CIRRELT

Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation

Effects of the Limit Order Book on Price Dynamics

Tolga Cenesizoglu
Georges Dionne
Xiaozhou Zhou

November 2014

CIRRELT-2014-60

Bureaux de Montréal :
Université de Montréal
Pavillon André-Aisenstadt
C.P. 6128, succursale Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :
Université Laval
Pavillon Palasis-Prince
2325, de la Terrasse, bureau 2642
Québec (Québec)
Canada G1V 0A6
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

Effects of the Limit Order Book on Price Dynamics

Tolga Cenesizoglu[1], Georges Dionne[1,2,*], Xiaozhou Zhou[1,2]

[1] Department of Finance, HEC Montréal, 3000 Côte-Sainte-Catherine, Montréal, Canada H3T 2A7

[2] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT)

Abstract. In this paper, we analyze whether the state of the limit order book affects future price movements in line with what recent theoretical models predict. We do this in a linear vector autoregressive system which includes midquote return, trade direction and variables that are theoretically motivated and capture different dimensions of the information embedded in the limit order book. We find that different measures of depth and slope of bid and ask sides as well as their ratios cause returns to change in the next transaction period in line with the predictions of Goettler, Parlour, and Rajan (2009) and Kalay and Wohl (2009). Limit order book variables also have significant long term cumulative effects on midquote return, which is stronger and takes longer to be fully realized for variables based on higher levels of the book. In a simple high frequency trading exercise, we show that it is possible in some cases to obtain economic gains from the statistical relation between limit order book variables and midquote return.

Keywords. High frequency limit order book, high frequency trading, high frequency transaction price, asset price, midquote return, high frequency return.

---

* Corresponding author: Georges.Dionne@cirrelt.ca

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2014

1 Introduction

Regardless of their original trading mechanism, almost all of the world's major exchanges now feature electronic limit order books. Some of them such as Euronext Paris have completely abandoned any form of floor trading and operate as pure electronic limit order markets without any designated market makers. Others such as NASDAQ also had to adapt their trading mechanisms to reflect the growing importance of electronic limit order books originating from alternative trading systems such as Electronic Communication Networks (ECN).

As the importance of electronic limit order books in financial markets increases, so does the demand for information embedded in them. Most exchanges such as those operated by NYSE Euronext now offer investors access to historical and real-time data on their limit order books for a fee while others such as Frankfurt Borse make their electronic limit order book data available with a minor delay on their websites. More importantly, historical and real-time data on limit order books are available at ever increasing frequencies, thanks to recent technological advancements in electronic trading systems. For example, Frankfurt Borse offers historical data on its electronic limit order book including trades and quotes up to 20 levels with millisecond time stamps. Thus, there is an immense wealth of historical and real-time information embedded in high frequency limit order books available to investors.

Whether information embedded in the limit order book should have any effect on future price movements is a theoretical question. Earlier microstructure models such as Kyle (1985), Glosten and Milgrom (1985), Rock (1996) and Glosten (1994) treated limit orders as free options provided by uninformed investors to the market and susceptible to being picked off by later better informed investors. To put differently, most earlier microstructure models implicitly assumed that the limit order book cannot possibly be informative for future price movements. However, recent theoretical models allow informed investors to choose between limit and market orders and show that they indeed use not only market, as assumed in the previous literature, but also limit orders in rational expectations equilibria.[1] Regardless of the channel through which information gets embedded in the limit order book, the common prediction of these models is that limit orders should contain relevant information for the true value of the underlying asset and, thus, affect future price movements.

In this paper, we analyze whether the state of the limit order book affects future price movements in line with what the theory predicts. To this end, we reconstruct the first 20 levels of the historical limit order book every millisecond for several companies traded at Frankfurt Stock Exchange in July 2010 and June 2011 based on the data from the Xetra electronic trading system. The information embedded in high frequency limit order book is quite rich and is not easy to summarize with a single variable. Instead, we consider variables such as measures of depth and slope that are theoretically motivated and capture different dimensions of the information embedded in the limit order book. Based on existing theoretical models, we then develop our hypotheses on how depth and slope of bid and ask sides should

[1] For example, informed investors might use limit orders to avoid detection (as in Kumar and Seppi (1994)), to insure themselves against the price they might obtain for their market orders (as in Chakravarty and Holden (1995)), to take advantage of their sufficiently persistent private information (as in Kaniel and Liu (2006) and Kalay and Wohl (2009)). There is also a more recent literature on dynamic limit order markets with strategic traders, such as Foucault, Kadan, and Kandel (2005), Rosu (2009) and Goettler, Parlour, and Rajan (2009). Foucault, Kadan, and Kandel (2005) show that patient trades tend to submit limit orders while impatient ones submit market orders in equilibrium. Rosu (2009) shows that fully strategic, symmetrically informed liquidity traders can choose between market and limit orders based on their trade off between execution price and waiting costs. Goettler, Parlour, and Rajan (2009) find that limit orders tend to be submitted mostly by speculators and competition among them results in their private information being reflected in the limit order book.

affect future returns. Specifically, we argue based on Goettler, Parlour, and Rajan (2009) that an increase in depth at lower levels of the ask (bid) side results in lower (higher) future prices, while an increase in the depth at higher levels of the ask (bid) side results in higher (lower) future prices. Similarly, we expect based on Kalay and Wohl (2009) that an increase in the slope of the ask (bid) side results in higher (lower) future prices, regardless of the levels considered to measure it.

To test these hypotheses, we consider data in transaction, rather than calendar period, and calculate midquote returns and limit order book variables right after a trade, as in Hasbrouck (1991). We then estimate a linear vector autoregressive system (VAR) that includes midquote return, trade direction and each limit order book variable one at a time, while controlling for the contemporaneous effect of trade direction on returns and limit order book variables. This empirical specification allows us to analyze the effect of limit order book variables on return while still controlling for the effect of trade direction and autocorrelation in returns.

In this framework, we first focus on the coefficient estimates on the first lags of limit order book variables in the return equation, which reveal the initial effect of limit order book variables on return. Most limit order book variables considered in this paper have significant parameter estimates on their first lags in the return equation, even after controlling for lagged values of returns, trade directions and the limit order book variable itself. This in turn suggests that limit orders contain relevant information about future price movements in line with the predictions of recent theoretical models. More importantly, the coefficient estimates on the first lag of most limit order book variables have signs predicted by our hypotheses based on Goettler, Parlour, and Rajan (2009) and Kalay and Wohl (2009). To be more precise, depth, especially of the ask side, has coefficient estimates of different signs on its first lag depending on the levels used to measure it, in line with our hypothesis based on Goettler, Parlour, and Rajan (2009). On the other hand, slopes of both sides have coefficient estimates of the same sign regardless of levels used to measure them, confirming our hypothesis based on Kalay and Wohl (2009).

The coefficient estimates on the first lags of limit order book variables in the return equation provide information about the initial effects of these variables on return. However, they do not immediately reveal the long term cumulative effects of these variables on returns given that return, trade direction and limit order book variables have significant dynamics of their own. To take this into account, we calculate the impulse response functions of return to each limit order book variable. The impulse response functions suggest that the long term cumulative effects of most limit order book variables are generally in the same direction as their initial effects with few exceptions. For example, the long term cumulative effect of bid side depth depends closely on the levels considered. Bid side depth based on the first two, five and twenty levels have positive long term cumulative effects on return. On the other hand, bid side depth excluding the first level have negative long term cumulative effects on return. More importantly, our results on the long term cumulative effects of limit order book variables on return are mostly in line with our hypotheses and, thus, provide further empirical evidence from a different angle in their support.

We then analyze whether this relation is a causal one. To this end, we analyze causality between limit order book variables and return in the sense of Granger (1969). Most measures of ask and bid side depth as well as their ratio causes return to change at the next transaction period. The only exception is the bid side depth between the second and fifth levels, for which we fail to reject the null hypothesis that it does not cause return based on the F-statistic. Similarly, most measures of ask and bid side slope as well as their ratio causes return to change at the next transaction period. Once again, the exception to this is the slope measure between the second and fifth levels of ask and bid sides as well as their ratio. Furthermore, the test statistics tend to be much higher for limit order book variables measured based on lower levels (up to and including the fifth level) of the limit order book, suggesting that there is stronger empirical evidence for the relevance of information embedded in the lower levels of the limit order book.

Having analyzed the statistical relation between return and the information embedded in the limit order book, we then ask whether an investor can use this statistical relation to obtain economic gains. To this end, we consider an in-sample trading exercise similar to those considered in Kozhan and Salmon (2012). Specifically, having observed a transaction (and, thus, the return, trade direction and limit order book variables), we calculate our forecast of the midquote return at the next transaction period based on the estimated coefficients of the VAR model. We then calculate our trading signal for the next transaction period based on whether our forecast is greater or less than a threshold. In particular, we consider a forecast greater than a positive threshold to be a buy opportunity or signal and a forecast less than a symmetric negative threshold to be a sell opportunity or signal while a forecast between these two thresholds is not considered to be a strong enough signal. Tick-by-tick transaction prices tend to be quite noise. Hence, similar to a technical analysis trading strategy, we only consider trading signals when they are further confirmed by the relative movements of short- and long-run moving averages of recent transaction prices. To be more precise, we take a long position based on our trading signal, only when the short-run moving average crosses from below the long-run moving average by more than a specified amount. Similarly, we take a short position based on our trading signal, only when the short-run moving average crosses from above the long-run moving average by more than 0.06 euros. Otherwise, we do not consider a trading signal to be strong enough and, thus do not trade, if it is not confirmed by the relative movements of short- and long-run moving averages.

We start our trading exercise at the beginning of July 2010. We do not take any position until we observe a signal strong enough. When we receive such a signal, we then take a long or short position of one share depending on the signal. At every transaction period, we calculate our trading signal and reevaluate our position. Specifically, if we have a long (short) position and receive a strong enough buy (sell) signal or a signal that is not strong enough, we continue to keep our long (short) position of one share. On the other hand, if we have a long (short) position and receive a strong enough sell (buy) signal, we then close our long (short) position and hold a short (long) position of one share. We continue in this fashion until the last transaction in July 2010 and keep the position in the last transaction till trading terminates in July 2010. We evaluate the performance of trading strategies based on their cumulative returns when we sell at bid and buy at ask prices.

Trading strategies based on most limit order book variables outperform a benchmark model that does not utilize the information embedded in the limit order book. To be more specific, trading strategies based on 22 out of 30 limit order book variables outperform the benchmark model. Of this 22 trading strategies, 21 provide positive returns. Most trading strategies based on ask side or ratio variables outperform the benchmark strategy while only half of the trading strategies based on bid side variables do so. Furthermore, trading strategies based on ask side variables tend to outperform those based on the corresponding bid side variable. Also, trading strategies based on variables that capture the information embedded in the lower levels of the limit order book tend to outperform those based on variables that do not include this information. Finally, these results hold when we consider a latency of 0, 500 or 1000 milliseconds, suggesting the robustness of our results to a more realistic assumption of 500 or 1000 milliseconds latency. As one would expect, profits based on a latency of 500 or 1000 milliseconds are lower than those based on a latency of 0 milliseconds for the same threshold parameter. Furthermore, profits tend to decrease monotonically as we consider higher latencies. However, we should note that the performance of trading strategies depends on a set of chosen parameters. There are some sets of parameters for which trading strategies using limit order book information outperform the benchmark and there are others for which the opposite holds. In other words, it is feasible to obtain economic gains using information embedded in the limit order book for some sets of parameters but this is not always the case.

Our paper is related to a growing body of papers which focus on the relation between high frequency returns and the information embedded in the limit order book. Biais, Hillion, and Spatt (1995) are among the first to analyze the dynamics of limit order markets and document many interesting facts. Specifically, they find that price revisions tend to move in the direction of previous limit order flows, suggesting that the limit order book contains some relevant information for the future path of prices. In contrast, Griffiths, Smith, Turnbull, and White (2000) find that limit orders tend to have a negative impact on prices in the Toronto Stock Exchange due to the possibility that limit orders can be "picked off" by better informed investors. This finding in turn suggests that limit orders are placed by less informed investors and, thus, do not convey much relevant information about prices. On the other hand, using data from the Australian Stock Exchange, Cao, Hansch, and Wang (2009) find that the limit order book is somewhat informative with a contribution of approximately 22% to price discovery. They also find that order imbalances between the demand and supply schedules along the book are significantly related to future short-term returns, even after controlling for the autocorrelations in return, the inside spread, and the trade imbalance. Similarly, using data from NYSE's TORQ (Trades, Orders, Reports, and Quotes), Kaniel and Liu (2006) find that the informed traders prefer limit orders to market orders and, thus, limit orders are more informative than market orders. More recently, Beltran-Lopez, Giot, and Grammig (2009) also finds that factors extracted from the limit order book have non-negligible information relevant for the long run evolution of prices in the German Stock Exchange. Specifically, they find that shifts and rotations of the order book can explain between 5% to 10% of the long run evolution of prices depending on the liquidity of the asset.

Our paper differs from the previous literature in many dimensions. First of all, most papers consider a single variable that is supposed to summarize all the information embedded in the limit order book. However, as we demonstrate in this paper, the dynamic relation between return, trade direction and the state of the limit order book is more complex than what can be captured by a single variable. Instead, we consider a wide range of variables that capture different dimensions of the information embedded in the limit order book. More importantly, we show that different variables, and even sometimes the same variable measured based on different levels of the limit order book, can have different effects on the short and long term dynamics of returns. Second, most of our variables are closely related to and motivated by the recent dynamic models of limit order markets and have not been previously considered in a similar context. Furthermore, we provide some preliminary empirical evidence in support of the causal relation between limit order book variables and return as implied by these theoretical models. Third, to the best of our knowledge, we are among the first to report the impulse response function of return to limit order book variables, which allows us to analyze the long term cumulative effect of limit order book variables on return after several transactions. Fourth, most papers focus on the statistical relation between return, trade direction and the state of the limit order book and do not provide much evidence whether this statistical relation is economically important for investors. In this paper, we do this by providing some preliminary evidence from a simple trading strategy based on the statistical relation between return, trade direction and the state of the limit order book. Finally, we should note that our data set is relatively unique and allows us to capture the state of the limit order book at a higher frequency than the ones that have been used in the previous empirical literature, providing a much finer analysis. Hence, part of our analysis can be considered as new out-of-sample tests of existing theoretical models and our findings provide empirical support for these models while the rest of our analysis is novel empirical evidence that might provide some guidance to new generation of models.

The rest of the paper is organized as follows. Section 2 presents the details of our data set. Section 3 discusses the variable definitions. Section 4 develops the empirical hypotheses based on theoretical models of limit order markets. Section 6 presents the coefficient estimates of limit order book variables in the return equation, Section 7 presents impulse response functions of return to limit order book variables and Section 8 presents results on whether limit order book variables Granger cause return to change. Section 9 analyzes the economic value of information embedded in the limit order book based on a trading strategy. Section 10 discusses out-of-sample results based on another company and time period. Section 11 discusses some further robustness checks. Section 12 concludes.

2 Data

Our data comes from the automated order-driven trading system Xetra operated by the Deutsche Borse Group at Frankfurt Stock Exchange (FSE). It is the main German trading platform accounting for more than 90% of total transactions at all German exchanges. The trading and order processing (entry, revision, execution and cancellation) of Xetra system is highly computerized. Since September 20, 1999, the normal trading hours are from 9h00 to 17h30 CET (Central European Time).

The raw dataset contains all events that are tracked and sent through the data streams. We first process the raw dataset using a software called XetraParser developed by Bilodeau (2013).[2] We then reconstruct the first twenty levels of the limit order book at millisecond time intervals. The limit order book can change when either a trade is executed or a limit order is placed, modified or canceled. In the unlikely event that these two types of events have the same millisecond time stamp, we need to make an assumption on the sequence of events given that we do not observe which one of them arrived earlier. We assume that a trade is always executed before any other change to the limit order book with the same millisecond time stamp. Thus, we first modify the limit order book to reflect the trade execution

[2]We thank Yann Bilodeau for his help in constructing the dataset and comments.

before taking its snapshot. In other words, if a trade is executed at a given millisecond, then the snapshot of the limit order book for that millisecond already reflects the executed trade. To avoid any problems due to this assumption, we ignore the state of the limit order book when a trade is executed and use its snapshots either one millisecond before or after a trade. Based on these snapshots, we measure variables summarizing the state of the limit order book every millisecond. We also eliminate all data corresponding to the three call auctions during a trading day since the price during these auctions is determined based on a certain set of rules and not by trading activity.

We consider data for two blue chip stocks in the DAX30 index, namely Merck (MRK) and SAP (SAP). The two stocks are in completely different industries: Merck is world's oldest operating chemical and pharmaceutical company while SAP specializes in enterprise software and related services to manage business operations and customer relations. Merck is a relatively small company in the DAX30 index with a market capitalization of approximately 4 billion Euros at the end of June 2010 corresponding to 0.75% of total market capitalization of stocks in the DAX30 index. On the other hand, SAP had market capitalization of approximately 33 billion Euros at the end of June 2010, almost 8 times that of Merck, representing 6.25% of total market capitalization of stocks in the DAX30 index. Panel (a) of Table 1 shows that SAP is much more actively traded than Merck with an average daily volume of approximately 8 times that of Merck in July 2010. Specifically, SAP and Merck had, respectively, average daily volumes of approximately 4 million and 0.5 million shares in July 2010, corresponding to around 15% and 2% of total trading volume of the stocks in the DAX30 index. However, Merck with an average share turnover of 0.8% in July 2010 seems to be more liquid than SAP which has an average share turnover of 0.3%. Fourth, Panel (b) of Table 1 shows that the characteristics of their daily returns are also quite different in July 2010. For example, Merck had a positive return while SAP had a negative return with relatively lower volatility than that of Merck. This is also reflected in the range of returns for both companies.

[Insert Table 1 here]

Similarly, we consider two relatively different time periods, July 2010 and June 2011. Both companies are much bigger in June 2011 with market capitalizations of almost 5 and 50 million for Merck and SAP, respectively. Trading volume for both companies are lower in June 2011 as presented in Panel (a) of Table 1. This is despite of an increase in the total trading volume for all the stocks in the DAX30 index, suggesting that these stocks are relatively less traded in June 2011 compared to July 2010. Both of the companies had negative returns in June 2011 with Merck having a relatively better performance and lower volatility than SAP.

Although we analyze both companies in both periods, we present empirical results for Merck in July 2010 as our main results due to space limitations. However, given the differences between companies and time periods discussed above, we consider other results as an out-of-sample test of our main results. To this end, we discuss these additional results, which are available in an online appendix, in detail in Section 10.

3 Variable Definitions

The information embedded in the limit order book is quite rich and is not easy to summarize with a single variable. Hence, we consider several variables based on different levels of the limit order book to capture different dimensions of the information embedded in the limit order book. These variables can be categorized into two groups depending on whether they summarize information embedded in one or both sides of the limit order book.

We start with the variables that use information embedded in only one side of the limit order book. The first variable we consider is the depth between levels l_1 and l_2 of bid side , $D^B_{l_1,l_2,t}$, or the ask side, $D^A_{l_1,l_2,t}$, at period t. It is simply defined as the cumulative quantity available between levels l_1 and l_2 at period t:

$$D^B_{l_1,l_2,t} = \sum_{i=l_1}^{l_2} Q^B_{i,t}, \tag{1}$$

$$D^A_{l_1,l_2,t} = \sum_{i=l_1}^{l_2} Q^A_{i,t} \tag{2}$$

for $l_1 = 1, \ldots, 20$ and $l_2 \geq l_1$. $Q^B_{i,t}$ and $Q^A_{i,t}$ are the quantities available at the i^{th} level of the bid and ask side in period t, respectively. The depth measures the cumulative demand and supply for the stock at different levels of demand and supply schedules, respectively. In other words, the higher is the bid side depth, the higher is the overall demand for the stock and the higher is the ask side depth, the higher is the overall supply of the stock.

The second variable is the slope of bid, $S^B_{l_1,l_2,t}$, or ask, $S^A_{l_1,l_2,t}$, sides between levels l_1 and l_2 at t. It is defined as the change in the price relative to the cumulative quantity available between levels l_1 and l_2 in period t:

$$S^B_{l_1,l_2,t} = \frac{P^B_{l_2,t} - P^B_{l_1,t}}{D^B_{l_1+1,l_2,t}} \tag{3}$$

$$S^A_{l_1,l_2,t} = \frac{P^A_{l_2,t} - P^A_{l_1,t}}{D^A_{l_1+1,l_2,t}} \tag{4}$$

for $l_1 = 1, \ldots, 19$ and $l_2 > l_1$. $P^B_{l,t}$ and $P^A_{l,t}$ are the l^{th} best bid and ask prices, respectively in period t. The slope of the bid side is a measure of price sensitivity to changes in quantity demanded and is always negative. A high (in absolute value) bid-side slope coefficient implies that the bid price will decrease more for a given change in quantity demanded. Ceteris paribus, this in turn suggests that the investors are willing to buy at lower prices for the same total quantity demanded. Similarly, the slope of the ask side is a measure of price sensitivity to changes in quantity supplied and is always positive. A high ask-side slope coefficient implies that the ask price will increase more for a given change in quantity supplied, which in turn suggests that the investors are willing to sell at higher prices for the same total quantity supplied.

To combine information embedded in both sides of the limit order book, we consider the ratios of the variables introduced above. Specifically, the depth ratio is simply defined as the (normalized) difference in the cumulative quantity available at different levels of bid and ask sides:

$$DR_{l_1,l_2,t} = \frac{D^A_{l_1,l_2,t} - D^B_{l_1,l_2,t}}{D^A_{l_1,l_2,t} + D^B_{l_1,l_2,t}} \tag{5}$$

for $l_1 = 1, \ldots, 20$ and $l_2 \geq l_1$. This variable is bounded above by one and below by minus one. Positive values of depth ratio indicate that the total supply of the stock between two given levels of the ask side is greater than the total demand between the same two levels of the bid side. Furthermore, as the quantity supplied increases relative to the quantity demanded between the same two levels of the bid and ask sides, the depth ratio increases. Similarly, we define the slope ratio as the (normalized) difference in the slopes of bid and ask sides between different levels:

$$SR_{l_1,l_2,t} = \frac{S^A_{l_1,l_2,t} - |S^B_{l_1,l_2,t}|}{S^A_{l_1,l_2,t} + |S^B_{l_1,l_2,t}|} \tag{6}$$

for $l_1 = 1, \ldots, 19$ and $l_2 > l_1$. We take the absolute value of the bid slope since it is always negative by definition. This variable is also bounded between one and minus one. Positive values of slope ratio indicate that the supply schedule is steeper than the demand schedule. Furthermore, as the supply schedule becomes steeper relative to the demand schedule the same two levels of the bid and ask sides, the slope ratio increases.

[Insert Table 2 here]

Figure 1 presents two snapshots of the limit order book for Merck in July 2010. As it can easily be seen from Figure 1, both bid and ask sides of the limit order book can take on different shapes at different points in time. The limit order book variables are designed to summarize these different shapes that the limit order book can take on. To see what information different limit order book variables capture, Table 2 presents the values of the limit order book variables that correspond to these two snapshots. For example, it is easy to see from Figure 1 that there is more total demand and less total supply in the first snapshot compared to the second one. These facts are captured by $D^B_{1,20}$ and $D^A_{1,20}$. Furthermore, there seems to be more total demand and less total supply in the higher levels of the limit order book in the first snapshot compared to the second one. This can indeed be verified by comparing $D^B_{5,20}$ and $D^A_{5,20}$ of the first snapshot to those of the second one. As another example, compare the slopes of the bid side in these two snapshots. The bid side in the first snapshot appears to be overall steeper than that in the second snapshot, which is confirmed by comparing $S^B_{1,20}$'s. On the other hand, the bid side of the first snapshot seems to be flatter compared to the second snapshot when we focus on the first two and five levels. This turns out to be the case when we compare $S^B_{1,2}$ and $S^B_{1,5}$ of the two snapshots. Similarly, the ratio variables allow us to infer about the shapes of bid and ask sides with respect to each other. For example, the bid side appears to be overall steeper than the ask side in the first snapshot, which is confirmed by the fact that $SR_{1,20,t}$ is negative.

[Insert Figure 1 here]

Table 3 presents the transformations applied to the limit order book variables and summary statistics for the transformed limit order book variables as well as log returns. Although not presented in Table 3, we also analyze whether any of the limit order book variables have unit root based on augmented Dickey-Fuller test. We reject the existence of a unit root for all limit order book variables at 1% statistical significance level, suggesting the stationarity of these variables.

[Insert Table 3 here]

4 Hypotheses Development

In this section, based on recent theoretical models of limit order markets, we develop our hypotheses on the effect of limit order book variables on return. We first consider the effect of depth on return. Goettler, Parlour, and Rajan (2009) develop a theoretical model where traders optimally choose the type of order to submit and whether to acquire information about the asset. They solve for the equilibrium of this model and show that depth at different levels of the limit order book should not only be informative about future prices but also have different effects on them. Specifically, they show that there are only a few stale orders in the book since traders submitting limit orders revisit the market and resubmit orders, on average, twice as often as the true value of the asset changes. Thus, orders submitted in the higher levels of ask (bid) side suggest that the current best ask (bid) is too low (high) and hence lead to an upward (downward) revision in expectations about the true value of the asset. On the other hand, given that the transactions prices and traders' prices are, on average, equal to the true value of the asset, depth at lower levels of ask (bid) side lead to lower (higher) prices. Let us denote our hypothesis for the ask side with a and that for the bid side with b. Then, our first hypotheses on the effect of depth on return can be summarized as follows:

> *Hypothesis 1a (Goettler, Parlour, and Rajan (2009)): An increase in depth at lower levels of the ask side results in lower future prices, while an increase in the depth at higher levels of the ask side results in higher future prices.*

> *Hypothesis 1b (Goettler, Parlour, and Rajan (2009)): An increase in depth at lower levels of the bid side results in higher future prices, while an increase in the depth at higher levels of the bid side results in lower future prices.*

We then consider the effect of slope on return. Kalay and Wohl (2009) analyze the relation between slope and future returns in the Noisy Rational Expectations Equilibrium (NREE) models of Hellwig (1980), Kyle (1989), Admati (1985), and Easley and O'hara (2004). Specifically, they consider the framework of Hellwig (1980) and solve for the equilibrium. Given their assumption that limit orders can only be submitted by informed traders, as the number of informed traders on the bid side increases, the bid side becomes flatter, or equivalently, the slope of the bid side decreases. An increase in the number of informed traders on the bid side also implies a decrease in the relative importance of liquidity traders on the bid side. This in turn results in a decrease in the future price of the stock. The opposite model holds for the ask side. Thus, our hypotheses can be summarized as follows:

> *Hypothesis 2a (Kalay and Wohl (2009)): An increase in the slope of the ask side results in higher future prices.*

Hypothesis 2b (Kalay and Wohl (2009)): An increase in the slope of the bid side results in lower future prices.

We should note that the predictions of Kalay and Wohl (2009) on the effect of slope on return do not depend on the levels of the limit order book used to measure it, unlike the predictions of Goettler, Parlour, and Rajan (2009) for the effect of depth on return. However, we still consider slope measured based on different levels of the limit order book to analyze whether this prediction of Kalay and Wohl (2009) holds.

We now consider the effect of depth and slope ratios on return. Recall that the depth ratio is defined as the normalized difference between ask and bid side depths. Given our hypotheses 1a and 1b, one would expect to observe an effect of depth ratio on return similar to but stronger than those proposed in Hypothesis 1a and 1b. Similarly, recall that the slope ratio is defined as the normalized difference between ask and bid side slopes. Given our hypotheses 2a and 2b, one would also expect to observe an effect of slope ratio on return similar to but stronger than those proposed in Hypothesis 2a and 2b. Let *c* denote the combined versions of hypotheses a and b. Then our hypotheses on the effect of depth and slope ratios on return can then be summarized as follows:

Hypothesis 1c (Goettler, Parlour, and Rajan (2009)): An increase in depth ratio based on lower levels of the limit order book results in lower future prices while an increase in the depth ratio based on higher levels results in higher future prices.

Hypothesis 2c (Kalay and Wohl (2009)): An increase in the slope ratio results in higher future prices.

5 The Empirical Model

In this section, we discuss the empirical model that we use to test our hypotheses. To this end, we follow Hasbrouck (1991) who shows that a vector autoregressive system for the interactions between return and trade direction is consistent with stylized market microstructure models such as Glosten and Milgrom (1985). Specifically, Hasbrouck (1991) suggests using the following vector autoregressive model to analyze the effects of information embedded in trades on prices:

$$r_t = \sum_{\tau=1}^{\infty} \alpha_{r,\tau} r_{t-\tau} + \sum_{\tau=1}^{\infty} \alpha_{x,\tau} x_{t-\tau} + \varepsilon_{r,t} \tag{7a}$$

$$x_t = \sum_{\tau=1}^{\infty} \beta_{r,\tau} r_{t-\tau} + \sum_{\tau=1}^{\infty} \beta_{x,\tau} x_{t-\tau} + \varepsilon_{x,t} \tag{7b}$$

where t indexes trades, x_t is the sign of the trade in period t (+1 for a trade initiated by a buyer and -1 for a trade initiated by a seller), r_t is the midquote return defined as the change in the average of best bid and ask quotes between period $t-1$ and t, i.e. $r_t = \Delta q_t = q_t - q_{t-1}$ and q_t is the simple average of best bid and ask quotes in period t. This is a very general and flexible model that nests many of the standard microstructure models as special cases. The disturbances in this framework, $\varepsilon_{r,t}$ and $\varepsilon_{x,t}$, are generally modeled as white noise processes and can be interpreted as public information embedded in unexpected returns and private information embedded in unexpected trades, respectively.

In this paper, we assume that the dynamics of a limit order market can also be approximated by a linear vector autoregressive system similar to that proposed by Hasbrouck (1991). Specifically, we consider each variable summarizing different limit order book-related information separately as a third state variable in the VAR in addition to return and trade direction:

$$r_t = \sum_{\tau=1}^{\infty} \alpha_{r,\tau} r_{t-\tau} + \sum_{\tau=1}^{\infty} \alpha_{x,\tau} x_{t-\tau} + \sum_{\tau=1}^{\infty} \alpha_{z,\tau} z_{t-\tau} + \varepsilon_{r,t} \tag{8a}$$

$$x_t = \sum_{\tau=1}^{\infty} \beta_{r,\tau} r_{t-\tau} + \sum_{\tau=1}^{\infty} \beta_{x,\tau} x_{t-\tau} + \sum_{\tau=1}^{\infty} \beta_{z,\tau} z_{t-\tau} + \varepsilon_{x,t} \tag{8b}$$

$$z_t = \sum_{\tau=1}^{\infty} \gamma_{r,\tau} r_{t-\tau} + \sum_{\tau=1}^{\infty} \gamma_{x,\tau} x_{t-\tau} + \sum_{\tau=1}^{\infty} \gamma_{z,\tau} z_{t-\tau} + \varepsilon_{z,t} \tag{8c}$$

where z_t is a variable that summarizes a certain dimension of the information embedded in the limit order book.

This specification can be considered as a reduced form linear approximation that is designed to capture dynamics of limit order market models discussed in the introduction. Secondly, it is flexible and allows us to analyze the effect of limit order book-related information on prices while still controlling for trade-related information. For example, the immediate effect of limit order book-related information on prices is captured by $\alpha_{z,1}$. Finally, Goettler, Parlour, and Rajan (2009) argue that competition among speculators results in their private information being partially revealed in the limit order book. Hence, $\varepsilon_{z,t}$ in this framework can be interpreted as an unexpected private information shock embedded in the limit order book variable of interest.

As mentioned above, we can measure limit order book variables including best bid and ask prices every millisecond. However, a trade can only be matched to a millisecond interval and thus one needs to decide whether to take a snapshot of the limit order book right before or right after a trade. The theory does not provide much guidance on this issue. In this paper, we follow the previous literature, e.g. Hasbrouck (1991) and Dufour and Engle (2000), and measure the limit order book variables right after a trade occurs.[3]

This sampling approach implies that the midquote return and limit order book variables in period t are observed right after (less than a millisecond after) the trade, and thus trade direction. Thus, one can include trade direction in period t to control for its contemporaneous effect on return and limit order book variables in the estimated version of (8). Our results are similar regardless of whether or not we control for this contemporaneous effect. We choose to present results on the statistical relation between return and limit order book variables based on the specification that includes this contemporaneous effect.

However, we should note that any contemporaneous effects of trade direction on return and limit order book variables cannot be interpreted as a causal relation even if trade direction is observed right before these variables. This is mainly due to the fact that market participants (human or non-human) cannot be possibly reacting to and trading

[3]We also considered the alternative sampling approach of measuring the limit order book variables right before a trade occurs. Our results on the effects of limit order book variables on return do not change significantly when consider this alternative sampling approach.

based on any information embedded in contemporaneous return and limit order book variables which are only observed less than a millisecond before the trade occurs. Xetra reports that the average time required for an order to travel from the trading participant's system across the network to its back-end and for confirmation of receipt to be sent back to the participant is about 13 milliseconds.[4] Assuming that the two legs of this round trip are equally fast, it takes about 6.5 milliseconds for an order to travel from the trading participant's system across to the Xetra back-end. Hence, even if we make the unrealistic assumption that any necessary computations of an algorithmic trading strategy or the reaction of a human trader take less than a millisecond, it is physically impossible for their orders to arrive at the market within a millisecond. Hence, we choose to present our results on the trading strategy based on the empirical specification that does not include the contemporaneous effect of trade direction on return and limit order book variables.

Following Hasbrouck (1991) and Dufour and Engle (2000), we also assume that the infinite sums in the model in Equation (8) can be truncated at J lags. Furthermore, the timing convention discussed above is reflected in the starting points of the summations in the estimated version of (8). Specifically, the summations for trade direction in the equations for return and limit order book variables start at zero instead of one. Then, the estimated version of the model in Equation (8) can be expressed as follows:

$$r_t = \sum_{\tau=1}^{J} \alpha_{r,\tau} r_{t-\tau} + \sum_{\tau=0}^{J} \alpha_{x,\tau} x_{t-\tau} + \sum_{\tau=1}^{J} \alpha_{z,\tau} z_{t-\tau} + \varepsilon_{r,t} \tag{9a}$$

$$x_t = \sum_{\tau=1}^{J} \beta_{r,\tau} r_{t-\tau} + \sum_{\tau=1}^{J} \beta_{x,\tau} x_{t-\tau} + \sum_{\tau=1}^{J} \beta_{z,\tau} z_{t-\tau} + \varepsilon_{x,t} \tag{9b}$$

$$z_t = \sum_{\tau=1}^{J} \gamma_{r,\tau} r_{t-\tau} + \sum_{\tau=0}^{J} \gamma_{x,\tau} x_{t-\tau} + \sum_{\tau=1}^{J} \gamma_{z,\tau} z_{t-\tau} + \varepsilon_{z,t} \tag{9c}$$

Hasbrouck (1991) and Dufour and Engle (2000) consider $J = 5$ assuming that five lags are sufficient to capture the dynamics of the variables of interest. In this paper, we consider different lag structures up to a maximum of five lags. To be consistent with the previous literature, we present results based on a lag structure of five lags.[5]

We estimate the empirical model via ordinary least squares (OLS), which provides consistent parameter estimates for return and limit order book variables given that they are stationary. OLS also yields consistent estimates for the parameters in the trade equation even though we estimate a linear specification for a limited dependent variable. As discussed in detail in Dufour and Engle (2000), this is mainly due to the fact that the conditional mean of the trade sign is generally correctly specified given that the probability of a buy or sell is never far from 1/2. A similar argument holds for depth and slope ratios. Specifically, the OLS yields consistent estimates for the parameters in the equations for depth and slope ratios even though they are limited between -1 and +1. This is again due to the fact that their conditional mean is correctly specified since their means are not far from zero. However, these estimates are inefficient and standard errors are biased. Hence, we present heteroskedasticity and autocorrelation consistent standard errors (Newey and West (1987)). The results are presented in Table 4.

[4] http://deutsche-boerse.com/dbg/dispatch/en/binary/gdb_content_pool/imported_files/public_files/10_downloads/31_trading_member/10_Products_and_Functionalities/20_Stocks/BR_Xetra_Speed.pdf

[5] Results based on the model estimated based on different numbers of lags are similar to those presented in the paper and available from the authors upon request.

6 Coefficient Estimates

We start with the effect of ask side variables on returns. Ask side depth measured based on the first two, five and twenty levels have significantly negative coefficient estimates on their first lags. On the other hand, ask side depth between the second and fifth levels has a significantly positive coefficient estimates on its first lag. These results provide some empirical evidence in support of Hypothesis 1a that an increase in depth at lower levels of the ask side results in lower future prices at the next transaction period while an increase in the depth at higher levels results in higher future prices. Furthermore, ask side depth between the fifth and twentieth levels has a statistically insignificant coefficient estimate on its first lag. This in turn suggests that any information, that is relevant for future prices, embedded in the ask side depth is mostly available at the lower levels of the ask side. These results are not only statistically but also economically significant. To understand the economic magnitude of these coefficient estimates, one needs to take into account the fact that we consider the logarithm of the ask side depth in the VAR system. For example, assuming that everything else remains constant, the log stock price decreases by 0.18 basis points ($-0.180b.p. = -0.260 \times \log(2)$) at the next trade following a two-fold increase in the depth of the first two levels of the ask side. Although this might look like an economically insignificant change at first sight, it is indeed more than 18 times the mean log return between two trades, which is around 0.01 basis points.

[Insert Table 4 here]

Different measures of ask side slope have positive and significant coefficient estimates on their first lags with the exception of the slope between second and fifth levels of the ask side. This in turn confirms Hypothesis 2a that an increase in the slope of the ask side results in higher future prices. The economic magnitude of this effect is a 0.09 ($0.132 \times \log(2)$), 0.13 ($0.189 \times \log(2)$) and 0.40 ($0.575 \times \log(2)$) basis points increase in the price at the next trade following a two-fold increase in the slope of the first two, five and twenty levels of the ask side, respectively. Once again, considering that the mean log return between two trades is around 0.01 basis points, these effects are economically important. Furthermore, these results suggest that there is relevant information for future prices not only in lower but also in higher levels of the ask side in terms of its slope, as suggested by the significant coefficient estimate on the first lag of the ask side slope between the fifth and twentieth levels.

We now consider the bid side variables. The coefficient estimates on the first lag of most bid side variables have the opposite signs of those on the first lag of the corresponding ask side variable. This is in line with our hypotheses in Section 4. Furthermore, the economic interpretations of the coefficient estimates on the bid side variables are similar to those on the corresponding ask side variables. Thus, we only briefly discuss our results on the bid side variables.

First, the coefficient estimates on the first lag of bid side depth in the first two and five levels are both significantly positive. This in turn suggests that an increase in depth at lower levels of the bid side results in higher future prices at the next transaction period, in line with the first part of our Hypothesis 1b. However, the coefficient estimates on the first lag of depth in the higher levels of the bid side are either significantly positive for $D^B_{1,20}$ and $D^B_{5,20}$ or negative but insignificant for $D^B_{2,5}$. In other words, these coefficient estimates do not provide any supporting evidence for the second part of our Hypothesis 1b. Second, different measures of bid side slope have negative and significant coefficient estimates on their first lags with the exception of the slope between second and fifth levels of the ask side. These provide empirical evidence in support of Hypothesis 2b that an increase in the slope of the bid side results in lower future prices. Finally, the information embedded in both lower and higher levels of the bid side is relevant for future prices as suggested by the significant coefficient estimates on the first lag of bid side variables based on both lower and higher levels.

We now turn our attention to the variables that use information embedded in both sides of the limit order book. The coefficient estimates on the first lag of depth ratio are all negative regardless of the levels considered. These estimates are statistically significant with the exception of the depth ratio between second and fifth levels of the limit order book. These results suggest that an increase in depth ratio results in a lower future price regardless of the levels considered, providing evidence for the first part of Hypothesis 1c but against its second part. On the other hand, the coefficient estimates on the first lag of slope ratio are all significantly positive, with the exception of the slope ratio between the second and fifth levels of the limit order book. These results in turn suggest that an increase in the slope ratio results in higher future prices, providing evidence in support of Hypothesis 2c. Given that these variables are not transformed like variables based on one side of the limit order book, their economic interpretation is straightforward. For example, the log stock price decreases by 14 basis points at the next trade following a 10% increase in the depth ratio and increases by 19 basis points following a 10% increase in the slope ratio.

7 Impulse Response Functions

So far, we have focused on the coefficient estimates on the first lags of limit order book variables in the return equation. These coefficient estimates reveal the initial effects of each limit order book variable on return but not their long term cumulative effects. Given that most of the variables in the VAR system have also significant dynamics of their own, one needs to calculate the impulse response functions of return to limit order book variables to infer about long term cumulative effects of limit order book variables on returns. To do this, we first simulate the estimated VAR system for a long enough period setting all residual terms to zero, i.e. $\varepsilon_{z,t} = \varepsilon_{r,t} = \varepsilon_{x,t} = 0$, to obtain its steady state. Second, starting with the steady state, we simulate the VAR system once again but this time assuming that the initial residual of the limit order book variable of interest is +1 while all other residuals (initial or future) remain zero. The difference between these two simulations of the VAR system is the impulse response function of return to a unit shock to the limit order book variable of interest.

Figure 2 presents the impulse response functions of return to limit order book variables. The cumulative effect of an increase in depth between any levels of the ask side is negative. This is in line with the initial effects of ask side depth on return discussed above based on the coefficient estimates on the first lags. The only exception is the ask side depth between second and fifth levels, which has a significantly positive initial effect but a negative long term cumulative effect. These results provide further supporting empirical evidence for the first part of Hypothesis 1a. However, they also suggest that the second part of Hypothesis 1a might not hold in the data when one considers the long term cumulative effect of ask side depth rather than its initial effect.

[Insert Figure 2 here]

In contrast to ask side depth, the long term cumulative effect of bid side depth depends on the levels considered. On one hand, bid side depth based on the first two, five and twenty levels have positive long term cumulative effects on return, although their cumulative effects in the short and medium term tend to be higher than those in the long term. These findings are in line with the initial effects of these variables on return and provide further empirical support for the first part of Hypothesis 1b. On the other hand, measures of bid side depth that exclude the first level have negative long term cumulative effects on return, although they have positive initial effects as well as positive cumulative effects in the short and medium term. These findings provide empirical evidence in support of the second part of Hypothesis 1b, unlike the coefficient estimates discussed above. Hence, these impulse response functions of return to bid side depth based on different levels suggest that Hypothesis 1b holds in the data when one considers long term cumulative effects rather than initial or short to medium run effects.

We now turn our attention to the long term cumulative effects of slope measures. Ask side slope has a positive long term cumulative effect regardless of the levels considered to measure it. These results are in line with their initial effects discussed above based on the coefficient estimates on their first lags in the return equation and provides further empirical evidence in support of Hypothesis 2a. On the other hand, the long term cumulative effects of bid side slope measures are negative with the exception of bid side slope between the second and fifth levels. This is mostly in line with Hypothesis 2b but also suggests that bid side slope might have different long term cumulative effects on returns depending on the levels considered to measure it, in contrast to its initial effect. Finally, we consider the long term cumulative effects of depth and slope ratios on return. Regardless of the levels considered, depth ratio has a negative long term cumulative effects on return. This provides further empirical evidence for the first part of Hypothesis 1c. However, it also provides some evidence contrary to the second part of Hypothesis 1c. This in turn suggests that Hypothesis 1c might not hold when we consider the long term cumulative effects rather than initial effects. Measures of slope ratio based on different levels of the limit order have positive long term cumulative effects on return, providing further empirical support for Hypothesis 2c.

The impulse response functions can be economically interpreted as the change in the log price of the stock in basis points as a function of transaction periods in response to a one unit positive shock to the limit order book variable of interest while taking into account its transformation. For example, log stock price decreases by almost 0.1 basis points in the long term following almost a three-fold increase[6] in between the first and fifth levels of the ask side. Furthermore, the long term cumulative effects of limit order book variables based on higher levels of the limit order book tend to be stronger and take longer to be fully realized. This suggests that there might be more information (relevant for returns) embedded in the higher levels of the limit order book but it takes longer for this information to be fully incorporated in prices.

[6]Note that a one unit positive shock to a limit order book variable with a log transformation implies $e(\simeq 2.718)$-fold increase in the limit order book variable of interest.

8 Granger Causality

We have discussed whether limit order book variables have any significant effect on returns based on coefficient estimates and impulse response functions. However, we have not yet answered whether the state of the limit order book causes returns to change at the next transaction period. In this section, we address this question by analyzing the causal effect of limit order book variables on returns.

To do this, we consider the statistical test of causality in the sense of Granger (1969). Specifically, consider that we are interested in whether a given limit order book variable Granger causes return. We run a regression of return on its own five lags and five lags of the limit order book variable. We then conduct an F-test of the null hypothesis that the coefficients on the lags of the limit order book variable are jointly zero. Rejecting this null hypothesis suggests that the limit order book variable causes the return. To conduct the test, we consider the asymptotic F-statistic that has a χ^2 distribution with 5 degrees of freedom asymptotically. Last column of Table 4 presents statistics testing whether limit order book variables Granger cause return.

Our results can be summarized as follows: Most measures of ask and bid side depth as well as their ratio causes return to change at next transaction period. The only exception is the bid side depth between the second and fifth levels, for which we fail to reject the null hypothesis that it does not cause return based on the F-statistic. Similarly, most measures of ask and bid side slope as well as their ratio causes return to change at the next transaction period. Once again, the exception to this is the slope measure between the second and fifth levels of ask and bid sides as well as their ratio. Furthermore, the test statistics tend to be much higher for limit order book variables measured based on lower levels of the limit order book, suggesting that there is stronger empirical evidence for the relevance of information embedded in the lower levels of the limit order book. To sum up, these results provide further empirical evidence that the state of the limit order book affects returns at the next transaction period, although it is hard to interpret them in terms of individual hypotheses discussed in Section 4 as they do not provide any information on the directional effect of limit order book variables on return.

9 Economic Value of the Information in the Limit Order Book

So far, we have discussed the statistical relation between returns and the information embedded in the limit order book. In this section, we analyze whether an investor can use this significant statistical relation to obtain economic gains. To this end, we consider simple trading strategies similar to those discussed in Kozhan and Salmon (2012).

Specifically, at each transaction period t, we first obtain forecast of the midquote return at the next transaction period $t+1$, $\hat{r}^{(j)}_{t+1}$, based on a restricted version of the VAR model in Equation 9a that excludes the contemporaneous effect of trade direction on return.[7] We then calculate our trading signal for the transaction period based on whether our forecast is greater or less than a threshold. In particular, we consider a forecast greater than a positive threshold, i.e. $\hat{r}^{(j)}_{t+1} > \kappa$ where $\kappa > 0$, to be a buy opportunity or signal and a forecast less than a symmetric negative threshold,

[7] We also consider trading strategies based on a forecasting model that includes the contemporaneous effect of trade direction on returns. As discussed above, any contemporaneous effect of trade direction on return is not a causal relation even if trade direction is observed right before the return is calculated. Our results based on this VAR model are similar to those based on the restricted VAR model and, thus, are not presented.

i.e. $\hat{r}_t^{(j)} < -\kappa$ where $\kappa > 0$, to be a sell opportunity or signal while we do not trade on any forecast between these two thresholds, i.e. $|\hat{r}_{t+1}^{(j)}| \leq \kappa$. Similar to Kozhan and Salmon (2012), κ can be considered as a parameter to filter out potentially weak signals. In this paper, we consider a κ of 1 basis points. As we consider higher values for κ, we trade less frequently.

In addition, we remove further noise from the transaction data and, thus from our trading strategy, using a moving average filter. Specifically, for each transaction period t, we calculate a short-run moving average based on the last three and a long-run moving average based on the last forty midquote prices including the midquote price at transaction period t. Similar to a technical analysis trading strategy, we consider the relative movement of the short and long-run moving averages to confirm the strength of our trading signal discussed above. To be more precise, we take a long position based on our trading signal discussed above, only when the short-run moving average crosses from below the long-run moving average by more than 0.06 euros, which correspond to approximately 10 basis points for MRK in July 2010. Similarly, we take a short position based on our trading signal discussed above, only when the short-run moving average crosses from above the long-run moving average by more than 0.06 euros. Otherwise, we do not consider a trading signal to be strong enough and, thus do not trade, if it is not confirmed by the relative movements of short- and long-run moving averages.

We start our trading exercise at the beginning of July 2010. We do not take any position until we observe a signal strong enough based on both of the filters discussed above. When we receive such a signal, we then take a long or short position of one share depending on the signal. At every transaction period, we calculate our trading signal and reevaluate our position. Specifically, if we have a long (short) position and receive a strong enough buy (sell) signal or a signal that is not strong enough, we continue to keep our long (short) position of one share. On the other hand, if we have a long (short) position and receive a strong enough sell (buy) signal, we then close our long (short) position and hold a short (long) position of one share. We continue in this fashion until the last transaction in July 2010 and keep the position in the last transaction till trading terminates in July 2010.

Several remarks are in order about our trading strategy. First of all, we implicitly assume that our trading does not alter the dynamics of the relation between return, trade direction and limit order book. We believe that this is a reasonable assumption since we only consider trading one share at a time which should be negligible given the trading volume of Merck in July 2010. Secondly, it is an in-sample trading exercise since we use the coefficient estimates based on the whole sample to obtain our return forecasts at each transaction period. An out-of-sample trading exercise would require the estimation of the VAR model at each transaction period based on the information available only up to that transaction period, which is computationally quite intensive. Finally, in our trading exercise, we calculate our signal based on the snapshot of the limit order book right after (less than a millisecond after) a transaction and assume that we can trade at the prices in that snapshot of the limit order book. However, in reality, the computation of our trade signal after a transaction and the processing of our order at the exchange are not instantaneous. Hence, we might end up trading at prices different than those in the snapshot of the limit order book right after a transaction. Hence, we perform the same exercise when we assume a latency of 500 and 1000 milliseconds and trade at prices observed 500 and 1000 milliseconds after observing a transaction. This exercise allow us not only to check the robustness of our results to a more realistic assumption about latency but also to analyze the effect of speed on the profitability of our trading strategies. Given that our forecasting model is designed to predict returns with a zero delay, we expect profits based on a latency of zero millisecond to be higher than those based on a latency of 500 and 1000 milliseconds.

We consider different trading strategies based on different limit order book variables as separate predictors in the forecasting model in Equation 9. We evaluate the performance of these trading strategies based on their cumulative returns over the month. We also distinguish between returns based on trading at midquote and bid and ask prices. Returns based on midquote prices might not necessarily be attainable by traders. Returns based on bid and ask prices, on the other hand, provide a more realistic performance measure for the trading strategy since we consider actual prices at which the traders can buy and sell the stock. Thus, we present results based on trading at bid and ask rather than midquote prices.

As a benchmark, we consider a trading strategy based on a forecasting model that ignores information embedded in limit order book variables. Specifically, we obtain our trade signals in the benchmark trading strategy based on the estimation of the empirical model in Equation 9 with the restriction that $\alpha_{z,\tau} = 0$ for $\tau = 1, \ldots, 5$. This benchmark model allows us to analyze whether limit order book variables provide additional information useful to the trader above and beyond what is embedded in past returns and trade direction. One can think of this benchmark strategy as the equilibrium strategy of a trader who faces barriers to entry. For example, a trader might want to use the benchmark strategy if she cannot process or have access to the limit order book due to technological or financial reasons. Furthermore, we also consider this trading strategy with and without applying the moving average filter. This allows us to have an idea about the effect of the moving average filter on the profits from this trading strategy.

[Insert Table 5 here]

Table 8 presents returns from trading strategies when we sell at bid and buy at ask prices as well as the number of trades required to implement these trading strategies. We start by comparing results based on the benchmark strategies with and without the moving average filter. This allows us to provide some intuition on why we apply the additional moving average filter. First of all, as expected, the number of trades decreases significantly from 270 to 96 when we apply the moving average filter. Furthermore, although still negative, returns based on the benchmark strategy increases significantly, regardless of the assumed latency, when we apply the moving average filter. The negative returns on the benchmark strategy without the moving average filter is due to the large number of trades required to implement this strategy and the loss associated with each round trip transaction due to the bid-ask spread. Kozhan and Salmon (2012) also report large negative returns from a trading strategy based on a linear model, with a daily cumulative return of as low as -92% (corresponding to a compound monthly return of -100% assuming that there are 22 trading days in a month) based on actual bid and ask prices. As discussed above, these findings are due to the fact that transaction data and, thus, trading returns are noisy and the additional filter based on the difference between short- and long-run moving averages allows to us to decrease this noise further.

More importantly, most trading strategies based on limit order book information outperform the benchmark model. To be more specific, trading strategies based on 22 out of 30 limit order book variables outperform the benchmark model. Of this 22 trading strategies, 21 provide positive returns. Most trading strategies based on ask side or ratio variables outperform the benchmark strategy while only half of the trading strategies based on bid side variables do so. Furthermore, trading strategies based on ask side variables tend to outperform those based on the corresponding bid side variable. Also, trading strategies based on variables that capture the information embedded in the lower levels of the limit order book tend to outperform those based on variables that do not include this information. Finally, these results hold when we consider a latency of 0, 500 or 1000 milliseconds, suggesting the robustness of our results to a more realistic assumption of 500 or 1000 milliseconds latency. As one would expect, profits based on a latency of 500 or 1000 milliseconds are lower than those based on a latency of 0 milliseconds for the same threshold parameter. Furthermore, profits tend to decrease monotonically as we consider higher latencies. On the other hand, the relative performance of strategies based on limit order book variables with respect to the benchmark strategy generally increases as the latency increases from 0 millisecond to 1000 milliseconds, suggesting that the information embedded in the limit order book can still be economically important for traders with a relatively high latency.

Several remarks are in order concerning these results. First of all, there are four parameters that we had to choose to operationalize our trading strategy: the time period for short- and long-run moving averages, the threshold parameters for the moving average filter and our trading signal. We chose parameters which we considered to be reasonable. However, we should note that they were still chosen in an ad hoc fashion as there is no theoretical model that provides any guidance on how to choose these parameters. We considered other sets of parameters and results undoubtedly change. There are some sets of parameters for which trading strategies using limit order book information continue to outperform the benchmark and there are others for which this is no longer the case. In other words, our results suggest that it is feasible to obtain economic gains using information embedded in the limit order book for some sets of parameters but this is not always the case.

10 Robustness Checks based on Another Company and Time Period

In this paper, we have focused on a single company, i.e. Merck, in a single month, i.e. July 2010. However, our findings are not due to our choice of company and time period. To demonstrate this, we have also analyzed the effect of limit order book variables on return for Merck in June 2011 and SAP in July 2010 and June 2011. In this section, we briefly summarize these additional findings which are available in an online appendix.

First of all, regardless of the company and time period, most limit order book variables have significant coefficient estimates on their first lags in the return equation and these coefficient estimates have the same signs and similar magnitudes as our benchmark results. Furthermore, similar to our benchmark results, the coefficient estimates on higher order lags of most limit order book variables are not significantly different than zero. Second, the long term cumulative effects of most limit order book variables on returns are similar to those presented in Figure 2. Third, similar to our benchmark results, most limit order book variables significantly cause return to change at the next transaction period. Finally, the results from the trading exercise are also similar to our benchmark results. To sum up, these additional results suggest that the effect of limit order book variables do not change significantly with the stock and time period and are very similar to those in our benchmark analysis.

11 Further Robustness Checks

In this section, we discuss several checks we have implemented to test the robustness of our results. The results are available in an online appendix.

11.1 Daily Results

Our results are based on the estimation of the VAR system in Equation 9 over the whole sample period between 1st July 2010 and 31th July 2010. We have enough observations in a given day to estimate the VAR system for each day of our sample period separately. This allows us to analyze whether results based on the whole sample period continue to hold when we focus on a single day. Furthermore, it also provides some information about the stability of our results. Not surprisingly, daily coefficient estimates on the first lag of limit order book variables in the return equation tend to change from one day to another but most of them tend to be stable and move around the corresponding estimates based on the whole sample.

11.2 Controlling for Other Factors

In this section, we analyze the robustness of our results by controlling for other factors that might potentially affect the relation between return, trade direction and the state of the limit order book. One such variable is the duration measured as the waiting time between consecutive transactions. Dufour and Engle (2000) analyze the effect of duration on the relation between return and trade direction by considering the interaction between duration and trade direction as an additional variable in their empirical specification. In this framework, they find that as duration decreases, the price impact of trades, the speed of price adjustment to trade-related information, and the positive autocorrelation of signed trades all increase. They argue that markets have reduced liquidity when they are most active and there is an increased presence of informed trades. Our empirical approach differs from theirs in the sense that we include (log) duration itself in the empirical specification in Equation 9 as an additional state variable. Other variables that might also capture market activity are transaction volume and return volatility. We consider log number of shares traded at each transaction and log absolute midquote return (right before each transaction) to proxy for transaction volume and return volatility, respectively, and include these variables separately as additional state variables in the empirical specification in Equations 9. Our results available in an online appendix suggest that the relation between return, trade direction and the state of the limit order book do not change significantly when we control for duration, transaction volume and volatility.

12 Conclusion

In this paper, we analyze whether the state of the limit order book affects future price movements in line with what the theory predicts. To this end, we reconstruct the first 20 levels of the historical limit order book every millisecond for several companies traded at Frankfurt Stock Exchange in July 2010 and June 2011 based on the data from the

Xetra electronic trading system. We consider several variables that summarize different dimensions of the information embedded in the limit order book, such as depth and slope of both ask and bid sides separately as well as their ratios. Following Hasbrouck (1991), we estimate a linear vector autoregressive system (VAR) that includes midquote return, trade direction and a limit order book variable one at a time. In line with theoretical models of dynamic limit order markets, we find that the state of the limit order book help predict short run midquote return. Limit order book variables also have significant long term cumulative effects on midquote return and trade directions. This long term cumulative effect is stronger and takes longer to be fully realized for variables based on higher levels of the book. In a simple high frequency trading exercise, we show that it is possible to obtain economic gains from the relation between limit order book variables and midquote return.

References

Admati, Anat R., 1985, A noisy rational expectations equilibrium for multi-asset securities markets, *Econometrica* 53, pp. 629–658.

Beltran-Lopez, Héléna, Pierre Giot, and Joachim Grammig, 2009, Commonalities in the order book, *Financial Markets and Portfolio Management* 23, 209–242.

Biais, Bruno, Pierre Hillion, and Chester Spatt, 1995, An Empirical Analysis of the Limit Order Book and the Order Flow in the Paris Bourse, *The Journal of Finance* 50, 1655–1689.

Bilodeau, Y., 2013, Xetraparser, *[computer software], HEC Montreal*.

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

Chakravarty, Sugato, and Craig W. Holden, 1995, An integrated model of market and limit orders, *Journal of Financial Intermediation* 4, 213 – 241.

Dufour, Alfonso, and Robert F Engle, 2000, Time and the Price Impact of a Trade, *The Journal of Finance* 55, 2467–2498.

Easley, David, and Maureen O'hara, 2004, Information and the cost of capital, *The Journal of Finance* 59, 1553–1583.

Foucault, Thierry, Ohad Kadan, and Eugene Kandel, 2005, Limit order book as a market for liquidity, *The Review of Financial Studies* 18, pp. 1171–1217.

Glosten, Lawrence R, 1994, Is the electronic open limit order book inevitable?, *The Journal of Finance* 49, 1127–1161.

———, and Paul R Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 14, 71–100.

Goettler, Ronald L., Christine A. Parlour, and Uday Rajan, 2009, Informed traders and limit order markets, *Journal of Financial Economics* 93, 67 – 87.

Granger, C. W. J., 1969, Investigating causal relations by econometric models and cross-spectral methods, *Econometrica* 37, pp. 424–438.

Griffiths, Mark D, Brian F Smith, D.Alasdair S Turnbull, and Robert W White, 2000, The costs and determinants of order aggressiveness, *Journal of Financial Economics* 56, 65 – 88.

Hasbrouck, Joel, 1991, Measuring the information content of stock trades, *The Journal of Finance* 46, 179–207.

Hellwig, Martin F, 1980, On the aggregation of information in competitive markets, *Journal of Economic Theory* 22, 477 – 498.

Kalay, Avner, and Avi Wohl, 2009, Detecting liquidity traders, *Journal of Financial and Quantitative Analysis* 44, 29–54.

Kaniel, Ron, and Hong Liu, 2006, So what orders do informed traders use?, *The Journal of Business* 79, pp. 1867–1913.

Kozhan, Roman, and Mark Salmon, 2012, The information content of a limit order book: The case of an FX market, *Journal of Financial Markets* 15, 1–28.

Kumar, P., and D.J. Seppi, 1994, Limit and market orders with optimizing traders, *Working paper Carnegie Mellon University*.

Kyle, Albert S, 1985, Continuous Auctions and Insider Trading, *Econometrica* 53, 1315–1335.

——— , 1989, Informed Speculation with Imperfect Competition, *The Review of Economic Studies* 56, 317–355.

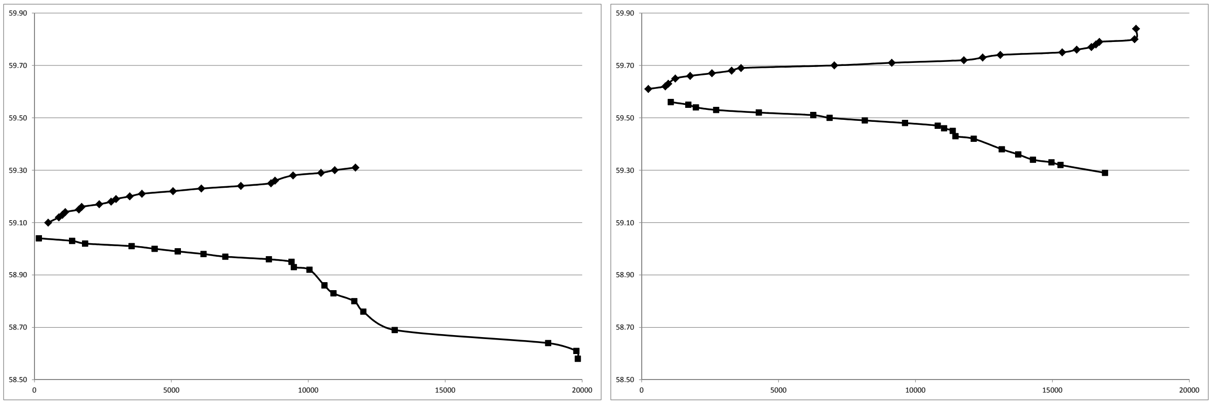
Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, pp. 703–708.

Rock, K, 1996, The specialists order book and price anomalies, *Working Paper,Harvard University*.

Rosu, I, 2009, A Dynamic Model of the Limit Order Book, *The Review of Financial Studies* 22, 4601.

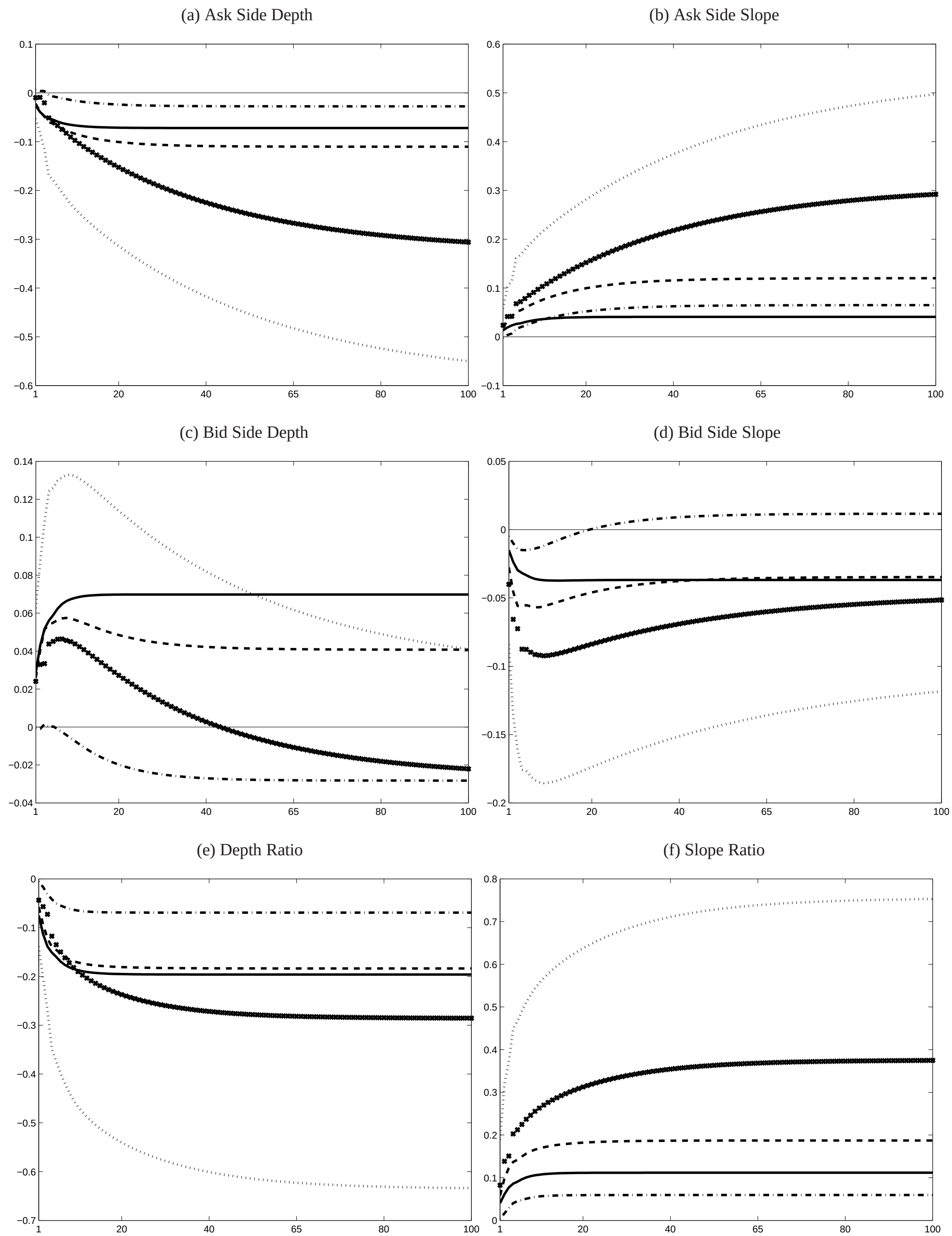
Figure 1: Two Snapshots of the Limit Order Book

(a) Snapshot at 9:10:27.948 on July 1, 2010 (b) Snapshot at 9:33:36.100 on July 1, 2010



Note: Panel (a) and (b) present the price and depth of the first 20 levels of the limit order book for Merck on July 1, 2010 at 9:10:27.948 and 9:33:36.100, respectively. The diamonds and squares represent different levels of the ask and bid sides, respectively.

Figure 2: Impulse Response Function of Return to Limit Order Book Variables



Note: This figure presents the impulse response functions of midquote return to a one unit positive shock to limit order book variables. The horizontal axis is transaction periods, i.e. the number of transactions since the initial shock, and the vertical axis is the response of midquote return in basis points. The solid, dashed, small dotted and dashed-dotted line correspond to the limit order book variable measured based on the first two, five, ten and twenty levels, respectively. The big dotted line corresponds to the limit order book variable measured between second and fifth levels.

Table 1: Summary Statistics for Daily Trading Volume and Returns

(a) Daily Trading Volume

	MRK		SAP	
	July-2010	June-2011	July-2010	June-2011
Mean	532,321	356,218	4,198,806	3,308,930
Median	478,219	316,050	4,185,447	3,070,861
Std. Deviation	294,149	121,066	1,547,879	1,491,676
Min	202,535	210,017	2,286,050	1,296,397
Max	1,629,225	681,365	8,676,879	8,618,285

(b) Daily Returns

	MRK		SAP	
	July-2010	June-2011	July-2010	June-2011
Mean	0.590%	-0.087%	-0.208%	-0.152%
Median	0.336%	-0.019%	-0.331%	-0.018%
Std. Deviation	1.662%	1.073%	1.140%	1.177%
Min	-2.867%	-2.849%	-2.266%	-2.608%
Max	4.507%	2.039%	2.626%	2.114%

Note: This table presents summary statistics for daily trading volume (Panel (a)) and returns (Panel (b)) for MRK and SAP in July 2010 and June 2011.

Table 2: Variables for the Snapshots of the Limit Order Book in Figure 1

LOB Variable	Snapshot 1	Snapshot 2
Ask Side Variables		
Depth (Levels 1-2)	890	864
Depth (Levels 1-5)	1624	1774
Depth (Levels 1-20)	11723	18056
Depth (Levels 2-5)	1124	1528
Depth (Levels 6-20)	10599	16826
Slope (Levels 1-2)	0.000051	0.000016
Slope (Levels 1-5)	0.000044	0.000033
Slope (Levels 1-20)	0.000019	0.000013
Slope (Levels 2-5)	0.000041	0.000044
Slope (Levels 6-20)	0.000016	0.000011
Bid Side Variables		
Depth (Levels 1-2)	1378	1705
Depth (Levels 1-5)	4390	4282
Depth (Levels 1-20)	19840	16923
Depth (Levels 2-5)	4228	3218
Depth (Levels 6-20)	16290	14200
Slope (Levels 1-2)	-0.000008	-0.000016
Slope (Levels 1-5)	-0.000009	-0.000012
Slope (Levels 1-20)	-0.000023	-0.000017
Slope (Levels 2-5)	-0.000010	-0.000012
Slope (Levels 6-20)	-0.000027	-0.000018
Ratio Variables		
Depth Ratio (Levels 1-2)	-0.215168	-0.327365
Depth Ratio (Levels 1-5)	-0.459927	-0.414135
Depth Ratio (Levels 1-20)	-0.257168	0.032391
Depth Ratio (Levels 2-5)	-0.579970	-0.356089
Depth Ratio (Levels 6-20)	-0.211648	0.084639
Slope Ratio (Levels 1-2)	0.723600	0.018268
Slope Ratio (Levels 1-5)	0.649243	0.449419
Slope Ratio (Levels 1-20)	-0.110834	-0.137310
Slope Ratio (Levels 2-5)	0.608115	0.581224
Slope Ratio (Levels 6-20)	-0.263582	-0.244090

Note: This table presents the values of the limit order book variables based on the snapshots of the limit order book for Merck on July 1, 2010 at 9:10:27.948 (Snapshot 1) and 9:33:36.100 (Snapshot 2) presented in Figure 1.

Table 3: Summary Statistics

	Transformation	Mean	Median	Std. Dev.	Min	Max
Log Return	$\times$ 10000					
Ask Side Variables						
Depth (Levels 1-2)	$\ln(x)$	6.599	6.640	0.852	0.693	10.017
Depth (Levels 1-5)	$\ln(x)$	7.973	8.008	0.531	5.521	10.188
Depth (Levels 1-20)	$\ln(x)$	9.616	9.628	0.310	8.231	10.667
Depth (Levels 2-5)	$\ln(x)$	7.555	7.609	0.626	4.466	10.140
Depth (Levels 6-20)	$\ln(x)$	9.367	9.378	0.350	7.742	10.579
Slope (Levels 1-2)	$\ln(x)$	-10.156	-10.233	1.170	-14.601	-3.219
Slope (Levels 1-5)	$\ln(x)$	-10.856	-10.932	0.684	-13.393	-7.367
Slope (Levels 1-20)	$\ln(x)$	-10.882	-10.911	0.462	-12.262	-8.865
Slope (Levels 2-5)	$\ln(x)$	-10.956	-11.054	0.737	-13.647	-6.359
Slope (Levels 6-20)	$\ln(x)$	-10.865	-10.879	0.513	-12.272	-8.658
Bid Side Variables						
Depth (Levels 1-2)	$\ln(x)$	6.570	6.627	0.792	2.398	9.264
Depth (Levels 1-5)	$\ln(x)$	7.883	7.912	0.532	4.754	9.778
Depth (Levels 1-20)	$\ln(x)$	9.558	9.585	0.332	8.001	10.869
Depth (Levels 2-5)	$\ln(x)$	7.459	7.506	0.632	4.220	9.379
Depth (Levels 6-20)	$\ln(x)$	9.322	9.353	0.360	7.745	10.745
Slope (Levels 1-2)	$\ln(\|x\|)$	-10.167	-10.272	1.116	-13.403	-2.813
Slope (Levels 1-5)	$\ln(\|x\|)$	-10.768	-10.831	0.685	-12.654	-7.371
Slope (Levels 1-20)	$\ln(\|x\|)$	-10.801	-10.870	0.525	-12.242	-8.879
Slope (Levels 2-5)	$\ln(\|x\|)$	-10.856	-10.942	0.744	-12.886	-7.212
Slope (Levels 6-20)	$\ln(\|x\|)$	-10.792	-10.857	0.572	-12.306	-8.615
Ratio Variables						
Depth Ratio (Levels 1-2)		0.012	0.012	0.444	-0.995	0.991
Depth Ratio (Levels 1-5)		0.040	0.047	0.320	-0.912	0.922
Depth Ratio (Levels 1-20)		0.028	0.022	0.154	-0.566	0.694
Depth Ratio (Levels 2-5)		0.041	0.050	0.339	-0.911	0.942
Depth Ratio (Levels 6-20)		0.021	0.016	0.177	-0.691	0.713
Slope Ratio (Levels 1-2)		0.005	0.004	0.551	-0.999	0.999
Slope Ratio (Levels 1-5)		-0.039	-0.048	0.364	-0.954	0.928
Slope Ratio (Levels 1-20)		-0.038	-0.039	0.219	-0.775	0.696
Slope Ratio (Levels 2-5)		-0.042	-0.052	0.386	-0.984	0.966
Slope Ratio (Levels 6-20)		-0.034	-0.034	0.273	-0.779	0.795

Note: This table presents the transformation applied to limit order book variables and the summary statistics for these transformed variables.

Table 4: Coefficient Estimates on Lagged Values of Limit Order Book Variables (z_{t-j}) and F-statistics for Granger Causality Tests

	$\alpha_{z,1}$	$\alpha_{z,2}$	$\alpha_{z,3}$	$\alpha_{z,4}$	$\alpha_{z,5}$	F-stat
Ask Side Variables						
Depth (Levels 1-2)	-0.260***	0.003	0.008	0.028	0.019	408.518***
Depth (Levels 1-5)	-0.224***	0.022	0.033	-0.005	0.062*	75.322***
Depth (Levels 1-20)	-0.532***	0.228	0.003	-0.226	0.398***	51.619***
Depth (Levels 2-5)	0.058*	-0.021	0.021	-0.074**	0.026	11.236**
Depth (Levels 6-20)	-0.100	0.130	-0.091	-0.215*	0.194*	16.472***
Slope (Levels 1-2)	0.132***	-0.009	0.001	-0.009	-0.018	193.198***
Slope (Levels 1-5)	0.189***	-0.028	-0.038	0.013	-0.074**	67.759***
Slope (Levels 1-20)	0.575***	-0.064	-0.325**	0.268	-0.381***	93.289***
Slope (Levels 2-5)	-0.019	0.030	-0.021	0.059*	-0.033	6.714
Slope (Levels 6-20)	0.238**	-0.009	-0.171	0.197*	-0.200**	32.749***
Bid Side Variables						
Depth (Levels 1-2)	0.286***	0.000	-0.002	-0.029	-0.020	443.729***
Depth (Levels 1-5)	0.260***	-0.040	-0.010	-0.064	-0.027	93.188***
Depth (Levels 1-20)	0.627***	-0.304	-0.047	-0.058	-0.129	52.157***
Depth (Levels 2-5)	-0.008	-0.005	0.016	-0.022	0.022	1.177
Depth (Levels 6-20)	0.240**	-0.123	-0.084	0.068	-0.055	11.778**
Slope (Levels 1-2)	-0.151***	-0.022	-0.011	0.022	0.007	303.520***
Slope (Levels 1-5)	-0.277***	0.014	0.045	0.094**	0.031	154.872***
Slope (Levels 1-20)	-0.840***	0.186	0.291**	0.152	0.130	172.180***
Slope (Levels 2-5)	-0.050	-0.006	0.003	0.022	0.000	9.005
Slope (Levels 6-20)	-0.400***	0.081	0.181*	-0.037	0.117	62.047***
Ratio Variables						
Depth Ratio (Levels 1-2)	-0.751***	0.003	0.009	0.092***	0.047	942.974***
Depth Ratio (Levels 1-5)	-0.579***	0.073	0.018	0.080	0.096*	182.466***
Depth Ratio (Levels 1-20)	-1.386***	0.571**	0.039	-0.170	0.460**	129.497***
Depth Ratio (Levels 2-5)	-0.060	0.006	-0.058	0.004	-0.014	11.557**
Depth Ratio (Levels 6-20)	-0.436***	0.264	-0.004	-0.280	0.188	31.213***
Slope Ratio (Levels 1-2)	0.404***	0.024	0.033	-0.035	-0.038	475.614***
Slope Ratio (Levels 1-5)	0.582***	-0.042	-0.080	-0.091	-0.096*	225.576***
Slope Ratio (Levels 1-20)	1.897***	-0.276	-0.724***	0.100	-0.561***	340.444***
Slope Ratio (Levels 2-5)	0.041	0.035	0.008	0.022	-0.032	7.520
Slope Ratio (Levels 6-20)	0.825***	-0.103	-0.407**	0.284*	-0.351***	115.302***

Note: This table presents the parameter estimates of lagged limit order book variables in the return equation ($\alpha_{z,\tau}$) and F-statistics for Granger Causality tests. ***, **, * represent statistical significance at 1%, 5%, 10%, respectively.

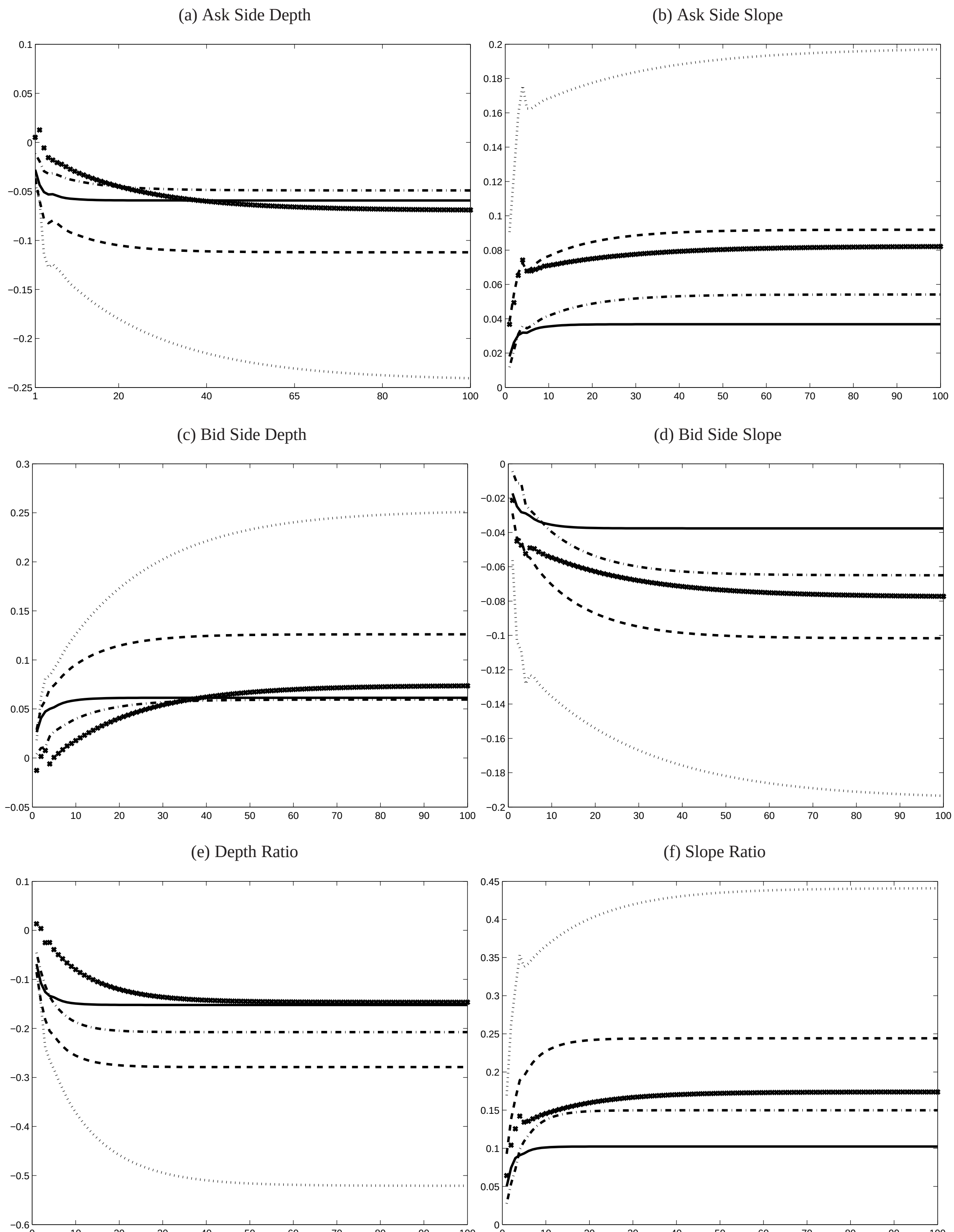
Table 5: Bid-Ask Return on Trading Strategies

	Number of Trades	Latency 0 ms	500 ms	1000 ms
Ask Side Variables				
Depth (Levels 1-2)	119	3.676%	3.351%	2.619%
Depth (Levels 1-5)	98	6.607%	6.330%	5.591%
Depth (Levels 1-20)	98	5.384%	4.954%	4.400%
Depth (Levels 2-5)	96	0.045%	-0.466%	-0.815%
Depth (Levels 6-20)	102	-3.666%	-4.108%	-4.582%
Slope (Levels 1-2)	120	2.229%	2.016%	1.478%
Slope (Levels 1-5)	99	2.119%	1.791%	1.289%
Slope (Levels 1-20)	100	0.284%	-0.176%	-0.469%
Slope (Levels 2-5)	88	1.557%	1.044%	0.718%
Slope (Levels 6-20)	100	-3.445%	-4.198%	-4.335%
Bid Side Variables				
Depth (Levels 1-2)	118	2.223%	0.910%	0.377%
Depth (Levels 1-5)	111	-4.928%	-5.478%	-5.857%
Depth (Levels 1-20)	93	-1.580%	-2.452%	-3.344%
Depth (Levels 2-5)	94	0.972%	0.311%	-0.157%
Depth (Levels 6-20)	93	0.599%	0.306%	-0.299%
Slope (Levels 1-2)	127	-6.685%	-7.805%	-8.461%
Slope (Levels 1-5)	120	-6.846%	-7.798%	-8.328%
Slope (Levels 1-20)	103	1.404%	0.632%	0.215%
Slope (Levels 2-5)	92	0.013%	-0.937%	-1.607%
Slope (Levels 6-20)	85	-2.272%	-3.057%	-3.388%
Ratio Variables				
Depth Ratio (Levels 1-2)	141	0.726%	-0.708%	-1.181%
Depth Ratio (Levels 1-5)	109	-0.578%	-2.034%	-2.571%
Depth Ratio (Levels 1-20)	97	5.393%	5.056%	4.414%
Depth Ratio (Levels 2-5)	94	1.491%	0.726%	0.134%
Depth Ratio (Levels 6-20)	89	0.582%	0.187%	-0.118%
Slope Ratio (Levels 1-2)	121	1.169%	0.673%	0.048%
Slope Ratio (Levels 1-5)	115	-4.944%	-5.900%	-6.420%
Slope Ratio (Levels 1-20)	112	2.937%	2.270%	1.944%
Slope Ratio (Levels 2-5)	90	1.471%	0.740%	0.150%
Slope Ratio (Levels 6-20)	104	4.778%	4.089%	3.848%
Benchmark with MA filter	96	-1.429%	-1.964%	-2.478%
Benchmark without MA filter	270	-3.305%	-6.425%	-8.061%

Note: This table presents the cumulative bid-ask returns on trading strategies over the whole trading period of July 2010. The trading strategy is based on the empirical model in Equation 9 that excludes the contemporaneous effect of trade direction on return and described in detail in Section 9. The parameter of the filter for the return forecasts, κ, is set to 1 basis points. The long and short-run moving average filters are calculated based on the last three and forty transaction prices, including the most recent one, respectively. The parameter of the moving average filter is set to 0.06 euros. Benchmark with and without moving average (MA) filter present results from trading strategies based on a forecasting model that ignores information embedded in limit order book variables.

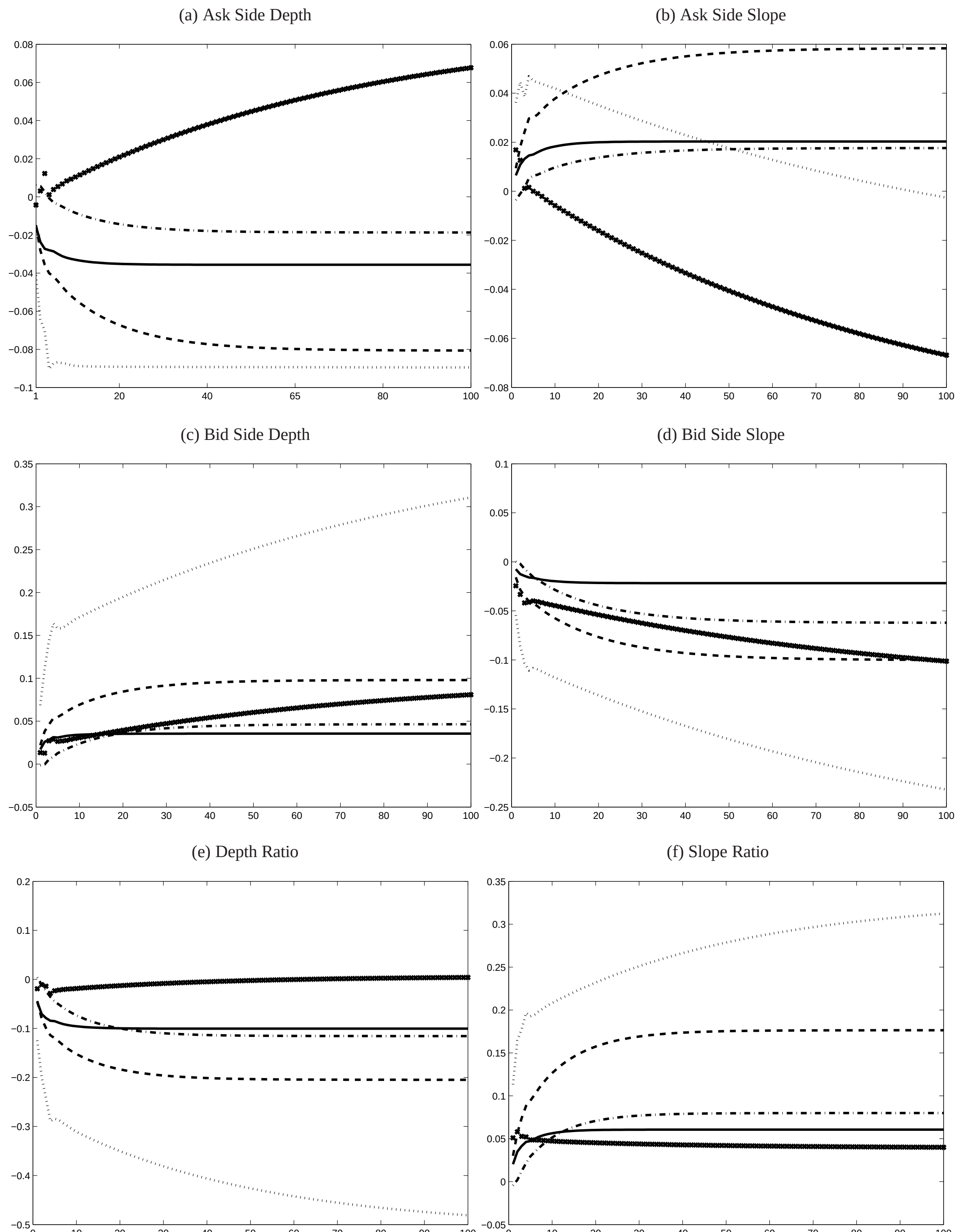
Online Appendix

Figure 2* (MRK-June 2011): Impulse Response Function of Return to Limit Order Book Variables



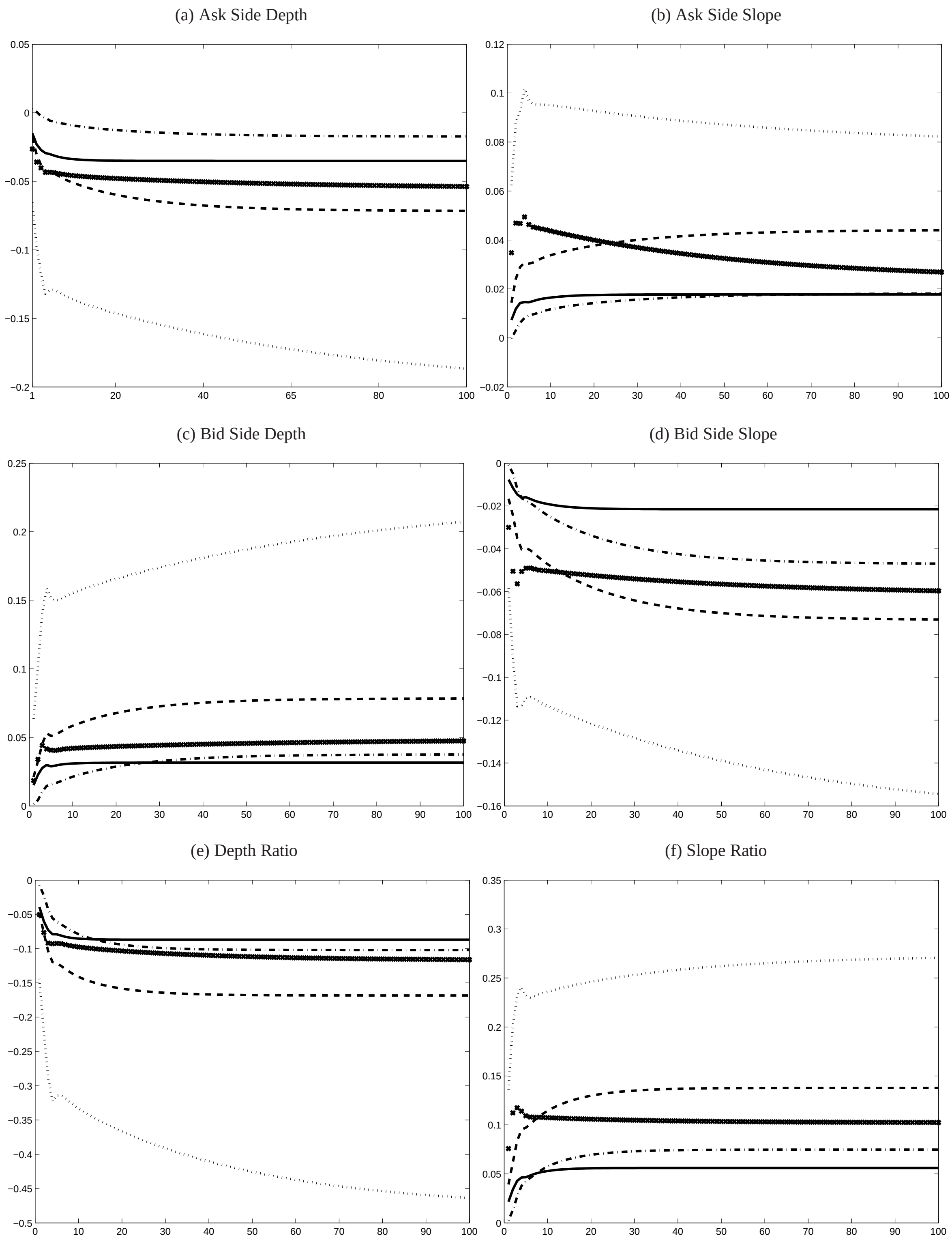
Note: This figure presents the impulse response functions of midquote return to a one unit positive shock to limit order book variables. The horizontal axis is transaction periods, i.e. the number of transactions since the initial shock, and the vertical axis is the response of midquote return in basis points. The solid, dashed, small dotted and dashed-dotted line correspond to the limit order book variable measured based on the first two, five, ten and twenty levels, respectively. The big dotted line corresponds to the limit order book variable measured between second and fifth levels.

Figure 2* (SAP-July 2010): Impulse Response Function of Return to Limit Order Book Variables



Note: This figure presents the impulse response functions of midquote return to a one unit positive shock to limit order book variables. The horizontal axis is transaction periods, i.e. the number of transactions since the initial shock, and the vertical axis is the response of midquote return in basis points. The solid, dashed, small dotted and dashed-dotted line correspond to the limit order book variable measured based on the first two, five, ten and twenty levels, respectively. The big dotted line corresponds to the limit order book variable measured between second and fifth levels.

Figure 2* (SAP-June 2011): Impulse Response Function of Return to Limit Order Book Variables



Note: This figure presents the impulse response functions of midquote return to a one unit positive shock to limit order book variables. The horizontal axis is transaction periods, i.e. the number of transactions since the initial shock, and the vertical axis is the response of midquote return in basis points. The solid, dashed, small dotted and dashed-dotted line correspond to the limit order book variable measured based on the first two, five, ten and twenty levels, respectively. The big dotted line corresponds to the limit order book variable measured between second and fifth levels.

Table 4* (MRK-June 2011): Coefficient Estimates on Lagged Values of Limit Order Book Variables (z_{t-j}) and F-statistics for Granger Causality Tests

	$\alpha_{z,1}$	$\alpha_{z,2}$	$\alpha_{z,3}$	$\alpha_{z,4}$	$\alpha_{z,5}$	F-stat
Ask Side Variables						
Depth (Levels 1-2)	-0.280***	-0.021	0.013	0.039*	0.029	484.462***
Depth (Levels 1-5)	-0.387***	0.031	-0.023	0.119***	0.086**	225.580***
Depth (Levels 1-20)	-0.460***	0.194	-0.368***	0.307**	0.163	73.831***
Depth (Levels 2-5)	-0.132***	0.019	-0.038	0.052	0.025	39.603***
Depth (Levels 6-20)	0.046	0.052	-0.253**	0.041	0.050	14.659**
Slope (Levels 1-2)	0.183***	-0.004	-0.008	-0.020	-0.016	347.733***
Slope (Levels 1-5)	0.398***	-0.100*	-0.017	-0.063	-0.100***	265.603***
Slope (Levels 1-20)	0.940***	-0.395***	-0.004	-0.172	-0.258**	241.749***
Slope (Levels 2-5)	0.134***	-0.003	0.020	-0.027	-0.050	51.356***
Slope (Levels 6-20)	0.377***	-0.149	0.032	-0.066	-0.132*	64.606***
Bid Side Variables						
Depth (Levels 1-2)	0.258***	0.034*	-0.013	-0.044**	-0.003	436.874***
Depth (Levels 1-5)	0.286***	0.071	-0.104**	-0.001	-0.042	160.880***
Depth (Levels 1-20)	0.231*	0.214	-0.170	-0.133	0.036	39.615***
Depth (Levels 2-5)	0.045	0.041	-0.055	0.080*	-0.012	28.334***
Depth (Levels 6-20)	-0.089	0.133	-0.017	-0.128	0.153*	10.034*
Slope (Levels 1-2)	-0.168***	-0.007	-0.001	0.032**	-0.004	330.725***
Slope (Levels 1-5)	-0.285***	0.022	0.097**	-0.009	0.040	159.143***
Slope (Levels 1-20)	-0.605***	0.039	0.305**	-0.033	0.203*	119.103***
Slope (Levels 2-5)	-0.050	-0.045	0.066	-0.089*	0.033	35.210***
Slope (Levels 6-20)	-0.249***	-0.001	0.140	0.003	0.074	32.606***
Ratio Variables						
Depth Ratio (Levels 1-2)	-0.682***	-0.077**	0.032	0.091**	0.036	958.045***
Depth Ratio (Levels 1-5)	-0.872***	-0.107	0.042	0.103	0.075	492.646***
Depth Ratio (Levels 1-20)	-0.864***	-0.031	-0.238	0.438**	0.043	129.374***
Depth Ratio (Levels 2-5)	-0.491***	-0.087	-0.011	0.013	-0.001	209.102***
Depth Ratio (Levels 6-20)	0.090	-0.076	-0.260*	0.168	-0.124	16.998***
Slope Ratio (Levels 1-2)	0.500***	0.028	0.007	-0.073**	-0.011	699.074***
Slope Ratio (Levels 1-5)	0.937***	-0.060	-0.045	-0.045	-0.090	556.052***
Slope Ratio (Levels 1-20)	1.780***	-0.479***	-0.330*	-0.138	-0.461***	381.203***
Slope Ratio (Levels 2-5)	0.302***	0.089	-0.008	0.101	-0.029	126.460***
Slope Ratio (Levels 6-20)	0.690***	-0.167	-0.104	-0.064	-0.201*	92.848***

Note: This table presents the parameter estimates of lagged limit order book variables in the return equation ($\alpha_{z,\tau}$) and F-statistics for Granger Causality tests. ***, **, * represent statistical significance at 1%, 5%, 10%, respectively.

Table 4* (SAP-July 2010): Coefficient Estimates on Lagged Values of Limit Order Book Variables (z_{t-j}) and F-statistics for Granger Causality Tests

	$\alpha_{z,1}$	$\alpha_{z,2}$	$\alpha_{z,3}$	$\alpha_{z,4}$	$\alpha_{z,5}$	F-stat
Ask Side Variables						
Depth (Levels 1-2)	-0.149***	-0.005	0.018**	0.028***	0.010	725.592***
Depth (Levels 1-5)	-0.170***	0.004	0.017	0.030	0.032	200.519***
Depth (Levels 1-20)	-0.435***	0.134	0.139	-0.070	0.213**	81.138***
Depth (Levels 2-5)	0.036**	-0.005	-0.035*	-0.017	0.004	20.061***
Depth (Levels 6-20)	-0.014	0.050	0.051	-0.172**	0.100	11.644**
Slope (Levels 1-2)	0.066***	0.010	-0.002	-0.005	-0.011*	292.209***
Slope (Levels 1-5)	0.119***	-0.019	-0.001	0.003	-0.051***	115.713***
Slope (Levels 1-20)	0.366***	-0.234***	-0.126	0.087	-0.096	72.913***
Slope (Levels 2-5)	-0.010	0.008	0.017	0.021	-0.021	9.634*
Slope (Levels 6-20)	0.138***	-0.126*	-0.093	0.074	-0.006	24.225***
Bid Side Variables						
Depth (Levels 1-2)	0.164***	0.006	-0.032***	0.001	-0.031***	866.889***
Depth (Levels 1-5)	0.233***	-0.004	-0.044*	-0.006	-0.066***	348.940***
Depth (Levels 1-20)	0.720***	-0.167	-0.086	-0.113	-0.287***	221.762***
Depth (Levels 2-5)	0.003	-0.001	0.046**	-0.008	0.010	34.827***
Depth (Levels 6-20)	0.123*	-0.084	0.099	-0.071	-0.044	14.388**
Slope (Levels 1-2)	-0.075***	-0.014**	0.010	0.003	0.011*	370.853***
Slope (Levels 1-5)	-0.174***	0.003	0.036	0.011	0.038*	248.443***
Slope (Levels 1-20)	-0.561***	0.162	0.153	0.104	0.103	210.339***
Slope (Levels 2-5)	-0.017	0.000	-0.029	-0.006	-0.003	49.262***
Slope (Levels 6-20)	-0.238***	0.094	0.030	0.080	0.013	54.028***
Ratio Variables						
Depth Ratio (Levels 1-2)	-0.445***	-0.013	0.077***	0.036*	0.053***	1,792.649***
Depth Ratio (Levels 1-5)	-0.484***	0.013	0.071	0.039	0.099***	598.490***
Depth Ratio (Levels 1-20)	-1.304***	0.326**	0.250*	0.037	0.523***	318.873***
Depth Ratio (Levels 2-5)	-0.028	-0.074*	-0.033	-0.026	0.013	90.826***
Depth Ratio (Levels 6-20)	-0.144	0.137	-0.031	-0.131	0.158	9.335*
Slope Ratio (Levels 1-2)	0.208***	0.045***	-0.020	-0.007	-0.032***	700.599***
Slope Ratio (Levels 1-5)	0.354***	-0.023	-0.021	-0.010	-0.077**	398.015***
Slope Ratio (Levels 1-20)	1.158***	-0.434***	-0.353**	-0.028	-0.242**	305.225***
Slope Ratio (Levels 2-5)	0.007	0.018	0.060	0.028	-0.006	55.836***
Slope Ratio (Levels 6-20)	0.463***	-0.250**	-0.168	-0.002	-0.021	62.026***

Note: This table presents the parameter estimates of lagged limit order book variables in the return equation ($\alpha_{z,\tau}$) and F-statistics for Granger Causality tests. ***, **, * represent statistical significance at 1%, 5%, 10%, respectively.

Table 4* (SAP-June 2011): Coefficient Estimates on Lagged Values of Limit Order Book Variables (z_{t-j}) and F-statistics for Granger Causality Tests

	$\alpha_{z,1}$	$\alpha_{z,2}$	$\alpha_{z,3}$	$\alpha_{z,4}$	$\alpha_{z,5}$	F-stat
Ask Side Variables						
Depth (Levels 1-2)	-0.145***	-0.011	0.010	0.011	0.017**	759.731***
Depth (Levels 1-5)	-0.181***	0.001	0.027	0.026	0.048**	227.392***
Depth (Levels 1-20)	-0.676***	0.262***	0.124	0.042	0.181**	250.467***
Depth (Levels 2-5)	0.022	-0.028	-0.014	0.012	-0.013	11.329**
Depth (Levels 6-20)	-0.276***	0.147**	0.049	0.012	0.040	55.613***
Slope (Levels 1-2)	0.075***	0.004	-0.004	-0.017**	-0.012*	298.485***
Slope (Levels 1-5)	0.150***	-0.020	-0.031	-0.034	-0.030	146.611***
Slope (Levels 1-20)	0.635***	-0.279***	-0.199**	-0.001	-0.131**	342.564***
Slope (Levels 2-5)	0.004	0.021	0.004	-0.005	-0.004	9.628*
Slope (Levels 6-20)	0.350***	-0.161***	-0.112	-0.005	-0.059	154.401***
Bid Side Variables						
Depth (Levels 1-2)	0.147***	0.007	0.005	-0.019**	-0.033***	784.009***
Depth (Levels 1-5)	0.212***	-0.027	0.030	-0.042	-0.093***	323.400***
Depth (Levels 1-20)	0.672***	-0.229**	0.045	-0.186*	-0.257***	271.469***
Depth (Levels 2-5)	0.019	0.009	0.035	-0.011	-0.026	22.445***
Depth (Levels 6-20)	0.210***	-0.058	-0.037	-0.094	-0.013	36.769***
Slope (Levels 1-2)	-0.077***	-0.001	-0.005	0.009	0.019***	336.439***
Slope (Levels 1-5)	-0.171***	0.044*	-0.035	0.046	0.064***	222.854***
Slope (Levels 1-20)	-0.608***	0.200**	0.094	0.218***	0.071	333.443***
Slope (Levels 2-5)	-0.016	-0.020	-0.036	0.016	0.027	32.190***
Slope (Levels 6-20)	-0.315***	0.068	0.129	0.116*	-0.007	142.415***
Ratio Variables						
Depth Ratio (Levels 1-2)	-0.384***	-0.020	-0.001	0.037*	0.059***	1,613.299***
Depth Ratio (Levels 1-5)	-0.476***	0.022	-0.024	0.065	0.138***	646.022***
Depth Ratio (Levels 1-20)	-1.503***	0.536***	0.086	0.230	0.453***	570.258***
Depth Ratio (Levels 2-5)	-0.093**	-0.058	-0.092	0.031	0.057*	121.941***
Depth Ratio (Levels 6-20)	-0.541***	0.242*	0.085	0.109	0.048	96.736***
Slope Ratio (Levels 1-2)	0.218***	0.013	0.013	-0.032**	-0.037***	666.664***
Slope Ratio (Levels 1-5)	0.405***	-0.066	0.035	-0.079	-0.096***	436.497***
Slope Ratio (Levels 1-20)	1.401***	-0.498***	-0.336*	-0.241**	-0.219**	716.291***
Slope Ratio (Levels 2-5)	0.041	0.056	0.054	0.001	-0.033	62.194***
Slope Ratio (Levels 6-20)	0.778***	-0.281**	-0.281	-0.129	-0.047	317.660***

Note: This table presents the parameter estimates of lagged limit order book variables in the return equation ($\alpha_{z,\tau}$) and F-statistics for Granger Causality tests. ***, **, * represent statistical significance at 1%, 5%, 10%, respectively.

Table 5* (MRK-June 2011): Bid-Ask Return on Trading Strategies

	Number of Trades	Latency 0 ms	500 ms	1000 ms
Ask Side Variables				
Depth (Levels 1-2)	25	2.633%	1.941%	1.996%
Depth (Levels 1-5)	19	2.266%	1.744%	1.796%
Depth (Levels 1-20)	21	-2.133%	-2.784%	-2.708%
Depth (Levels 2-5)	19	-7.106%	-8.311%	-8.251%
Depth (Levels 6-20)	11	2.148%	1.842%	1.816%
Slope (Levels 1-2)	24	1.633%	0.957%	0.435%
Slope (Levels 1-5)	25	1.312%	0.565%	0.073%
Slope (Levels 1-20)	22	1.545%	0.650%	0.131%
Slope (Levels 2-5)	15	-1.594%	-2.098%	-2.035%
Slope (Levels 6-20)	11	0.190%	-0.020%	-0.046%
Bid Side Variables				
Depth (Levels 1-2)	18	1.762%	1.167%	1.010%
Depth (Levels 1-5)	18	4.569%	3.709%	3.709%
Depth (Levels 1-20)	13	-5.807%	-5.795%	-5.598%
Depth (Levels 2-5)	13	3.497%	3.267%	3.267%
Depth (Levels 6-20)	13	-2.260%	-2.452%	-2.452%
Slope (Levels 1-2)	20	-0.763%	-1.760%	-1.833%
Slope (Levels 1-5)	19	3.063%	1.954%	1.796%
Slope (Levels 1-20)	18	-2.432%	-3.259%	-3.409%
Slope (Levels 2-5)	10	3.075%	2.764%	2.764%
Slope (Levels 6-20)	9	2.503%	2.116%	2.169%
Ratio Variables				
Depth Ratio (Levels 1-2)	35	1.262%	0.136%	0.032%
Depth Ratio (Levels 1-5)	35	1.379%	-0.107%	0.002%
Depth Ratio (Levels 1-20)	15	1.137%	0.500%	0.710%
Depth Ratio (Levels 2-5)	24	-2.799%	-2.929%	-2.903%
Depth Ratio (Levels 6-20)	13	-1.399%	-1.464%	-1.258%
Slope Ratio (Levels 1-2)	35	2.346%	1.097%	0.497%
Slope Ratio (Levels 1-5)	40	1.625%	-0.278%	-0.496%
Slope Ratio (Levels 1-20)	26	3.483%	3.103%	3.183%
Slope Ratio (Levels 2-5)	22	0.381%	-0.006%	-0.006%
Slope Ratio (Levels 6-20)	8	2.598%	2.344%	2.344%
Benchmark with MA filter	9	-0.233%	-0.402%	-0.377%
Benchmark without MA filter	86	-10.645%	-8.105%	-9.061%

Note: This table presents the cumulative bid-ask returns on trading strategies for MRK over the whole trading period of June 2011. The trading strategy is based on the empirical model in Equation 9 that excludes the contemporaneous effect of trade direction on return and described in detail in Section 9. The parameter of the filter for the return forecasts, κ, is set to 1 basis points. The long and short-run moving average filters are calculated based on the last transaction price and last forty transaction prices, including the most recent one, respectively. The parameter of the moving average filter is set to 0.12 euros. Benchmark with and without moving average (MA) filter present results from trading strategies based on a forecasting model that ignores information embedded in limit order book variables.

Table 6: Table 5* (SAP-July 2010): Bid-Ask Return on Trading Strategies

	Number of Trades	Latency 0 ms	500 ms	1000 ms
Ask Side Variables				
Depth (Levels 1-2)	31	8.102%	7.336%	6.717%
Depth (Levels 1-5)	31	-0.634%	-0.672%	-0.896%
Depth (Levels 1-20)	33	0.296%	0.081%	-0.253%
Depth (Levels 2-5)	29	2.757%	2.634%	2.404%
Depth (Levels 6-20)	27	1.210%	1.421%	1.329%
Slope (Levels 1-2)	31	6.881%	6.752%	6.267%
Slope (Levels 1-5)	32	4.988%	4.451%	4.004%
Slope (Levels 1-20)	27	-0.955%	-1.082%	-1.384%
Slope (Levels 2-5)	31	-3.292%	-3.482%	-3.880%
Slope (Levels 6-20)	27	2.902%	2.560%	2.246%
Bid Side Variables				
Depth (Levels 1-2)	37	1.456%	1.269%	1.232%
Depth (Levels 1-5)	39	1.417%	0.872%	0.237%
Depth (Levels 1-20)	29	1.979%	1.817%	1.760%
Depth (Levels 2-5)	31	0.574%	0.251%	-0.190%
Depth (Levels 6-20)	29	-0.467%	-0.288%	-0.404%
Slope (Levels 1-2)	35	-3.190%	-4.397%	-5.165%
Slope (Levels 1-5)	37	1.629%	1.093%	0.392%
Slope (Levels 1-20)	25	2.196%	1.829%	1.815%
Slope (Levels 2-5)	27	1.161%	1.358%	1.076%
Slope (Levels 6-20)	27	3.472%	3.366%	3.078%
Ratio Variables				
Depth Ratio (Levels 1-2)	45	3.822%	3.529%	3.377%
Depth Ratio (Levels 1-5)	39	5.662%	5.304%	5.038%
Depth Ratio (Levels 1-20)	30	7.201%	7.258%	6.824%
Depth Ratio (Levels 2-5)	29	-0.355%	-0.430%	-0.734%
Depth Ratio (Levels 6-20)	29	4.221%	4.322%	4.228%
Slope Ratio (Levels 1-2)	39	8.819%	9.001%	8.664%
Slope Ratio (Levels 1-5)	37	3.951%	3.130%	2.339%
Slope Ratio (Levels 1-20)	29	4.067%	4.087%	3.754%
Slope Ratio (Levels 2-5)	29	2.328%	2.237%	1.925%
Slope Ratio (Levels 6-20)	27	1.082%	0.729%	0.585%
Benchmark with MA filter	29	-2.231%	-2.001%	-2.115%
Benchmark without MA filter	536	-23.693%	-26.451%	-27.623%

Note: This table presents the cumulative bid-ask returns on trading strategies for SAP over the whole trading period of July 2010. The trading strategy is based on the empirical model in Equation 9 that excludes the contemporaneous effect of trade direction on return and described in detail in Section 9. The parameter of the filter for the return forecasts, κ, is set to 0.5 basis points. The long and short-run moving average filters are calculated based on the last two and forty transaction prices, including the most recent one, respectively. The parameter of the moving average filter is set to 0.05 euros. Benchmark with and without moving average (MA) filter present results from trading strategies based on a forecasting model that ignores information embedded in limit order book variables.

Table 7: Table 5* (SAP-June 2011): Bid-Ask Return on Trading Strategies

	Number of Trades	Latency 0 ms	500 ms	1000 ms
Ask Side Variables				
Depth (Levels 1-2)	22	2.546%	2.475%	2.388%
Depth (Levels 1-5)	24	3.391%	3.841%	3.383%
Depth (Levels 1-20)	24	0.297%	0.323%	0.264%
Depth (Levels 2-5)	24	3.292%	3.841%	3.383%
Depth (Levels 6-20)	26	2.753%	3.323%	2.844%
Slope (Levels 1-2)	25	0.373%	0.257%	0.222%
Slope (Levels 1-5)	24	3.490%	3.940%	3.482%
Slope (Levels 1-20)	24	0.777%	0.947%	0.864%
Slope (Levels 2-5)	24	3.292%	3.841%	3.383%
Slope (Levels 6-20)	26	2.802%	3.175%	2.696%
Bid Side Variables				
Depth (Levels 1-2)	27	4.693%	4.539%	4.783%
Depth (Levels 1-5)	27	4.618%	4.370%	4.384%
Depth (Levels 1-20)	26	5.597%	5.535%	5.460%
Depth (Levels 2-5)	25	3.102%	3.526%	3.070%
Depth (Levels 6-20)	25	4.137%	4.565%	4.105%
Slope (Levels 1-2)	25	1.597%	1.990%	1.516%
Slope (Levels 1-5)	27	3.938%	3.668%	3.631%
Slope (Levels 1-20)	25	5.268%	5.010%	4.910%
Slope (Levels 2-5)	25	3.102%	3.526%	3.070%
Slope (Levels 6-20)	24	2.877%	2.843%	2.794%
Ratio Variables				
Depth Ratio (Levels 1-2)	25	4.543%	4.264%	4.533%
Depth Ratio (Levels 1-5)	23	5.095%	5.044%	5.059%
Depth Ratio (Levels 1-20)	24	5.707%	5.811%	5.620%
Depth Ratio (Levels 2-5)	25	3.005%	3.576%	3.120%
Depth Ratio (Levels 6-20)	27	3.520%	4.094%	3.611%
Slope Ratio (Levels 1-2)	25	0.166%	-0.265%	-0.129%
Slope Ratio (Levels 1-5)	25	5.193%	5.172%	5.135%
Slope Ratio (Levels 1-20)	23	3.028%	2.913%	3.061%
Slope Ratio (Levels 2-5)	25	3.005%	3.576%	3.120%
Slope Ratio (Levels 6-20)	25	2.601%	2.960%	2.469%
Benchmark with MA filter	25	3.005%	3.428%	2.972%
Benchmark without MA filter	24536	-99.877%	-99.961%	-99.969%

Note: This table presents the cumulative bid-ask returns on trading strategies for SAP over the whole trading period of June 2011. The trading strategy is based on the empirical model in Equation 9 that excludes the contemporaneous effect of trade direction on return and described in detail in Section 9. The parameter of the filter for the return forecasts, κ, is set to 0.3 basis points. The long and short-run moving average filters are calculated based on the last five and forty transaction prices, including the most recent one, respectively. The parameter of the moving average filter is set to 0.06 euros. Benchmark with and without moving average (MA) filter present results from trading strategies based on a forecasting model that ignores information embedded in limit order book variables.

Table A1: Coefficient Estimates on Lagged Values of Limit Order Book Variables and Lagged Values of Volume and F-statistics for Granger Causality Tests for Limit Order Book Variables

	$\alpha_{z,1}$	$\alpha_{z,2}$	$\alpha_{z,3}$	$\alpha_{z,4}$	$\alpha_{z,5}$	$\theta_{z,1}$	$\theta_{z,2}$	$\theta_{z,3}$	$\theta_{z,4}$	$\theta_{z,5}$	F-stat
Ask Side Variables											
Depth (Levels 1-2)	-0.260***	-0.001	0.006	0.028	0.018	0.021**	0.010	-0.006	0.007	-0.001	415.482***
Depth (Levels 1-5)	-0.225***	0.019	0.033	-0.005	0.063*	0.012	0.007	-0.007	0.005	-0.004	76.81***
Depth (Levels 1-20)	-0.533***	0.225	0.004	-0.228	0.401***	0.010	0.006	-0.006	0.006	-0.004	52.36***
Depth (Levels 2-5)	0.057*	-0.021	0.022	-0.074**	0.026	0.007	0.006	-0.006	0.006	-0.004	11.16**
Depth (Levels 6-20)	-0.101	0.131	-0.091	-0.216*	0.194*	0.008	0.006	-0.006	0.006	-0.003	16.71***
Slope (Levels 1-2)	0.132***	-0.009	0.002	-0.009	-0.018	0.011	0.007	-0.007	0.005	-0.004	194.29***
Slope (Levels 1-5)	0.190***	-0.027	-0.039	0.014	-0.074**	0.010	0.005	-0.008	0.005	-0.005	68.35***
Slope (Levels 1-20)	0.575***	-0.066	-0.324**	0.267	-0.380***	0.007	0.005	-0.007	0.005	-0.004	92.95***
Slope (Levels 2-5)	-0.019	0.030	-0.022	0.059*	-0.033	0.008	0.006	-0.006	0.006	-0.004	6.79
Slope (Levels 6-20)	0.237**	-0.010	-0.169	0.196*	-0.199**	0.007	0.006	-0.006	0.005	-0.003	32.37***
Bid Side Variables											
Depth (Levels 1-2)	0.286***	0.001	-0.001	-0.028	-0.019	-0.006	0.000	-0.010	0.003	-0.008	444.45***
Depth (Levels 1-5)	0.260***	-0.042	-0.009	-0.065	-0.027	0.002	0.005	-0.007	0.006	-0.004	92.20***
Depth (Levels 1-20)	0.627***	-0.308	-0.046	-0.059	-0.128	0.006	0.006	-0.006	0.006	-0.004	51.74***
Depth (Levels 2-5)	-0.008	-0.006	0.017	-0.022	0.023	0.008	0.006	-0.006	0.006	-0.004	1.24
Depth (Levels 6-20)	0.240**	-0.123	-0.084	0.068	-0.056	0.008	0.006	-0.006	0.005	-0.004	11.67**
Slope (Levels 1-2)	-0.151***	-0.022	-0.011	0.022	0.007	0.003	0.005	-0.007	0.005	-0.004	302.62***
Slope (Levels 1-5)	-0.278***	0.015	0.045	0.095**	0.031	0.004	0.006	-0.005	0.008	-0.002	154.25***
Slope (Levels 1-20)	-0.841***	0.186	0.292**	0.152	0.131	0.008	0.007	-0.005	0.007	-0.003	172.45***
Slope (Levels 2-5)	-0.049	-0.006	0.002	0.022	0.000	0.007	0.006	-0.006	0.006	-0.003	8.76
Slope (Levels 6-20)	-0.401***	0.080	0.182*	-0.038	0.118	0.009	0.006	-0.006	0.006	-0.004	62.40***
Ratio Variables											
Depth Ratio (Levels 1-2)	-0.751***	0.003	0.008	0.092***	0.047	0.008	0.005	-0.008	0.005	-0.005	943.24***
Depth Ratio (Levels 1-5)	-0.579***	0.073	0.018	0.081	0.096*	0.007	0.005	-0.008	0.005	-0.004	182.23***
Depth Ratio (Levels 1-20)	-1.386***	0.573**	0.039	-0.170	0.460**	0.007	0.005	-0.007	0.006	-0.004	129.26***
Depth Ratio (Levels 2-5)	-0.060	0.007	-0.058	0.005	-0.014	0.008	0.006	-0.007	0.005	-0.004	11.46**
Depth Ratio (Levels 6-20)	-0.436***	0.265	-0.004	-0.280	0.188	0.008	0.006	-0.007	0.006	-0.004	31.14***
Slope Ratio (Levels 1-2)	0.404***	0.024	0.033	-0.035	-0.038	0.008	0.006	-0.007	0.005	-0.005	475.73***
Slope Ratio (Levels 1-5)	0.582***	-0.043	-0.080	-0.091	-0.096*	0.007	0.005	-0.008	0.005	-0.004	225.27***
Slope Ratio (Levels 1-20)	1.897***	-0.278	-0.724***	0.099	-0.561***	0.006	0.005	-0.006	0.006	-0.004	339.93***
Slope Ratio (Levels 2-5)	0.040	0.035	0.008	0.022	-0.032	0.008	0.006	-0.007	0.005	-0.004	7.41
Slope Ratio (Levels 6-20)	0.825***	-0.104	-0.407**	0.283*	-0.351***	0.007	0.006	-0.006	0.006	-0.004	115.06***

Note: This table presents the parameter estimates of lagged limit order book variables ($\alpha_{z,\tau}$) and lagged volume ($\theta_{z,\tau}$) in the return equation and F-statistics for Granger Causality tests for limit order book variables. ***, **, * represent statistical significance at 1%, 5%, 10%, respectively.

Table A2: Coefficient Estimates on Lagged Values of Limit Order Book Variables and Lagged Values of Volatility and F-statistics for Granger Causality Tests for Limit Order Book Variables

	$\alpha_{z,1}$	$\alpha_{z,2}$	$\alpha_{z,3}$	$\alpha_{z,4}$	$\alpha_{z,5}$	$\theta_{z,1}$	$\theta_{z,2}$	$\theta_{z,3}$	$\theta_{z,4}$	$\theta_{z,5}$	F-stat
Ask Side Variables											
Depth (Levels 1-2)	-0.261***	0.004	0.007	0.028	0.019	-0.005	-0.013	0.015	-0.001	0.007	408.561***
Depth (Levels 1-5)	-0.224***	0.024	0.032	-0.006	0.061*	-0.006	-0.013	0.015	0.000	0.006	74.764***
Depth (Levels 1-20)	-0.532***	0.236	-0.001	-0.227	0.395***	-0.006	-0.013	0.015	-0.001	0.006	51.014***
Depth (Levels 2-5)	0.058**	-0.019	0.022	-0.074**	0.025	-0.006	-0.012	0.016	0.001	0.007	11.359**
Depth (Levels 6-20)	-0.101	0.131	-0.089	-0.215*	0.192*	-0.006	-0.013	0.015	0.000	0.007	16.162***
Slope (Levels 1-2)	0.132***	-0.010	0.001	-0.008	-0.018	-0.006	-0.014	0.014	-0.001	0.006	192.846***
Slope (Levels 1-5)	0.190***	-0.028	-0.041	0.015	-0.073**	-0.007	-0.014	0.014	-0.001	0.006	67.657***
Slope (Levels 1-20)	0.581***	-0.058	-0.332**	0.267	-0.383***	-0.010	-0.014	0.014	-0.001	0.007	95.005***
Slope (Levels 2-5)	-0.020	0.030	-0.022	0.059*	-0.031	-0.006	-0.013	0.015	0.000	0.007	6.517
Slope (Levels 6-20)	0.241**	-0.005	-0.171	0.193*	-0.202**	-0.008	-0.014	0.014	0.000	0.007	33.211***
Bid Side Variables											
Depth (Levels 1-2)	0.286***	0.001	-0.003	-0.030	-0.020	-0.006	-0.012	0.016	0.002	0.008	445.273***
Depth (Levels 1-5)	0.262***	-0.037	-0.012	-0.067	-0.027	-0.007	-0.013	0.016	0.002	0.008	95.265***
Depth (Levels 1-20)	0.627***	-0.298	-0.048	-0.062	-0.129	-0.006	-0.012	0.016	0.001	0.008	52.519***
Depth (Levels 2-5)	-0.007	-0.004	0.016	-0.023	0.022	-0.006	-0.013	0.015	0.000	0.007	1.123
Depth (Levels 6-20)	0.237**	-0.122	-0.081	0.068	-0.056	-0.005	-0.012	0.015	0.000	0.007	11.546**
Slope (Levels 1-2)	-0.151***	-0.022	-0.012	0.022	0.007	-0.004	-0.010	0.017	0.003	0.010	306.058***
Slope (Levels 1-5)	-0.278***	0.013	0.043	0.095**	0.030	-0.004	-0.011	0.017	0.002	0.009	156.625***
Slope (Levels 1-20)	-0.837***	0.188	0.287**	0.145	0.131	-0.001	-0.010	0.016	0.000	0.007	170.746***
Slope (Levels 2-5)	-0.050*	-0.008	0.003	0.023	0.001	-0.005	-0.012	0.016	0.001	0.008	9.553*
Slope (Levels 6-20)	-0.397***	0.083	0.179*	-0.044	0.117	-0.003	-0.011	0.016	0.001	0.007	61.430***
Ratio Variables											
Depth Ratio (Levels 1-2)	-0.752***	0.004	0.008	0.093***	0.047	-0.006	-0.014	0.015	0.000	0.008	944.711***
Depth Ratio (Levels 1-5)	-0.581***	0.074	0.019	0.083	0.095*	-0.007	-0.014	0.015	0.000	0.007	183.617***
Depth Ratio (Levels 1-20)	-1.385***	0.571**	0.039	-0.167	0.457**	-0.006	-0.013	0.014	-0.001	0.007	129.066***
Depth Ratio (Levels 2-5)	-0.062	0.007	-0.056	0.005	-0.014	-0.006	-0.013	0.015	0.000	0.007	11.469**
Depth Ratio (Levels 6-20)	-0.432***	0.264	-0.005	-0.281	0.186	-0.006	-0.013	0.015	0.000	0.007	30.852***
Slope Ratio (Levels 1-2)	0.404***	0.023	0.032	-0.034	-0.038	-0.005	-0.012	0.015	0.001	0.008	475.647***
Slope Ratio (Levels 1-5)	0.582***	-0.041	-0.082	-0.090	-0.096*	-0.006	-0.013	0.015	0.001	0.007	225.613***
Slope Ratio (Levels 1-20)	1.893***	-0.275	-0.725***	0.105	-0.562***	-0.005	-0.012	0.015	0.000	0.007	339.446***
Slope Ratio (Levels 2-5)	0.040	0.037	0.007	0.021	-0.031	-0.006	-0.013	0.015	0.000	0.007	7.473
Slope Ratio (Levels 6-20)	0.823***	-0.104	-0.406**	0.287*	-0.352***	-0.006	-0.012	0.015	0.000	0.007	114.885***

Note: This table presents the parameter estimates of lagged limit order book variables ($\alpha_{z,\tau}$) and lagged volatility ($\theta_{z,\tau}$) in the return equation and F-statistics for Granger Causality tests for limit order book variables. ***, **, * represent statistical significance at 1%, 5%, 10%, respectively.

Table A3: Coefficient Estimates on Lagged Values of Limit Order Book Variables and Lagged Values of Duration and F-statistics for Granger Causality Tests for Limit Order Book Variables

	$\alpha_{z,1}$	$\alpha_{z,2}$	$\alpha_{z,3}$	$\alpha_{z,4}$	$\alpha_{z,5}$	$\theta_{z,1}$	$\theta_{z,2}$	$\theta_{z,3}$	$\theta_{z,4}$	$\theta_{z,5}$	F-stat
Ask Side Variables											
Depth (Levels 1-2)	-0.261***	0.001	0.008	0.027	0.022	0.008**	0.000	0.007**	-0.003	0.000	411.239***
Depth (Levels 1-5)	-0.225***	0.017	0.035	-0.010	0.064*	0.006*	0.000	0.009**	-0.003	0.001	78.238***
Depth (Levels 1-20)	-0.536***	0.224	-0.001	-0.226	0.402***	0.006*	0.000	0.008**	-0.003	0.001	53.656***
Depth (Levels 2-5)	0.057*	-0.024	0.023	-0.076**	0.025	0.005*	-0.001	0.008**	-0.003	0.001	11.375**
Depth (Levels 6-20)	-0.105	0.134	-0.100	-0.208*	0.192*	0.006*	-0.001	0.008**	-0.002	0.001	17.265***
Slope (Levels 1-2)	0.133***	-0.006	0.002	-0.007	-0.018	0.008**	0.001	0.009**	-0.002	0.001	200.595***
Slope (Levels 1-5)	0.190***	-0.021	-0.039	0.021	-0.076**	0.007**	0.001	0.009***	-0.003	0.002	73.024***
Slope (Levels 1-20)	0.585***	-0.062	-0.316**	0.264	-0.385***	0.008**	0.001	0.008**	-0.002	0.001	98.034***
Slope (Levels 2-5)	-0.017	0.033	-0.022	0.062*	-0.032	0.006*	0.000	0.008**	-0.003	0.001	8.238
Slope (Levels 6-20)	0.248**	-0.015	-0.160	0.187	-0.197**	0.007**	0.000	0.008**	-0.002	0.001	34.776***
Bid Side Variables											
Depth (Levels 1-2)	0.286***	0.000	-0.002	-0.031	-0.019	0.002	-0.003	0.007**	-0.003	0.000	439.588***
Depth (Levels 1-5)	0.259***	-0.044	-0.008	-0.070	-0.024	0.004	-0.002	0.007**	-0.004	0.000	90.746***
Depth (Levels 1-20)	0.623***	-0.309	-0.046	-0.062	-0.125	0.005	-0.001	0.007**	-0.003	0.000	50.207***
Depth (Levels 2-5)	-0.009	-0.008	0.017	-0.025	0.022	0.006*	-0.001	0.008**	-0.003	0.001	1.348
Depth (Levels 6-20)	0.235**	-0.122	-0.088	0.071	-0.058	0.005	-0.001	0.008**	-0.003	0.000	10.840*
Slope (Levels 1-2)	-0.151***	-0.022	-0.012	0.023*	0.006	0.001	-0.004	0.006*	-0.004	-0.001	299.381***
Slope (Levels 1-5)	-0.277***	0.018	0.042	0.101**	0.027	0.003	-0.002	0.007**	-0.004	0.000	151.560***
Slope (Levels 1-20)	-0.837***	0.185	0.296**	0.151	0.128	0.002	-0.002	0.007**	-0.003	0.001	167.886***
Slope (Levels 2-5)	-0.048	-0.004	0.002	0.025	0.000	0.005	-0.001	0.007**	-0.003	0.000	7.551
Slope (Levels 6-20)	-0.394***	0.076	0.189*	-0.046	0.120	0.004	-0.002	0.007**	-0.004	0.000	59.089***
Ratio Variables											
Depth Ratio (Levels 1-2)	-0.751***	0.004	0.009	0.092***	0.047	0.004	-0.002	0.007**	-0.004	0.000	939.840***
Depth Ratio (Levels 1-5)	-0.578***	0.074	0.019	0.082	0.095*	0.005	-0.001	0.008**	-0.003	0.001	181.279***
Depth Ratio (Levels 1-20)	-1.383***	0.575**	0.036	-0.165	0.459**	0.005	-0.001	0.007**	-0.003	0.000	127.238***
Depth Ratio (Levels 2-5)	-0.060	0.006	-0.057	0.006	-0.014	0.005	-0.001	0.008**	-0.003	0.001	11.056*
Depth Ratio (Levels 6-20)	-0.432***	0.267	-0.008	-0.276	0.189	0.005	-0.001	0.007**	-0.003	0.000	29.960***
Slope Ratio (Levels 1-2)	0.404***	0.024	0.033	-0.035	-0.038	0.005	-0.001	0.007**	-0.003	0.001	474.658***
Slope Ratio (Levels 1-5)	0.582***	-0.042	-0.081	-0.092	-0.096*	0.005	-0.001	0.008**	-0.003	0.001	225.031***
Slope Ratio (Levels 1-20)	1.893***	-0.278	-0.722***	0.098	-0.560***	0.004	-0.002	0.006*	-0.004	0.000	337.104***
Slope Ratio (Levels 2-5)	0.041	0.035	0.007	0.021	-0.032	0.005	-0.001	0.008**	-0.003	0.001	7.387
Slope Ratio (Levels 6-20)	0.821***	-0.105	-0.404**	0.282*	-0.351***	0.005	-0.001	0.007**	-0.003	0.000	113.145***

Note: This table presents the parameter estimates of lagged limit order book variables ($\alpha_{z,\tau}$) and lagged duration ($\theta_{z,\tau}$) in the return equation and F-statistics for Granger Causality tests for limit order book variables. ***, **, * represent statistical significance at 1%, 5%, 10%, respectively.

Table A4: Daily Coefficient Estimates on the First Lag of Limit Order Book Variables in the Return Equation

	Trading Days in July 2010																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Ask Side Variables																						
Depth (Levels 1-2)	-0.307***	-0.146	-0.492***	-0.307***	-0.339***	-0.202***	-0.284***	-0.173**	-0.286***	-0.177***	-0.206***	-0.227***	-0.100	-0.284***	-0.358***	-0.192*	-0.346***	-0.493***	-0.165***	-0.431***	-0.303***	-0.196***
Depth (Levels 1-5)	-0.522**	-0.031	-0.357*	-0.192	-0.218	-0.013	-0.092	-0.151	-0.023	0.028	-0.201*	-0.128	-0.097	-0.264	-0.404***	0.186	-0.313*	-0.441**	-0.257**	-0.990***	-0.396**	-0.202
Depth (Levels 1-20)	-1.704***	-0.018	-1.122	0.005	-2.135***	-0.179	-0.742	0.078	0.048	-0.687*	0.093	-0.430	-0.567	0.214	-1.789*	0.136	-1.295**	-2.101***	-0.459	-1.533***	-0.103	-0.385
Depth (Levels 2-5)	-0.113	0.132	0.020	0.067	0.066	0.142	0.174*	0.034	0.239**	0.163**	0.008	0.082	-0.097	0.100	-0.158	0.364*	-0.067	0.005	-0.025	-0.370**	0.078	0.046
Depth (Levels 6-20)	-0.965	0.166	-0.013	0.270	-1.095**	-0.268	-0.405	0.264	0.063	-0.669*	0.450	-0.093	0.202	0.647	-0.661	-0.202	-0.716	-1.523**	0.050	-0.215	0.423	-0.189
Slope (Levels 1-2)	0.267***	0.080	0.315***	0.088*	0.129**	0.134***	0.124***	0.060	0.119**	0.128***	0.077*	0.086	-0.027	0.196***	0.251***	0.075	0.241***	0.314***	0.054**	0.219***	0.158***	0.084**
Slope (Levels 1-5)	0.511***	0.224	0.305	0.086	0.140	0.053	-0.032	0.105	0.002	-0.039	0.169	0.129	0.158	0.278	0.382***	-0.168	0.424***	0.282	0.140	0.709***	0.249**	0.158
Slope (Levels 1-20)	-0.081	0.829**	0.406	-0.228	0.224	0.512	-0.631	0.098	1.174**	0.724*	-0.489	0.956**	0.838*	0.593	1.249*	0.413	1.973***	1.130**	0.365	1.409***	0.672*	0.968***
Slope (Levels 2-5)	0.094	-0.038	-0.048	-0.056	-0.020	-0.068	-0.185**	0.003	-0.210*	-0.153**	0.001	0.019	0.193	0.046	0.115	-0.283	0.120	-0.135	0.072	0.322*	-0.061	0.017
Slope (Levels 6-20)	-0.357	0.367	-0.225	-0.249	-0.008	0.311	-0.348	-0.064	0.708*	0.592*	-0.622*	0.462	0.357	0.179	0.364	0.608**	1.000*	0.897**	0.096	0.328	0.184	0.750**
Bid Side Variables																						
Depth (Levels 1-2)	0.279***	0.169*	0.240**	0.386***	0.337***	0.309***	0.281***	0.248***	0.272***	0.156***	0.268***	0.200***	0.371***	0.376***	0.308***	0.334***	0.345***	0.268***	0.343***	0.136	0.388***	0.356***
Depth (Levels 1-5)	0.449**	-0.094	0.118	0.455***	-0.020	0.191	0.177	0.327**	0.335**	0.153	0.315***	-0.075	0.213	0.465**	0.061	0.467**	0.097	0.110	0.508***	0.256	0.534***	0.349***
Depth (Levels 1-20)	-0.344	0.366	1.087	0.791	-0.302	0.046	0.917**	1.048**	0.746	0.609	1.017***	1.211**	1.631*	1.874***	1.029	1.260*	1.198	1.279**	0.860**	-0.532	0.426	0.298
Depth (Levels 2-5)	0.116	-0.193	-0.057	0.119	-0.293	-0.163	-0.009	0.109	0.163	-0.024	-0.047	-0.297***	-0.060	0.096	-0.165	0.156	-0.228	-0.150	0.220***	0.058	0.038	0.057
Depth (Levels 6-20)	-1.057	0.327	0.692	-0.347	-0.230	-0.343	0.375	0.338	0.110	0.200	0.644**	1.059***	1.238*	1.219**	1.261***	0.205	0.904	0.928*	0.033	-0.659	0.094	0.059
Slope (Levels 1-2)	-0.112*	-0.119**	-0.174**	-0.226***	-0.199***	-0.144***	-0.166***	-0.112**	-0.185***	-0.068*	-0.043	-0.050	-0.132	-0.106*	-0.099**	-0.129*	-0.132**	-0.277***	-0.158***	-0.103*	-0.258***	-0.192***
Slope (Levels 1-5)	-0.384**	-0.063	-0.080	-0.376***	0.045	-0.183	-0.162	-0.286**	-0.327**	-0.201*	-0.069	0.189*	-0.255	-0.278	-0.298	-0.261	-0.095	-0.445**	-0.395***	-0.538***	-0.474***	-0.386***
Slope (Levels 1-20)	-0.906	-0.739	-0.624	0.023	-0.986*	-1.117**	-0.667**	-0.559	-0.435	-0.436	-0.942***	-1.473***	-1.189*	-1.164*	-1.753***	-0.989***	-1.057	-0.728*	-0.627**	-0.355	-0.751**	-0.865***
Slope (Levels 2-5)	-0.083	0.048	0.033	-0.060	0.281	0.128	0.004	-0.103	-0.165	0.020	-0.055	0.320***	-0.040	-0.169	-0.051	-0.081	0.254*	0.098	-0.204***	-0.301**	-0.123	-0.113
Slope (Levels 6-20)	-0.349	-0.478	-0.449	0.384	-0.616	-0.445	-0.270	-0.026	-0.031	0.020	-0.721***	-1.191***	-0.423	-0.913*	-1.148***	-0.414	-0.797	-0.122	-0.193	0.211	-0.200	-0.455*
Ratio Variables																						
Depth Ratio (Levels 1-2)	-0.806***	-0.451***	-0.937***	-0.917***	-0.880***	-0.765***	-0.798***	-0.513***	-0.738***	-0.487***	-0.637***	-0.603***	-0.546***	-0.869***	-0.910***	-0.780***	-0.942***	-1.127***	-0.645***	-0.822***	-0.939***	-0.699***
Depth Ratio (Levels 1-5)	-1.034***	0.070	-0.672*	-0.762***	-0.219	-0.194	-0.336	-0.554**	-0.412	-0.146	-0.576***	-0.035	-0.268	-0.778***	-0.617**	-0.500	-0.503*	-0.773**	-1.061***	-1.396***	-1.060***	-0.650***
Depth Ratio (Levels 1-20)	-1.335	-0.457	-2.732**	-0.701	-2.093**	-0.402	-1.880**	-1.094	-0.787	-1.518**	-1.285**	-2.080***	-2.398**	-2.100**	-3.484**	-1.077	-2.576**	-3.529***	-1.502***	-0.744	-0.780	-0.752
Depth Ratio (Levels 2-5)	-0.483*	0.195	-0.329	-0.151	0.313	0.198	0.100	-0.167	-0.051	0.155	0.054	0.489**	0.239	-0.191	-0.092	0.281	0.005	-0.215	-0.413***	-0.640**	-0.183	-0.134
Depth Ratio (Levels 6-20)	0.440	-0.211	-0.717	0.633	-1.027	-0.067	-0.859	-0.054	-0.029	-1.079**	-0.345	-1.562***	-0.852	-0.533	-2.121*	-0.423	-1.698*	-2.152**	0.013	0.612	0.100	-0.243
Slope Ratio (Levels 1-2)	0.562***	0.312**	0.688***	0.439***	0.392***	0.475***	0.421***	0.224**	0.433***	0.280***	0.152*	0.185	0.190	0.395***	0.446***	0.277**	0.509***	0.783***	0.263***	0.503***	0.565***	0.399***
Slope Ratio (Levels 1-5)	0.948***	0.372	0.646*	0.533**	0.101	0.278	0.133	0.459**	0.348	0.150	0.311	-0.204	0.510	0.732***	0.876***	0.166	0.652**	1.080***	0.756***	1.539***	0.894***	0.652***
Slope Ratio (Levels 1-20)	1.039	2.073***	1.571	-0.118	1.460	2.415***	0.090	0.767	2.143***	1.426**	1.025*	2.990***	2.275**	2.138***	3.483***	2.016***	3.656***	1.962***	1.180**	1.852**	1.610***	2.023***
Slope Ratio (Levels 2-5)	0.144	-0.102	0.016	0.014	-0.316	-0.254	-0.245	0.118	-0.103	-0.236*	-0.018	-0.455**	0.241	0.157	0.223	-0.209	-0.106	-0.089	0.381***	0.800***	0.135	0.110
Slope Ratio (Levels 6-20)	0.141	0.990	0.104	-0.650	0.688	1.077**	-0.054	-0.035	1.080*	0.729	0.383	2.086***	0.924	1.211*	1.613	1.341**	2.064**	0.911	0.317	0.130	0.451	1.324***

Note: This table presents the parameter estimates on the first lag of limit order book variables in the return equation ($\alpha_{z,1}$) in each trading day in July 2010. ***, **, * represent statistical significance at 1%, 5%, 10%, respectively.