On the (Market Microstructure) Origins of the Return Distribution

January 15, 2025

Abstract

We propose a reduced-form microstructure model of price formation in an order-driven market. In this framework, the shape of the return distribution is determined jointly by the distribution of market orders and the shape of the limit order book. Our model implies: (1) the mean and skewness of the return distribution are increasing functions of the LOB imbalance (the ratio of the slopes of the ask and bid sides); (2) the return variance is a decreasing function of the LOB convexity (the ratio of the slopes of the higher and lower levels); (3) the return kurtosis is an increasing function of the LOB convexity; (4) a higher LOB imbalance shifts both the left (below the median) and right (above the median) parts of the return distribution to the right, while leaving the median fixed at zero; (5) all quantiles below (above) the median increase (decrease) as LOB convexity increases; (6) ask-side convexity primarily affects the right side of the return distribution, while bid-side convexity influences the left side. We test and provide empirical support for the predictions of our model using comprehensive ultra-high-frequency limit order book data for a sample of NYSE stocks from 2002 to 2012. We establish the causal effect of the shape of the limit order book on the distribution of stock returns using RegSHO as an exogenous shock to the shape of the limit order book.

Key words: Limit Order Book, Imbalance, Convexity, Skewness, Kurtosis, Quantiles, Market Order Distribution, RegSHO

JEL Classification: G12, G18

1 Introduction

In order-driven markets, non-marketable limit orders are placed in a queue prioritized by price and time, forming what is known as the limit order book. Transaction prices are established when market orders (and marketable limit orders) arrive and execute against the existing limit orders in the book. The limit order book, in this sense, acts as a mirror reflecting incoming market orders. Consequently, the distribution of market orders and the shape of the limit order book jointly determine the distribution of returns.

Figure 1 provides an intuitive framework for visualizing the relationship between the distribution of market orders, the shape of the limit order book, and the price formation. In this simplified framework, we assume that the average shape of both the ask and bid sides of the limit order book over a given interval – say, a trading day – can be represented by a continuous price impact function centered around a midquote price (set at $15 in this example) as depicted in the top-right panel, labeled "Price Impact Function $P(Q)$." Here, positive values in the market order distribution correspond to market buy orders executed against the ask side of the limit order book, while negative values correspond to market sell orders executed against the bid side. Furthermore, we assume that, over this time interval, market orders arrive according to a normal distribution, shown in the lower-right panel, titled "Quantity PDF." We further assume in this simplified framework that there is no market order imbalance – defined as the difference in volume (in number of shares) between buy and sell orders – yielding a market order distribution with a mean of zero. As market orders arrive and execute (or, figuratively speaking, reflected) against the limit order book, transaction prices are observed, as represented in the top-left panel, "Price PDF."

[Figure 1 about here.]

Panel (a) of Figure 1 illustrates the baseline scenario with a balanced, linear limit order book, where the price impact functions for both bid and ask sides are linear and have identical slopes. In this case, the return distribution is a scale transformation of the market order distribution, where the scaling factor is the reciprocal of the limit order book slope. Thus, if the market orders follow a normal distribution with zero mean and some variance, the returns will also exhibit a normal distribution with zero mean, and the return variance will equal the market order variance multiplied by the slope of the limit order book. This is the well-known fact that higher transaction costs or greater illiquidity lead to greater return variance, given that the limit order book slope can be considered as a measure of transaction costs, or equivalently, market illiquidity.

Panel (b) of Figure 1 illustrates the impact of an imbalanced limit order book, where the ask side's price impact function has a steeper slope than the bid side, though both remain linear. Here, the limit order book (LOB) imbalance, defined as the (log) ratio of the ask and bid side slopes, can be considered as an indicator of buying or selling pressure from limit orders. When the ask side slope exceeds that of the bid side, as in this example, there is relatively more liquidity on the bid side, creating a relative buying pressure from limit orders. Consequently, a positive limit order book imbalance produces a positively skewed return distribution in the absence of market order (MO) imbalance, suggesting a strong link between return skewness and LOB imbalance.

Panel (c) of Figure 1 demonstrates the effects of a balanced but convex limit order book, where the price impact functions for both ask and bid sides have equal overall slopes, indicating no LOB imbalance, but exhibit convexity – meaning that slopes farther from the midquote are steeper than those closer to it. LOB convexity, defined as the ratio of slopes at higher versus lower levels, can be viewed as an indicator of the relative price impact of large versus small market orders. A convex (or concave) limit order book occurs when the slope farther from the midquote, i.e., at higher levels, is steeper (or less steep) than the slope near the midquote, i.e., at lower levels. In a convex limit order book, as seen in Panel (c), large market orders exert a disproportionately stronger price impact due to reduced liquidity at levels away from the midquote. Consequently, this shape leads to a return distribution with fatter tails and higher kurtosis than a normal distribution, indicating a direct relationship between LOB convexity and the tails and kurtosis of the return distribution.

Building on our intuition from this simple framework, we develop a reduced-form model of price formation in an order driven market and conduct a formal analysis of how the shape of the limit order book influences the first four central moments and the quantiles of the return distribution. Specifically, under assumptions similar to those used in Figure 1, we demonstrate that returns can be represented as the sum of truncated normal variables in our reduced-form price formation model. Consequently, the non-central and central moments of the return distribution can be derived in closed form, expressed as functions of the parameters governing the market order distribution and the shape of the limit order book. We also obtain the quantiles of the return distribution either based on Cornish-Fisher expansion using its first four central moments or based on simulations.

The implications of our reduced-form price formation model are as follows: both the mean and skewness of the return distribution are monotonically increasing functions of LOB imbalance, regardless of MO imbalance or LOB convexity. The impact of LOB imbalance on return variance and kurtosis, however, is contingent on the sign of MO imbalance. Specifically, with positive MO imbalance, return variance increases with LOB imbalance, while return kurtosis decreases. Conversely, with negative MO imbalance, return variance decreases, and return kurtosis increases as LOB imbalance increases.

While the first four central moments reveal significant insights into how LOB imbalance influences the return distribution's shape, they do not provide a complete picture. We further examine how the quantiles of the return distribution vary with LOB imbalance through simulations based on our reduced-form model. All quantiles, with the exception of the median, of the return distribution are higher than the corresponding quantiles of a normal distribution when there is positive LOB imbalance regardless of the LOB convexity. This implies that, with zero MO imbalance, a higher LOB imbalance shifts both the left (below the median) and right (above the median) parts of the return distribution to the right, while leaving the median fixed at zero. This suggests that the effects of the LOB imbalance on return mean and skewness dominate its effects on return variance and kurtosis.

Turning to LOB convexity, we find that its impact on the return variance and kurtosis is straightforward when considering symmetric convexity, where the convexities of the ask and bid sides are equal. Specifically, return variance decreases, while kurtosis increases with LOB convexity. However, the effects of LOB convexity on return mean and skewness are more complex and depend closely on the signs and magnitudes of both MO and LOB imbalances. For instance, when both LOB and MO imbalances are positive, the return mean decreases with increasing LOB convexity, while return skewness increases.

The combined effect of these findings on the return distribution is that all quantiles below (above) the median increase (decrease) as LOB convexity rises, irrespective of LOB imbalance. In line with the framework discussed earlier, our reduced-form price formation model further suggests that higher LOB convexity leads to fatter tails and higher kurtosis compared to a normal distribution, after accounting for its effect on return variance. Moreover, when considering bid- and ask-side convexities separately, our model implies that ask-side convexity primarily affects the right side of the return distribution, while bid-side convexity influences the left side.

Our reduced-form price formation model formalizes not only the intuition derived from our simple framework but also the testable implications for the effect of LOB shape parameters on the return distribution. We evaluate these implications using data from the Trade and Quote (TAQ) for the parameters of the return and market order distributions and Thomson Reuters (now Refinitiv) Tick History (TRTH) databases for the LOB shape parameters. In our analysis, we aggregate data at the daily level to reduce the sensitivity of higher moments to intraday noise. Specifically, we compute the first four central moments of the return distribution using five minute midquote returns and those of the market order distribution using the number of shares of each trade for each stock-day pair in our sample. We also compute the LOB shape parameters for each snapshot and use their time-weighted averages over a given day as our daily proxies.

We estimate separate daily fixed effects panel regressions of the first four central moments of the return distribution on lagged values of LOB shape parameters. In these regressions we control for the contemporaneous effects of the first four central moments of the market order distribution, as well as for the three central moments of the return distribution other than the one being analyzed. Our choice to aggregate the variables at a daily level implies that the return distribution on day t is determined by the market order distribution on day t executed against the limit order book on the previous trading day $t-1$. This approach also helps to mitigate endogeneity concerns in interpreting the effect of LOB shape parameters on the shape of the return distribution.

Our results provide strong empirical support for the main predictions of our model. Specifically, we observe that return skewness increases with LOB imbalance, independent of MO imbalance and LOB convexity, consistent with our theoretical expectations. Although the economic magnitude of this effect is somewhat smaller than predicted by our model, LOB imbalance ranks as the second most economically significant factor among the MO and LOB variables considered. Similarly, we find that variance decreases and kurtosis increases with LOB convexity, aligning with our theoretical predictions. Notably, LOB convexity emerges as the second most economically significant variable influencing return variance and kurtosis.

By considering lagged LOB shape parameters in our regressions for return moments, we argue that our results establish a causal effect of LOB shape on return moments. To further support causality, we leverage the exogenous shock to LOB shape caused by Regulation SHO (Reg SHO). Our approach closely follows Diether et al. (2009), who examined Reg SHO's effect on market quality. Specifically, we assess the causal impact of LOB parameters on return moments by using the exogenous changes to LOB shape induced by Reg SHO.

Our results indicate that LOB convexity is the primary driver of changes in return variance and kurtosis. Specifically, LOB convexity for pilot stocks decreases significantly more than for control stocks after the uptick rule's suspension. This relative decrease in LOB convexity accounts for the observed increase in return variance and decrease in return kurtosis. We further show that other factors, such as shifts in the market order distribution, cannot explain these changes in the return distribution.

Our paper contributes to both the theoretical and empirical literature on limit order books in many dimensions. Theoretically, our reduced-form price formation model can be considered an extension of Kyle's framework (Kyle (1985)) where we allow not only for distinct price impact functions (lambdas) on the buy and sell sides but also different price impacts of large versus small market orders. Both Rosu (2009) and Goettler, Parlour, and Rajan (2009) develop theoretical models of the limit order book where the price impact functions can be nonlinear. However, these and other theoretical papers, such as Foucault, Kadan, and Kandel (2005) and Chakrabarty, Li, Nguyen, and Van Ness (2007) are mostly interested in the theoretical relation between the LOB and future price movements, i.e. the first moment of returns. To the best of our knowledge, we are among the first to consider how the shape of the LOB determines the whole distribution of returns and not just its first moment.

Empirically, there is a growing body of literature analyzing the information content of LOB for future price movements. Biais, Hillion, and Spatt (1995) are among the first to explore the dynamics of limit order markets and observe that price adjustments tend to follow the direction of prior limit order flows, implying that the limit order book contains valuable information regarding future price trends. Among others, Hasbrouck (1991), Hasbrouck (1992), Keim and Madhavan (1996), and Knez and Ready (1996), Kaniel and Liu (2006), Wuyts (2008), Kalay and Wohl (2009), Cao, Hansch, and Wang (2009), Beltran-Lopez, Giot, and Grammig (2009), Kozhan and Salmon (2012), Roesch and Kaserer (2014), Cenesizoglu, Dionne, and Zhou (2022), Cenesizoglu and Grass (2018) examine LOB data from markets around the world, such as US, Spain, Israel, Australia, and Germany, and show that LOBs are generally nonlinear and contain information about future price movements. Once again, most of these papers focus on the empirical relation between the LOB and future price movements, i.e. the first moment of returns. The closest paper to ours is Brennan, Chordia, Subrahmanyam, and Tong (2012) who conduct an empirical analysis of buy- and sell-side liquidity pricing using monthly transaction data. Our high-frequency data provides a unique opportunity to uncover new empirical relationships between LOB shape and return distribution, contributing significantly to the empirical literature on the effect of LOB on returns. More importantly, we are among the first to analyze how the shape of the LOB affects the whole return distribution and not just its first moment.

The rest of the paper is organized as follows: Section 2 presents our reduced-form price formation model. Section 3 discusses the implications of our reduced-form price formation model for the effects of the shapes of limit order book and market order distribution on the return distribution. Section 4 presents the data and the variable definitions. Section 5 presents the results from our regression analysis. Section 6 attempts to establish a causal effect of LOB shape on the return distribution based differences-in-differences analysis using RegSHO as an exogenous shock to LOB shape. Finally, Section 8 concludes.

2 Reduced-Form Model of Price Formation

In this section, we first develop a reduced-form model of how transaction prices form in an order-driven market when market orders are executed against the existing limit order book. We then derive the implications of this reduced-form price formation model for the effects of the market order distribution and, more importantly, the limit order book shape on the first four central moments and quantiles of the return distribution. For the rest of the paper, small case letters indicate variables in log.

We assume that at any instant t, the market orders denoted by Q arrive with a normal distribution with mean μ and standard deviation σ, i.e. $q \sim N(\mu, \sigma^2)$. Positive (negative) values of Q correspond to market buy (sell) orders and are matched and executed against the prevailing limit orders available in the ask (bid) side of the limit order book. We refer to the mean of the market order distribution as market order (MO) imbalance.

We assume that ask and bid sides of the limit order book can be described by piecewise linear price impact functions. To be more precise, we assume that the log prices in the ask and bid sides of the limit order book can be written as piecewise linear functions of their respective cumulative depths, D_A and D_B:

$$
\begin{aligned}
p_A(D_A) &= m + (a-b)/2 + S_{A,low} D_A 1_{\{0<D_A<K_A\}} + ((S_{A,low} - S_{A,high})K_A + S_{A,high} D_A) 1_{\{D_A \geq K_A\}} \\
p_B(D_B) &= m - (a-b)/2 - S_{B,low} D_B 1_{\{0<D_B<K_B\}} - ((S_{B,low} - S_{B,high})K_B + S_{B,high} D_B) 1_{\{D_B \geq K_B\}}
\end{aligned}
$$

where a and b are the logarithm of the best ask and bid prices, respectively, and m is the midquote price defined as $(a+b)/2$. We assume that there is total cumulative depth of $\bar{D}_A$ available in the ask side of the limit order book and use K_A to denote the threshold for cumulative depth distinguishing between low and high levels. Any cumulative depth less than K_A is considered as corresponding to the lower levels of the ask side. $S_{A,low}$ and $S_{A,high}$ are the slopes of the lower and higher levels of the ask side, respectively. The slopes, total cumulative depth and threshold for the bid side are defined similarly.

The overall slopes of the ask and bid sides, S_A and S_B can be defined as the quantity-weighted averages of the slopes of their corresponding low and high levels as follows:

$$S_A = \frac{K_A}{\bar{D}_A} S_{A,low} + \frac{\bar{D}_A - K_A}{\bar{D}_A} S_{A,high} \tag{1}$$

$$S_B = \frac{K_B}{\bar{D}_B} S_{B,low} + \frac{\bar{D}_B - K_B}{\bar{D}_B} S_{B,high} \tag{2}$$

We can also define the overall slope of the limit order book, S_{avg}, as the simple arithmetic average of the overall slopes of the ask and bid sides, i.e. $S_{avg} = (S_A + S_B)/2$.[1]

There are several intuitive ways to think about the slope variables. We discuss them for the overall slope of the ask side but they are very similar for the other slope variables. First of all, mathematically, the overall slope of the ask side measures how the ask prices increase as a function of the quantity demanded, i.e. the size of the market buy order to be executed. Thus, when the ask side is steeper, the ask prices increase faster as a function of the quantity demanded compared to a less steep ask side. Second, the slope can be considered as a measure of transaction costs in the limit order book. Specifically, a trader who wants to buy a certain number of shares will have to pay a higher

[1] We can also define the overall LOB slope, S_{avg}, as a weighted average of the overall slopes of the ask and bid sides as $S_{avg} = \bar{D}_A/(\bar{D}_A + \bar{D}_B)S_A + \bar{D}_B/(\bar{D}_A + \bar{D}_B)S_B)$. As we will discuss below, the overall LOB slope acts as a scaling factor for the mean and variance of the return distribution but does not affect its skewness and kurtosis. Hence, using this alternative definition of the overall LOB slope will slightly change the mean and variance implied by our reduced-form model of price formation but will not alter skewness or kurtosis.

(quantity-weighted) average price when the ask side is steeper. Third, the slope can be also considered as a measure of liquidity with a steeper slope indicating that there is less liquidity in the limit order book. Finally, one can think of the overall slope of the ask side as a measure of selling pressure from the limit orders. To be more precise, if the overall slope of the ask side is lower, there is more liquidity on the ask side and thus more selling pressure from the limit orders. In the rest of the paper, we use these different ways to think about the LOB slopes interchangeably, i.e. a steeper slope indicates higher transaction costs and less liquidity in the limit order book.

Suppose now that a market order Q arrives at time t and is executed against the corresponding side. Assuming that the spread $(a-b)$ remains unchanged after the market order is executed, the new midquote price, $\tilde{m}$, is given by:

$$\tilde{m} = \begin{cases} m + (a-b)/2 + S_{A,low}Q1_{\{0<Q<k_A\}} + ((S_{A,low} - S_{A,high})K_A + S_{A,high}Q)1_{\{Q \geq K_A\}}, & \text{if } Q > 0; \\ m - (a-b)/2 - S_{B,low}|Q|1_{\{0<|Q|<k_B\}} - ((S_{B,low} - S_{B,high})K_B + S_{B,high}|Q|)1_{\{|Q| \geq K_B\}}, & \text{if } Q < 0. \end{cases} \tag{3}$$

Equation 3 follows from our assumption that positive (negative) values of Q correspond to market buy (sell) orders and are matched and executed against the limit orders on the ask (bid) side of the limit order book. More specifically, when a market buy (sell) order arrives, it is matched against the ask (bid) side and the price impact function of the ask (bid) side determines the new midquote price. For market sell orders, we need to use their absolute values since the price impact function of the bid side are defined over cumulative depth, which is positive by definition. The midquote return is then given by:

$$\begin{aligned} r \equiv \tilde{m} - m &= \begin{cases} S_{A,low}Q1_{\{0<Q<K_A\}} + ((S_{A,low} - S_{A,high})K_A + S_{A,high}Q)1_{\{Q \geq K_A\}}, & \text{if } Q > 0; \\ -(S_{B,low}|Q|1_{\{0<|Q|<K_B\}} + ((S_{B,low} - S_{B,high})K_B + S_{B,high}|Q|)1_{\{|Q| \geq K_B\}}), & \text{if } Q < 0. \end{cases} \\ &= S_{A,low}Q1_{\{0<Q<K_A\}} + ((S_{A,low} - S_{A,high})K_A + S_{A,high}Q)1_{\{Q \geq K_A\}} \\ &+ S_{B,low}Q1_{\{-K_B<Q<0\}} + ((S_{B,high} - S_{B,low})K_B + S_{B,high}Q)1_{\{Q \leq -K_B\}} \end{aligned} \tag{4}$$

It is easy to see from Equation 4 the intuition behind the relation between the shape of the limit order book and returns. First of all, it is the ask side of the book that determines the return when a market buy order arrives, i.e. $Q > 0$. If the size of market buy order is small, smaller than the threshold K_A to be more precise, then return is determined only by the slope of the lower levels. On other hand, if the size of the market buy order is larger than the threshold K_A, then the return is determined by the slopes of both the higher and lower levels. Furthermore, steeper is the overall slope, higher is the price impact and, thus, return in absolute value. A similar intuition holds when for the bid side for market sell orders, i.e. negative values of Q.

Several remarks are in order before proceeding to the implications of this model for the relation between the shape of the limit order book and the return distribution. First of all, we make the implicit assumption that no other order (limit or market) arrives when computing returns based on a given market order and shape of the limit order book. Second, our approach is intentionally reduced-form in the sense that we do not consider the underlying economic mechanism that give rise to a specific market order and/or shape of the limit order book. In reality, they are jointly determined by the interactions of strategic economic agents and might depend on past market and limit orders as well as returns among other potential factors. In our model, we abstract from these complications in favor of a simple analysis on the relation between the shape of the limit order book and the return distribution.

Before deriving the moments of the return distribution based on this reduced-form model of price formation, we define several variables which summarize the shape of the limit order book. We start with the LOB imbalance, denoted as I and defined as the ratio of the overall ask and bid side slopes, i.e. $I = S_A/S_B$. In order to have a symmetric measure centered around zero, we consider the log transformation of imbalance when analyzing its relation with the return distribution in Section 3. A negatively infinite (log) imbalance indicates that the ask side is infinitely more liquid than the bid side and the price does not change following a market buy order of any size. Similarly, a positively infinite (log) imbalance indicates that the bid side is infinitely more liquid than the ask side and the price does not change following a market sell order of any size. Between these two extremes, a positive (log) imbalance indicates that the slope of the ask side is higher than that of the bid side and, thus, a relative buying pressure from limit orders given that there is relatively more liquidity on the bid side. A (log) imbalance of zero indicates a balanced limit order book, or equivalently, no buying or selling pressure from the limit orders.

We define ask (bid) side convexity, C_A (C_B), as the ratio of the slopes of its higher levels to those of its lower levels as follows:

$$C_A = S_{A,high}/S_{A,low} \tag{5}$$

$$C_B = S_{B,high}/S_{B,low} \tag{6}$$

Once again, in order to have a symmetric measure centered around zero, we consider the log transformation of convexity when analyzing its relation with the return distribution in Section 3. A negatively infinite (log) convexity indicates the lower levels are infinitely more liquid than the higher levels and the price does not change following a market order of any size. Similarly, a positively infinite (log) convexity indicates that higher levels are infinitely more liquid than the lower levels and the price change is only determined by the lower levels. Between these two extremes, a positive (log) convexity indicates there is less liquidity in the higher levels of the ask side and a large (greater than K_A to be more precise) market buy order would have a disproportionately larger market impact than a small (smaller than K_A to be more precise) market buy order. A negative (log) convexity, or a concave limit order book, implies the opposite. We refer to a side with a (log) convexity of zero as linear.

The following proposition derives the j^{th} non-central moment of the return distribution, $E[r^j]$:

Proposition 1. *The j^{th} non-central moment of the return distribution, $E[r^j]$, can be written as follows:*

$$\begin{aligned} E[r^j] \quad &= \quad (2S_{avg}\sigma)^j \Bigg[\Bigg(\frac{I}{I+1} \frac{\tilde{\tilde{D}}_A}{\tilde{K}_A + (\tilde{\tilde{D}}_A - \tilde{K}_A)C_A} \Bigg)^j \Bigg(M_j(\tilde{\mu}, 1, 0, \tilde{K}_A)(\Phi(\tilde{K}_A, \tilde{\mu}, 1) - \Phi(0, \tilde{\mu}, 1)) \\ &+ \quad M_j(C_A\tilde{\mu} + (1 - C_A)\tilde{K}_A, C_A, \tilde{K}_A, +\infty)(1 - \Phi(\tilde{K}_A, C_A\tilde{\mu} + (1 - C_A)\tilde{K}_A, C_A)) \Bigg) \\ &+ \quad \Bigg(\frac{1}{I+1} \frac{\tilde{\tilde{D}}_B}{\tilde{K}_B + (\tilde{\tilde{D}}_B - \tilde{K}_B)C_B} \Bigg)^j \Bigg(M_j(\tilde{\mu}, 1, -\tilde{K}_B, 0)(\Phi(0, \tilde{\mu}, 1) - \Phi(-\tilde{K}_B, \tilde{\mu}, 1)) \\ &+ \quad M_j(C_B\tilde{\mu} + (C_B - 1)\tilde{K}_B, C_B, -\infty, -\tilde{K}_B)\Phi(-\tilde{K}_B, C_B\tilde{\mu} + (C_B - 1)\tilde{K}_B, C_B) \Bigg) \Bigg] \end{aligned} \tag{7}$$

where $\tilde{\mu}$, $\tilde{K}_A$, $\tilde{K}_B$, $\tilde{\tilde{D}}_A$ and $\tilde{\tilde{D}}_B$ are the corresponding variables normalized by dividing them by the market order standard deviation σ, i.e. $\tilde{\mu} = \mu/\sigma$, $\tilde{K}_A = K_A/\sigma$, $\tilde{K}_B = K_B/\sigma$, $\tilde{\tilde{D}}_A = Q_A/\sigma$ and $\tilde{\tilde{D}}_B = Q_B/\sigma$. $M_j(\nu, \omega, l, u)$ denotes the j^{th} non-central moment of a truncated normal with a mean ν and standard deviation ω and lower and upper truncation points of l and u, respectively. $M_j(\nu, \omega, l, u)$ can be computed numerically based on the recursive *formula described in the appendix.* $\Phi(x, \nu, \omega)$ *denotes the cumulative density function (CDF) of a normal distribution with mean* ν *and* ω *evaluated at* x*.*

Proposition 1 shows that the non-central moments of the return distribution can be expressed analytically as functions of the market order distribution parameters, μ and σ, and the LOB shape parameters, S_{avg}, K_A, K_B, $\bar{D}_A$, $\bar{D}_B$, I, C_A and C_B. This in turn implies that the central moments of the return distribution can also be expressed as analytical functions of these variables. Proposition 1 also shows that the average slope of the limit order book, S_{avg}, is a scaling factor for non-central moments. For example, if the average LOB slope is twice higher, the first four non-central moments of the return distribution are higher by factors of 2, 4, 8, and 16, respectively. This finding in turn implies that the third and fourth central moments, i.e. skewness and kurtosis, respectively, given their definitions, do not depend on the average slope of the ask and bid sides. Second, the standard deviation of the market order distribution is not only a scale factor like the average LOB slope but also a variable against which all other quantity variables are measured. In other words, the non-central moments depend on the values of quantity variables (such as the mean of the market order distribution or the cumulative depths in the ask and bid sides) relative to the standard deviation of the market order distribution but not on their absolute values.

3 Implications

In this section, we analyze the implications of our reduced-form price formation model on the effects of the market order distribution and LOB shape on the first four central return moments and return distribution. We start with the former before turning our attention to the latter, which is our main interest.

3.1 The Effects of Market Order Distribution

The following lemma characterizes the return distribution for the special case of a balanced and linear limit order book and provides the implications of our reduced-form price formation model for the unconditional effects of market order distribution on return distributions.

Lemma 1. *Assume that the limit order book is balanced and linear, i.e.* $S_{A,low} = S_{A,high} = S_{B,low} = S_{B,high} = S$

and, thus, $I = 1$ *and* $C_A = C_B = 1$*. The return can be expressed as* $r = SQ$ *and are normally distributed with a mean of* $S\mu$ *and a variance of* $(S\sigma)^2$ *and, thus, have zero skewness and (excess) kurtosis.*

In Lemma 1, our assumption that the market orders are normally distributed implies that returns are normally distributed when the limit order book is balanced and linear. One can easily generalize the result in Lemma 1 under a more general assumption for the distribution of market orders. Specifically, if market orders have a probability density function f, i.e. $Q \sim f(x)$, returns would have a distribution function given by $1/Sf(x/S)$. In other words, the return distribution reflects (up to a multiplicative factor which is the reciprocal of the limit order book slope) the market order distribution when the limit order book is balanced and linear. This has important implications for our empirical analysis as it shows the importance of controlling for the distribution of market orders when analyzing the effect of LOB shape parameters on the return distribution.

Lemma 1 has several other intuitive implications. First, the expected returns are positive (negative) if there is expected buying (selling) pressure from the market orders, i.e. the MO imbalance is positive ($\mu > 0$). This is intuitive since one expects prices to increase, and thus positive expected returns, when there are more market buy than market sell orders. Second, returns are more volatile when market orders are more volatile. This is due to the fact that larger market orders are more likely when the distribution of market orders are more volatile. This in turn makes the returns more volatile given that the price impact of larger market orders is higher.

3.2 The Effects of Primary LOB Shape Variables

In this section, we discuss the effects of LOB shape on the first four central return moments and the return distribution. We distinguish between two sets of LOB shape variables. We consider LOB imbalance and LOB convexity as primary LOB shape variables because as we will discuss they are important determinants of the return distribution. The total quantity available in the ask and bid sides or their ratio and the average slope of the limit order book do not have important direct effects on the return distribution and are thus considered as secondary LOB shape variables.

3.2.1 The Effects of LOB Imbalance

As we discuss in the introduction, the LOB imbalance is one of the important determinants of the LOB shape and, thus, the return distribution. In this section, we analyze the implications of our reduced-form price formation model on the effects of LOB imbalance on the return distribution. To do this, we first set the mean of the market order distribution to zero so that there is no market order imbalance. As we will discuss in more detail in Section 3.2.4, the MO imbalance can significantly alter certain effects of the LOB imbalance on the return distribution. Setting MO imbalance to zero here allows us to analyze the unconditional effect of LOB imbalance on the return distribution. Second, we use the medians (over all stock-day pairs in our sample) of the empirical counterparts of all other model parameters to calibrate them. In addition to the mean of the market order distribution, the only other variable that is not calibrated to the median of its sample counterpart is the LOB convexity, the other primary LOB shape parameter. To be more precise, we consider three different values of (log) LOB convexity to analyze whether the effect of the LOB imbalance on the return distribution changes as a function of LOB convexity. For the sake of intuition, we assume in this section that the ask and bid side convexities are always equal to each other, i.e. $c_A = c_B = c$[2] and can take on three different values: $c = -0.4226, 0, 0.4226$ where 0.4226 is the average of the sample standard deviations (std) of c_A and c_B, i.e. $(std(c_A)+std(c_B))/2 = 0.4226$. Table 1 summarizes the calibrated values of the theoretical variables used to obtain the figures in Section 3.

[Table 1 about here.]

Figure 2 presents how the first four central moments of the return distribution change as a function of (log) LOB imbalance ($i = \log(I)$) under the assumptions discussed above. We consider LOB imbalance values between minus and plus one times the standard deviation of its sample counterpart (0.7044), i.e. the limits of the x-axis in Figure 2 are -0.7044 and 0.7044. This choice allows us to compare the theoretical magnitudes of the effect of LOB imbalance on return moments to those from our empirical analysis where we present the economic significance of each variable defined as its regression coefficient estimate times its sample standard deviation.

[Figure 2 about here.]

[2] We analyze the effect of asymmetric ask and bid side convexities in Section 3.2.2.

We start our discussion with mean presented in panel (a). The return mean is a monotonically increasing function of LOB imbalance regardless of the LOB convexity. Furthermore, the mean is positive when LOB imbalance is positive and negative otherwise under the assumption that the MO imbalance is zero. In other words, the sign of the LOB imbalance determines the sign of the expected returns when there is no MO imbalance. These results follow from the fact that a positive LOB imbalance implies that the bid side is relatively more liquid than the ask side, indicating a relative buying pressure from the limit orders in the book. In the absence of any MO imbalance, a buying pressure from the limit orders in the book implies higher expected prices and thus higher expected returns. Figure 2 also shows that the magnitude of the effect of LOB imbalance on return mean increases as LOB convexity decreases.

Panel (b) presents how the return variance changes as a function of the LOB imbalance. The return variance is a U-shaped function of LOB imbalance and is at its minimum when the limit order book is balanced, i.e. the LOB imbalance is zero. Furthermore, the U-shaped effect of the LOB imbalance on the return variance is relatively more pronounced when the LOB convexity is lower. The economic magnitude of this effect is relatively small under the chosen calibration of other model parameters. For example, the return variance is about 0.00101 (corresponding to a daily volatility of 3.16%) when the LOB imbalance is set to minus one times its standard deviation while it is about 0.00098 (corresponding to a daily volatility of 3.14%) at its minimum when the LOB imbalance is set to zero. This corresponds to a 3.1% change in the return variance in response to a one standard deviation change in the LOB imbalance.

Panel (c) presents the effect of LOB imbalance on the return skewness. Similar to the return mean, in the absence of any MO imbalance, the skewness is a monotonically increasing function of the LOB imbalance and its sign is determined by the sign of the LOB imbalance regardless of the LOB convexity. In other words, when there is no MO imbalance, returns are positively skewed if and only if LOB imbalance is positive. This implication confirms our intuition based on the simple intuitive framework discussed in the introduction. A positive LOB imbalance implies a buying pressure from limit orders and, thus, results in a positively skewed return distribution in the absence of any buying or selling pressure from market orders. Furthermore, this positive effect of LOB imbalance on return skewness is amplified when LOB convexity is high.

Finally, panel (d) presents the effect of LOB imbalance on the return kurtosis. Our model predicts that any LOB imbalance results in excess kurtosis, with a more imbalanced (positive or negative) LOB resulting in higher excess kurtosis. In other words, the kurtosis increases monotonically starting from a value of three - corresponding to a balanced LOB - as we consider more positive or negative LOB imbalances. Our model also predicts that the U-shape of this effect of LOB imbalance on return kurtosis is more pronounced when LOB convexity is higher. On the other hand, the effect of LOB imbalance on the return kurtosis is completely negligible when LOB convexity is low, i.e. set to minus one times the sample standard deviation of its empirical counterpart.

Although the first four central moments reveal important information about the effect of the LOB imbalance on the shape of the distribution, they do not provide a full picture. To this end, we analyze how the quantiles of the return distribution change as a function of the LOB imbalance. Given that we do not have closed form expressions for the quantiles of the return distribution unlike the first four central moments, we do this by simulating 1,000,000 returns based on Equation 4 under the same assumptions used above to obtain our results for the first four central moments.[3]

Figure 3 presents the inverse cumulative density function (CDF) of simulated returns, i.e. the quantiles from 0.01 to 0.99 with increments of 0.01, for low, zero and high values of LOB imbalance under the assumption that the LOB convexity is low (panel (a)), zero (panel (b)) and high (panel (c)) in addition to the same assumptions used above. Recall that the returns are normally distributed with a zero mean when MO imbalance, LOB imbalance and LOB convexity are zero. Thus, the quantiles of the simulated returns when LOB imbalance is zero correspond to the quantiles of the normal distribution (the solid line in panel (b) of Figure 3).

[Figure 3 about here.]

All quantiles, with the exception of the median, are higher (lower) when the LOB imbalance is high (low) compared to when the LOB imbalance is zero. This is true regardless of the LOB convexity although the magnitude of the effect of the LOB imbalance on return quantiles depend on the LOB convexity, which we will discuss later. Furthermore, recall that the mean and skewness are increasing functions of LOB imbalance while variance and kurtosis are U-shaped functions. These effects of LOB imbalance on the first four moments imply a shift to the right for the right part (quantiles above the median) of the return distribution. On the other hand, as we consider more negative values of the LOB imbalance, the effects of LOB imbalance on the mean and skewness imply a shift to the right for the left

[3] We also considered approximating the return distribution based on a Gram-Charlier type expansion using the first four central moments and the relations based on this approximation are very similar to those based on simulations.

part (quantiles below the median) of the return distribution while its effects on the variance and kurtosis imply a shift to the left. Our results on the left part of the return distribution suggest that the effects of LOB imbalance on mean and skewness dominate its effects on variance and kurtosis. This is an important finding confirming once again that the LOB imbalance is an important determinant of the return mean and skewness while it might not play important role in determining the return variance and kurtosis.

3.2.2 The Effects of LOB Convexity

In this section, we analyze the implications of our reduced-form price formation model on the effects of LOB convexity, the other primary LOB shape parameter, on the return distribution. Similar to our analysis for the LOB imbalance, we set the mean of the market order distribution to zero so that there is no MO imbalance and consider three different values of (log) LOB imbalance, $i = -0.7044, 0, +0.7044$ to analyze whether the effect of the LOB convexity on the return distribution changes as a function of LOB imbalance. We also start our analysis by considering the effect of symmetric convexity where we set the ask and bid side convexities equal to each other, i.e. $c_A = c_B = c$. Figure 4 presents how the first four central moments of the return distribution changes as a function of (log) LOB convexity ($i = \log(c)$) under these assumptions. We consider LOB convexity values between minus and plus one times the standard deviation of its sample counterpart (0.4226), i.e. the limits of the x-axis in Figure 4 are -0.4226 and 0.4226.

[Figure 4 about here.]

We start with the return mean. The effect of LOB convexity on the return mean depends crucially on the LOB imbalance. First of all, as discussed above, in the absence of any MO imbalance the sign of the return mean is completely determined by the sign of the LOB imbalance independent of the LOB convexity. Second, the return mean is a decreasing function of LOB convexity when there is positive LOB imbalance and increasing function when there is negative LOB imbalance. Furthermore, the returns have zero mean regardless of LOB convexity when there is no LOB imbalance. Third, a higher (in magnitude) LOB imbalance amplifies the effect of LOB convexity on the mean of the return distribution. This effect is more pronounced for negative values of LOB convexity.

We now turn our attention to the return variance. Regardless of LOB imbalance, the return variance is a monotonically decreasing function of LOB convexity. Furthermore, as expected from our discussion in Section 3.2.1, the effect of LOB convexity on the return variance does not change significantly for different values of LOB imbalance. As the limit order book becomes more convex, the price impact of small market orders decreases because these small market orders are executed against the lower levels of the book, which are more liquid than the higher levels when the book is more convex.

We now consider the return skewness. Similar to the return mean, the effect of LOB convexity on the return skewness depends crucially on the LOB imbalance. In the absence of any MO imbalance, the sign of the return skewness is completely determined by the sign of the LOB imbalance independent of the LOB convexity. However, differently from the return mean, the return skewness is an increasing function of LOB convexity when there is positive LOB imbalance and a decreasing function when there is negative LOB imbalance. Furthermore, the returns have zero skewness regardless of LOB convexity when there is no LOB imbalance. A higher (in magnitude) LOB imbalance amplifies the effect of LOB convexity on the skewness and this effect is more pronounced for positive values of convexity.

Finally, the return kurtosis is a monotonically increasing function of LOB convexity regardless of the LOB imbalance. Furthermore, in the absence of any MO and LOB imbalance, returns are leptokurtic when the book is convex and platykurtic when the book is concave, confirming our intuition based on the simple intuitive framework discussed in the introduction. As the book becomes more convex or equivalently less concave, the price impact of large market orders increases because they are executed against the higher levels of the book, which are less liquid than the lower levels when the book is more convex. This higher price impact of large market orders than creates fatter tails in the return distribution. The opposite intuition holds when the book is concave.

Figure 5 presents the inverse cumulative density function (CDF) of simulated returns, i.e. the quantiles from 0.01 to 0.99 with increments of 0.01, for low, zero and high values of LOB convexity under the assumption that the LOB imbalance is low (panel (a)), zero (panel (b)) and high (panel (c)) in addition to the same assumptions used above. Once again, the returns are normally distributed with a zero mean when MO imbalance, LOB imbalance and LOB convexity are zero. Thus, the quantiles of the simulated returns when LOB convexity is zero correspond to the quantiles of the normal distribution (the solid lines in panels (a), (b), and (c) of Figure 5).

[Figure 5 about here.]

Figure 5 shows that a high LOB convexity results in the quantiles below (above) the median to be higher (lower) than the corresponding quantiles when the LOB convexity is zero. A low LOB convexity results in the opposite, i.e. quantiles below (above) the median are lower (higher) than the corresponding quantiles when LOB convexity is zero. These results hold regardless of the LOB imbalance. In other words, all quantiles below (above) the median are increasing (decreasing) functions of the LOB convexity, regardless of th LOB imbalance. Recall from Figure 4 that the variance is a decreasing function of the LOB convexity and kurtosis is an increasing function. Thus, a higher convexity results in a tighter return distribution through its effect on the return variance while it results in a wider return distribution through its effect on the return kurtosis. Figure 5 shows that the former of these two opposing effects dominate and a higher convexity results in a tighter return distribution. That said, there is a kink in the inverse CDF of simulated returns when LOB convexity is different from zero. This kink corresponds to the cutoff point where the slope changes between the lower and higher levels and drives the effect of LOB convexity on the tails of the return distribution. To see this effect more clearly, Figure 6 presents the inverse CDF of the simulated returns and a normal distribution with zero mean and variance corresponding to the variance of simulated returns when the LOB convexity is high or low.

[Figure 6 about here.]

Panel (a) of Figure 6 shows that the tails (quantiles below 5% and above 95%) of the return distribution are further away from zero that those of a normal distribution with the same variance when the LOB convexity is positive, i.e. the LOB is convex. This in turn implies that a higher LOB convexity results in fatter tails and higher kurtosis than a normal distribution once we control for its effect on the return variance. The opposite holds when the LOB convexity is negative, i.e. the LOB is concave. This result confirms our intuition from the introduction that the LOB convexity determines the tails and thus the kurtosis of the return distribution.

We now consider the effect of asymmetric convexity on the first four central moments of the return distribution. We do this by setting the (log) ask side convexity to zero, i.e. the ask side is linear, and allowing the bid side convexity to change in the range considered above and vice versa. Figure 7 and 8 present how the first four central moments of the return distribution changes as a function of the bid- and ask-side convexities, respectively. First of all, similar to the effect of symmetric convexity, the return variance is a decreasing function and the return kurtosis is an increasing function of both bid- and ask-side convexities regardless of the LOB imbalance. The effects of bid- and ask-side convexities on the return mean and skewness are slightly different than what we observe when we consider symmetric convexity. Here, we discuss the results and the intuition for the effect of the bid-side convexity on the return mean and skewness. They are very similar for the ask-side convexity with opposite signs. The return mean is an increasing function of the bid-side convexity regardless of the LOB imbalance considered. Everything else equal, the slopes of the lower levels of the bid side decreases and that of its higher levels increases as bid-side convexity increases. In other words, when we hold everything else constant, the buying pressure from the lower levels of the bid side increases as bid-side convexity increases. This in turn results in a higher return mean. The effect of bid-side convexity on skewness is more nuanced. The skewness is an increasing function of the bid-side convexity when the LOB imbalance is zero or high. On the other hand, it is a decreasing function when the LOB imbalance is low. The LOB imbalance plays a role on how fast buying pressure from the lower levels of the bid side increases as bid-side convexity increases. Buying pressure from the lower levels of the bid side increases much slower when LOB imbalance is low compared to when LOB is zero or high. This in turn results in a lower skewness as bid-side convexity increases when LOB imbalance is low.

[Figure 7 about here.]

[Figure 8 about here.]

Finally, Figure 9 presents the effects of bid- (panels (a)-(c)) and ask-convexities (panels (d)-(f)) on the quantiles of the return distribution. Similar to Figure 5, panels (a) and (d) present the inverse CDF of simulated returns for low, zero and high values of LOB bid- and ask-side convexities, respectively, under the same assumptions as above. Once again, the quantiles of the simulated returns correspond to the quantiles of the normal distribution (the solid lines in panels (a) and (d)) because the returns are normally distributed with a zero mean when MO imbalance, LOB imbalance, ask- and bid-side convexities are all zero. It is easy to see from panels (a) and (d) that the ask-side convexity affects the right, but not the left, part of the return distribution while the opposite holds for the bid-side convexity. Furthermore, the right (left) part of the return distribution shifts to the left (right) as ask-side (bid-side) convexity increases similar to what is observed for symmetric convexity in Figure 5. However, unlike the effect of symmetric convexity in Figure 5, the effects of ask- or bid-side convexity on the return quantiles through their effects on the first four central moments are more nuanced. To provide some intuition on these effects, we consider the ask-side convexity when LOB imbalance is zero. Everything else equal, an increase in the ask-side convexity implies a shift to the left of the return distribution through its effects on the return mean and skewness presented in Figure 8. On the other hand, a higher ask-side convexity results in a tighter return distribution through its effect on the return variance while it results in a wider return distribution through its effect on the return kurtosis. Our results in panel (d) show that the effect of the ask-side convexity on quantiles higher than the median through its effects on return mean, variance and skewness dominate that through its effect on the return kurtosis and we observe a shift to the left for the right part of the return distribution when ask-side convexity is high. These opposing effects cancel each other perfectly for the left part of the return distribution which remains unchanged as ask-side convexity changes. Once we control for the effect of the ask-side convexity on the return quantiles through its effect on the return variance as in Figure 6, a higher ask-convexity results in a return distribution whose quantiles between 1% and 99% are all less than (with the exception of the median which is equal) to the quantiles of the normal distribution the same variance as shown in panel (e) of Figure 9. The opposite holds when ask-side convexity is low as shown in panel (f). The results for the bid-side are the same but with opposite signs and are presented in Panels (b) and (c) of Figure 9.

[Figure 9 about here.]

3.2.3 The Effects of Secondary LOB Shape Variables

In this section, we briefly discuss the effects of the secondary LOB shape variables on the return moments and distribution. We start our discussion with the average slope of the limit order book. Recall from the discussion following Proposition 1 that the average slope of the limit order book, S_{avg}, is a scaling factor for non-central moments, which implies that if the average LOB slope is twice higher, the first four non-central moments of the return distribution are higher by factors of 2, 4, 8, and 16, respectively. This in turn implies the following for its effect on the return distribution: (1) The return mean is zero regardless of the average slope of the limit order book when both MO and LOB imbalances are zero. Otherwise, the return mean can be an increasing or a decreasing function of the average slope of the limit order book depending on the signs and magnitudes of the MO and LOB imbalances. (2) The return variance is an increasing function. (3) The return skewness and kurtosis do not depend on the average slope of the limit order book. (4) The effect of the average slope on the return distribution depends closely on the signs and magnitudes of the MO and LOB imbalances.

We now consider the effects of the quantities available in the bid and ask sides. When we change the quantities available in the bid and ask sides symmetrically, the effects of the total quantity available in the limit order book on the return moments and distribution are similar to those for the average slope: (1) The return mean is zero regardless of the total quantity available in the limit order book when both MO and LOB imbalances are zero. Otherwise, the return mean can be an increasing or a decreasing function of the total quantity available in the limit order book depending on the signs and magnitudes of the MO and LOB imbalances. (2) The return variance is an increasing function of the total quantity available if and only if LOB convexity is positive regardless of the signs and magnitudes of the MO and LOB imbalances. (3) The return skewness and kurtosis do not depend on the average slope of the limit order book. (4) The effect of the average slope on the return distribution depends closely on the signs and magnitudes of the MO and LOB imbalances.

Finally, when we change the quantity available in the bid side while keeping the quantity available in the ask side constant, or vice versa, our model predicts the following: (1) If the LOB convexity is zero, the total quantity available in the ask or the bid side does not affect the return moments or distribution regardless of the MO or LOB imbalances. (2) Otherwise, the return moments can be increasing or decreasing functions of the total quantity available in the ask or the bid side depending on other determinants.

3.2.4 The Conditional Effects of Primary LOB Shape Variables

So far, we have analyzed the effects of the primary LOB shape parameters on the return distribution, while keeping all other model parameters constant at their calibrated values. In this section, we discuss how the effects of these primary LOB shape variables change as a function of the parameters of the market order distribution and the secondary LOB shape parameters.

We start with the effect of MO imbalance, which was assumed zero in our discussions so far. Figure 10 presents the effect of LOB imbalance on the first four central moments for low, zero and high values for the MO imbalance where low and high values are minus and plus one standard deviation of the market order distribution. We set the LOB convexity to zero while all other parameters are calibrated to the values presented in Table 1. The return mean and skewness are increasing functions of the LOB imbalance regardless of the MO imbalance. In other words, the sign of the effect of the LOB imbalance on the return mean and skewness does not depend on the MO imbalance. That said, MO imbalance affects the magnitudes of these effects of LOB imbalance on the return mean and skewness. For example, the effect of the LOB imbalance on the return skewness is smaller in magnitude when the MO imbalance is not zero. On the other hand, MO imbalance significantly changes the effect of LOB imbalance on the return variance and kurtosis. When MO imbalance is different from zero, the return variance and kurtosis are no longer U-shaped functions of LOB imbalance. When MO imbalance is low, the return variance is a monotonically decreasing function of the LOB imbalance and the return kurtosis is a monotonically increasing function. The opposite holds when MO imbalance is high.

[Figure 10 about here.]

Figure 11 presents the effect of LOB convexity on the first four central moments for zero LOB imbalance and low, zero and high MO imbalance. First of all, the mean is a decreasing function of LOB convexity when MO imbalance is positive and an increasing function when it is negative. The opposite holds for skewness. It is easy to see from these results that in the absence of LOB imbalance the MO imbalance plays its role in determining the effect of the LOB convexity on the return mean and skewness. Related to this point, recall that the mean and skewness do not depend on LOB convexity when the LOB imbalance is zero. The results in panels (a) and (c) show that even in the absence of LOB imbalance, the mean and skewness can be increasing or decreasing functions of LOB convexity depending on the sign of the MO imbalance. Of course, these results also suggest that the effect of LOB imbalance on the relation between the mean and skewness and LOB convexity can be undone or amplified by MO imbalance. Panel (b) shows that the MO imbalance does not significantly alter the effect of the LOB convexity on the variance. To be more specific, the variance is a decreasing function of LOB convexity, regardless of the MO imbalance. Panel (d) of 11 shows that the kurtosis is an increasing function of the LOB convexity when MO imbalance is zero, as already observed in panel (d) of Figure 4. More importantly, panel (d) of 11 shows that MO imbalance changes the effect of LOB convexity on the kurtosis, unlike the LOB imbalance. When MO imbalance is different from zero, the kurtosis first decreases and reaches three when LOB convexity is zero and then starts to increase for higher positive values of LOB convexity. Finally, panels (b) and (d) show that the MO imbalance changes the effect of LOB convexity on the return variance and kurtosis symmetrically. In other words, the effects of LOB convexity on the return variance and kurtosis are identical when MO imbalance is positive or negative with the same magnitude. The dashed and dotted lines in panels (b) and (d) of Figure 11, corresponding to the cases where MO imbalance set to plus and minus one times its sample standard deviation respectively, overlap.

[Figure 11 about here.]

Figure 12 presents how the effects of LOB imbalance and convexity on return quantiles change with MO imbalance. First of all, Lemma 1 implies that everything else equal an increase in the MO imbalance results in a right shift for the return distribution without changing its shape. This location shift changes the reference point for the effects of LOB imbalance and convexity. To be more precise, recall from Figures 3 and 5 that the reference point is the 50% quantile, i.e. the median, when MO imbalance is zero. For example, Figure 5 shows that a high LOB convexity results in the quantiles below (above) the median to be higher (lower) than the corresponding quantiles when the LOB convexity is zero. Under our calibration, this reference point becomes between 15% and 16% quantiles when MO imbalance is high and between 84% and 85% quantiles when MO imbalance is low. To put it differently, when MO imbalance is high, a high LOB convexity results in the quantiles below (above) 15% to be higher (lower) than the quantiles of the corresponding normal distribution. When MO imbalance is low, a high LOB convexity results in the quantiles below (above) 84% to be higher (lower) than the quantiles of the corresponding normal distribution.

[Figure 12 about here.]

Next, we consider the effect of the market order standard deviation. Recall our discussion from above that the standard deviation of the market order distribution is a scale factor. This in turn implies that (1) the effects of LOB shape variables on the return mean and variance are simply amplified when the standard deviation of the market order distribution is high; (2) the effect of LOB shape variables on skewness and kurtosis do not depend on the standard deviation of the market order distribution. In other words, our theory predicts that ceteris paribus the standard deviation of the market order should not significantly alter the effect of LOB shape variables on return moments. That said, the standard deviation of the market orders is also a reference variable against which all other variables are measured. For example, the definition of higher and lower levels of the limit order book in our model is based on the market order standard deviation. In our baseline analysis, we define any market order one standard deviation away from zero as a large market order hitting the higher levels of the book. This in turn implies that the effect of LOB convexity on variance, skewness but especially on kurtosis to be more pronounced when the standard deviation of the market order distribution is high. This is due to the fact that the probability of large market orders hitting the higher levels of the order book is higher when the market orders are more volatile.

Finally, we briefly discuss how the effects of primary LOB shape variables, i.e. LOB imbalance and convexity, on the return distribution changes with secondary LOB shape variables such as the total quantities available in the bid and ask sides as well as the average slope of the limit order book. We start our discussion with the total quantities available in the bid and ask sides and their ratio. We first assume that the total quantities available in the bid and ask sides are equal and consider three values, low, median, high for this variable. We then assume that the total quantity available in the ask side is equal to that used in our main calibration presented in Table 1 and consider three values, low, median, high, for the total quantity available in the bid side. We find that neither the quantities nor their ratio does not alter the effect of LOB imbalance on the return distribution. The effect of the total quantity available and the ratio on how return moments change with LOB convexity is a little bit more nuanced. When both LOB and MO imbalances are zero, the LOB convexity does not have any effect on the return mean and skewness. The total quantity available does not change this fact. Furthermore, the total quantity available does not significantly alter the effects of LOB convexity on the return variance and kurtosis. Consequently, the quantity does not significantly alter the effect of LOB convexity on the return distribution. When we consider the quantity ratio, we find that both the return mean and skewness are decreasing (increasing) functions of the LOB convexity if there is more (less) total quantity available on the ask side. On the other hand, the quantity ratio does not alter the effect of LOB convexity on the return variance and kurtosis. When combined, we do not find the quantity ratio alters the effect of LOB convexity on the return quantiles. Turning our attention to the average slope of the limit order book, S_{avg}, recall from our discussion following Proposition 1 that the return skewness and kurtosis do not depend on the average slope of the limit order book. This in turn implies that the average slope does not alter at all the effects of LOB imbalance or LOB convexity on return skewness and kurtosis but amplifies their effects on the return mean and variance. Specifically, the effects of LOB imbalance and LOB convexity discussed above become more pronounced when the average slope of the limit order book is higher, i.e. when there is less liquidity in the limit order book.

4 Data and Variable Description

To empirical analyze the empirical relation between the shapes of the limit order book, market order and return distribution, we need empirical proxies for the variables in our reduced-form model of price formation. We use data from TAQ database to obtain empirical proxies for the moments of the return and market order distributions. As mentioned above, we need to choose a time interval to compute these moments. In our empirical analysis, we opt for aggregating the data at the daily frequency mainly to minimize the sensitivity of higher moments to noisy intraday data.[4]

Following the literature on realized variance, we compute realized non-central moments based on 5-minute (log) midquote returns in a given trading day. Specifically, we compute return for the m^{th} 5-minute interval during the trading day t for stock i, r_{itm}, as the difference between the (log) prices at the end m^{th} and $(m-1)^{th}$ 5-minute intervals for $m = 1, \ldots, 78$. We compute j^{th} non-central moment of returns $E[r_{it}^j] = \sum_{m=1}^{78} r_{itm}^j$ for $j = 2, 3, 4$. We obtain the corresponding central moments as per their usual definitions based on non-central moments assuming that 5-minute returns have zero mean. We use the daily return from Center for Research in Security Prices (CRSP) as our proxy for the first central moment, i.e. the daily expected return. Finally, for a given stock-day pair, we classify each trade as a buy (+1), a sell (-1) or unclassified (0) based on the Lee and Ready algorithm. We then estimate the first four central moments of the signed volume in shares for each stock-day pair and use them as our proxies for the first four central moments of the market order distribution.

We now turn our attention to the limit order book shape parameters. Our limit order book (LOB) data is from the Thomson Reuters Tick History (TRTH) Market Depth (MD) New York Stock Exchange (NYSE) files. This dataset includes the snapshots of the first ten levels of the limit order book, i.e. the prices and the number of shares (quantity)

[4]Chordia, Roll, and Subrahmanyam (2002) study the relationship among liquidity, order imbalances, and returns and argue that daily time intervals balance the trade-off between reducing problems related to very-high-frequency data and capturing short-term effects among the variables of interest.

available at the first ten price level for bid and ask sides. Each snapshot is identified by a Reuters Instrument Code (RIC) and a millisecond timestamp. As soon as there is a change in price or quantity at any level of the book due to a newly placed, withdrawn, or executed order, a new snapshot of the entire book is created. We clean this data and match it to TAQ and CRSP using the procedures described in Cenesizoglu and Grass (2018). Briefly, we require that each LOB snapshot includes price and volume data for all levels and that prices increase monotonically throughout the book. We further delete observations with Level 1 (Level 10) bid-ask spreads above 25% (250%), midprices below $1 or above $1000. Finally, deleting all firm-days with fewer than 100 snapshots in the TRTH data or fewer than 100 trades in the matched TAQ dataset as well as observations with incomplete CRSP data results in LOB snapshots for 2,050 stocks over 2,736 trading days and 3.58 million stock-day observations between 2002 and 2012.[5] Our sample period begins on January 2002 – when the NYSE began making level-two LOB data available to market participants outside the trading floor and TRTH MD data begin – and ends in December 2012.[6] We focus on this period for two main reasons. First is the sheer size of this data. For example, our initial data before cleaning includes 52.44 billion LOB snapshots up to 20 price-quantity values for both bid and ask sides. The size increases almost exponentially during and after the Global Financial Crisis. The second reason is that this sample period includes two important periods: Securities and Exchange Commission's (SEC) temporary suspension of short-sale price tests for a set of pilot securities between May 2005 and August 2007, which we use to establish a causal link, and the Global Financial Crisis between 2007 and 2009.

For a given LOB snapshot s with a timestamp t_s, we compute the slope of the bid (ask) between levels l_1 and l_2, $S^B_{l_1,l_2,s}$ ($S^A_{l_1,l_2,s}$), as follows:

$$S^B_{l_1,l_2,s} = -\frac{P^B_{l_2,s} - P^B_{l_1,s}}{D^B_{l_1+1,l_2,s}}, \tag{8}$$

$$S^A_{l_1,l_2,s} = \frac{P^A_{l_2,s} - P^A_{l_1,s}}{D^A_{l_1+1,l_2,s}} \tag{9}$$

where $P^B_{l,s}$ and $P^A_{l,s}$ are, respectively, bid and ask prices at level l and $D^B_{l_1+1,l_2,s}$ and $D^A_{l_1+1,l_2,s}$ are, respectively,

[5] Although the TRTH MD data are also available for stocks listed on NASDAQ, we follow Korajczyk and Sadka (2008) and restrict our analysis to stocks listed on the NYSE due to differences in trading mechanisms between the two exchanges.

[6] We exclude days for which the number of book snapshots per trading hour is less than 50% of the monthly average, namely January 28th, February 1st, October 31st and November 20th, 2002, January 21st and October 15th, 2003, January 13th, 2004, and May 4th, 2009. TRTH confirms technical problems related to data collection for these days.

cumulative quantity available between levels l_1 and l_2 in the LOB snapshot s for $l_1 = 1, \ldots, 9$ and $l_2 > l_1$. The slopes of the bid and ask sides between levels l_1 and l_2 at time t_s, $S^B_{l_1,l_2,s}$ and $S^A_{l_1,l_2,s}$, are defined as the change in the price relative to the cumulative quantity available between levels l_1 and l_2. Given that the prices in the bid side decrease as the level increases, the change in the bid price between levels l_1 and l_2 for $l_2 > l_1$ is always negative. Hence, we consider the negative of the price change so that bid side slope is always positive.

We then compute LOB imbalance with a cutoff level of l for snapshot s, $I_{l,s}$, as the ratio of the ask- and bid-sides slopes between levels 1 and l as follows:

$$I_{l,s} = \frac{S^A_{1,l,s}}{S^B_{1,l,s}} \tag{10}$$

Similarly, we compute ask and bid side convexity with a cutoff level of l for snapshot s, $C^A_{l,s}$ and $C^B_{l,s}$, as the ratio of the slope between levels l and 20 to the slope between levels 1 and l as follows:

$$C^A_{l,s} = \frac{S^A_{l,10,s}}{S^A_{1,l,s}} \tag{11}$$

$$C^B_{l,s} = \frac{S^B_{l,10,s}}{S^B_{1,l,s}} \tag{12}$$

Having computed these measures for each snapshot, we aggregate them to daily level for each stock-day pair to match with their corresponding daily return moments. Specifically, we consider a time-weighted aggregation where each LOB snapshot and any measure computed based on that snapshot gets the weight corresponding to the time that this snapshot was in effect, i.e. the time until the next LOB snapshot. This aggregation scheme puts more weight on snapshots that were in effect for a longer period. The computation is similar for all variables. The following equation illustrates this aggregation scheme mathematically for the bid-side slope of a given stock-day pair:

$$S^B_{l_1,l_2} = \sum_{s=1}^{S-1} S^B_{l_1,l_2,s} \frac{t_{s+1} - t_s}{t_S - t_1} \tag{13}$$

where S is the total number of LOB snapshot for that stock-day pair, t_s is the timestamp of the s^{th} LOB snapshot as before.

Several remarks are in order regarding our theoretical and empirical measures of slopes, imbalance and convexity.

First, in our theoretical framework, we define convexity of the limit order book based on the cumulative depth. To be more precise, we refer to levels with a cumulative depth less than a given amount (k_A for the ask side or k_B for the bid side) as low levels and those with a cumulative depth higher than these thresholds as higher levels. We then show in our theoretical framework that low and high levels could be defined by these threshold normalized by the standard deviation of market order distribution. Given that it is not always feasible to match the observed quantities in LOB to the same multiple of market order standard deviation for all stock-day pairs, our empirical measures of convexity as presented above are based on levels instead of quantities. To mitigate the effects of this discrepancy between the empirical and theoretical definitions of convexity, we consider three convexity measures with cutoff levels $l = 2, 3, 5$. For a given stock-day pair, we use the one based on the level where the cumulative quantity is closest to one market order standard deviation of that stock on the previous trading day. Specifically, for a given stock-day pair, we find the level that minimizes the absolute value of the difference between k times the convexity cutoff, $\tilde{K}$, and the average cumulative quantity available in the first l levels of bid and ask sides, both normalized by market order standard deviation of the previous trading day, i.e. $l^* = \arg\min_{l=1,2,\ldots,10} |(Q_{A,l}/\sigma + Q_{B,l}/\sigma)/2 - k \times \tilde{K}|$. If l^* is one, we use $c_{avg,2} = (c_2^A + c_2^B)/2$. If l^* is greater than one but less than or equal to three, we use $c_{avg,3} = (c_3^A + c_3^B)/2$. If l^* is greater than three, we use $c_{avg,5} = (c_5^A + c_5^B)/2$. Finally, as mentioned above, our theoretical model predicts that the instantaneous return distribution over the period s and $s + \epsilon$ is determined by the market orders received over the same period and executed against the limit order at time s. In other words, the shape of the limit order book should be considered with a lag. Our choice to aggregate the variables at a daily level implies that the return distribution on day t is determined by the market order distribution on day t and limit order book on the previous trading day $t - 1$. There is also underlying assumption that the limit order book is resilient and thus have some level of autocorrelation over trading days.

The following table provides summary statistics for different variables used in our empirical analysis.

[Table 2 about here.]

Daily returns exhibit the standard properties documented in the literature. They have a zero mean, a slightly positive skewness, a kurtosis higher than three and are slightly negatively autocorrelated. The variance of returns are highly positively autocorrelated while the skewness and kurtosis are only slightly positively autocorrelated. Market orders have a slightly positive mean and median but are not significantly different from zero. They also have an average autocorrelation of around 15%. The standard deviation of market orders have a mean of 1291 and a median of 559 and exhibits a positive autocorrelation. Market orders are also slightly positively skewed and have relatively high kurtosis. The variance and kurtosis of market order distribution both exhibit positive autocorrelation while their skewness has practically zero autocorrelation.

Regarding the limit order book variables, there is on average more quantity available on the ask than bid side. The bid side has on average a higher slope than the ask side. This is also reflected in the (log) LOB imbalance, which has a positive mean, suggesting that there is on average more buying than selling pressure in the limit order book. The three measures of average (log) convexity all have positive means, suggesting that the slope of the higher levels is on average greater than that of the lower levels. More importantly, all limit order book variables are highly autocorrelated, suggesting that the limit order book shape is highly persistent from one trading day to the next.

5 Empirical Results

In this section, we test the empirical implications of our reduced form price formation model. To this end, we first consider the following linear regression of a given central moment of the return distribution on contemporaneous values of all other central moments and the first four central moments of the market order distribution and lagged values of LOB shape parameters.

5.1 Unconditional Effects of LOB Imbalance and Convexity on Return Moments

$$\mathrm{M}_{j,it} = \alpha_i + \sum_{k\neq j} \beta_k^j M_{j,it} + \sum_{k=1}^{4} \gamma_k^j \mathrm{MO}_{k,it} + \delta_1 \log(Q_{it}^{avg}) + \delta_2 \log(S_{it}^{avg}) + \delta_3 \log(Imb_{it}) + \delta_4 \log(Conv_{it}^{avg}) + \varepsilon_{it} \quad (14)$$

where $\mathrm{M}_{j,it}$ is the j^{th} central moment of the return distribution and $\mathrm{MO}_{k,it}$ is the k^{th} central moment of the market order distribution. The LOB shape variables are as defined above. We estimate this specification with firm fixed effects and standard errors clustered at the firm level.

Several remarks are in order regarding the empirical specification. First, we include other moments of the return distribution as control variables. We are not interested in the relation between different return moments and hence do not present the corresponding estimates. Second, we use the lagged values of the LOB shape parameters. As mentioned above, the shape of the limit order book should be considered with a lag and our choice to aggregate the variables at a daily level implies that the return distribution on day t is determined by the market order distribution on day t and limit order book on the previous trading day $t-1$. This also allows us to avoid endogeneity problems in interpreting the effect of LOB shape parameters on the shape of the return distribution.

Table 3 presents the economic significance (coefficient estimate times one standard deviation of the corresponding independent variable) of market order distribution and LOB shape parameters. These numbers can be interpreted as the effect of a one standard deviation increase in the independent variable on the return moment of interest and can be directly compared to Figures 2 and 4 which also show the effect of a one standard deviation increase in certain variables.

[Table 3 about here.]

We start our discussion with the LOB imbalance and LOB convexity, our main variables of interest, before turning our attention to the other variables. In line with our predictions, the LOB imbalance has a significantly positive coefficient estimate suggesting that return mean is higher following days with higher than normal LOB imbalance. More precisely, a one standard deviation increase in the LOB imbalance increases return mean on the following day by two basis points. This is of relatively small economic importance given that mean return has a standard deviation of 0.0304 in our sample. Furthermore, LOB imbalance is economically the least important variable among the market order and limit order book variables considered. Although the sign of this effect is in line with our predictions, its magnitude is much smaller than what our theory predicts. Panel (a) of Figure 2 shows that a one standard deviation increase in LOB imbalance increases mean return by about 1%. This discrepancy in the empirical and theoretical magnitudes of this effect might be due to inherent noise in daily return and the control variables in our empirical specification, which are not in our theoretical framework, such as other return moments or higher moments of the market order distribution. Turning our attention to the LOB convexity, we find that it has a significantly positive effect on return mean. A one standard deviation increase in LOB convexity increases return mean on the following day by almost 3.5 basis points. It is not straightforward to say whether this finding is in line with our theoretical predictions because our model predicts that the effect of LOB convexity on return mean depends closely on the sign of the MO and LOB imbalances (see panels (a) of Figures 4 and 11). In the next section, we consider an empirical specification with interactions between LOB imbalance, MO imbalance and LOB convexity and analyze empirically whether the effect of LOB convexity on return mean conditional on LOB and MO imbalances are in line with our predictions.

We now discuss the effects of primary LOB shape variables on return variance, starting with LOB imbalance. We find that the LOB imbalance has a significantly positive effect on the return variance on the following day. A one standard deviation increase in the LOB imbalance increases the return variance by about 0.00026, which is economically important given that it corresponds to about 10% of the standard deviation of the return variance in our sample. Our theory predicts that the return variance is an increasing (decreasing) function of LOB imbalance when MO imbalance is positive (negative) as presented in Panel (b) of Figure 2. Hence, it is not straightforward to say whether this unconditional effect is in line with what our theory predicts. In the next section, we analyze whether the effect of LOB imbalance conditional on MO imbalance is in line with our predictions. That said, the magnitude of this effect is similar to what our theory predicts. Panel (b) of Figure 10) shows that when MO imbalance is high the return variance increases from about 0.00023 to about 0.00035, an increase of 0.00012 (compared to an increase of 0.00026 observed in the data) following a one standard deviation increase in the LOB imbalance. Turning our attention to LOB convexity, we find that a one standard deviation increase in the LOB convexity decreases return variance by 0.00042 in the following day. This effect is also economically important, corresponding to about 16% of the standard deviation of return variance in our sample. Furthermore, among the LOB and MO variables considered LOB convexity is the second most economically important variable for return variance. The sign and magnitude of this effect provide strong empirical evidence for our theoretical predictions. Recall that our theory predicts that return variance is a decreasing function of LOB convexity regardless of the sign and magnitudes of LOB imbalance (Panel (b) of Figure 4) and MO imbalance (Panel (b) of Figure 11). The magnitude of this effect in our theory depends on both LOB and MO imbalances. For example, in the case of zero MO and LOB imbalances, the return variance in our model decreases from about 0.00023 to about 0.000008 following an increase of LOB convexity from zero to one standard deviation. This is comparable to the empirical magnitude of this effect (0.00042) observed in the data.

Regarding return skewness, we find that it increases by 0.01340 on the day following a one standard deviation increase in LOB imbalance. The sign of this effect is in line with our theory, which predicts that return skewness is an increasing function of LOB imbalance regardless of MO imbalance and LOB convexity. The economic magnitude of this effect is relatively small considering the fact that the standard deviation of return skewness in our sample is 1.2313. Furthermore, the empirical magnitude of this effect is also lower than what our model predicts. For example, in the case of zero LOB convexity and MO imbalance (the solid line in panel (c) Figure 2), the return skewness in our model increases from zero to about 0.75 when LOB imbalance increases from zero to one standard deviation. This difference in magnitudes might be due to the highly noisy nature of daily skewness estimates and the control variables in our empirical specification which do not exist in our theoretical model. All this said, we should emphasize the fact that the LOB imbalance is economically the second most important variable among the MO and LOB variables considered. This is in line with our theory which predicts that the LOB imbalance should be one of the most important determinants of the return skewness. Turning our attention to LOB convexity, we find that it does not have a significant effect on return skewness. Our theory predicts that the effect of LOB convexity on return skewness depends crucially on MO and LOB imbalances. It is then probably not surprising to find an insignificant unconditional effect of LOB convexity on return skewness.

Our empirical findings on the effects of LOB imbalance and convexity on return kurtosis can be summarized as follows. LOB imbalance has a significantly negative effect on the return kurtosis. Our theory predicts that regardless of LOB convexity return kurtosis is an increasing (decreasing) function of LOB imbalance when MO imbalance is negative (positive) and a U-shaped function of LOB imbalance when there is no MO imbalance (see panels (d) of Figures 10 and 2, respectively). This finding is in line with the predictions of our model when MO imbalance is positive, which it is on average in our sample. Regarding the economic magnitude of this effect, a one standard increase in LOB imbalance decreases return kurtosis by 0.13162, which is relatively small considering that return kurtosis has a standard deviation of 5.4989 in our sample. That said, LOB imbalance is economically the third most important determinant of return kurtosis and the magnitude of this effect is comparable to that in our theory. When MO imbalance is positive, return kurtosis decreases from 3 to about 2.7 following a one standard deviation increase in LOB imbalance (see dotted line in panel (d) of Figure 4). The effect of LOB convexity on return kurtosis is significantly positive and thus in line with our theory which predicts that return kurtosis is an increasing function of LOB convexity regardless of LOB and MO imbalance. The magnitude of this effect is small compared to the standard deviation of return kurtosis in our sample (5.4989) and to what our theory predicts. Panel (d) of Figure 4 shows that a one standard deviation increase in LOB convexity results in an increase of return kurtosis from 3 to about 6. More importantly, LOB convexity is economically the second most important determinant of the return kurtosis, in line with our theory and intuition which predicts that the LOB convexity should be one of the most important determinants of the return kurtosis.

Finally, our theory predicts that the effects of primary LOB shape variables, especially that LOB imbalance, on variance and kurtosis can be slightly nonlinear depending on other determinants (see panels (b) and (d) of Figures 2 and 10). We analyze the nonlinear effect of LOB shape variables on returns moments by adding the squared LOB imbalance and convexity in the empirical specification in Equation 14. Our results, not presented here for the sake of brevity, do not provide much evidence of any nonlinear effect of primary LOB shape variables on returns moments.

5.2 Unconditional Effects of Other Variables on Return Moments

In this section, we summarize the effects of the first four moments of the MO distribution and the secondary LOB shape variables, i.e. total cumulative quantity in and the slope of the LOB. Recall that the mean and standard deviation of market order distribution and the secondary LOB shape variables are directly implied by our theory and thus we can easily compare whether their observed empirical effects on the return moments are in line with the implications of our model. On the other hand, the skewness and kurtosis of the market order distribution are not directly considered in our model. That said, as discussed in Section 3, our model predicts that the third and fourth moments of the MO distribution might have significant effects on the corresponding moments of the return distribution.

We start our discussion with the first four moments of the MO distribution before turning our attention to secondary LOB shape variables. An increase in the market order imbalance is positively related to return mean on the same day. A one standard deviation increase in market order mean is associated with almost a 20 basis points increase in return mean, which corresponds to about 6% of the standard deviation of daily returns in our sample. The sign of this effect is intuitive and more importantly in line with our theory. However, the magnitude of this effect is smaller than what our theory predicts. Panel (a) of Figure 10 shows that a one standard deviation increase in MO mean increases return mean by a little more than 1% when there is no LOB imbalance. That said, MO mean is the most important determinant of the return mean. Turning our attention to the higher moments of the MO imbalance, we find that the effect of MO imbalance standard deviation is neither statistically nor economically significant. This is in line with our theory which predicts that the market order standard deviation should not be an important determinant of the return mean. On the other hand, the MO imbalance skewness and kurtosis have significantly negative and positive effects on the return mean, respectively. It is hard to say whether these effects of market order skewness and kurtosis are in line with our theory, which assumes that the market order imbalance is normally distributed.

The MO imbalance mean does not have a significant effect on the return variance while MO imbalance standard deviation has a significantly positive effect. These findings are in line with our expectations based on Lemma 1, which states that the return distribution should reflect the MO imbalance distribution in the absence of LOB imbalance. That said, the economic magnitude and importance of the MO imbalance standard deviation are smaller than we expected. It is only the sixth most important out of eight determinants. As mentioned before, our model is relatively silent on the effect of MO imbalance skewness and kurtosis on the return variance. Intuitively, one would expect the market order skewness to be negatively related to the return variance and kurtosis to be positively related. Our empirical results suggest that this is indeed the case and that the MO imbalance kurtosis is economically the third most important determinant of the return variance.

The MO imbalance and skewness have significantly positive effects on the return skewness, in line with our expectations based on Lemma 1. The MO imbalance is by far the most important variable in determining the return skewness while MO skewness is only the fifth most important. We do not find any significant effects of MO imbalance variance and kurtosis on the return skewness. These are also in line with our expectations since we do not expect MO imbalance variance and kurtosis to be related to return skewness based on Lemma 1.

The MO imbalance standard deviation has a significant positive effect on the return kurtosis. This is in line with our expectations based on Lemma 1 and the positive relation between variance and kurtosis of a distribution. That said, the MO imbalance standard deviation is economically only the fourth most important determinant of the return kurtosis. It is not surprising to find that the MO imbalance mean and skewness do not have any significant effects given that mean and skewness of a distribution are not generally related to its kurtosis. What is surprising is the significantly negative effect of the MO imbalance kurtosis on the return kurtosis. One would expect this effect to be positive based on Lemma 1. That said, the implications of Lemma 1 hold only in the absence of any LOB imbalance. The fact that LOB shape variables are by for the most important determinants of the return kurtosis might change the effect of MO imbalance kurtosis on the return kurtosis.

Finally, we find that the average LOB slope has significantly negative effects on the return mean and skewness and significantly positive effects on the return variance and kurtosis. Recall that our model predicts that the return variance is an increasing function the average slope of the limit order book while the return skewness and kurtosis do not depend on it. Thus, the significantly positive effect of the average LOB slope on the return variance is in line with our model while its significant effects on the return skewness and kurtosis are not. Regarding its effect on the return mean, it is hard to say whether it is in line with our model, which does not provide a clear prediction on this effect.

We also find that the average total quantity available in the limit order book has significantly positive effects on the return mean and variance and significantly negative effects on skewness and kurtosis. Also, it is the economically most important determinant of the return variance and kurtosis. Our conclusions whether these results are in line with our model are similar to our conclusions regarding the effects of the average slope. Specifically, recall that our model predicts that the return variance is an increasing function of the total quantity available if and only if LOB convexity is positive regardless of the signs and magnitudes of the MO and LOB imbalances and the return skewness and kurtosis do not depend on the total quantity available. Thus, the significantly positive effect of the average LOB slope on the return variance can be considered in line with our model while its significant effects on the return skewness and kurtosis are not. Regarding its effect on the return mean, it is hard to say whether it is in line with our model, which does not provide a clear prediction on this effect.

5.3 Conditional Effects of LOB Imbalance and Convexity on Return Moments

In this section, we analyze the empirical effects of LOB Imbalance and convexity conditional on each other and also on MO imbalance and MO standard deviation. To do this, we consider an empirical specification similar to that in Equation 14 and include the interaction terms between LOB imbalance and convexity and MO imbalance. To be more precise, we estimate the following specification with firm fixed effects and standard errors clustered at the firm level.

$$
\begin{aligned}
\mathrm{M}_{j,it} &= \alpha_i + \sum_{k\neq j} \beta_k^j M_{j,it} + \sum_{k=1}^{4} \gamma_k^j \mathrm{MO}_{k,it} + \delta_1 \log(Q_{it}^{avg}) + \delta_2 \log(S_{it}^{avg}) + \delta_3 \log(Imb_{it}) \\
&+ \delta_4 \log(Conv_{it}^{avg}) + \delta_5 \log(Imb_{it}) \times \mathrm{MO}_{1,it} + \delta_6 \log(Conv_{it}^{avg}) \times \mathrm{MO}_{1,it} + \delta_7 \log(Imb_{it}) \times \log(Conv_{it}^{avg}) \\
&+ \delta_8 \log(Imb_{it}) \times \mathrm{MO}_{1,it} + \delta_9 \log(Conv_{it}) \times \mathrm{MO}_{1,it} \\
&+ \delta_{10} \log(Imb_{it}) \times \mathrm{MO}_{2,it} + \delta_{11} \log(Conv_{it}) \times \mathrm{MO}_{2,it} + \varepsilon_{it} \qquad (15)
\end{aligned}
$$

We start with how the effect of LOB imbalance on return moments changes depending on MO imbalance and standard deviation and LOB convexity. The effect of LOB imbalance on return moments conditional on market imbalance is given by $\delta_3 + \delta_5 \times \mathrm{MO}^1$ from the above specification. To understand this conditional effect, we set all variables other than LOB Imbalance and MO imbalances to zero. We then set MO imbalance to its high and low values using its standard deviation($\sigma(MOImb)$), i.e. High MO Imbalance = $+\sigma(MOImb)$ and Low MO Imbalance = $-\sigma(MOImb)$. We choose one standard deviation to correspond to our theoretical results presented in Figures 2 and 10. Finally, to understand the economic magnitude of this conditional effect, we multiply this coefficient by standard deviation of LOB imbalance. To summarize, the economic magnitude of the LOB Imbalance's effect on return mean when MO imbalance is high is given by $(\hat{\delta}_3 + \hat{\delta}_5 \times \sigma(\text{MO Imb})) \times \sigma(\text{LOB Imb})$. Table 4 presents these conditional effects.

[Table 4 about here.]

We start with the effect of LOB imbalance conditional on MO imbalance. First note that the conditional effect of LOB imbalance on returns moments when MO imbalance is set to zero should be similar to its unconditional effect presented in Table 3. The differences between the two are due to the fact the regression for the conditional effects in Equation 15 includes few variables that are not in the regression for the unconditional effect in Equation 14. LOB imbalance has a significant positive effect on return mean only when MO imbalance is high. This is in line with our theory and the unconditional effect of LOB imbalance on return mean. That said, this effect changes sign and becomes significantly negative when MO imbalance is low. This is in contrast to our theory which predicts that the return mean should be an increasing function of LOB imbalance regardless of the MO imbalance. Furthermore, the effect of LOB imbalance on return mean becomes insignificant when we consider a zero order imbalance, suggesting that controlling for the interaction terms makes the unconditional effect of LOB imbalance insignificant. The effect of LOB imbalance on higher moments do not depend strongly on the MO imbalance. Regardless of MO imbalance, variance and skewness are increasing functions while kurtosis is a decreasing function of LOB imbalance. These are in line with our results on the unconditional effects of LOB imbalance on higher moments. That said, the results for variance and kurtosis provide at best mixed support for our theory, which predicts that variance (kurtosis) should be an increasing function of LOB imbalance when MO imbalance is high (low) and vice versa.

We now turn our attention to how the effect of LOB imbalance changes with MO standard deviation. Recall that our theory predicts that the effects of LOB shape variables on return moments, especially on skewness and kurtosis, are amplified when market orders are more volatile. Our findings in panel (b) of Table 4 are broadly consistent with this prediction. To be more precise, we find that the LOB imbalance has a significantly positive effect on return mean only when market orders exhibit more volatility. The effect of LOB imbalance on return variance does not seem to depend on market order standard deviation. More importantly, in line with our predictions, we find that the effects of LOB imbalance on both skewness and kurtosis are much more pronounced when market orders are more volatile.

Regarding how the effect of LOB imbalance on return moments changes with LOB convexity, we find that this effect is much more pronounced for all moments except for return mean when the LOB convexity is high. The findings for return mean and skewness are in line with our theory which predicts the effect of LOB imbalance on returns mean (skewness) to be more pronounced when LOB convexity is low (high). Our theory predicts that the effect of LOB imbalance on variance and kurtosis are U-shaped and this is much more pronounced when LOB convexity is high. As mentioned above at the end of Section 5.1, we do not find much empirical evidence on this nonlinearity. Instead, our results on the unconditional effect of LOB imbalance on variance and kurtosis suggests that skewness is an increasing function while kurtosis is a decreasing function of LOB imbalance. Our results in panel (c) of Table 4 suggests that this continues to hold and more importantly, these effects are more pronounced when LOB convexity is high in line with our predictions.

Table 5 presents the results on the effects of LOB convexity on return moments conditional on MO imbalance and standard deviation and LOB imbalance. We start our discussion with how the effects of LOB convexity on return moments changes with MO imbalance. We find that the effect of LOB convexity on return mean changes sign depending on the MO imbalance. Specifically, return mean is an increasing function of LOB convexity when MO imbalance is low while the opposite holds when MO imbalance is high, exactly in line with our predictions. Variance is a decreasing function of LOB convexity regardless of MO imbalance with a higher slope when MO imbalance is low, also broadly consistent with our predictions. LOB convexity has a positive but insignificant effect on return skewness regardless of MO imbalance. This effect although insignificant is more pronounced when MO imbalance is low. This is not consistent with our theory which predicts that return skewness is an increasing function of LOB convexity when MO imbalance is high and vice versa. That said, the fact that this effect is statistically insignificant is broadly consistent with its insignificant unconditional effect, which is in line with our predictions. Kurtosis is an increasing function of LOB convexity regardless of MO imbalance with a higher slope when MO imbalance is low, broadly consistent with our predictions. Finally, turning our attention to the effect of LOB convexity on return moments conditional MO standard deviation, we find that the effect of LOB convexity on return is much more pronounced and significant when market orders are more volatile, broadly consistent with our predictions.

[Table 5 about here.]

5.4 Unconditional Effects of LOB Shape and MO Distribution on Return Quantiles

In this section, we turn our attention to the unconditional effects of market order distribution and LOB shape on the quantiles of the return distribution. To do this, we first compute p^{th} quantile of the return distribution for each stock-day pair based on the following Cornish-Fisher expansion:

$$q_p \approx M_1 + \sqrt{M_2}\left(z_p + \frac{M_3}{6}\left(z_p^2 - 1\right) + \frac{M_4}{24}\left(z_p^3 - 3z_p\right) - \frac{M_3^2}{36}\left(2z_p^3 - 5z_p\right)\right) \tag{16}$$

where M_k is the k^{th} central moment of the return distribution, z_p is the p^{th} quantile of a standard normal distribution and $p = 5\%, 10\%25\%, 50\%, 75\%, 90\%$ and 95%. We then estimate the following empirical specification for each quantile separately:

$$q_{p,it} = \alpha_i + \sum_{k=1}^{4} \gamma_k^j \text{MO}_{k,it} + \delta_1 \log(Q_{it}^{avg}) + \delta_2 \log(S_{it}^{avg}) + \delta_3 \log(Imb_{it}) + \delta_4 \log(Conv_{it}^{avg}) + \varepsilon_{it} \tag{17}$$

for $p = 5\%, 10\%25\%, 50\%, 75\%, 90\%$ and 95%. This empirical specification is similar to that in Equation 14 with the main difference that we do not include the return moments as control variables because the quantiles based on the Cornish-Fisher expansion are already functions of the return moments. We estimate this empirical specification with firm fixed effects and standard errors clustered at the firm level. Table 6 presents the economic significance (coefficient estimate times one standard deviation of the corresponding independent variable) of market order distribution and LOB shape parameters. These numbers can be interpreted as the effect of a one standard deviation increase in the independent variable on the return moment of interest and can be directly compared to Figures 3 and 5 which also show the effect of a one standard deviation increase in certain variables.

[Table 6 about here.]

We start our discussion with the LOB imbalance and LOB convexity, our main variables of interest, before turning our attention to the other variables. In line with our predictions, an increase in the LOB imbalance on day $t-1$ has significantly positive effects on all return quantiles considered, except the 5% quantile, on day t. In other words, an increase in the LOB imbalance results in a shift to the right for the whole return distribution. Recall from Table 3 that the LOB imbalance has significantly positive effects on the return mean, variance and skewness and a significantly negative effect on the return kurtosis. Our results for the left shoulder of the return distribution, i.e. the 25% quantile, suggest that the effect of the LOB imbalance on the return mean and skewness dominate its effect on the return variance. Similarly, our results for the right tail of the return distribution, i.e. 90% and 95% quantiles, suggest that the effect of the LOB imbalance on the return mean and skewness dominate its effect on the return kurtosis. The economic magnitude of these effects are smaller than what our model predicts. For example, Figure 3 shows that a one standard deviation increase in the LOB imbalance increases the 90% quantile by approximately 0.005 from 0.02 to 0.025. Our results in Table 6 for 90% shows that this effect is around 0.00059, almost an order of magnitude smaller. Furthermore, the LOB imbalance is economically the least important variable for the left part of the return distribution. That said, it is economically the third and fourth most important variable for the right part of the return distribution. Overall, these empirical findings confirm our prediction that the LOB imbalance is an important determinant of the return mean and skewness while it might not play important role in determining the return variance and kurtosis.

LOB convexity has significantly positive effects on the left part, the median as well as the right shoulder of the return distribution while its effects on the right tail are statistically insignificant. This is also reflected in the economic importance of LOB convexity in determining the quantiles of the return distribution. LOB convexity is economically the third most important variable for the left part and the median while it is the least important variable for the right tail. Furthermore, its effect on the left tail (5% and 10% quantiles) is more pronounced than its effect on the left shoulder (25% quantile). These results suggest that the empirical effects of the LOB convexity on the return mean and variance dominate its effect on the return kurtosis presented in Table 3. Overall, the sign and magnitude of these effects are broadly consistent with our theory. Figure 5 shows that an increase in the average LOB convexity increases the quantiles below the median and decreases the quantiles above the median and that this effect should be more pronounced for the tails than the shoulders.

Turning our attention to the secondary LOB shape parameters, the average LOB slope has significantly negative effects on all quantiles and is economically the second most important variable. Furthermore, its effect on the left part of the distribution is more pronounced than the its effect on the right part. These results in turn imply that an increase in the LOB slope results in a shift to the left for the whole return distribution and also possible make the return distribution more negatively skewed in line with our results in Table 3. These results are also intuitive since an increase in the average LOB slope implies a decrease in liquidity. The average LOB quantity has significantly positive effects on the quantiles above the median and significantly negative effects on the quantiles below the median. The magnitude of its effects on the left and right part of the distribution are similar, suggesting a widening of the return distribution. This in turn suggests that its positive effect on the return variance is stronger than its negative effect on the return kurtosis.

Finally, the MO imbalance is economically the most important variable for the return distribution and has significantly positive effects on all quantiles. In other words, an increase in the MO imbalance implies not surprisingly a shift to the right for the whole return distribution. The MO standard deviation does not have a significant effect on the return quantiles. As one would expect, an increase in the MO skewness implies a shift to the left for the whole return distribution while an increase in the MO kurtosis implies a widening of the return distribution. That said, the MO variance, skewness and kurtosis are on average economically less important than the LOB shape variables in determining the return distribution.

6 Establishing Causality

Given that we consider LOB shape parameters with a lag in the regressions for return moments, we argue that our results above establish the causal effect of LOB shape on return moments. In this section, we provide further empirical evidence on this causal effect from a different angle. Our approach follows closely Diether et al. (2009), who analyze the effect on market quality of Reg SHO but our objective is different from theirs. We establish the causal effect of LOB parameters on return moments by using the exogenous shock to the LOB shape caused by the Reg SHO.

Short-selling restrictions in the US were introduced in the 1930s following the stock market crash of 1929. They are designed to make short-selling more difficult. For example, SEC introduced the uptick rule in 1938 to restrict the short sell of a given stock on the NYSE only if the most recent change in its price was strictly positive. Regulation SHO was introduced on September 7, 2004 to "update short sale regulation in light of numerous market developments since short sale regulation was first adopted in 1938 and to address concerns regarding persistent failures to deliver and potentially abusive naked short selling". Reg SHO includes several sets of rules which became effective on January 3, 2005 and were amended several times. Our focus in this paper is on Rule 202T of Regulation SHO (17 CFR 242.202T), called the Pilot Program. The objective of the Pilot Program was to evaluate the effectiveness of price test restrictions on short sales and their effects on market volatility, price efficiency, and liquidity, as in Diether et al. (2009). Following the adaptation of Rule 202T, SEC excluded about 1,000 designated securities, called Pilot Stocks, from the operation of any short sale price test rule, such as the uptick rule. The pilot stocks began trading without any short sale restrictions on May 2, 2005. Every third stock in the Russell 3000 index sorted by volume was chosen as a pilot stock while the remaining 2,000 stocks constitute the control stocks.

In this paper, given that we only have LOB data for NYSE stocks, we focus on the effect of the Pilot Program on the NYSE stocks. Diether et al. (2009) argue that the uptick rule of the NYSE significantly affects how the specialist adjust non-compliant short sell orders so that they can be presented to the market. Specifically, non-compliant short-sale orders, i.e. those that are not submitted after an uptick, are effectively changed into nonmarketable limit orders and added to the limit order book. Once the uptick rule is suspended, short-sale orders can no longer be classified as non-compliant and thus are not changed into nonmarketable limit orders or added to the book. As a result, Diether et al. (2009) hypothesize that the *first-order effect* of the suspension of the uptick rule for pilot stocks would be a decrease in the depth at the best ask price for these stocks relative to the control stocks. They also predict that a reduction in the buy order imbalance, a widening of quoted and effective spreads, and an increase in short-term volatility for NYSE-listed Pilot stocks relative to control stocks following the suspension of the uptick rule. Using differences-in-differences analysis, they provide empirical evidence in support of this first-order effect.

However, we are interested in how the suspension of the uptick rule would affect the overall shape of the limit order book and market order distribution and as a result the overall return distribution. Hence, we first analyze how the first four central moments of the return distribution change as a result of the suspension of the uptick rule based on a difference-in-differences analysis. Given the econometric problems associated with having control variable in a DiD setting, we decided to perform similar differences-in-differences analyses for MO distribution and LOB shape parameters and discuss based on the results what variables might be the driving factors behind the observed change in the return distribution.

Panel (a) in Table 7 presents the results of the DiD analysis on the first four central moments of the return distribution. Our results show that the mean of pilot and control stocks increase significantly with similar magnitudes resulting in an insignificant effect of the suspension of the uptick rule on pilot stocks relative control stocks. This result is similar to what we observe for the skewness. To be more precise, although the skewness of both pilot and control stocks increase following the suspension of the uptick rule, they do so with similar magnitudes resulting in an insignificant difference between them. On the other, we find statistically significant effect of the suspension of the uptick rule on the variance and kurtosis of pilot stocks relative to the control stocks. More precisely, the variance of both pilot and control stocks decrease significantly but that of pilot stocks decrease significantly less than that of control stocks following the suspension of the uptick rule, resulting in a significantly positive difference between the two. The results for kurtosis are little bit more nuanced as the average kurtosis of pilot stocks significantly decreases following the suspension of the uptick rule while that of control stocks significantly increases and thus resulting in a significant difference between them. Overall, these results show that the suspension of the uptick rule increases the variance and decreases the kurtosis of pilot stocks relative to control stocks while it does not seem to affect the mean and skewness of pilot and control stocks differently.

Panel (b) and (c) of Table 7 present the results of the DiD analysis on the main determinants of the return distribution according to our model, i.e. the first four moments of the market order distribution and the LOB shape parameters, respectively. We use these results to explain what factors might be driving the observed effect of the suspension of the uptick rule on the first four moments of the return distribution. We start our discussion with the mean. Recall from our theory that the main determinants of the return mean are the MO and LOB imbalances. Our results suggest that the buying pressure from market orders, i.e. the market order imbalance, decreases for pilot stocks significantly more than for control stocks following the suspension of the uptick rule. On the other hand, the opposite holds for the buying pressure from limit orders, i.e. the limit order book imbalance, which decreases significantly more for control stocks than for pilot stocks. The magnitudes of these two effects are such that they cancel each other out and we do not observe any significantly different effects of the suspension of the uptick rule between the return means of pilot and control stocks.

Turning our attention to the return variance, we argue that the change in the LOB convexity is the main driving factor behind the observed change in the return variance following the suspension of the uptick rule. To see this, recall from our theory that the return variance is mainly determined by the market order variance, the average slope and convexity of the limit order book and is an increasing functions of market order variance and the average slope of the limit order book and a decreasing function of the LOB convexity. The market order variance cannot be the driving factor behind the observed relative change in the return variances because we do not observe any significantly different reactions between the return variances of pilot and control stocks to the suspension of the uptick rule. Similarly, we argue that the observed relative change in the return variances cannot be attributed to the average slope of the limit order book. The average slope of the limit order book decreases significantly more for the pilot stocks than for control stocks. This should have resulted in a significantly more pronounced decrease in the return variance of pilot stocks than that of control stocks, which is the opposite of what we observed in panel (a). On the hand, the LOB convexity decreases significantly more for pilot stocks than for control stocks. Given that the return variance is inversely related to LOB convexity, the relatively more pronounced decrease in the LOB convexity for pilot stocks compared to control stocks explains why the return variance of pilot stocks increases relative to the control stocks following the suspension of the uptick rule.

We now consider skewness. Note that the main determinants of skewness are the LOB imbalance and MO skewness. As mentioned above, the LOB imbalance of pilot stocks decreases significantly less than that of control stocks following the suspension of the uptick rule. This positive relative effect should have resulted in an increase in the return skewness, which should be an increasing function of the LOB imbalance according to our theory. That said, this effect seems to be cancelled out by the change in the MO skewness following the suspension of the uptick rule. To be more precise, we find that MO skewness of the pilot stocks decrease significantly more than that of control stocks, which should have resulted in a decrease of the return skewness of pilot stocks relative to control stocks.

Finally, we consider kurtosis. Recall that the main determinants of kurtosis are LOB convexity and MO kurtosis. We find that the kurtosis of pilot stocks decrease while that of control stocks increases, resulting in a significantly different effect of the suspension of the uptick rule on return kurtosis of pilot and control stocks. This result cannot be explained by the change in the MO kurtosis following the suspension of the uptick rule since the MO kurtosis of pilot stocks increase significantly more than that of control stocks, which should have resulted in a relative increase in their kurtosis. On the other than, LOB convexity of pilot stocks decreases significantly more than that of control stocks. Given that the return kurtosis is an increasing function of LOB convexity, the observed effect of the suspension of the uptick rule on return kurtosis must be due to the decrease in LOB convexity, despite the opposing effect of the change in the MO kurtosis.

[Table 7 about here.]

7 Robustness Checks

In this section, we discuss the robustness of our results in Tables 6 and 6 to using alternative variable definitions and empirical choices. For the sake of brevity, we summarize the robustness check and the corresponding results while the detailed results are presented in the online appendix. f

7.1 Using different levels for LOB variables

In our main results, we measure LOB variables based on the first ten levels of the bid and ask sides. We do this to use all the available information at our disposition. However, if investors engage in aggressive order splitting, these small market orders might not climb all ten levels of the limit order book and might be executed against the lower levels of the limit order book. This in turn implies that any LOB measure based on the first ten levels might not be relevant given that only a few market orders would be executed against all the first ten levels. To this end, we consider LOB measures based on the first three or five instead of the first ten levels of the bid and ask side. The definitions of all LOB variables except the LOB convexity can be easily adjusted to use the first three or five levels of the limit order book. We need to adjust our measure of LOB convexity to take into account the fact that we are using fewer levels. To be more precise, we define the ask side convexity based on the first l levels, $l = 3, 5$ for a given snapshot s as $C^A_{2,l,s} = \frac{S^A_{2,l,s}}{S^A_{1,2,s}}$ where the cutoff between lower and higher levels in both cases is level two. The bid side convexity is also defined similarly and the average LOB convexity is defined as their simple average. As in our main empirical results, we then compute the time-weighted average of this average LOB convexity measure and use it as our proxy for LOB convexity in our analysis. Tables OA.1 and OA.2 present our main results when we use the first five instead of ten levels of the limit order book. Our main results based on the first three levels of the limit order book are very similar to those based on the first five levels. More importantly, Tables OA.1 and OA.2 show that our main results on the effect of LOB variables on the first four moments and the quantiles of the return distribution presented in Tables 6 and 6 do not change significantly when we use only the first five instead of the first ten levels to compute the LOB variables.

7.2 Alternative Cutoff Point for LOB Convexity

Recall that in our main empirical results we consider three convexity measures with cutoff levels $l = 2, 3, 5$. For a given stock-day pair, we use the one based on the level where the cumulative quantity is closest to *one* market order standard deviation of that stock on the previous trading day. In this section, we change the definition of large market orders from those that are greater (in absolute value) than one standard deviation of the market order distribution to those that are greater (in absolute value) than two or three standard deviations. The results for two standard deviations are presented in Tables OA.3 and OA.4 are very similar to our main results, suggesting that our main conclusions are robust to using alternative definitions of LOB Convexity.

Similar to our robustness check above, we also considered a LOB convexity measure based on the first five levels. To be more precise, we compute two convexity measures based on the first five levels but with cutoff levels of two and three. Specifically, these two convexity measures for a given snapshot s of the ask side are defined as $C^A_{2,5,s} = \frac{S^A_{2,5,s}}{S^A_{1,2,s}}$ and $C^A_{3,5,s} = \frac{S^A_{3,5,s}}{S^A_{1,3,s}}$. For a given stock-day pair, we find the level that minimizes the absolute value of the difference between k times the convexity cutoff, $\tilde{K}$, and the average cumulative quantity available in the first l levels of bid and ask sides, both normalized by market order standard deviation of the previous trading day, i.e. $l^* = \arg\min_{l=1,2,\ldots,5} |(Q_{A,l}/\sigma + Q_{B,l}/\sigma)/2 - k \times \tilde{K}|$. If l^* is one, we use $c_{avg,2,5} = (C^A_{2,5,s} + C^B_{2,5,s})/2$. If l^* is greater than one but less than or equal to three, we use $c_{avg,3,5} = (C^A_{3,5,s} + C^B_{3,5,s})/2$. As in our main empirical results, we then compute the time-weighted average of this average LOB convexity measure and use it as our proxy for LOB convexity in our analysis. Our results not presented for the sake brevity are very similar to our main results, suggesting the robustness of our results using alternative definitions of LOB convexity.

7.3 Excluding the Global Financial Crisis

Finally, we exclude the period between 1 July 2007 and 31 December 2008 (inclusive) corresponding to Global Financial Crisis (GFC) from our sample. As it is well known, this is a period with relatively high volatility and more importantly includes a short-sale ban on financial stocks, which might affect our results. Tables OA.5 and OA.6 show that excluding the GFC from our sample do not significantly our main conclusions regarding the effect of LOB variables on the return distribution.

8 Conclusion

In this paper, we developed a reduced-form microstructure model of price formation to examine the role of market and limit orders in shaping the return distribution in an order-driven market. Our theoretical framework posits that the distribution of market orders and the structure of the limit order book (LOB) jointly determine the moments of the return distribution. By deriving closed-form expressions for the first four central moments, we established clear predictions on the relationship between LOB imbalance, convexity, and the shape of the return distribution.

Our empirical analysis, using high-frequency data from NYSE stocks, provides strong support for the key theoretical predictions. We find that LOB imbalance plays a significant role in influencing the mean and skewness of the return distribution, independent of the market order distribution. The convexity of the limit order book, on the other hand, has a pronounced effect on return variance and kurtosis, supporting the notion that market liquidity and transaction costs impact the tails of the return distribution.

Furthermore, by leveraging Reg SHO as an exogenous shock to liquidity, we were able to establish a causal relationship between LOB shape and return moments. Specifically, we identified that changes in LOB convexity were the primary driver behind the observed changes in return variance and kurtosis following the suspension of the uptick rule, further reinforcing the importance of market structure in determining return dynamics.

In sum, our findings contribute to the broader literature on price formation and market microstructure by offering a comprehensive model that links investor behavior, as reflected in order book characteristics, to asset return distributions. These results have practical implications for both market participants and regulators, highlighting the importance of limit order book dynamics in influencing market outcomes. Future research may extend this model to incorporate more complex market conditions or alternative order types, enhancing our understanding of how market structure interacts with trading behavior to shape price dynamics.

References

Beltran-Lopez, H., Giot, P., Grammig, J., 2009. Commonalities in the order book. Financial Markets and Portfolio Management 23 (3), 209–242.

Biais, B., Hillion, P., Spatt, C., 1995. An empirical analysis of the limit order book and the order flow in the paris bourse. Journal of Finance 50 (5), 1655–1689.

Brennan, M. J., Chordia, T., Subrahmanyam, A., Tong, Q., 2012. Sell-order liquidity and the cross-section of expected stock returns. Journal of Financial Economics 105 (3), 523–541.

Cao, C., Hansch, O., Wang, X., 2009. The information content of an open limit-order book. Journal of Futures Markets 29 (1), 16–41.

Cenesizoglu, T., Dionne, G., Zhou, X., 2022. Asymmetric effects of the limit order book on price dynamics. Journal of Empirical Finance 65, 77–98.

Cenesizoglu, T., Grass, G., 2018. Bid- and ask-side liquidity in the nyse limit order book. Journal of Financial Markets 38, 14–38.

Chakrabarty, B., Li, B., Nguyen, V., Van Ness, R. A., 2007. Trade classification algorithms for electronic communications network trades. Journal of Banking & Finance 31 (12), 3806–3821.

Chordia, T., Roll, R., Subrahmanyam, A., 2002. Order imbalance, liquidity, and market returns. Journal of Financial Economics 65 (1), 111–130.

Foucault, T., Kadan, O., Kandel, E., 2005. Limit order book as a market for liquidity. The Review of Financial Studies 18 (4), 1171–1217.

Goettler, R. L., Parlour, C. A., Rajan, U., 2009. Informed traders and limit order markets. Journal of Financial Economics 93 (1), 67–87.

Hasbrouck, J., 1991. Measuring the information content of stock trades. The Journal of Finance 46 (1), 179–207.

Hasbrouck, J., 1992. Assessing the quality of a security market: A new approach to transactions cost measurement. Review of Financial Studies 5 (2), 191–212.

Kalay, A., Wohl, A., 2009. Detecting liquidity traders. Journal of Financial and Quantitative Analysis 44 (1), 29–54.

Kaniel, R., Liu, H., 2006. So what orders do informed traders use? The Journal of Business 79 (4), 1867–1913.

Keim, D. B., Madhavan, A., 1996. The upstairs market for large-block transactions: Analysis and measurement of price effects. The Review of Financial Studies 9 (1), 1–36.

Knez, P. J., Ready, M. J., 1996. Estimating the profits from trading strategies. Review of Financial Studies 9, 1121–1163.

Korajczyk, R. A., Sadka, R., 2008. Pricing the commonality across alternative measures of liquidity. Journal of Financial Economics 87 (1), 45–72.

Kozhan, R., Salmon, M., 2012. The information content of a limit order book: The case of an fx market. Journal of Financial Markets 15 (1), 1–28.

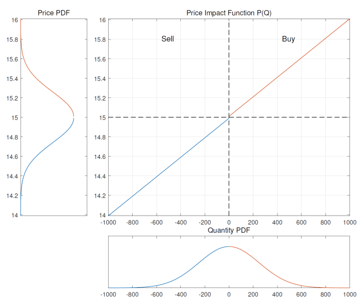
Kyle, A. S., 1985. Continuous auctions and insider trading. Econometrica 53 (6), 1315–1335.

Roesch, C. G., Kaserer, C., 2014. Reprint of: Market liquidity in the financial crisis: The role of liquidity commonality and flight-to-quality. Journal of Banking and Finance 45 (1), 152–170.

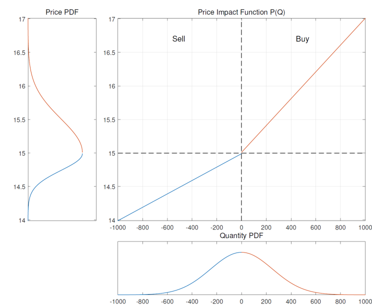
Rosu, I., 2009. A dynamic model of the limit order book. The Review of Financial Studies 22 (11), 4601–4641.

Wuyts, G., 2008. The impact of liquidity shocks through the limit order book. Unpublished working paper. University of Leuven.

(a) Symmetric Limit Order Book



(b) Imbalanced Limit Order Book



(c) Convex Limit Order Book

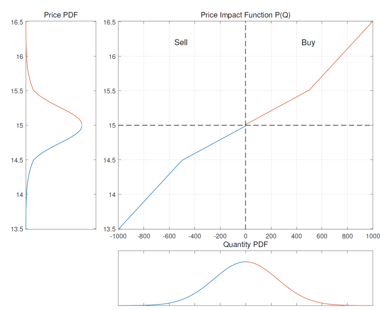


Figure 1: *Market Order Distribution, Limit Order Book and Return Distribution*
Note: This figure illustrates the relationship between the distribution of market orders, the shape of the limit order book, and the distribution of returns. In all three panels, the figure labeled "Quantity PDF" (lower right) displays the assumed normal distribution of market orders, centered at zero, indicating no market order imbalance. Positive values (shown in red) represent buy market orders matched with the ask side of the limit order book, while negative values (shown in blue) correspond to sell market orders matched with the bid side. The "Price Impact Function $P(Q)$" (top right) represents the shape of the ask and bid sides of the limit order book, modeled as a continuous price impact function, against which market orders are executed. The total shares available in both bid and ask sides are assumed to be 1,000 and the midquote is assumed to be \$15 per share. The "Price PDF" (top left) presents the distribution of prices around the midquote of \$15, resulting from the execution of market orders with different number of shares (presented in "Quantity PDF" (lower right)) against the limit order book (presented in "Price Impact Function P(Q)"). The blue (red) line in "Price PDF" corresponds to negative (positive) returns resulting from the execution of sell (buy) market orders. Panel (a) presents the case with a balanced, linear limit order book where both the bid and ask sides have identical slopes. Panel (b) presents the case with an imbalanced, linear limit order book, with a steeper slope on the ask side compared to the bid side. Panel (c) presents the case with a balanced but convex limit order book, where the higher levels on both bid and ask sides have steeper slopes than their lower levels.

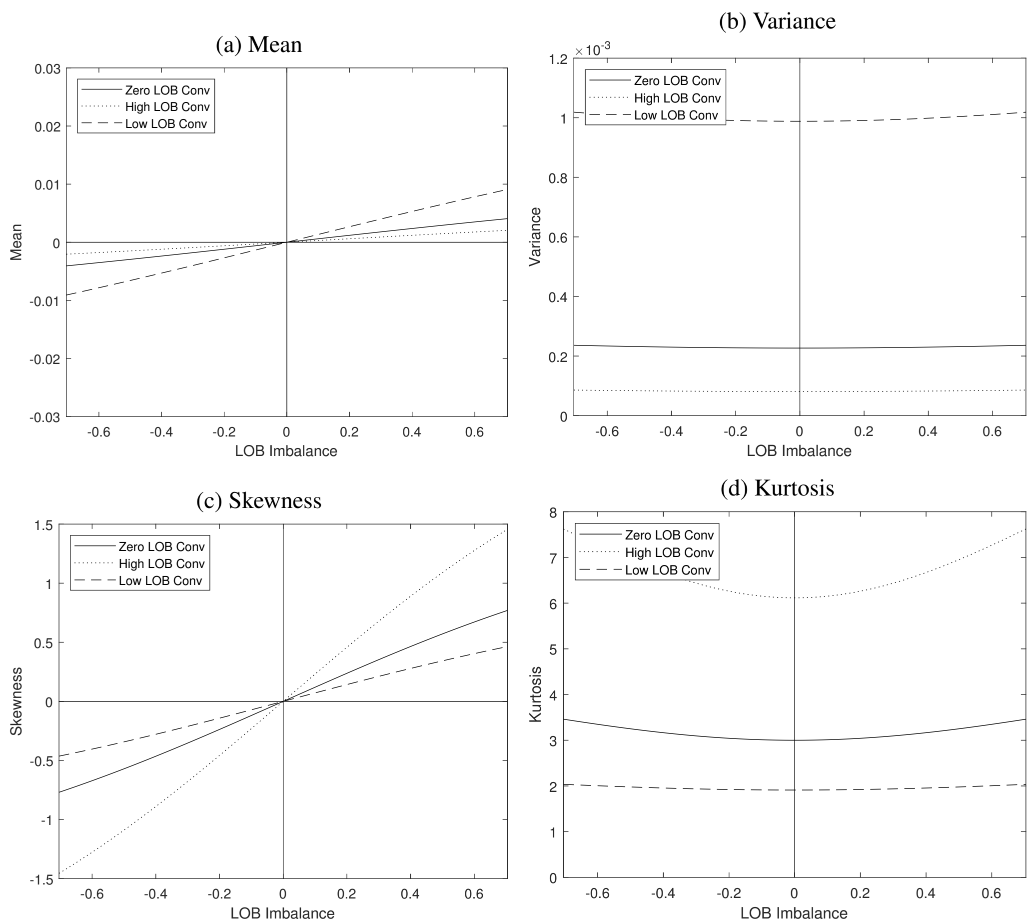


Figure 2: *The Effects of the LOB Imbalance on the Return Mean, Variance, Skewness and Kurtosis*
Note: This figure presents the effect of the LOB Imbalance on the first four central moments of the return distribution for three different values of the LOB convexity, high (dotted line), zero (solid line) and low (dashed line). The high and low values of LOB convexity are 0.4226 and - 0.4226, corresponding respectively to plus and minus one times the average of the sample standard deviations (std) of the empirical counterparts of c_A and c_B, i.e. $(std(c_A) + std(c_B))/2 = 0.4226$. All other variables are calibrated to the values presented in Table 1. We consider LOB imbalance values between minus and plus one times the standard deviation of its sample counterpart (0.7044), i.e. the limits of the x-axis are -0.7044 and 0.7044.

(a) Low LOB Convexity (b) Zero LOB Convexity (c) High LOB Convexity

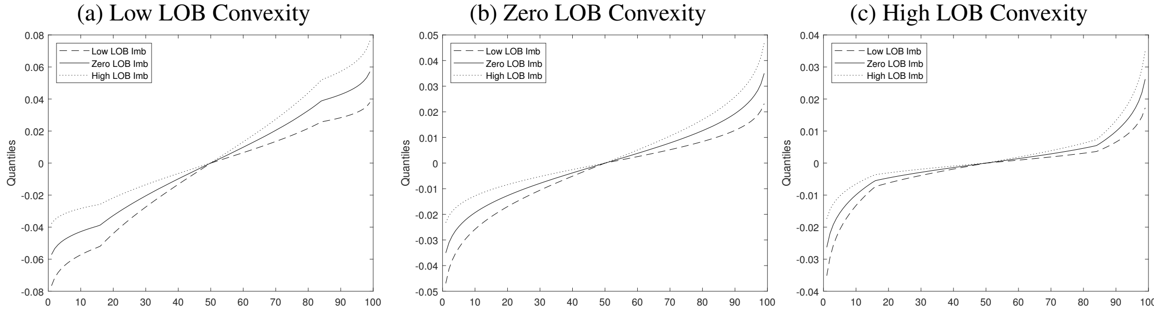


Figure 3: *The Effects of the LOB Imbalance on the Return Distribution*
Note: This figure presents the quantiles between 1% and 99% of the return distribution for three different values of LOB imbalance, high (dotted line), zero (solid line) and low (dashed line) and for three different values of LOB convexity, low (panel (a)), zero (panel (b)) and high (panel (c)). The high and low values of LOB imbalance are 0.7044 and -0.7044 and the high and low values of LOB convexity are 0.4226 and - 0.4226, corresponding respectively to plus and minus one times the sample standard deviations of their empirical counterparts. The market order imbalance is set to zero and all other variables are calibrated to the values presented in Table 1.

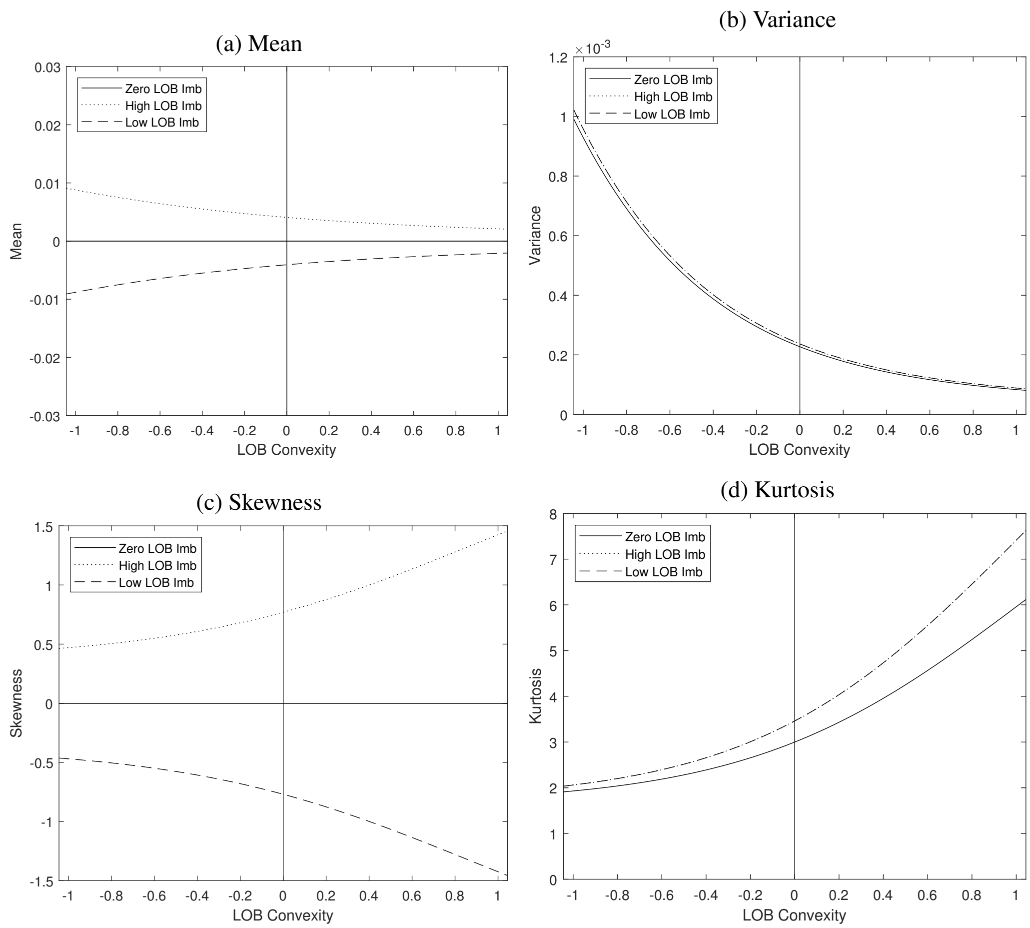


Figure 4: *The Effects of the Symmetric LOB Convexity on the Return Mean, Variance, Skewness and Kurtosis*
Note: This figure presents the effect of the LOB Convexity on the first four central moments of the return distribution for three different values of the LOB Imbalance, high (dotted line), zero (solid line) and low (dashed line). The high and low values of LOB imbalance are 0.7044 and -0.7044 corresponding respectively to plus and minus one times the sample standard deviation of its empirical counterpart. All other variables are calibrated to the values presented in Table 1. We consider LOB convexity values between minus and plus one times the standard deviation of its sample counterpart (0.7044), i.e. the limits of the x-axis are -0.4226 and 0.4226.

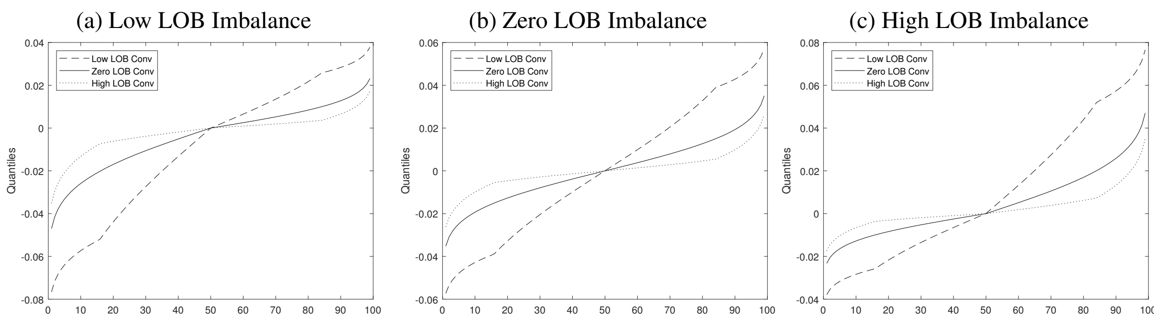


Figure 5: *The Effects of the Symmetric LOB Convexity on the Return Distribution*
Note: This figure presents the quantiles between 1% and 99% of the return distribution for three different values of LOB convexity, high (dotted line), zero (solid line) and low (dashed line) and for three different values of LOB imbalance, low (panel (a)), zero (panel (b)) and high (panel (c)). The high and low values of LOB imbalance are 0.7044 and -0.7044 and the high and low values of LOB convexity are 0.4226 and - 0.4226, corresponding respectively to plus and minus one times the sample standard deviations of their empirical counterparts. The market order imbalance is set to zero and all other variables are calibrated to the values presented in Table 1.

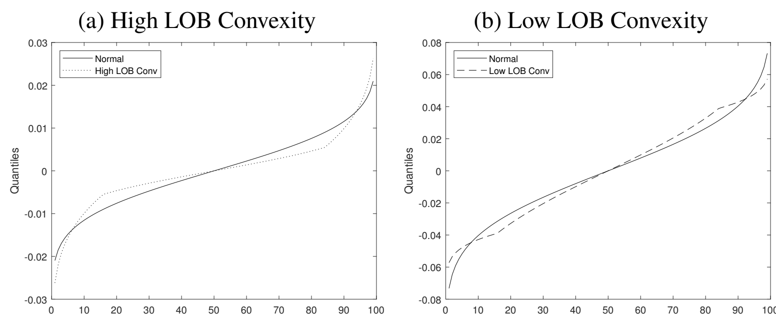


Figure 6: *The Effects of the Symmetric LOB Convexity on the Tails of the Return Distribution*
Note: This figure presents the quantiles between 1% and 99% of the simulated return distribution when LOB convexity is high in panel (a) and low in panel (b). The solid line in both panels represent the quantiles of the normal distribution with zero mean and variance given by the variance of the corresponding simulated returns. The high and low values of LOB convexity are 0.4226 and - 0.4226, corresponding respectively to plus and minus one times the sample standard deviation of its empirical counterparts. The LOB imbalance and the MO imbalance are both set to zero and all other variables are calibrated to the values presented in Table 1.

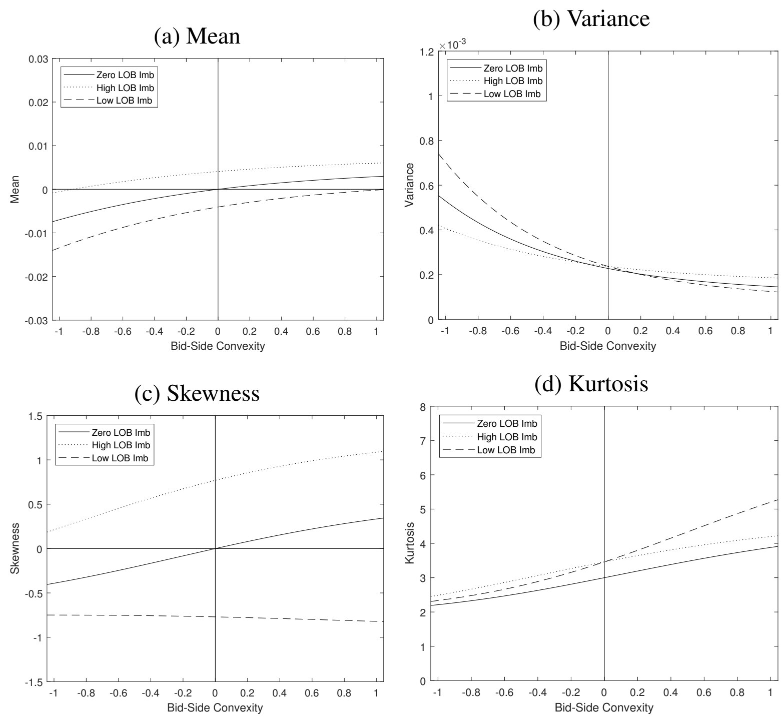


Figure 7: *The Effects of the Bid-Side Convexity on the Return Mean, Variance, Skewness and Kurtosis*
Note: This figure presents the effect of the bid-side convexity on the first four central moments of the return distribution for three different values of the LOB Imbalance, high (dotted line), zero (solid line) and low (dashed line). The high and low values of LOB imbalance are 0.7044 and -0.7044 corresponding respectively to plus and minus one times the sample standard deviation of its empirical counterpart. All other variables are calibrated to the values presented in Table 1. We consider bid-side convexity values between minus and plus one times the standard deviation of its sample counterpart (0.7044), i.e. the limits of the x-axis are -0.4226 and 0.4226.

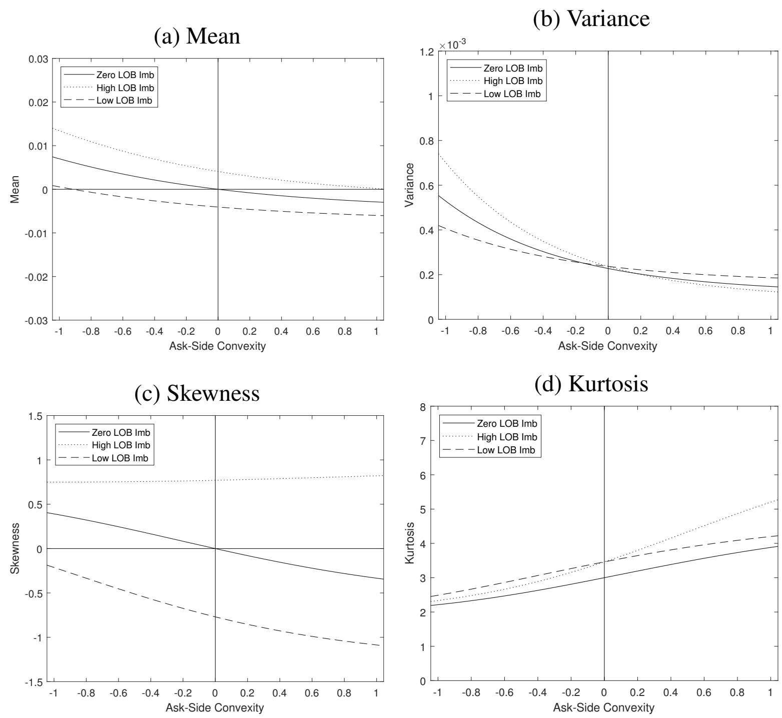


Figure 8: *The Effects of the Bid-Side Convexity on the Return Mean, Variance, Skewness and Kurtosis*
Note: This figure presents the effect of the ask-side convexity on the first four central moments of the return distribution for three different values of the LOB Imbalance, high (dotted line), zero (solid line) and low (dashed line). The high and low values of LOB imbalance are 0.7044 and -0.7044 corresponding respectively to plus and minus one times the sample standard deviation of its empirical counterpart. All other variables are calibrated to the values presented in Table 1. We consider ask-side convexity values between minus and plus one times the standard deviation of its sample counterpart (0.7044), i.e. the limits of the x-axis are -0.4226 and 0.4226.

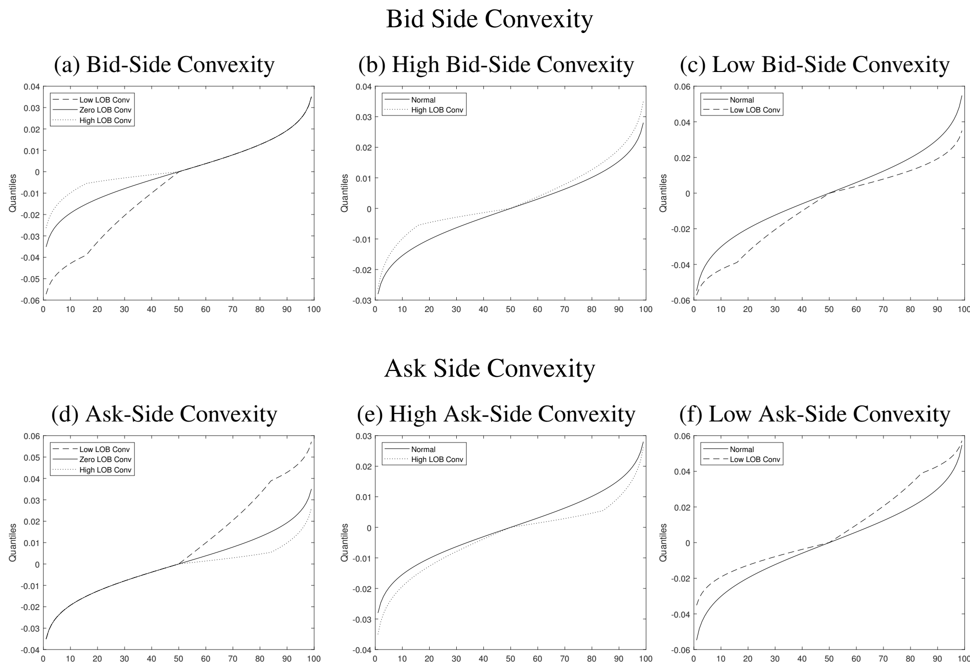


Figure 9: *The Effects of the Ask- and Bid-Side Convexities on the Return Distribution*
Note: Panels (a) and (d) of this figure present the quantiles between 1% and 99% of the return distribution for high (dotted line), zero (solid line) and low (dashed line) values of bid- and ask-side convexities, respectively. Panels (b) and (e) present the quantiles between 1% and 99% of the simulated return distribution for high values of bid- and ask-side convexities, respectively. Panels (c) and (f) present the quantiles between 1% and 99% of the simulated return distribution for high values of bid- and ask-side convexities, respectively. The solid lines panels (b), (c), (e) and (f) represent the quantiles of the normal distribution with zero mean and variance given by the variance of the corresponding simulated returns. The high and low values of LOB convexity in all panels are 0.4226 and - 0.4226, corresponding respectively to plus and minus one times the sample standard deviation of its empirical counterparts. The LOB imbalance and the MO imbalance are both set to zero and all other variables are calibrated to the values presented in Table 1.

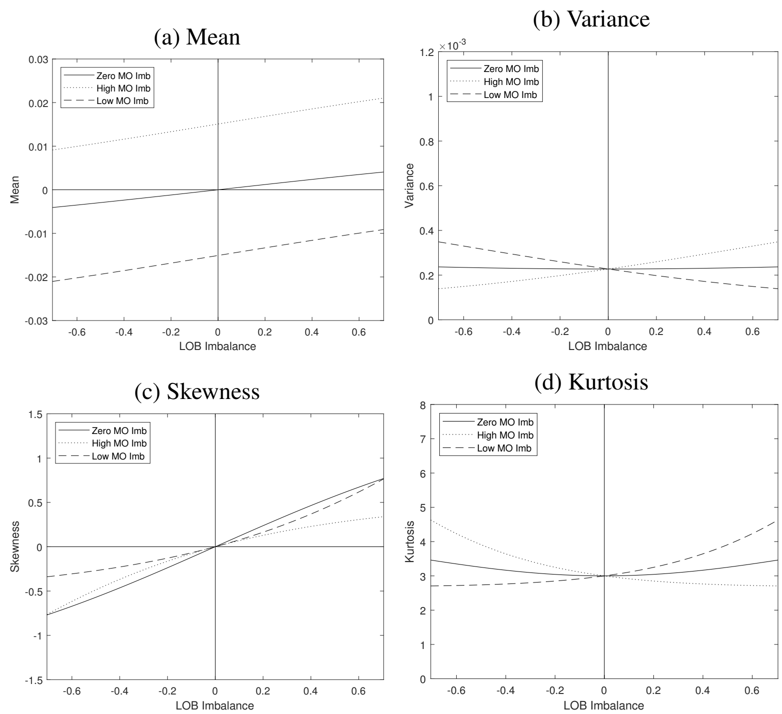


Figure 10: *The Effects of the LOB Imbalance on the Return Mean, Variance, Skewness and Kurtosis conditional on MO Imbalance*

Note: This figure presents the effect of the LOB Imbalance on the first four central moments of the return distribution for three different values of the MO imbalance, high (dotted line), zero (solid line) and low (dashed line). The high and low values of MO imbalance 559 and -559 corresponding respectively to plus and minus one times the average of the sample standard deviations (std) of its empirical counterparts. LOB convexity is set to zero. All other variables are calibrated to the values presented in Table 1. We consider LOB imbalance values between minus and plus one times the standard deviation of its sample counterpart (0.7044), i.e. the limits of the x-axis are -0.7044 and 0.7044.

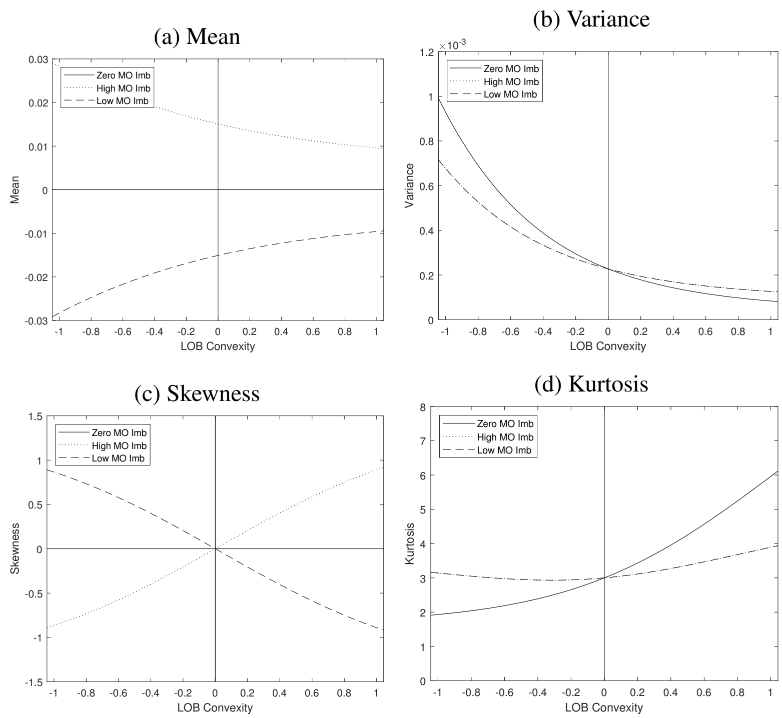


Figure 11: *The Effects of the LOB Convexity on the Return Mean, Variance, Skewness and Kurtosis conditional on MO Imbalance*
Note: This figure presents the effect of the LOB Convexity on the first four central moments of the return distribution for three different values of the MO imbalance, high (dotted line), zero (solid line) and low (dashed line). The high and low values of MO imbalance 559 and -559 corresponding respectively to plus and minus one times the average of the sample standard deviations (std) of its empirical counterparts. LOB Imbalance is set to zero. All other variables are calibrated to the values presented in Table 1. We consider LOB Convexity values between minus and plus one times the standard deviation of its sample counterpart (0.7044), i.e. the limits of the x-axis are -0.7044 and 0.7044.

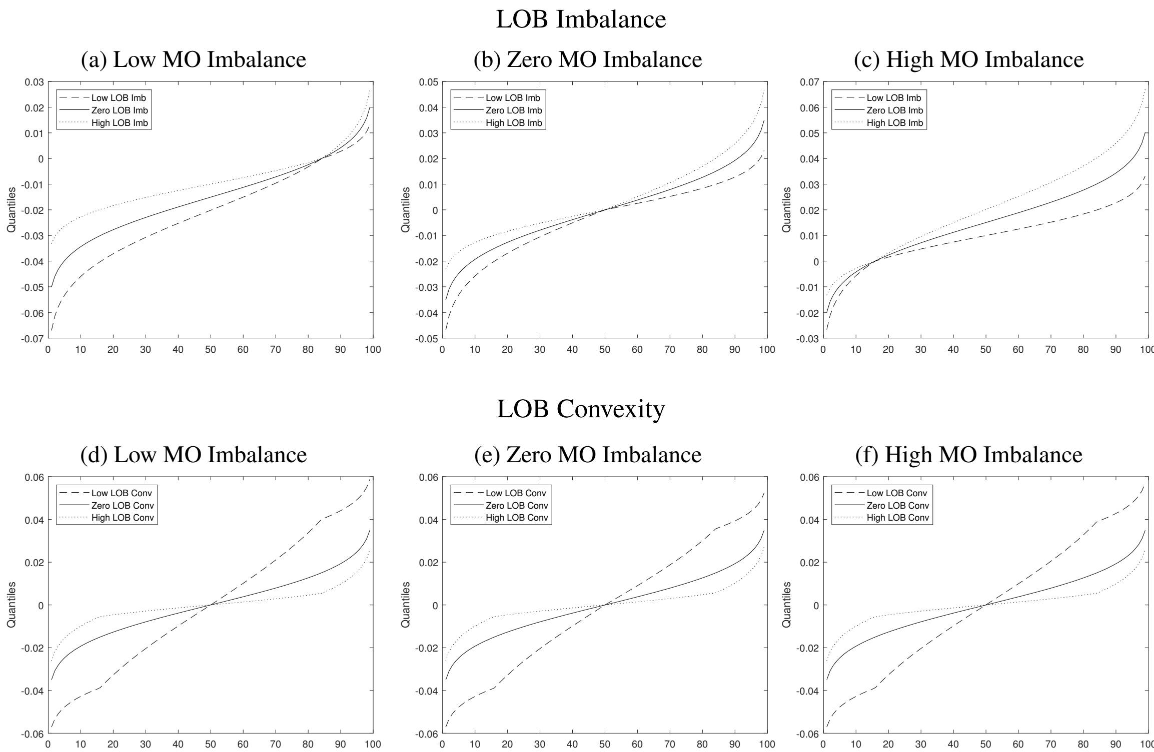


Figure 12: *The Effects of the Ask- and Bid-Side Convexities on the Return Distribution*
Note: Panels (a) and (d) of this figure present the quantiles between 1% and 99% of the return distribution for high (dotted line), zero (solid line) and low (dashed line) values of bid- and ask-side convexities, respectively. Panels (b) and (e) present the quantiles between 1% and 99% of the simulated return distribution for high values of bid- and ask-side convexities, respectively. Panels (c) and (f) present the quantiles between 1% and 99% of the simulated return distribution for high values of bid- and ask-side convexities, respectively. The solid lines panels (b), (c), (e) and (f) represent the quantiles of the normal distribution with zero mean and variance given by the variance of the corresponding simulated returns. The high and low values of LOB convexity in all panels are 0.4226 and - 0.4226, corresponding respectively to plus and minus one times the sample standard deviation of its empirical counterparts. The LOB imbalance and the MO imbalance are both set to zero and all other variables are calibrated to the values presented in Table 1.

Table 1: Calibration of Model Parameters

Parameter	Symbol	Value
Market Order Mean	μ	0.0000
Market Order Std. Dev.	σ	559
Avg. LOB Slope ($\times 10000$)	S_{avg}	0.2694
LOB Cutoff Quantity Ask	K_A	559
LOB Cutoff Quantity Bid	K_B	559
LOB Total Quantity Ask	$\bar{D}_A$	10481
LOB Total Quantity Bid	$\bar{D}_B$	10481
(log) LOB Imbalance	i	0.0542
(log) LOB Convexity Ask	c_A	0.4426
(log) LOB Convexity Bid	c_B	0.4426

Note: This table presents the names and symbols of the model parameters as well as their calibrated values. We compute the empirical counterparts of each model parameter for each stock and trading day in our sample presented in Section 4, and use the median of these variables over all stocks and trading days to calibrate these model parameters. The mean of the market order distribution is not calibrated but rather set to zero. In addition to the mean of the market order distribution, the only other variable that is not calibrated to the median of its sample counterpart is the LOB convexity, the other primary LOB shape parameter. We consider three different values of (log) LOB convexity to analyze whether the effect of the LOB imbalance on the return distribution changes as a function of LOB convexity. We assume in this section that the ask and bid side convexities are always equal to each other, i.e. $c_A = c_B = c$ and can take on three different values: $c = -0.4226, 0, 0.4226$ where 0.4226 is the average of the sample standard deviations (std) of c_A and c_B, i.e. $(std(c_A) + std(c_B))/2 = 0.4226$.

Table 2: Summary Statistics

Variable	Symbol	Mean	Median	Std. Dev.	5% Perc.	95% Perc.	Autocorr
Return Mean	M^1	0.0006	0.0000	0.0304	-0.0405	0.0426	-0.0217
Return Variance	M^2	0.0035	0.0028	0.0026	0.0012	0.0083	0.7848
Return Skewness	M^3	0.0075	-0.0002	1.2313	-1.8598	1.9035	0.0152
Return Kurtosis	M^4	6.9787	5.3303	5.4984	3.0136	16.3946	0.0821
MO Imb Mean	μ	17.7227	2.9238	147.8520	-52.8244	141.4816	0.1468
MO Imb Std. Dev.	θ	1291	559	3621	148	4409	0.2576
MO Imb Skewness	-	0.6834	0.3327	35.4439	-49.2652	50.5662	0.0087
MO Imb Kurtosis	-	1983	236	7954	9	8974	0.6837
LOB Cum Quant Ask (Levels 1-10)	Q_A	21193	11104	114431	2924	60692	0.9502
LOB Cum Quant Bid (Levels 1-10)	Q_B	19628	9858	111737	2731	55644	0.9519
LOB Slope Ask (Levels 1-10) ($\times 10000$)	S_A	1.8611	0.2575	25.5464	0.0224	5.0411	0.8241
LOB Slope Bid (Levels 1-10) ($\times 10000$)	S_B	2.0945	0.2813	23.4200	0.0237	6.9309	0.8409
(Log) LOB Imb (Levels 1-10)	I	0.1226	0.0542	0.7040	-0.9094	1.3425	0.6575
(Log) Avg. LOB Conv (Levels 2-10)	C_{avg}	0.4422	0.1412	1.3494	-1.2305	3.0996	0.9237
(Log) Avg. LOB Conv (Levels 3-10)	C_{avg}	0.4433	0.1768	1.1850	-1.0074	2.8192	0.9117
(Log) Avg. LOB Conv (Levels 5-10)	C_{avg}	0.6648	0.4426	1.0431	-0.6151	2.7447	0.8916

Note: This table presents some summary statistics on the variables used in our empirical analysis. Symbol presents the corresponding symbol of the variable in our reduced-form price formation model. If the symbol is missing, the variable is not in our model but used in our empirical analysis. The summary statistics are computed over all stock-day pairs in our sample. The return mean is the daily (log) returns from CRSP. The j^{th} non-central moment of returns $E[r_{it}^j] = \sum_{m=1}^{78} r_{itm}^j$ for $j = 2, 3, 4$ are computed using 5-minute (log) midquote returns from TAQ in a given trading day. The central moments are obtained as per their usual definitions based on non-central moments assuming that 5-minute returns have zero mean. The first four central moments of the market order distribution are, respectively, the first four central moments of the signed volume in shares based on the Lee-Ready algorithm. LOB Cum Quant Ask (Levels 1-10) and LOB Cum Quant Bid (Levels 1-10) are the total quantities available in the first ten levels of the ask and bid sides of the limit order book, respectively. LOB Slope Ask (Levels 1-10) ($\times 10000$) and LOB Slope Bid (Levels 1-10) ($\times 10000$) are the 10000 times the slopes of the first ten levels of the ask and bid sides, respectively, defined in Equation 4. (Log) LOB Imb (Levels 1-10) is the (log) limit order book imbalance defined as the ratio of the slopes of the first ten levels of the bid and ask sides as in Equation 10. (Log) Avg. LOB Conv (Levels x-10) is the (log) limit order book convexity defined as the average of the bid and ask sides convexities, which are computed as the ratios of the slopes of a given side between levels 1-x and x-10 as in Equations 11 and 12.

Table 3: The Effect of Limit Order Book Shape and Market Order Distribution

(a) Economic Importance

	Mean	Variance	Skewness	Kurtosis
MO Imb Mean	0.00196***	0.00000	0.02645***	0.00044
MO Imb Std. Dev.	0.00023*	0.00004***	-0.00057	0.10216***
MO Imb Skewness	-0.00044***	-0.00001***	0.00911***	0.00145
MO Imb Kurtosis	0.00022***	0.00032***	-0.00139	-0.08671***
LOB Avg. Cum Quant (t-1)	0.00027***	0.00047***	-0.01083***	-0.18896***
LOB Avg. Slope (t-1)	-0.00086***	0.00031***	-0.00917***	0.06307**
LOB Imb (t-1)	0.00021***	0.00026***	0.01340***	-0.13162***
LOB Conv (t-1)	0.00034***	-0.00042***	0.00333*	0.18609***

(b) Ranking

	Mean	Variance	Skewness	Kurtosis
MO Imb Mean	1	8	1	8
MO Imb Std. Dev.	6	6	8	4
MO Imb Skewness	3	7	5	7
MO Imb Kurtosis	7	3	7	5
LOB Avg. Cum Quant (t-1)	5	1	3	1
LOB Avg. Slope (t-1)	2	4	4	6
LOB Imb (t-1)	8	5	2	3
LOB Conv (t-1)	4	2	6	2

Note: Panel (a) of this table presents the economic significance of market order distribution and LOB shape parameters. The economic significance is computed as the coefficient of a given variable from the estimation of the regression model in Equation 14 with firm fixed effects times the standard deviation of that variable computed over all stock-day pairs in our sample. The standard errors are clustered at the firm level. Panel (b) presents the ranking of each variable among the eight variables of interest, i.e. market order distribution and LOB shape parameters, excluding the other control variables, i.e. the other three moments of the return distribution. A smaller ranking indicates a higher economic importance, with one being the most important and eight being the least important variable among the eight variables.

Table 4: The Effect of LOB Imbalance Conditional on Market Order Distribution and LOB Convexity

(a) Conditional on Market Order Imbalance

	Mean	Variance	Skewness	Kurtosis
Low MO Imb	-0.00022***	0.00048***	0.02058***	-0.16877***
Zero MO Imb	0.00004	0.00048***	0.01910***	-0.16551***
High MO Imb	0.00030***	0.00048***	0.01761***	-0.16224***

(b) Conditional on Market Order Variance

	Mean	Variance	Skewness	Kurtosis
Low MO Std	-0.00008	0.00049***	0.01635***	-0.15811***
Zero MO Std	0.00004	0.00048***	0.01910***	-0.16551***
High MO Std	0.00017***	0.00047***	0.02184***	-0.17291***

(c) Conditional on LOB Convexity

	Mean	Variance	Skewness	Kurtosis
Low LOB Conv	0.00001	0.00059***	0.02234***	-0.18316***
Zero LOB Conv	0.00004	0.00048***	0.01910***	-0.16551***
High LOB Conv	0.00007*	0.00038***	0.01585***	-0.14785***

Note: This table presents the economic significance of the effects of the LOB Imbalance on the first four moments of the return distribution conditional on the MO imbalance mean in panel (a), MO imbalance variance in panel (b) and LOB convexity in panel (c). The empirical specification with interaction terms in Equation 15 is estimated with firm fixed effects. The economic significance of the effect of LOB imbalance conditional on, say high MO imbalance, is computed as $(\hat{\delta}_3 + \hat{\delta}_5 \times \sigma(\text{MO Imb})) \times \sigma(\text{LOB Imb})$ where σ(MO Imb and σ(LOB Imb are the sample standard deviations of MO and LOB imbalances, respectively, computed over all stock-day pairs in our sample. The high and low values of MO Imbalance Mean and Variance and LOB convexity are set to plus and minus one times their standard deviations, e.g. high MO Imbalance = $+\sigma(MOImb)$ and Low MO Imbalance = $-\sigma(MOImb)$. The standard errors are clustered at the firm level.

Table 5: The Effect of LOB Convexity Conditional on Market Order Distribution and LOB Imbalance

(a) Conditional on Market Order Imbalance

	Mean	Variance	Skewness	Kurtosis
Low MO Imb	0.00079***	-0.00038***	0.00271	0.17620***
Zero MO Imb	0.00028***	-0.00034***	0.00228	0.16199***
High MO Imb	-0.00022**	-0.00030***	0.00185	0.14779***

(b) Conditional on Market Order Variance

	Mean	Variance	Skewness	Kurtosis
Low MO Std	0.00027*	-0.00042***	-0.00126	0.21422***
Zero MO Std	0.00028***	-0.00034***	0.00228	0.16199***
High MO Std	0.00030***	-0.00026***	0.00582***	0.10977***

(c) Conditional on LOB Convexity

	Mean	Variance	Skewness	Kurtosis
Low LOB Imb	0.00025***	-0.00023***	0.00553***	0.14434***
Zero LOB Imb	0.00028***	-0.00034***	0.00228	0.16199***
High LOB Imb	0.00032***	-0.00045***	-0.00097	0.17965***

Note: This table presents the economic significance of the effects of the LOB Imbalance on the first four moments of the return distribution conditional on the MO imbalance mean in panel (a), MO imbalance variance in panel (b) and LOB convexity in panel (c). The empirical specification with interaction terms in Equation 15 is estimated with firm fixed effects. The economic significance of the effect of LOB imbalance conditional on, say high MO imbalance, is computed as $(\hat{\delta}_3 + \hat{\delta}_5 \times \sigma(\text{MO Imb})) \times \sigma(\text{LOB Imb})$ where σ(MO Imb and σ(LOB Imb are the sample standard deviations of MO and LOB imbalances, respectively, computed over all stock-day pairs in our sample. The high and low values of MO Imbalance Mean and Variance and LOB convexity are set to plus and minus one times their standard deviations, e.g. high MO Imbalance = $+\sigma(MOImb)$ and Low MO Imbalance = $-\sigma(MOImb)$. The standard errors are clustered at the firm level.

Table 6: The Effect of Limit Order Book Shape and Market Order Distribution

(a) Economic Importance

	q_5	q_{10}	q_{25}	q_{50}	q_{75}	q_{90}	q_{95}
MO Imb Mean	0.00270***	0.00264***	0.00257***	0.00255***	0.00257***	0.00263***	0.00268***
MO Imb Std. Dev.	0.00022	0.00024	0.00026	0.00026	0.00026	0.00028	0.00030*
MO Imb Skewness	-0.00039***	-0.00040***	-0.00042***	-0.00043***	-0.00043***	-0.00043***	-0.00042***
MO Imb Kurtosis	-0.00033***	-0.00019**	0.00000	0.00016**	0.00032***	0.00050***	0.00064***
LOB Avg. Cum Quant (t-1)	-0.00070***	-0.00046***	-0.00017***	0.00007	0.00029***	0.00057***	0.00080***
LOB Avg. Slope (t-1)	-0.00172***	-0.00150***	-0.00128***	-0.00117***	-0.00107***	-0.00091***	-0.00074***
LOB Imb (t-1)	-0.00004	0.00007***	0.00020***	0.00031***	0.00043***	0.00059***	0.00072***
LOB Conv (t-1)	0.00121***	0.00101***	0.00076***	0.00055***	0.00034***	0.00011*	-0.00008

(b) Ranking

	q_5	q_{10}	q_{25}	q_{50}	q_{75}	q_{90}	q_{95}
MO Imb Mean	1	1	1	1	1	1	1
MO Imb Std	7	6	5	6	8	7	7
MO Imb Skew	5	5	4	4	3	6	6
MO Imb Kurt	6	7	8	7	6	5	5
(Log) Cum Quant Ask (t-1)	4	4	7	8	7	4	2
(Log) Avg LOB Slope (t-1)	2	2	2	2	2	2	3
(Log) LOB Imb (t-1)	8	8	6	5	4	3	4
(Log) LOB Conv (t-1)	3	3	3	3	5	8	8

Note: Panel (a) of this table presents the economic significance of market order distribution and LOB shape parameters in determining the quantiles of the return distribution. The economic significance is computed as the coefficient of a given variable from the estimation of the regression model in Equation 17 with firm fixed effects times the standard deviation of that variable computed over all stock-day pairs in our sample. The standard errors are clustered at the firm level. Panel (b) presents the ranking of each variable among the eight variables of interest, i.e. market order distribution and LOB shape parameters, excluding the other control variables, i.e. the other three moments of the return distribution. A smaller ranking indicates a higher economic importance, with one being the most important and eight being the least important variable among the eight variables.

Table 7: Differences-in-differences Analysis of Return Moments, Market Order Distribution and LOB Shape Parameters around RegSHO

(a) Moments of the Return Distribution

	Pilot			Control			
Ret	Pre	Post	Diff	Pre	Post	Diff	Diff-Diff
Mean	-0.0001	0.0023	0.0023***	-0.0003	0.0024	0.0027***	-0.0004
Variance	26.4197	25.2810	-1.1387***	26.0983	23.9229	-2.1754***	1.0367***
Skewness	0.0157	0.0752	0.0595***	0.0087	0.0663	0.0576***	0.0019
Kurtosis	6.6908	6.3131	-0.3777***	6.6627	6.8309	0.1682***	-0.5459***

(b) Moments of the Market Order Distribution

	Pilot			Control			
	Pre	Post	Diff	Pre	Post	Diff	Diff-Diff
Mean	42.2960	7.9547	-34.3413	46.3519	38.5226	-7.8292***	-26.5121***
Variance	1,842.6188	1,788.4930	-54.1258	1,927.1735	1,814.0361	-113.1373	59.0115
Skewness	1.6515	0.7004	-0.9510	1.7343	1.6325	-0.1018***	-0.8492***
Kurtosis	345.3823	375.0246	29.6423	345.2019	355.7659	10.5640***	19.0783***

(c) Shape of the LOB

	Pilot			Control			
LOB	Pre	Post	Diff	Pre	Post	Diff	Diff-Diff
Log Imb	0.1197	0.0304	-0.0893***	0.1338	-0.0110	-0.1448***	0.0555***
Log Conv	1.4181	1.1714	-0.2467***	1.4317	1.2638	-0.1679***	-0.0788***
Avg Cum Quant	5.0935	5.0856	-0.0079***	5.1069	5.1158	0.0089***	-0.0168***
Avg Slope	-9.6307	-9.8340	-0.2032***	-9.6657	-9.8252	-0.1595***	-0.0437**

Note: This table presents the results of a differences-in-differences analysis around the suspension of the uptick rule following the adaptation of Rule 202T on May 2, 2005. The pilot stocks are the stocks that were exclude from the operation of any short sale price test rule, such as the uptick rule. Every third stock in the Russell 3000 index sorted by volume was chosen as a pilot stock while the remaining 2,000 stocks constitute the control stocks. Our sample includes the pilot and control stocks traded on the NYSE and excludes other stocks. The numbers represent the averages of the corresponding variable in the row over stocks and period specified in the columns. The columns titled "Pre" and "Post" cover respectively the three month period before and after the suspension of the uptick rule on May 2, 2005. The columns titled "Pilot" and "Control" present the results for the pilot and control stocks. The columns titled "Diff" presents the differences in means between the pre and post periods for pilot and control stocks. The column titled "Diff-Diff" presents the differences in differences between pilot and control stocks. The standard errors are HAC standard errors.

Proof of Proposition 1

Recall from Equation 3 that the return is given by

$$\begin{aligned} r &= S_{A,low}Q1_{\{0<Q<K_A\}} + ((S_{A,low} - S_{A,high})K_A + S_{A,high}Q)1_{\{Q\geq K_A\}} \\ &+ S_{B,low}Q1_{\{-K_B<Q<0\}} + ((S_{B,high} - S_{B,low})K_B + S_{B,high}Q)1_{\{Q\leq -K_B\}} \end{aligned}$$

Dividing the first two components by $S_{A,low}$ and the last two components by $S_{B,low}$ and using the definitions of ask and bid side convexities, i.e. $C_A = S_{A,high}/S_{A,low}$ and $C_B = S_{B,high}/S_{B,low}$ respectively, yields

$$\begin{aligned} r &= S_{A,low}\Big[Q1_{\{0<Q<K_A\}} + ((1 - C_A)K_A + C_AQ)1_{\{Q\geq K_A\}}\Big] \\ &+ S_{B,low}\Big[Q1_{\{-K_B<Q<0\}} + ((C_B - 1)K_B + C_BQ)1_{\{Q\leq -K_B\}}\Big] \end{aligned}$$

These equations imply that the return is composed of four mutually exclusive components defined by the indicator functions of the market order size (Q). For example, the first component $S_{A,low}Q1_{\{0<Q<K_A\}}$ is a normal variable with a mean $S_{A,low}\mu$ and a variance $(S_{A,low}\theta)^2$ truncated below at zero and above at K_A. This in turn implies that the j^{th} non-central moment of the return can be written as follows as the sum of the non-central moments of each component which themselves are truncated normal variables:

$$\begin{aligned} E[r^j] &= S^j_{A,low}\Big[M_j(\mu,\sigma,0,K_A)\Big(\Phi(K_A,\mu,\sigma) - \Phi(0,\mu,\sigma)\Big) \\ &+ M_j(C_A\mu + (1-C_A)K_A, C_A\sigma, K_A, +\infty)\Big(1 - \Phi(K_A, C_A\mu + (1-C_A)K_A, C_A\sigma)\Big)\Big] \\ &+ S^j_{B,low}\Big[M_j(\mu,\sigma,-K_B,0)\Big(\Phi(0,\mu,\sigma) - \Phi(-K_B,\mu,\sigma)\Big) \\ &+ M_j(C_B\mu + (C_B-1)K_B, C_B\sigma, -\infty, -K_B)\Phi(-K_B, C_B\mu + (C_B-1)K_B, C_B\sigma)\Big] \end{aligned}$$

where the first four non-central moments of a truncated normal distribution can be obtained as follows:

$$\begin{aligned} M_1(\mu,\sigma,a,b) &= \mu - \sigma\frac{\phi\left(\frac{b-\mu}{\sigma}\right) - \phi\left(\frac{a-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)} && (18)\\ M_2(\mu,\sigma,a,b) &= \mu^2 + \sigma^2 - \sigma\frac{(\mu+b)\phi\left(\frac{b-\mu}{\sigma}\right) - (\mu+a)\phi\left(\frac{a-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)} && (19)\\ M_3(\mu,\sigma,a,b) &= 2\sigma^2 M_1(\mu,\sigma,a,b) + \mu M_2(\mu,\sigma,a,b) - \sigma\frac{b^2\phi\left(\frac{b-\mu}{\sigma}\right) - a^2\phi\left(\frac{a-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)} && (20)\\ M_4(\mu,\sigma,a,b) &= 3\sigma^2 M_2(\mu,\sigma,a,b) + \mu M_3(\mu,\sigma,a,b) - \sigma\frac{b^3\phi\left(\frac{b-\mu}{\sigma}\right) - a^3\phi\left(\frac{a-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)} && (21) \end{aligned}$$

Normalizing the variables in the above equation by dividing them by σ yields:

$$\begin{aligned} E[r^j] &= (S_{A,low}\sigma)^j\Big[M_j(\tilde{\mu},1,0,\tilde{K}_A)\Big(\Phi(\tilde{K}_A,\tilde{\mu},0) - \Phi(0,\tilde{\mu},1)\Big)\\ &+ M_j(C_A\tilde{\mu} + (1-C_A)\tilde{K}_A, C_A, \tilde{K}_A, +\infty)\Big(1 - \Phi(\tilde{K}_A, C_A\tilde{\mu} + (1-C_A)\tilde{K}_A, C_A)\Big)\Big]\\ &+ (S_{B,low}\sigma)^j\Big[M_j(\tilde{\mu},1,-\tilde{K}_B,0)\Big(\Phi(0,\tilde{\mu},1) - \Phi(-\tilde{K}_B,\tilde{\mu},1)\Big)\\ &+ M_j(C_B\tilde{\mu} + (C_B-1)\tilde{K}_B, C_B, -\infty, -\tilde{K}_B)\Phi(-\tilde{K}_B, C_B\tilde{\mu} + (C_B-1)\tilde{K}_B, C_B)\Big] && (22) \end{aligned}$$

where $\tilde{\mu} = \mu/\sigma$, $\tilde{K}_A = K_A/\sigma$ and $\tilde{K}_B = K_B/\sigma$ and

Using the definitions of definitions of average LOB slope ($S_{avg} = (S_A + S_B)/2$), LOB imbalance ($I = S_A/S_B$), $S_{A,low}$ and $S_{B,low}$ can be written as follows:

$$\begin{aligned} S_{A,low} &= \frac{2S_{avg}I}{I+1}\frac{\tilde{\bar{D}}_A}{\tilde{K}_A + (\tilde{\bar{D}}_A - \tilde{K}_A)C_A}\\ S_{B,low} &= \frac{2S_{avg}}{I+1}\frac{\tilde{\bar{D}}_B}{\tilde{K}_B + (\tilde{\bar{D}}_B - \tilde{K}_B)C_B} \end{aligned}$$

where $\tilde{\bar{D}}_A = \bar{D}_A/\sigma$ and $\tilde{\bar{D}}_B = \bar{D}_B/\sigma$.

Plugging these into Equation 8 yields the equation in Proposition 1.

Online Appendix

.1 Robustness Checks

[Table 8 about here.]

[Table 9 about here.]

[Table 10 about here.]

[Table 11 about here.]

[Table 12 about here.]

[Table 13 about here.]

Table OA.1: The Effect of Limit Order Book Shape and Market Order Distribution

(a) Economic Importance

	Mean	Variance	Skewness	Kurtosis
MO Imb Mean	0.00196***	0.00000	0.02696***	-0.00329
MO Imb Std. Dev.	0.00024*	0.00004***	0.00006	0.09547***
MO Imb Skewness	-0.00045***	-0.00001***	0.00910***	0.00135
MO Imb Kurtosis	0.00021***	0.00031***	-0.00216	-0.08170***
LOB Avg. Cum Quant (t-1)	0.00014***	0.00064***	-0.02345***	-0.09253***
LOB Avg. Slope (t-1)	-0.00086***	0.00062***	-0.02400***	0.08519***
LOB Imb (t-1)	0.00011***	0.00024***	0.01409***	-0.15598***
LOB Conv (t-1)	0.00017***	-0.00058***	0.00143	0.21540***

(b) Ranking

	Mean	Variance	Skewness	Kurtosis
MO Imb Mean	1	8	1	7
MO Imb Std	4	6	8	3
MO Imb Skew	3	7	5	8
MO Imb Kurt	5	4	6	6
(Log) Cum Quant Ask (t-1)	7	1	3	4
(Log) Avg LOB Slope (t-1)	2	2	2	5
(Log) LOB Imb (t-1)	8	5	4	2
(Log) LOB Conv (t-1)	6	3	7	1

Note: Panel (a) of this table presents the economic significance of market order distribution and LOB shape parameters. The economic significance is computed as the coefficient of a given variable from the estimation of the regression model in Equation 14 with firm fixed effects times the standard deviation of that variable computed over all stock-day pairs in our sample. The standard errors are clustered at the firm level. Panel (b) presents the ranking of each variable among the eight variables of interest, i.e. market order distribution and LOB shape parameters, excluding the other control variables, i.e. the other three moments of the return distribution. A smaller ranking indicates a higher economic importance, with one being the most important and eight being the least important variable among the eight variables.

Table OA.2: The Effect of Limit Order Book Shape and Market Order Distribution

(a) Economic Importance

	q_5	q_{10}	q_{25}	q_{50}	q_{75}	q_{90}	q_{95}
MO Imb Mean	0.00271***	0.00265***	0.00258***	0.00256***	0.00258***	0.00264***	0.00269***
MO Imb Std. Dev.	0.00022	0.00025	0.00027	0.00028*	0.00028*	0.00030*	0.00032*
MO Imb Skewness	-0.00039***	-0.00040***	-0.00042***	-0.00043***	-0.00043***	-0.00043***	-0.00042***
MO Imb Kurtosis	-0.00033***	-0.00019**	-0.00001	0.00015**	0.00030***	0.00047***	0.00061***
LOB Avg. Cum Quant (t-1)	-0.00133***	-0.00097***	-0.00055***	-0.00023***	0.00007	0.00044***	0.00075***
LOB Avg. Slope (t-1)	-0.00249***	-0.00211***	-0.00168***	-0.00139***	-0.00114***	-0.00081***	-0.00050***
LOB Imb (t-1)	-0.00012***	-0.00002	0.00010***	0.00021***	0.00034***	0.00048***	0.00060***
LOB Conv (t-1)	0.00127***	0.00101***	0.00066***	0.00036***	0.00006	-0.00027***	-0.00052***

(b) Ranking

	q_5	q_{10}	q_{25}	q_{50}	q_{75}	q_{90}	q_{95}
MO Imb Mean	1	1	1	1	1	1	1
MO Imb Std	7	6	6	5	6	7	8
MO Imb Skew	5	5	5	3	3	6	7
MO Imb Kurt	6	7	8	8	5	4	3
(Log) Cum Quant Ask (t-1)	3	4	4	6	7	5	2
(Log) Avg LOB Slope (t-1)	2	2	2	2	2	2	6
(Log) LOB Imb (t-1)	8	8	7	7	4	3	4
(Log) LOB Conv (t-1)	4	3	3	4	8	8	5

Note: Panel (a) of this table presents the economic significance of market order distribution and LOB shape parameters in determining the quantiles of the return distribution. The economic significance is computed as the coefficient of a given variable from the estimation of the regression model in Equation 17 with firm fixed effects times the standard deviation of that variable computed over all stock-day pairs in our sample. The standard errors are clustered at the firm level. Panel (b) presents the ranking of each variable among the eight variables of interest, i.e. market order distribution and LOB shape parameters, excluding the other control variables, i.e. the other three moments of the return distribution. A smaller ranking indicates a higher economic importance, with one being the most important and eight being the least important variable among the eight variables.

Table OA.3: The Effect of Limit Order Book Shape and Market Order Distribution

(a) Economic Importance

	Mean	Variance	Skewness	Kurtosis
MO Imb Mean	0.00196***	0.00000	0.02643***	0.00050
MO Imb Std. Dev.	0.00023*	0.00004***	-0.00061	0.10185***
MO Imb Skewness	-0.00044***	-0.00001***	0.00911***	0.00148
MO Imb Kurtosis	0.00022***	0.00033***	-0.00134	-0.08702***
LOB Avg. Cum Quant (t-1)	0.00027***	0.00045***	-0.01223***	-0.19677***
LOB Avg. Slope (t-1)	-0.00083***	0.00025***	-0.01170***	0.05648*
LOB Imb (t-1)	0.00021***	0.00026***	0.01330***	-0.13219***
LOB Conv (t-1)	0.00032***	-0.00037***	0.00549***	0.19386***

(b) Ranking

	Mean	Variance	Skewness	Kurtosis
MO Imb Mean	1	8	1	8
MO Imb Std	6	6	8	4
MO Imb Skew	3	7	5	7
MO Imb Kurt	7	3	7	5
Cum Quant Ask (t-1)	5	1	3	1
Avg LOB Slope (t-1)	2	5	4	6
LOB Imb (t-1)	8	4	2	3
LOB Conv (t-1)	4	2	6	2

Note: Panel (a) of this table presents the economic significance of market order distribution and LOB shape parameters. The economic significance is computed as the coefficient of a given variable from the estimation of the regression model in Equation 14 with firm fixed effects times the standard deviation of that variable computed over all stock-day pairs in our sample. The standard errors are clustered at the firm level. Panel (b) presents the ranking of each variable among the eight variables of interest, i.e. market order distribution and LOB shape parameters, excluding the other control variables, i.e. the other three moments of the return distribution. A smaller ranking indicates a higher economic importance, with one being the most important and eight being the least important variable among the eight variables.

Table OA.4: The Effect of Limit Order Book Shape and Market Order Distribution

(a) Economic Importance

	q_5	q_{10}	q_{25}	q_{50}	q_{75}	q_{90}	q_{95}
MO Imb Mean	0.00270***	0.00264***	0.00258***	0.00255***	0.00257***	0.00263***	0.00268***
MO Imb Std. Dev.	0.00022	0.00024	0.00026	0.00026	0.00026	0.00028	0.00030*
MO Imb Skewness	-0.00039***	-0.00040***	-0.00042***	-0.00043***	-0.00043***	-0.00043***	-0.00042***
MO Imb Kurtosis	-0.00034***	-0.00019**	-0.00001	0.00016**	0.00031***	0.00050***	0.00064***
LOB Avg. Cum Quant (t-1)	-0.00068***	-0.00044***	-0.00016***	0.00007	0.00028***	0.00055***	0.00076***
LOB Avg. Slope (t-1)	-0.00161***	-0.00142***	-0.00123***	-0.00114***	-0.00108***	-0.00095***	-0.00081***
LOB Imb (t-1)	-0.00003	0.00007***	0.00020***	0.00031***	0.00043***	0.00059***	0.00072***
LOB Conv (t-1)	0.00114***	0.00096***	0.00073***	0.00054***	0.00035***	0.00014***	-0.00002

(b) Ranking

	q_5	q_{10}	q_{25}	q_{50}	q_{75}	q_{90}	q_{95}
MO Imb Mean	1	1	1	1	1	1	1
MO Imb Std	7	6	5	6	8	7	7
MO Imb Skew	5	5	4	4	3	6	6
MO Imb Kurt	6	7	8	7	6	5	5
(Log) Cum Quant Ask (t-1)	4	4	7	8	7	4	3
(Log) Avg LOB Slope (t-1)	2	2	2	2	2	2	2
(Log) LOB Imb (t-1)	8	8	6	5	4	3	4
(Log) LOB Conv (t-1)	3	3	3	3	5	8	8
(Log) LOB Conv (t-1)	3	3	3	3	5	8	8

Note: Panel (a) of this table presents the economic significance of market order distribution and LOB shape parameters in determining the quantiles of the return distribution. The economic significance is computed as the coefficient of a given variable from the estimation of the regression model in Equation 17 with firm fixed effects times the standard deviation of that variable computed over all stock-day pairs in our sample. The standard errors are clustered at the firm level. Panel (b) presents the ranking of each variable among the eight variables of interest, i.e. market order distribution and LOB shape parameters, excluding the other control variables, i.e. the other three moments of the return distribution. A smaller ranking indicates a higher economic importance, with one being the most important and eight being the least important variable among the eight variables.

Table OA.5: The Effect of Limit Order Book Shape and Market Order Distribution

(a) Economic Importance

	Mean	Variance	Skewness	Kurtosis
MO Imb Mean	0.00204***	0.00001*	0.02290***	-0.00198
MO Imb Std. Dev.	0.00020	0.00006***	-0.00003	0.08131***
MO Imb Skewness	-0.00042***	-0.00001***	0.00401***	-0.00280
MO Imb Kurtosis	0.00017***	0.00019***	-0.00090	-0.05628***
LOB Avg. Cum Quant (t-1)	0.00027***	0.00036***	-0.01250***	-0.15913***
LOB Avg. Slope (t-1)	-0.00065***	0.00008**	-0.00979***	0.15360***
LOB Imb (t-1)	0.00031***	0.00013***	0.01311***	-0.11446***
LOB Conv (t-1)	0.00008*	-0.00015***	0.00366*	0.09367***

(b) Ranking

	Mean	Variance	Skewness	Kurtosis
MO Imb Mean	1	8	1	8
MO Imb Std	6	6	8	5
MO Imb Skew	3	7	5	7
MO Imb Kurt	7	2	7	6
(Log) Cum Quant Ask (t-1)	5	1	3	1
(Log) Avg LOB Slope (t-1)	2	5	4	2
(Log) LOB Imb (t-1)	4	4	2	3
(Log) LOB Conv (t-1)	8	3	6	4

Note: Panel (a) of this table presents the economic significance of market order distribution and LOB shape parameters. The economic significance is computed as the coefficient of a given variable from the estimation of the regression model in Equation 14 with firm fixed effects times the standard deviation of that variable computed over all stock-day pairs in our sample. The standard errors are clustered at the firm level. Panel (b) presents the ranking of each variable among the eight variables of interest, i.e. market order distribution and LOB shape parameters, excluding the other control variables, i.e. the other three moments of the return distribution. A smaller ranking indicates a higher economic importance, with one being the most important and eight being the least important variable among the eight variables.

Table OA.6: The Effect of Limit Order Book Shape and Market Order Distribution

(a) Economic Importance

	q_5	q_{10}	q_{25}	q_{50}	q_{75}	q_{90}	q_{95}
MO Imb Mean	0.00277***	0.00271***	0.00265***	0.00263***	0.00266***	0.00273***	0.00279***
MO Imb Std. Dev.	0.00015	0.00018	0.00021	0.00023	0.00024	0.00027	0.00030
MO Imb Skewness	-0.00044***	-0.00044***	-0.00045***	-0.00045***	-0.00046***	-0.00046***	-0.00047***
MO Imb Kurtosis	-0.00015***	-0.00007	0.00004	0.00014***	0.00023***	0.00034***	0.00042***
LOB Avg. Cum Quant (t-1)	-0.00050***	-0.00030***	-0.00007	0.00010**	0.00027***	0.00048***	0.00066***
LOB Avg. Slope (t-1)	-0.00107***	-0.00096***	-0.00088***	-0.00088***	-0.00089***	-0.00086***	-0.00078***
LOB Imb (t-1)	0.00031***	0.00035***	0.00041***	0.00046***	0.00053***	0.00062***	0.00070***
LOB Conv (t-1)	0.00043***	0.00035***	0.00025***	0.00018***	0.00011**	0.00001	-0.00006

(b) Ranking

	q_5	q_{10}	q_{25}	q_{50}	q_{75}	q_{90}	q_{95}
MO Imb Mean	1	1	1	1	1	1	1
MO Imb Std	8	7	6	5	6	7	7
MO Imb Skew	4	3	3	4	4	5	5
MO Imb Kurt	7	8	8	7	7	6	6
(Log) Cum Quant Ask (t-1)	3	6	7	8	5	4	4
(Log) Avg LOB Slope (t-1)	2	2	2	2	2	2	2
(Log) LOB Imb (t-1)	6	4	4	3	3	3	3
(Log) LOB Conv (t-1)	5	5	5	6	8	8	8

Note: Panel (a) of this table presents the economic significance of market order distribution and LOB shape parameters in determining the quantiles of the return distribution. The economic significance is computed as the coefficient of a given variable from the estimation of the regression model in Equation 17 with firm fixed effects times the standard deviation of that variable computed over all stock-day pairs in our sample. The standard errors are clustered at the firm level. Panel (b) presents the ranking of each variable among the eight variables of interest, i.e. market order distribution and LOB shape parameters, excluding the other control variables, i.e. the other three moments of the return distribution. A smaller ranking indicates a higher economic importance, with one being the most important and eight being the least important variable among the eight variables.