (a13), $\sigma_L = off$, $\bar{x}_L = 0$ and $\bar{X} = x$ hold. Together with (a15), we have

$$\begin{aligned} Var[R_{t_0}^T] &= Var\left[\frac{1}{2}\beta(\bar{N}_H + N_C - \bar{X})\right] \\ &= Var\left[\frac{1}{2}\beta(\bar{N}_H + N_C - x)\right] \\ &= \frac{1}{4}\beta^2 \underbrace{Var[x]}_{\lambda^2} - \frac{1}{2}\beta^2 \underbrace{Cov[x, \bar{N}_H]}_{=0} + \frac{1}{4}\beta^2 \underbrace{Var[\bar{N}_H]}_{=q(1-q)N_H} \\ &= \frac{1}{4}\beta^2\lambda^2 + \frac{1}{4}\beta^2 q(1-q)N_H^2, \end{aligned}$$

where in the third equation we used the fact that N_C is a constant and in the fifth equation we used the stochastic properties of x and $\bar{N}_H$ as of (a12) and (a14).

Now we consider the case $0 < \Phi_L \leq D_H$. According to Proposition 2, we have $\sigma_L = on$ when the hidden trader arrives and $\sigma_L = off$ when not. Together with (a13) and (a15), we obtain

$$\begin{aligned} Var[R_{t_0}^T] &= Var\left[\frac{1}{2}\beta(\bar{N}_H + N_C - \bar{X})\right] \\ &= \frac{1}{4}\beta^2 Var[\bar{N}_H + N_C - x - \bar{x}_L] \\ &= \frac{1}{4}\beta^2 Var[\bar{N}_H - x - \bar{x}_L] \\ &= \frac{1}{4}\beta^2 \underbrace{Var[\bar{N}_H]}_{=q(1-q)N_H} + \frac{1}{4}\beta^2 \underbrace{Var[x]}_{=\lambda^2} + \frac{1}{4}\beta^2 \times \underbrace{Var[\bar{x}_L]}_{=q(1-q)(D_H+N_C)^2} \\ &\qquad - \frac{1}{2}\beta^2 \underbrace{Cov[\bar{N}_H, x]}_{=0} - \frac{1}{2}\beta^2 Cov[\bar{N}_H, \bar{x}_L] + \frac{1}{2}\beta^2 \underbrace{Cov[x, \bar{x}_L]}_{=0}, \end{aligned} \tag{a22}$$

where we used the fact that N_C is fix in the third equation and stochastic independence in the forth equation as of (a14). Moreover, since the latent trader only trades when the hidden trader arrives, $\bar{x}_L$ is itself a binomial random variable taking value $D_H + N_C$ with probability q and 0 otherwise. In particular, we have $Var[\bar{x}_L] = q(1-q)(D_H + N_C)^2$ and one can easily show that $Cov[\bar{N}_H, \bar{x}_L] = q(1-q)N_H(D_H + N_C)$ holds. Hence, the return variance obeys

$$\begin{aligned} Var[R^T_{t_0}] &= \frac{1}{4}\beta^2 q(1-q)N_H^2 + \frac{1}{4}\beta^2\lambda^2 + \frac{1}{4}\beta^2 q(1-q)(D_H + N_C)^2 - \frac{1}{2}\beta^2 \underbrace{Cov[\bar{N}_H, \bar{x}_L]}_{=q(1-q)N_H(D_H+N_C)} \\ &= \frac{1}{4}\beta^2\lambda^2 + \frac{1}{4}\beta^2 q(1-q)\Big(N_H - D_H - N_C\Big)^2. \end{aligned} \tag{a23}$$

Using the same arguments, we consider the case $\Phi_L \leq 0 \leq D_H$. In this case, the latent trader trades always the amount $\bar{x}_L = \bar{D}_H + N_C$. Hence,

$$\begin{aligned} Var[R^T_{t_0}] &= Var[\beta(\bar{N}_H + N_C - \bar{X})] \\ &= \frac{1}{4}\beta^2 Var[\bar{N}_H + N_C - x - \bar{x}_L] \\ &= \frac{1}{4}\beta^2 Var[\bar{N}_H - x - \bar{D}_H] \\ &= \frac{1}{4}\beta^2 \underbrace{Var[\bar{N}_H]}_{=q(1-q)N_H^2} + \frac{1}{4}\beta^2 \underbrace{Var[x]}_{=\lambda^2} + \frac{1}{4}\beta^2 \underbrace{Var[\bar{D}_H]}_{=q(1-q)D_H^2} \\ &\qquad - \frac{1}{2}\beta^2 \underbrace{Cov[\bar{N}_H, x]}_{=0} - \frac{1}{2}\beta^2 \underbrace{Cov[\bar{N}_H, \bar{D}_H]}_{=q(1-q)N_H D_H} + \frac{1}{2}\beta^2 \underbrace{Cov[x, \bar{D}_H]}_{=0} \\ &= \frac{1}{4}\beta^2\lambda^2 + \frac{1}{4}\beta^2 q(1-q)\Big(N_H - D_H\Big)^2. \end{aligned}$$

(ii) Follows directly from (i) and Proposition 2. The cases $\Phi_{L_a} < 0$ and $N_H \geq \Phi_{L_a}$ follow immediately from $\sigma^*_H = D^*_H = N_H$. While the third case is independent of D_H. □

Proof of Proposition 4. The proof is analogously to the proof of Proposition 3. It follows directly from Proposition A2 by applying the definition of excess volatility (see Definition 3 on page 20). □

References

Note: For ease of reading, appended back-references are provided for each reference.

ADMANTI, A. AND P. PFLEIDERER (1991): "Sunshine Trading and Financial Market Equilibrium," *Review of Financial Studies*, 4, 443–481. 4

AHN, H.-J., K.-H. BAE, AND K. CHAN (2001): "Limit Orders, Depth, and Volatility: Evidence from the Stock Exchange of Hong Kong," *The Journal of Finance*, 56, 767–788. 21

AITKEN, M. J., H. BERKMAN, AND D. MAK (2001): "The Use of Undisclosed Limit Orders on the Australian Stock Exchange," *Journal of Banking & Finance*, 25, 1589–1603. 2, 22

BESSEMBINDER, H., M. PANAYIDES, AND K. VENKATAMARAN (2009): "Hidden Liquidity: An Analysis of Order Exposure Strategies in Electronic Stock Markets," *Journal of Financial Economics*, 94, 361–383. 2, 3, 4, 13, 14

BESSEMBINDER, H. AND P. J. SEGUIN (1993): "Price Volatility, Trading Volume, and Market Depth: Evidence from Futures Markets," *The Journal of Financial and Quantitative Analysis*, 28, pp. 21–39. 21

BIAIS, B. (1993): "Price Formation and Equilibrium Liquidity in Fragmented and Centralized Markets," *Journal of Finance*, 48, 157–184. 4

BIAIS, B., P. HILLION, AND C. SPATT (1995): "An Empirical Analysis of the Limit Order Book and the Order Flow in the Paris Bourse," *The Journal of Finance*, 50, 1655–1689. 7, 14, 32

BOLLERSLEV, T. AND I. DOMOWITZ (2012): "Trading Patterns and Prices in the Interbank Foreign Exchange Market," *The Journal of Finance*, 48, 1421–1443. 22

BOLLERSLEV, T. AND M. MELVIN (1994): "Bid-ask Spread and Volatility in the Foreign Exchange Market," *Journal of International Economics*, 36, 355–372. 22

BOOTH, G. G., J.-C. LIN, T. MARTIKAINEN, AND T. YIUMAN (2002): "Trading and Pricing in Upstairs and Downstairs Stock Markets," *Review of Financial Studies*, 15, 1111–1135. 8

BUTI, S. AND B. RINDI (2013): "Undisclosed Orders and Optimal Submission Strategies in a Limit Order Market," *Journal of Financial Economics*, 109, 797 – 812. 5

CAO, C., O. HANSCH, AND X. WANG (2009): "The Information Content of an Open Limit-order Book," *Journal of Futures Markets*, 29, 16–41. 14, 32

CHORDIA, T., R. ROLL, AND A. SUBRAHMANYAM (2002): "Order Imbalance, Liquidity, and Market Returns," *Journal of Financial Economics*, 65, 111 – 130. 32

CONRAD, J., K. M. JOHNSON, AND S. WAHAL (2003): "Institutional Trading and Alternative Trading Systems," *Journal of Financial Economics*, 70, 99–134. 10

COPELAND, T. E. AND D. GALAI (1983): "Information Effects on the Bid-Ask Spread," *The Journal of Finance*, 38, 1457–1469. 5

DE WINNE, R. AND C. D'HONDT (2009): "Hidden Liquidity in a Pure Order-Driven Market," *Finance Letters, forthcoming*. 14, 22

DEGRYSE, H., F. D. JONG, M. V. RAVENSWAAIJ, AND G. WUYTS (2005): "Aggressive Orders and the Resiliency of a Limit Order Market," *Review of Finance*, 9, 201–242. 7

DOMOWITZ, C. AND MADHAVAN (2003): "Dynamics of Liquidity in an Electronic Limit Order Book Market," Working paper, Duke University. 7

ENGLE, R. F. AND A. J. PATTON (2004): "Impacts of Trades in an Error-correction Model of Quote Prices," *Journal of Financial Markets*, 7, 1–25. 32

FOUCAULT, T., O. KADAN, AND E. KANDEL (2005): "Limit Order Book as a Market for Liquidity," *Review of Financial Studies*, 18, 1171–1217. 4, 7

FREY, S. AND P. SANDAS (2009): "The Impact of Iceberg Orders in Limit Order Books," Working paper, Center for Financial Research. 2, 4, 13

FUNG, H.-G. AND G. A. PATTERSON (1999): "The Dynamic Relationship of Volatility, Volume, and Market Depth in Currency Futures Markets," *Journal of International Financial Markets, Institutions and Money*, 9, 33–59. 21

GLOSTEN, L. R. AND P. R. MILGROM (1985): "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics*, 14, 71 – 100. 5, 7

GRIFFITHS, M. D., D. ALASDAIR, S. TURNBULL, AND R. W. WHITE (2001): "Upstairs Market for Principal and Agency Trades: Analysis of Adverse Information and Price Effects," *Journal of Finance*, 56, 65 – 88. 15

GRIFFITHS, M. D., B. F. SMITH, D. A. S. TURNBULL, AND R. W. WHITE (1998): "The Role of Tick Size in Upstairs Trading and Downstairs Trading," *Journal of Financial Intermediation*, 7, 393–417. 15

GROSSMANN, S. (1992): "The Informational Role of Upstairs and Downstairs Markets," *Journal of Business*, 65, 509–529. 3, 10

HALL, A. D. AND N. HAUTSCH (2006): "Order Aggressiveness and Order Book Dynamics," *Empirical Economics*, 30, 973–1005. 4, 14, 32

HARRIS, L. (1994): "Minimum Price Variations, Discrete Bid-Ask Spreads, and Quotation Sizes," *Review of Financial Studies*, 7, 149–178. 14

——— (1996): *Does a Large Minimum Price Variation Encourage Order Exposure?*, vol. 96, New York Stock Exchange. 14, 22, 25

——— (1997): "Order Exposure and Parasitic Traders," Working paper series, University of Southern California. 3, 5

——— (2003): *Trading & Exchanges. Market Microstructure for Practitioners*, Oxford University Press. 8

HARRIS, L. AND J. HASBROUCK (1996): "Market vs. Limit Orders: the SuperDOT Evidence on Order Submission Strategy," *Journal of Financial and Quantitative Analysis*, 31. 8

HASBROUCK, J. (1991): "Measuring the Information Content of Stock Trades," *The Journal of Finance*, 46, 179–207. 31

HASBROUCK, J. AND G. SAAR (2001): "Limit Orders and Volatility in a Hybrid Market: The Island ECN," . 22

——— (2009): "Technology and Liquidity Provision: The Blurring of Traditional Definitions," *Journal of Financial Markets*, 12, 143 – 172. 3

HAUTSCH, N. AND R. HUANG (2012): "The Market Impact of a Limit Order," *Journal of Economic Dynamics and Control*, 36, 501–522. 29, 31, 32

HENDERSHOTT, T. AND C. M. JONES (2005): "Island Goes Dark: Transparency, Fragmentation, Liquidity Externalities, and Multimarket Regulation," *Review of Financial Studies*, 18, 743–793. 2, 15

Hendershott, T. and H. Mendelson (2002): "Crossing Networks and Dealer Markets: Competition and Performance," *The Journal of Finance*, 55, 2071–2115. 4

Jorda, O. (2005): "Estimation and Inference of Impulse Responses by Local Projections," *The American Economic Review*, 95, 161–182. 33

Kalimipalli, M. and D. A. Warga (2002): "Bid-Ask Spread, Volatility, and Volume in the Corporate Bond Market," *The Journal of Fixed Income*, 11, 31–42. 22

Keim, D. B. and A. Madhavan (1996): "The Upstairs Market for Large-Block Transactions: An Analysis and Measurement of Price Effects," *The Review of Financial Studies*, 9, 1–36. 7, 8

Kyle, R. (1985): "Continous Auctions and Insider-Trading," *Econometrica*, 53, 1315–1335. 4, 7, 25

Madhavan, A. and M. Cheng (1997): "In Search of Liquidity: Block Trades in the Upstairs and Downstairs Markets," *The Review of Financial Studies*, 10, 175–203. 7

Mas-Colell, A., M. D. Whinston, and J. Green (1995): *Microeconomic Theory*, Oxford University Press. 11

Obizhaeva, A. and J. Wang (2013): "Optimal Trading Strategy and Supply/Demand Dynamics," *Journal of Financial Markets*, 16, 1–32. 30

Perold, A. F. (1988): "The Implementation Shortfall: Paper versus Reality," *Journal of Portfolio Management*, 14, 4–9. 8

Pesaran, H. H. and Y. Shin (1998): "Generalized Impulse Response Analysis in Linear Multivariate Models," *Economics Letters*, 17–29. 32, 33

Plerou, V., P. Gopikrishnan, and H. E. Stanley (2005): "Quantifying Fluctuations in Market Liquidity: Analysis of the Bid-ask Spread," *Physical Review E*, 71, 046131. 22

Ragunathan, V. and A. Peker (1997): "Price Variability, Trading Volume and Market Depth: Evidence from the Australian Futures Market," *Applied Financial Economics*, 7, 447–454. 21

Ranaldo, A. (2004): "Order Aggressiveness in Limit Order Book Markets," *Journal of Financial Markets*, 7, 53 – 74. 4, 14, 32

WANG, G. H. K. AND J. YAU (2000): "Trading volume, Bid-ask Spread, and Price Volatility in Futures Markets," *Journal of Futures Markets*, 20. 22

WATANABE, T. (2001): "Price Volatility, Trading Volume, and Market Depth: Evidence from the Japanese Stock Index Futures Market," *Applied Financial Economics*, 11, 651–658. 21

CFS Working Paper Series

Recent Issues

All CFS Working Papers are available at *www.ifk-cfs.de*.

No.	Authors	Title
467	Frank Betz, Nikolaus Hautsch, Tuomas A. Peltonen, Melanie Schienle	*Systemic Risk Spillovers in the European Banking and Sovereign Network*
466	Christiane Baumeister, Lutz Kilian	*A General Approach to Recovering Market Expectations from Futures Prices with an Application to Crude Oil*
465	Christopher D. Carroll, Jonathan A. Parker, Nicholas S. Souleles	*The Benefits of Panel Data in Consumer Expenditure Surveys*
464	Christopher D. Carroll	*Representing Consumption and Saving Without A Representative Consumer*
463	Piergiorgio Alessandri Sergio Masciantonio Andrea Zaghini	*Everything you always wanted to know about systemic importance (but were afraid to ask)*
462	Marti G. Subrahmanyam Dragon Yongjun Tang Sarah Qian Wang	*Credit Default Swaps and Corporate Cash Holdings*
461	Nils Friewald Rainer Jankowitsch Marti G. Subrahmanyam	*To Disclose or not to Disclose: Transparency and Liquidity in the Structured Product Market*
460	Lutz Kilian Robert J. Vigfusson	*The Role of Oil Price Shocks in Causing U.S. Recessions*
459	Simone M. Sepe Charles K. Whitehead	*Paying for Risk: Bankers, Compensation, and Competition*
458	Andrej Gill Matthias Heinz Heiner Schumacher	*Trust, Trustworthiness and Selection into the Financial Industry*

Center for Financial Studies | Goethe University Frankfurt | Grüneburgplatz 1 | 60323 Frankfurt am Main | Germany
Phone +49 69 798 30050 | Fax +49 69 798 30077 | info@ifk-cfs.de | www.ifk-cfs.de