Determinants of immediate price impacts at the trade level in an emerging order-driven market

Wei-Xing Zhou\textsuperscript{1,2,3}

\textsuperscript{1} School of Business, East China University of Science and Technology, Shanghai 200237, People’s Republic of China
\textsuperscript{2} School of Science, East China University of Science and Technology, Shanghai 200237, People’s Republic of China
\textsuperscript{3} Research Center for Econophysics, East China University of Science and Technology, Shanghai 200237, People’s Republic of China

E-mail: wxzhou@ecust.edu.cn

New Journal of Physics 14 (2012) 023055 (20pp)
Received 27 August 2011
Published 24 February 2012
Online at http://www.njp.org/
doi:10.1088/1367-2630/14/2/023055

Abstract. Common wisdom argues that, in general, large trades cause large price changes, whereas small trades cause small price changes. However, for extremely large price changes, the trade size and news play a minor role, while liquidity (especially price gaps on the limit order book) is a more influential factor. Hence, there might be other factors influencing the immediate price impacts of trades. In this paper, through mechanical analysis of price variations before and after a trade of arbitrary size, we identify that the trade size, the bid–ask spread, the price gaps and the outstanding volumes at the bid and ask sides of the limit order book have an impact on the changes in prices. We propose two regression models to investigate the influence of these microscopic factors on the price impact of buyer-initiated partially filled trades, seller-initiated partially filled trades, buyer-initiated filled trades and seller-initiated filled trades. We find that they have quantitatively similar explanatory powers and these factors can account for up to 44\% of the price impacts. Large trade sizes, wide bid–ask spreads, high liquidity at the same side and low liquidity at the opposite side will cause a large price impact. We also find that the liquidity at the opposite side has a more influential impact than the liquidity at the same side. Our results shed new light on the determinants of immediate price impacts.
1. Introduction

Econophysics is an interdisciplinary of physics and economics that applies the methods and ideas of statistical physics to economic and financial systems [1]. The distribution of assets is one of the most studied topics in econophysics [1] and in traditional finance as well [2], since it is the cornerstone of risk management and option pricing. The first work on this topic is usually attributed to Bachelier in 1900, where Brownian motion was used to model the evolution of stock prices [3]. The Gaussian distribution of stock returns was rediscovered by Osborne in 1959 [4], which is a key assumption in the Black–Scholes option pricing model [5]. A pioneering work was done by Mandelbrot in 1963, who argued that speculative price variations of cotton conform to the Lévy distribution [6], which is a natural followup of his studies on income distributions [7–10]. Using high-frequency data, Mantegna and Stanley [11] found that a truncated Lévy distribution is a better model for the S&P 500 index returns. Further research showed that the returns have power-law tails with the exponent close to three, known as the inverse cubic law [12–15]. A number of empirical investigations have been conducted on financial returns at different time scales in different stock markets over different time periods, and the distributions are found to have power-law tails with different tail exponents at short time intervals [16–31], which can be explained by the variational theory for turbulent signals [32, 33]. In addition, it has been shown that the exponent values are universal for three mature markets that do not display variations with respect to market capitalization or industrial sector [34, 35], which is however not the case for emerging markets [36].

So what causes the power-law tails in stock returns? A well-known adage on Wall Street states that it takes trading volume to move stock prices. The correlation between price fluctuation $r_{\Delta t}$ and trading volume $V_{\Delta t}$ has been extensively studied [37]. Dominating evidence shows that the magnitude of price fluctuation is linearly correlated with the trading volume at
different time scales $\Delta t$ from 1 min to 1 month \[38-46\]:

$$r_{\Delta t} = f(V_{\Delta t}) = \begin{cases} 
V_{\Delta t}/\lambda^+, & \text{for } V_{\Delta t} > 0, \\
V_{\Delta t}/\lambda^-, & \text{for } V_{\Delta t} < 0,
\end{cases} \quad (1)$$

where $f$ is the price impact function and $\lambda$ is a measure of market liquidity. In addition, the price–volume relation is asymmetric such that $\lambda^+ > \lambda^- > 0$; that is, the price impact of a selling volume is larger than a buying volume of the same size \[37\]. At the transaction level, theoretical analyses and empirical evidence show that the price impact function is nonlinear \[47–60\].

Concerning the volume–return relation at the transaction level, we are interested in the determination of the immediate price impact $r_{k,t+1}$ of trade size $\omega_{k,t}$ for a given stock $k$. The immediate price impact can be calculated as follows:

$$r_{k,t+1} = \frac{p_{k,t+1} - b_{k,t}}{p_{k,t}} = \frac{a_{k,t+1} + b_{k,t+1} - a_{k,t} - b_{k,t}}{a_{k,t} + b_{k,t}}, \quad (2)$$

where $p_{k,t}$, $a_{k,t}$, and $b_{k,t}$ are the mid-price, the best ask price and the best bid price right before the trade and $p_{k,t+1}$, $a_{k,t+1}$ and $b_{k,t+1}$ are the mid-price, the best ask price and the best bid price immediately after the trade. Using the Trade and Quote (TAQ) database, Lillo et al \[54\] unveiled a master curve for the price impact function such that the data collapse onto a single curve:

$$r \left( \frac{\omega}{\langle \omega \rangle}, C \right) = C^{-\gamma} g \left( \frac{\omega}{\langle \omega \rangle} \frac{1}{C^\delta} \right), \quad (3)$$

where $r$ is the shift of logarithmic mid-quote prices right before and after a trade of size $\omega$ occurs, $\langle \omega \rangle$ is the average volume per trade for each stock, $C$ is the stock capitalization, $\gamma$ and $\delta$ are two scaling parameters and $g(x)$ is found to be a concave function and has a power-law form for large $\omega$. Lim and Coggins \[55\] performed a similar analysis of the Australian Stock Exchange and a similar scaling was obtained, where the return and trade size were defined in a slightly different way by taking into account the intraday pattern effect. Using order flow data in the Chinese market, Zhou \[60\] found that trades resulting from filled and partially filled limit orders have very different price impacts. The price impact of trades from partially filled orders is constant when the volume is not too large, whereas that of filled orders shows power-law behavior. When the returns and the trade sizes are normalized by their stock-dependent averages, capitalization-independent scaling laws emerge for both buyer-initiated and seller-initiated trades from filled orders \[60\]:

$$r_k \langle r_k \rangle = A_k \left( \frac{\omega_k}{\langle \omega_k \rangle} \right)^\alpha, \quad (4)$$

where $\alpha \approx 2/3$ and the prefactor $A_k = A$ is independent of stock $k$ and its capitalization.

In the aforementioned studies, the trading volume or trade size is considered to be the sole driving factor of price fluctuations. For larger price fluctuations over a fixed clock time scale $\Delta t$, more determinants of price impact have been investigated. Farmer et al \[61\] investigated the transaction data of 16 companies traded on the London Stock Exchange (LSE) in the 4-year period 1999–2002 and found that large price fluctuations are essentially independent of the volume of order, but driven by liquidity fluctuations characterized by the gaps between the first few price levels on the opposite limit order book \[61\]. Using the TAQ data for the year 1997 and the order book data from the Island ECN for the year 2002, Weber and Rosenow \[62\] argued that a large trading volume alone is not sufficient to explain large price changes and a lack of liquidity is a necessary prerequisite for the occurrence of large price fluctuations.
Joulin et al [63] analyzed the one-minute data of 163 USA stocks and found that news and trading volume play a minor role in causing large price changes, which is thus conjectured to be caused by the vanishing of liquidity. Næs and Skjeltorp [64] studied the order flow data from the Oslo Stock Exchange in Norway and found that price fluctuations are positively correlated with trade number, a component of trading volume, and negatively correlated with different liquidity measures.

In this work, we research into the determinants of immediate price impact defined in equation (2) in an emerging stock market. This research is designed to have many differences compared to the literature. We focus on the immediate price impact after a transaction occurs and investigate filled and partially filled orders separately. We use three measures for liquidity based on the bid–ask spread, the standing volume on the limit order book and the gaps between successive price levels.

The motivation for this research is the following. As was briefly reviewed above, the determinants of immediate price impacts have not been investigated for the Chinese stock market, which is a representative emerging market. Our analysis here goes further by investigating separately filled and partially filled orders since these two kinds of orders have nontrivially very different behavior [60]. We expect that the influential factors will have different impacts on the price movement. This research is empirical and not intended for practical use. The new empirical results obtained in this work thus deepen our understanding of the dynamics of this specific market. For a given order, when all the independent variables are known, we can calculate the immediate price impact for sure. However, this treatment for individual orders cannot be generalized to the macroscopic level for a population of orders. It is thus necessary to perform regression analysis of the data.

The rest of this paper is organized as follows. Section 2 describes the data used in our study. Section 3 discusses the possible determinants of immediate price impact and specifies the model. The results are shown in section 4 for individual stocks and in section 5 for all stocks. We summarize our findings in section 6.

2. Data sets

We use the order flow data of 23 A-share stocks traded on the Shenzhen Stock Exchange (SZSE) of China in 2003. The A-shares are common stocks issued by mainland Chinese companies, subscribed and traded in the Chinese currency Renminbi, listed on mainland Chinese stock exchanges, bought and sold by Chinese nationals and approved foreign investors. The A-share market was open only to domestic investors in 2003. Note that the Chinese stock market is the largest emerging market in the world and became the second largest stock market after the USA market in 2009. Our sample stocks were part of the 40 constituent stocks included in the Shenzhen Stock Exchange component index in 2003.

The SZSE is open for trading from Monday to Friday except for public holidays and other dates as announced by the China Securities Regulatory Commission. On each trading day, the market opens at 9:15 and holds the opening call auction till 9:25, during which the trading system accepts order submission and cancellation, and all matched transactions are executed at 9:25. It is followed by a cooling period from 9:25 to 9:30. During cooling periods, the Exchange is open to orders routing from members, but does not process orders or process cancellation of orders. The information released to trading terminals also does not change during cooling periods. All matched transactions are executed at the end of cooling periods. The continuous
Table 1. Basic statistics for the 23 SZSE stocks. The columns give the codes of stocks, the average returns ($\langle r_{PB} \rangle$, $\langle r_{PS} \rangle$, $\langle r_{FB} \rangle$ and $\langle r_{FS} \rangle$, multiplied by 1000), the numbers of trades ($N$, 1000 shares), the fractions of partially filled orders ($F$) and the industrial sectors.

| Code   | $\langle r_{PB} \rangle$ | $\langle r_{PS} \rangle$ | $\langle r_{FB} \rangle$ | $\langle r_{FS} \rangle$ | $N$   | $F$   | Industry          |
|--------|---------------------------|---------------------------|---------------------------|---------------------------|-------|-------|-------------------|
| 000001 | 1.19                      | 1.21                      | 0.03                      | 0.05                      | 889.7 | 0.07  | Financial         |
| 000002 | 1.55                      | 1.54                      | 0.04                      | 0.05                      | 509.4 | 0.06  | Real estate       |
| 000009 | 2.23                      | 2.25                      | 0.05                      | 0.06                      | 448.0 | 0.08  | Conglomerates     |
| 000012 | 1.63                      | 1.61                      | 0.08                      | 0.10                      | 290.4 | 0.14  | Metals and nonmetals |
| 000016 | 1.86                      | 1.87                      | 0.08                      | 0.10                      | 188.6 | 0.15  | Electronics       |
| 000021 | 1.35                      | 1.37                      | 0.07                      | 0.09                      | 411.6 | 0.12  | Electronics       |
| 000024 | 1.65                      | 1.63                      | 0.09                      | 0.09                      | 133.6 | 0.16  | Real estate       |
| 000027 | 1.63                      | 1.63                      | 0.06                      | 0.05                      | 313.9 | 0.08  | Utilities         |
| 000063 | 1.27                      | 1.24                      | 0.07                      | 0.06                      | 265.5 | 0.10  | IT                |
| 000066 | 1.57                      | 1.59                      | 0.08                      | 0.09                      | 277.7 | 0.14  | Electronics       |
| 000088 | 1.63                      | 1.60                      | 0.10                      | 0.09                      | 97.2  | 0.13  | Transportation   |
| 000089 | 1.63                      | 1.64                      | 0.06                      | 0.07                      | 189.1 | 0.11  | Transportation   |
| 000406 | 1.56                      | 1.58                      | 0.05                      | 0.07                      | 271.4 | 0.13  | Petrochemicals    |
| 000429 | 2.35                      | 2.37                      | 0.08                      | 0.09                      | 117.4 | 0.14  | Transportation   |
| 000488 | 1.72                      | 1.62                      | 0.14                      | 0.17                      | 120.1 | 0.21  | Paper and printing |
| 000539 | 1.88                      | 1.85                      | 0.10                      | 0.14                      | 114.7 | 0.14  | Utilities         |
| 000541 | 1.56                      | 1.54                      | 0.09                      | 0.09                      | 68.7  | 0.18  | Electronics       |
| 000550 | 1.60                      | 1.59                      | 0.08                      | 0.09                      | 346.7 | 0.12  | Manufacturing     |
| 000581 | 1.84                      | 1.80                      | 0.10                      | 0.10                      | 94.0  | 0.17  | Manufacturing     |
| 000625 | 1.60                      | 1.61                      | 0.08                      | 0.10                      | 397.6 | 0.12  | Manufacturing     |
| 000709 | 2.02                      | 2.04                      | 0.04                      | 0.05                      | 207.8 | 0.10  | Metals and nonmetals |
| 000720 | 0.98                      | 1.15                      | 0.05                      | 0.08                      | 132.2 | 0.16  | Utilities         |
| 000778 | 1.38                      | 1.33                      | 0.07                      | 0.07                      | 157.3 | 0.15  | Manufacturing     |

double auction operates from 9:30 to 11:30 and from 13:00 to 15:00 and transaction occurs based on automatic one-to-one matching due to price-time priority. The market freezes in the time period from 11:30 to 13:00 and no actions could be taken. At daily closure of the market, unexecuted orders will be automatically removed. Only the trades during the continuous double auction are considered in this work.

Following [65], the trades are differentiated into four types according to their directions (whether a trade is seller-initiated or buyer-initiated) and aggressiveness: buyer-initiated partially filled (PB) trades resulting from partially filled buy orders, seller-initiated partially filled (PS) trades resulting from partially filled sell orders, buyer-initiated filled (FB) trades resulting from filled buy orders and seller-initiated filled (FS) trades resulting from filled sell orders. The basic statistics for the 23 SZSE stocks are presented in table 1 [60]. We note that the 23 stocks investigated in this work are representative, belonging to a variety of industrial sectors as shown in table 1.

Several interesting features can be observed from table 1. Firstly, the majority of trades have resulted from filled orders, which means that these investors are subjectively more impatient. Secondly, we find that the immediate price impacts of buyer-initiated and
seller-initiated trades are roughly symmetric since
\[ \langle r_{PB} \rangle \approx -\langle r_{PS} \rangle \quad \text{and} \quad \langle r_{FB} \rangle \approx -\langle r_{FS} \rangle \] (5)
for each stock. Thirdly, partially filled trades have a much larger impact on the price than filled trades [60]:
\[ \langle r_{PB} \rangle \gg \langle r_{FB} \rangle \quad \text{and} \quad -\langle r_{PS} \rangle \gg -\langle r_{FS} \rangle, \] (6)
which means that, on average, partially filled trades are much more aggressive than filled trades, which is confirmed by the fact that partially filled trades have larger sizes. A mechanical illustration of the third feature will be shown in section 3.

3. Model specification

In order to determine the factors influencing immediate price impact, we show in figure 1 the evolution of the limit order book and the price formation process of PS orders and filled orders.

Consider the situation where a sell order with price \( P = \pi \) and size \( V = 7 \) is submitted at time \( t \). Figure 1(a) shows the first five price levels of both sides of the limit order book right before the placement of a sell order, in which the ask prices, the bid prices and their volumes are denoted by \( a_i(t^-) \), \( b_i(t^-) \), \( V^A_i(t^-) \) and \( V^B_i(t^-) \), respectively. Assume that \( b_4(t^-) < \pi < b_3(t^-) \).

At time \( t \), the buy limit orders on the first three price levels are executed, whose trading volume is \( \omega_t = V^B_1(t^-) + V^B_2(t^-) + V^B_3(t^-) = 6 \). The unexecuted part of the sell order, whose quantity is \( V - \omega_t = 1 \), is stored on the sell limit order book at time \( t^+ \), as shown in figure 1(b), where we...
have \( a_i(t^+) = \pi, a_{i+1}(t^+) = a_i(t^-) \) for \( i = 1, 2, \ldots \), and \( b_i(t^+) = b_{i+1}(t^-) \). More generally, if we have

\[
b_{n+1}(t^-) < \pi < b_n(t^-) \quad \text{and} \quad \sum_{i=1}^{n} V_i^B(t^-) < V,
\]

then the sell order eats buy limit orders waiting on the order book and the new best ask and bid prices are \( b_1(t^+) = b_{n+1}(t^-) \) and \( a_1(t^+) = \pi < a_1(t^-) \). According to equation (2), the immediate price impact is

\[
r_{PS} = \frac{\pi + b_{n+1}(t^-) - a_1(t^-) - b_1(t^-)}{a_1(t^-) + b_1(t^-)}
\]

\[
= -\frac{a_1(t^-) - b_1(t^-)}{a_1(t^-) + b_1(t^-)} - \frac{b_1(t^-) - b_{n+1}(t^-)}{a_1(t^-) + b_1(t^-)} \frac{b_1(t^-) - \pi}{a_1(t^-) + b_1(t^-)}.
\]

The first term in the right-hand side of equation (9) indicates that the immediate price impact \( r_{PS} \) of partially filled orders is negatively proportional to the bid–ask spread \( S(t^-) \). The second term implies that \( r_{PS} \) is negatively proportional to the first \( n \) price gaps

\[
G_i^B(t^-) = b_i(t^-) - b_{i+1}(t^-), \quad i = 1, \ldots, n,
\]

and correlated with the volumes \( V_i^B(t^-) \) on the first \( n \) levels \( i = 1, 2, \ldots, n \). The correlation between \( r_{PS} \) and \( V_i^B(t^-) \) is due to the fact that the value of \( n \) is determined by the order book shape according to equation (7). The third term also indicates that \( r_{PS} \) is correlated with \( G_i^B(t^-) \) on the first \( n \) levels \( i = 1, 2, \ldots, n \). Note that the correlation between \( r_{PS} \) and \( \omega \) can also be inferred from equation (7). Taking into consideration the finding expressed in equation (4) [60], we can propose the following model:

\[
r_{PS}(t) = a_0 + a[\omega(t^-)]^\alpha + bS(t^-) + \sum_{i=1}^{n} d_i f_i G_i^B(t^-) + u_t,
\]

where the coefficients \( a, b \) and \( f_i \) are expected to be negative and the \( d_i \) coefficients are expected to be positive. Similarly, for partially filled buy orders, we can propose the following model:

\[
r_{PB}(t) = a_0 + a[\omega(t^-)]^\alpha + bS(t^-) + \sum_{i=1}^{n} c_i f_i G_i^A(t^-) + u_t,
\]

where the coefficients \( a, b \) and \( e_i \) are expected to be positive and the \( c_i \) coefficients are expected to be negative.

Panels (c) and (d) of figure 1 illustrate the evolution of the limit order book around a filled sell order. If the sell order is filled such that

\[
\pi \leq b_n(t^-) \quad \text{and} \quad \sum_{i=1}^{n} V_i^B(t^-) = V
\]

or

\[
\pi \leq b_{n+1}(t^-) \quad \text{and} \quad \sum_{i=1}^{n} V_i^B(t^-) < V < \sum_{i=1}^{n+1} V_i^B(t^-),
\]
then the new best ask and bid prices are \( b_1(t^+) = b_{n+1}(t^-) \) and \( a_1(t^+) = a_1(t^-) \). Hence, the immediate price impact is

\[
r_{FS} = \frac{b_{n+1}(t^-) - b_1(t^-)}{a_1(t^-) + b_1(t^-)}.
\] (15)

The equation implies that \( r_{FS} \) is negatively proportional to the first \( n \) price gaps \( G^B_i(t^-) \) and correlated with the volumes \( V^B_i(t^-) \) on the first \( n \) levels \( i = 1, 2, \ldots, n \), but independent of the bid–ask spread. We can thus propose the following model:

\[
r_{FS}(t) = a_0 + a[\omega(t^-)]^a + \sum_{i=1}^{n} d_i [V^B_i(t^-)]^\beta + \sum_{i=1}^{n} f_i G^B_i(t^-) + u_t,
\] (16)

where the coefficients \( a \) and \( f_i \) are expected to be negative and the \( d_i \) coefficients are expected to be positive. Similarly, for filled buy orders, we can propose the following model:

\[
r_{FB}(t) = a_0 + a[\omega(t^-)]^a + \sum_{i=1}^{n} c_i [V^A_i(t^-)]^\beta + \sum_{i=1}^{n} e_i G^A_i(t^-) + u_t,
\] (17)

where the coefficients \( a \) and \( e_i \) are expected to be positive and the \( c_i \) coefficients are expected to be negative.

It follows immediately from equations (9) and (15) that

\[
r_{FS} - r_{FS} = -\frac{a_1(t^-) - \pi}{a_1(t^-) + b_1(t^-)} = -\frac{a_1(t^-) - b_1(t^-)}{a_1(t^-) + b_1(t^-)} - \frac{b_1(t^-) - \pi}{a_1(t^-) + b_1(t^-)}
\] (18)

whose absolute value is not less than half of the relative bid–ask spread \( [a_1(t^-) - b_1(t^-)]/[a_1(t^-) + b_1(t^-)] \). Our data show that 91.05% of the FS trades have \( r_{FS} = 0 \), and all price movements caused by the PS trades are negative, which is not surprising in view of the limit order book mechanism. These considerations explain why PS trades have much larger price impacts than FS trades. This analysis applies to buyer-initiated trades as well. We note that 89.42% of the FB trades have \( r_{FB} = 0 \).

There are possible missing determinants of immediate price impact in the above discussions. One factor is the liquidity on the same side of the limit order book for executed buy or sell orders, since the order book shape and the price gaps on the buy and sell sides are statistically symmetric [66]. Another factor is the intraday patterns found in many financial quantities including the volatility, the bid–ask spread and the trading volume. Hence, the immediate price impact of each type of trade for stock \( k \) can be modeled as follows:

\[
r_{k,t+1} = f(\omega_{k,t}, S_{k,t}, V^A_{k,i=1:L,t}, V^B_{k,i=1:L,t}, G^A_{k,i=1:L,t}, G^B_{k,i=1:L,t}, D_{i=123,t}) + u_t,
\] (19)

where \( D_{i=123,t} \) are dummy variables characterizing the intraday pattern in a resolution of 10 min. We consider two types of models. The first type is the following:

\[
r_{k,t+1}/\langle r_k \rangle = a_0 + a \left( \frac{\omega_{k,t}}{\langle \omega_k \rangle} \right)^a + b S_{k,t} + \sum_{i=1}^{L} c_i \left( \frac{V^A_{k,i,t}}{\langle \omega_k \rangle} \right)^\beta + \sum_{i=1}^{L} d_i \left( \frac{V^B_{k,i,t}}{\langle \omega_k \rangle} \right)^\beta
\]

\[
\quad + \sum_{i=1}^{L} e_i \frac{G^A_{k,i,t}}{\langle r_k \rangle} + \sum_{i=1}^{L} f_i \frac{G^B_{k,i,t}}{\langle r_k \rangle} + \sum_{i=1}^{23} g_i D_{i,t} + u_t,
\] (20)
where the variables are normalized so that we can put together all the data of different stocks in the cross-sectional analysis in section 5, following [60]. The second model is expressed as follows:

\[
\frac{r_{k,t+1}}{(r_k)} = a_0 + a \left( \frac{\omega_{k,t}}{\omega_k} \right)^{\alpha} + b S_{k,t} + \sum_{i=1}^{L} c_i \ln \left( \frac{V_{k,i,t}^A}{\omega_k} \right) + \sum_{i=1}^{L} d_i \ln \left( \frac{V_{k,i,t}^B}{\omega_k} \right) + \sum_{i=1}^{23} e_i G_{k,i,t}^A |(r_k)| + \sum_{i=1}^{23} f_i G_{k,i,t}^B |(r_k)| + \sum_{i=1}^{23} g_i D_{i,t} + u_t. \tag{21}
\]

For simplicity, we term the model expressed in equation (20) as the \textit{power-law model} and the model in equation (21) as the \textit{logarithmic model}.

Since the size of the data is huge, we preprocess the data for each stock before calibrating the two models expressed in equations (20) and (21) to reduce the computational time. For each stock, the average values of \( r_{k,t+1}, S_{k,t}, V_{k,i,t=1:L,t}^A, V_{k,i,t=1:L,t}^B, G_{k,i,t=1:L,t}^A \) and \( G_{k,i,t=1:L,t}^B \) of the trades with the same size \( \omega_{k,t} \) are used. The parameters \( a_0, a, b, c_i, d_i, e_i \) and \( f_i \) are determined by minimizing the following objective function:

\[
Q(\alpha, \beta; a_0, a, b, c_i, d_i, e_i, f_i) = \sum_i u_i^2. \tag{22}
\]

This least-squares estimation is equivalent to the maximum likelihood estimation only if the measurement errors are independent and identically distributed Gaussian [67]. However, this condition may not hold. For stock price movements, it is known that both limit orders for liquidity provision and market orders for liquidity taking have long memory properties [68–70], and these two effects usually reconcile to produce diffusive price fluctuations [68, 69, 71], be the price impact either temporary or permanent [72, 73]. On the other hand, we do not know \textit{a priori} the distribution of the measurement errors. Therefore, we use this least-squares estimation in this work. We note that it might underestimate the \( p \)-values, which however should not change the quantitative conclusions.

In order to achieve the global optimization of the power-law model, we scan \( \alpha \) and \( \beta \) in a two-dimensional grid from 0.05 to 0.95 with a spacing step of 0.05. For each pair of \( \alpha \) and \( \beta \), we can perform a linear least-squares regression to determine the local optimal parameters and the adjusted \( R \)-square, \( R^2 \)-adj. For the logarithmic model, we need to scan \( \alpha \) only. The set of parameters that maximize the \( R^2 \)-adj is considered to be the globally optimal solution.

4. Results for individual stocks

4.1. The case of five levels of liquidity \( L = 5 \)

For each stock, we have calibrated the power-law model in equation (20) with \( L = 5 \) for the four types of trades, i.e. PB, PS, FB and FS trades. For each type of trade, the optimal values of \( \alpha \) and \( \beta \) are obtained by maximizing the adjusted \( R \)-square (\( R^2 \)-adj).

Taking stock 000001 as an example, we have \( \alpha = 0.25 \) and \( \beta = 0.15 \) for PB trades, \( \alpha = 0.30 \) and \( \beta = 0.20 \) for PS trades, \( \alpha = 0.55 \) and \( \beta = 0.10 \) for FB trades and \( \alpha = 0.45 \) and \( \beta = 0.05 \) for FS trades. The regression results are shown in table 2. The \( F \)-tests show that the model is significant for all four types of trades and the explanatory power of the variables is very high. For all four types of trade, the normalized immediate price impact \( r/(r) \) is positively
Table 2. Calibration of the power-law model in (20) with $L = 5$ for stock 000001. We obtain that $\alpha = 0.25$ and $\beta = 0.15$ for PB trades, $\alpha = 0.30$ and $\beta = 0.20$ for PS trades, $\alpha = 0.55$ and $\beta = 0.10$ for FB trades, and $\alpha = 0.45$ and $\beta = 0.05$ for FS trades. The numbers in the parentheses are the $p$-values. The estimates of the coefficients in red indicate that the associated variables are significant at the 5% level.

| Variable | PB       | PS       | FB       | FS       |
|----------|----------|----------|----------|----------|
| $R^2$-adj| 0.475    | 0.501    | 0.444    | 0.468    |
| $p$-value| 0.000    | 0.000    | 0.000    | 0.000    |
| $a_0$    | 0.98 (0.000) | 0.51 (0.000) | 31.59 (0.000) | 30.76 (0.000) |
| $a$ for $\omega$ | 1.15 (0.000) | 1.06 (0.000) | 1.85 (0.000) | 2.86 (0.000) |
| $b$ for $S$ | 20.01 (0.000) | 31.26 (0.000) | 227.19 (0.000) | 3.38 (0.635) |
| $c_1$ for $V_i^A$ | $-1.82$ (0.000) | $0.08$ (0.000) | $-30.83$ (0.000) | 1.08 (0.072) |
| $c_2$ for $V_i^A$ | $-0.28$ (0.000) | $0.01$ (0.405) | $-8.73$ (0.000) | 1.26 (0.072) |
| $c_3$ for $V_i^A$ | 0.01 (0.585) | $-0.03$ (0.065) | $-1.80$ (0.042) | 1.35 (0.063) |
| $c_4$ for $V_i^A$ | 0.08 (0.000) | $-0.01$ (0.430) | 2.09 (0.021) | 1.03 (0.138) |
| $c_5$ for $V_i^A$ | 0.01 (0.486) | $-0.03$ (0.032) | 3.46 (0.000) | 1.30 (0.053) |
| $d_1$ for $V_i^B$ | 0.18 (0.000) | $-1.43$ (0.000) | 2.30 (0.005) | $-36.45$ (0.000) |
| $d_2$ for $V_i^B$ | $-0.05$ (0.049) | $-0.26$ (0.000) | $-1.37$ (0.143) | $-11.77$ (0.000) |
| $d_3$ for $V_i^B$ | 0.00 (0.944) | $-0.03$ (0.048) | 0.66 (0.509) | 0.24 (0.751) |
| $d_4$ for $V_i^B$ | 0.03 (0.147) | 0.02 (0.148) | $-0.60$ (0.538) | 2.61 (0.000) |
| $d_5$ for $V_i^B$ | $-0.15$ (0.000) | 0.00 (0.852) | 1.80 (0.057) | 4.82 (0.000) |
| $10^2e_1$ for $G_i^A$ | 4.56 (0.000) | $-0.19$ (0.187) | 0.51 (0.000) | 0.12 (0.160) |
| $10^2e_2$ for $G_i^A$ | 1.77 (0.000) | 0.40 (0.099) | 0.92 (0.002) | $-0.05$ (0.652) |
| $10^2e_3$ for $G_i^A$ | 0.58 (0.022) | 0.77 (0.009) | $-0.81$ (0.031) | 0.24 (0.077) |
| $10^2e_4$ for $G_i^A$ | 0.22 (0.376) | $-0.28$ (0.292) | 2.37 (0.000) | 0.19 (0.216) |
| $10^2e_5$ for $G_i^A$ | $-0.06$ (0.826) | 0.51 (0.030) | 0.43 (0.429) | 0.41 (0.006) |
| $10^2f_1$ for $G_i^B$ | $-0.73$ (0.000) | 4.50 (0.000) | $-0.07$ (0.759) | 0.54 (0.000) |
| $10^2f_2$ for $G_i^B$ | $-1.06$ (0.001) | 2.96 (0.000) | $-0.29$ (0.346) | 1.03 (0.000) |
| $10^2f_3$ for $G_i^B$ | $-0.07$ (0.855) | $-0.29$ (0.426) | 0.23 (0.595) | 0.46 (0.024) |
| $10^2f_4$ for $G_i^B$ | 0.57 (0.155) | $-0.01$ (0.972) | 0.46 (0.304) | 0.51 (0.015) |
| $10^2f_5$ for $G_i^B$ | 0.35 (0.282) | $-0.48$ (0.173) | 0.33 (0.416) | $-0.21$ (0.321) |

correlated with the normalized trade size $\omega$ and the normalized bid–ask spread $S$. However, the relation between $r$ and $S$ is insignificant for the FS trades, which is consistent with the simple explanation in equation (15). For each type of trade, roughly speaking, the normalized immediate price impact is positively correlated with the standing volumes on the same side of the limit order book and the price gaps on the opposite side, and negatively correlated with the standing volumes on the opposite side and the price gaps on the same side. In other words, the magnitude of the price impact is large if the liquidity is high on the same side (large standing volumes and narrow price gaps) or if the liquidity is low on the opposite side (small standing volumes and wide price gaps). Closer scrutiny unveils that the liquidity on the same side of the limit order book has a smaller influence on the price impact. For instance, all coefficients for $V_i^A$ and $G_i^A$ are statistically insignificant for FS trades.

The results for other stocks are quite similar, as summarized in figure 2, where a plus or minus symbol (+ or −) indicates that the estimated coefficient is positive or negative, and the
symbol is marked in red if the associated coefficient is significant at the 5% level. It is evident that a larger order causes a greater price impact in magnitude for all four types of trades and for all individual stocks. For partially filled trades, the price impact is larger if the bid–ask spread is larger for all stocks, while for filled trades, the spread could have a positive or negative influence on the price impact. Again, there are more coefficients on the ask side ($V^B_i$ and $G^A_i$) that are significant for buyer-initiated trades, and there are more coefficients on the bid side ($V^B_i$ and $G^B_i$) that are significant for seller-initiated trades. This finding is consistent with the mechanical analysis expressed in equations (11), (12), (16) and (17).

Note that we do not show the results of the logarithmic model for individual stocks. The dependent matrices for several stocks are rank deficient within machine precision (singular, close to singular or badly scaled) such that no regression results can be obtained. For the stocks with a nonsingular dependence matrix, the results are qualitatively similar to those obtained from the power-law model.
Table 3. Calibration of the power-law model with \( L = 2 \) for Stock 000001. We obtain that \( \alpha = 0.25 \) and \( \beta = 0.15 \) for PB trades, \( \alpha = 0.30 \) and \( \beta = 0.20 \) for PS trades, \( \alpha = 0.55 \) and \( \beta = 0.10 \) for FB trades and \( \alpha = 0.45 \) and \( \beta = 0.10 \) for FS trades. The numbers in parentheses are the \( p \)-values. The estimates of the coefficients in red indicate that the associated variables are significant at the 5% level.

| Variable | PB       | PS       | FB       | FS       |
|----------|----------|----------|----------|----------|
| \( R^2 \)-adj | 0.473    | 0.500    | 0.438    | 0.464    |
| \( p \)-value | 0.000    | 0.000    | 0.000    | 0.000    |
| \( a_0 \) | 1.08 (0.000) | 0.44 (0.000) | 44.80 (0.000) | 19.82 (0.000) |
| \( a \) for \( \alpha \) | 1.16 (0.000) | 1.06 (0.000) | 1.82 (0.000) | 2.88 (0.000) |
| \( b \) for \( \alpha \) | 20.43 (0.000) | 31.37 (0.000) | 222.18 (0.000) | 3.19 (0.654) |
| \( c_1 \) for \( V^A_i \) | -1.83 (0.000) | 0.07 (0.000) | -30.60 (0.000) | 0.99 (0.000) |
| \( c_2 \) for \( V^A_2 \) | -0.25 (0.000) | -0.00 (0.932) | -8.52 (0.000) | 0.96 (0.001) |
| \( d_1 \) for \( V^B_i \) | 0.18 (0.000) | -1.45 (0.000) | 2.94 (0.000) | -16.81 (0.000) |
| \( d_2 \) for \( V^B_2 \) | -0.04 (0.052) | -0.26 (0.000) | -1.06 (0.233) | -4.83 (0.000) |
| \( 10^2 e_1 \) for \( G^A_1 \) | 4.54 (0.000) | -0.17 (0.238) | 0.46 (0.000) | 0.12 (0.182) |
| \( 10^2 e_2 \) for \( G^A_2 \) | 1.82 (0.000) | 0.55 (0.022) | 0.89 (0.003) | -0.04 (0.708) |
| \( 10^2 f_1 \) for \( G^B_1 \) | -0.71 (0.000) | 4.50 (0.000) | -0.04 (0.853) | 0.60 (0.000) |
| \( 10^2 f_2 \) for \( G^B_2 \) | -0.89 (0.007) | 2.95 (0.000) | -0.19 (0.533) | 1.06 (0.000) |

4.2. The case of two levels of liquidity \( L = 2 \)

Figure 2 shows that most coefficients of \( V^A_i \) and \( G^A_i \) (respectively \( V^B_i \) and \( G^B_i \)) for buyer-initiated (respectively seller-initiated) trades are insignificant when \( i \geq 3 \). We thus set \( L = 2 \) and regress the power-law model to the data again. The results for stock 000001 are presented in table 3. In this case, we find that \( \alpha = 0.25 \) and \( \beta = 0.15 \) for PB trades, \( \alpha = 0.30 \) and \( \beta = 0.20 \) for PS trades, \( \alpha = 0.55 \) and \( \beta = 0.10 \) for FB trades and \( \alpha = 0.45 \) and \( \beta = 0.10 \) for FS trades. The \( F \)-tests show that the model is significant for all four types of trades and the explanatory power of the variables is very high. We observe that the values of the power-law exponents \( \alpha \) and \( \beta \) are close to their counterparts in the case of \( L = 5 \) in table 2.

In table 3, the \( G^A_2 \) coefficient with the \( p \)-value being 0.022 for the PS trades is regarded as insignificant since the \( G^A_1 \) coefficient is insignificant. Comparing table 3 with table 2, we find that the corresponding coefficients are close to each other with similar significance levels. However, the coefficients of \( V^A_1 \) and \( V^A_2 \) become significant as \( L \) decreases from 5 to 2. This is not surprising since the two \( p \)-values for \( V^A_1 \) and \( V^A_2 \) are only slightly greater than 5% when \( L = 5 \), and the removal of variables in the regression model changes the \( p \)-values slightly.

The regression results for all individual stocks are illustrated in figure 3, where the meanings of the symbols are the same as in figure 2. Similar patterns are observed. We can reach a conclusion that the trade size, the bid–ask spread (a liquidity measure known as market width) and the market depth at the first few price levels on the opposite limit order book play a major role in determining the price impact of a trade.
4.3. Asymmetrical impacts of buyer- and seller-initiated trades

It has long been known that the market reacts differently to buy and sell orders, and sell orders have a larger price impact than purchases [74]. The regression results enable us to investigate the possible buy–sell asymmetry in the impacts of different influential factors. Figure 4 shows a comparison of the absolute values of the coefficients $a$, $b$, $c_1$, $d_1$, $e_1$ and $f_1$ among the four types of trades for individual stocks.

According to figure 4(a), there is no buy–sell asymmetry caused by trade sizes for partially filled trades since the $a$ values for PB and PS trades are close to each other. In contrast, FS trades have larger $a$ values than FB trades for most stocks, such that an FS trade has a larger price impact than an FB trade of the same size.

According to figure 4(b), there is no buy–sell asymmetry caused by bid–ask spreads for partially filled trades. Concerning filled trades, the majority of the stocks exhibit buy–sell asymmetry where a lack of liquidity (i.e. wide bid–ask spread) results in larger price impacts for FB trades than for FS trades.
Figures 4(c) and (e) show that the depth of the limit order book on the ask side, \( V_A^i \) and \( G_A^i \), has larger price impacts on buyer-initiated trades than on seller-initiated trades. In contrast, figures 4(d) and (f) show that the depth of the limit order book on the bid side, \( V_B^i \) and \( G_B^i \), has larger price impacts on seller-initiated trades than on buyer-initiated trades. These observations are consistent with the results obtained in previous sections stating that the liquidity on the opposite side has a greater influence on the price impact than the liquidity on the same side.

5. Cross-sectional regression for all stocks

In section 4, we have performed an analysis of individual stocks. In this section, we perform a cross-sectional analysis of all stocks. Both the power-law model and the logarithmic model are considered. In the regression, the normalized data for each variable are put together correspondingly and we use \( L = 5 \) in the two models.

5.1. The power-law model

Regressing the power-law model, we obtain that \( \alpha = 0.25 \) and \( \beta = 0.20 \) for PB trades, \( \alpha = 0.30 \) and \( \beta = 0.20 \) for PS trades, \( \alpha = 0.55 \) and \( \beta = 0.05 \) for FB trades and \( \alpha = 0.50 \) and \( \beta = 0.05 \) for FS trades. The estimated coefficients are presented in table 4, where the estimates that are significant at the 5% level are marked in red. The \( R \)-squares from the \( F \)-tests show that the model is significant for all four types of trades. We observe that the determinants (\( \omega, S, V_A^i, V_B^i, G_A^i \) and \( G_B^i \)) and the dummy variables can account for 43.6% of the price impact for PB trades, 22.2% for PS trades, 40.2% for FB trades and 39.6% for FS trades.
Table 4. Cross-sectional regression results of the power-law model in (20) with $L = 5$ for all stocks. We obtain that $\alpha = 0.25$ and $\beta = 0.20$ for PB trades, $\alpha = 0.30$ and $\beta = 0.20$ for PS trades, $\alpha = 0.55$ and $\beta = 0.05$ for FB trades and $\alpha = 0.50$ and $\beta = 0.05$ for FS trades. The numbers in parentheses are the $p$-values. The estimates of the coefficients in red indicate that the associated variables are significant at the 5% level.

| Variable       | PB   | PS   | FB   | FS   |
|----------------|------|------|------|------|
| $R^2$-adj      | 0.436| 0.222| 0.402| 0.396|
| $p$-value      | 0.000| 0.000| 0.000| 0.000|
| $a_0$          | 0.40 (0.000)| 0.90 (0.000)| 34.74 (0.000)| 18.67 (0.000)|
| $a$ for $\omega$| 1.26 (0.000)| 1.15 (0.000)| 1.81 (0.000)| 2.49 (0.000)|
| $b$ for $S$     | 20.48 (0.000)| 20.96 (0.000)| 6.45 (0.000)| −1.39 (0.149)|
| $c_1$ for $V_1^A$ | −1.423 (0.000)| 0.084 (0.000)| −48.380 (0.000)| 1.242 (0.000)|
| $c_2$ for $V_2^A$ | −0.139 (0.000)| 0.062 (0.000)| −10.806 (0.000)| 1.623 (0.000)|
| $c_3$ for $V_3^A$ | 0.027 (0.000)| 0.009 (0.364)| 3.947 (0.000)| 1.421 (0.000)|
| $c_4$ for $V_4^A$ | 0.048 (0.000)| 0.028 (0.004)| 5.312 (0.000)| 1.150 (0.000)|
| $c_5$ for $V_5^A$ | 0.024 (0.000)| −0.010 (0.273)| 4.390 (0.000)| 1.454 (0.000)|
| $d_1$ for $V_1^B$ | 0.099 (0.000)| −1.532 (0.000)| 3.273 (0.000)| −31.612 (0.000)|
| $d_2$ for $V_2^B$ | 0.062 (0.000)| −0.194 (0.000)| 1.645 (0.000)| −5.179 (0.000)|
| $d_3$ for $V_3^B$ | 0.047 (0.000)| −0.000 (0.984)| 1.978 (0.000)| 2.195 (0.000)|
| $d_4$ for $V_4^B$ | 0.042 (0.000)| 0.031 (0.001)| 1.453 (0.000)| 3.703 (0.000)|
| $d_5$ for $V_5^B$ | 0.022 (0.000)| 0.015 (0.111)| 3.123 (0.000)| 4.052 (0.000)|
| $10^2 e_1$ for $G_1^A$ | 3.785 (0.000)| −0.360 (0.000)| 0.957 (0.000)| 0.053 (0.000)|
| $10^2 e_2$ for $G_2^A$ | 0.569 (0.000)| −0.367 (0.000)| 0.361 (0.000)| 0.041 (0.012)|
| $10^2 e_3$ for $G_3^A$ | −0.195 (0.000)| −0.411 (0.000)| 0.361 (0.000)| 0.038 (0.036)|
| $10^2 e_4$ for $G_4^A$ | −0.431 (0.000)| −0.354 (0.000)| −0.055 (0.259)| 0.004 (0.822)|
| $10^2 e_5$ for $G_5^A$ | −0.411 (0.000)| −0.391 (0.000)| 0.030 (0.531)| 0.042 (0.016)|
| $10^2 f_1$ for $G_1^B$ | −0.239 (0.000)| 3.475 (0.000)| 0.142 (0.000)| 0.721 (0.000)|
| $10^2 f_2$ for $G_2^B$ | −0.362 (0.000)| 0.508 (0.000)| 0.052 (0.125)| 0.217 (0.000)|
| $10^2 f_3$ for $G_3^B$ | −0.487 (0.000)| −0.263 (0.000)| −0.077 (0.040)| 0.028 (0.156)|
| $10^2 f_4$ for $G_4^B$ | −0.249 (0.000)| −0.551 (0.000)| −0.021 (0.583)| −0.037 (0.062)|
| $10^2 f_5$ for $G_5^B$ | −0.386 (0.000)| −0.413 (0.000)| −0.024 (0.501)| −0.042 (0.008)|

For PB trades, the normalized price impact is positively correlated with the trade size $\omega$, the bid–ask spread $S$ and the limit order volumes $V_i^B$ at the bid side, and negatively correlated with the price gaps $G_i^B$ on the bid limit order book. For the two liquidity measures at the ask side, the situation is slightly complicated. We observe that $c_1$ and $c_2$ are negative and $c_3$, $c_4$ and $c_5$ are positive for the ask limit order volumes, and $e_1$ and $e_2$ are positive and $e_3$, $e_4$ and $e_5$ are negative for the ask price gaps. However, it is found that $|c_1 + c_2| \gg |c_3 + c_4 + c_5|$ and $|e_1 + e_2| \gg |e_3 + e_4 + e_5|$, which means that the first two levels dominate. We thus argue that the normalized price impact is negatively correlated with the ask order volume and positively correlated with the ask price gaps. Furthermore, table 4 shows that $|c_1 + c_2 + c_3 + c_4 + c_5| \gg |d_1 + d_2 + d_3 + d_4 + d_5|$ and $|e_1 + e_2 + e_3 + e_4 + e_5| \gg |f_1 + f_2 + f_3 + f_4 + f_5|$, which implies that the liquidity at the ask side of the limit order book has a major influence on the price impact in comparison with the liquidity at the bid side.
Similarly, for PS trades, the normalized price impact is positively correlated with the trade size $\omega$, the bid–ask spread $S$, the ask volumes and the bid price gaps on the book, and is negatively correlated with the bid volumes and the ask price gaps on the limit order book. In addition, the liquidity on the opposite side of the order book has a greater influence than the liquidity on the same side. For filled trades, a large trade size, a wide bid–ask spread, a large market depth on the same order book side, and a small market depth on the opposite order book side may cause a larger normalized price impact. There are three exceptions. Firstly, the coefficient $f_1$ for FB trades is positive, which is nevertheless much smaller than $e_1 + e_2 + e_3$. Secondly, the coefficients $e_1$, $e_2$ and $e_3$ for FS trades are positive, whose sum is however much smaller than $f_1 + f_2$. Thirdly, the coefficient $b$ for FS trades is insignificant.

5.2. The logarithmic model

We also perform cross-sectional regression of the logarithmic model expressed in equation (21). We obtain that $\alpha = 0.15$ for PB trades, $\alpha = 0.20$ for PS trades, $\alpha = 0.55$ for FB trades and $\alpha = 0.50$ for FS trades. The results are presented in table 5. The $p$-values of the $F$-tests indicate that the model is significant for all four types of trades. The determinants ($\omega$, $S$, $V_i^A$, $V_i^B$, $G_i^A$ and $G_i^B$) and the dummy variables can account for 42.5% of the price impact for PB trades, 21.6% for PS trades, 40.1% for FB trades and 39.6% for FS trades.

Comparing the estimated coefficients (excluding $a_0$) in table 5 with those in table 4, we find that they have exactly the same pattern, except that $b_3$ for PS trades becomes significant in the logarithmic model. Therefore, the two models lead to the same conclusions.

5.3. The connection between the two models

The difference between the two models is that there are power-law terms of the volumes in the power-law model (20) and alternative logarithmic terms of the volumes in the logarithmic model (21). If $0 < \beta \ll 1$, we have

\[
V^P = e^{\beta \ln V} \approx 1 + \beta \ln V.
\]

It follows that equation (20) becomes

\[
\frac{r_{k,t+1}}{\langle r_k \rangle} = a_0 + \sum_{i=1}^{L} (c_i + d_i) + a \left( \frac{\omega_{k,t}}{\langle \omega_k \rangle} \right)^{\alpha} + b S_{k,t} + \sum_{i=1}^{L} c_i \beta \ln \left( \frac{V_i^A}{\langle V_i^A \rangle} \right) + \sum_{i=1}^{L} d_i \beta \ln \left( \frac{V_i^B}{\langle V_i^B \rangle} \right) + \sum_{i=1}^{L} e_i \left( \frac{G_i^A}{\langle G_i^A \rangle} \right) + \sum_{i=1}^{L} f_i \left( \frac{G_i^B}{\langle G_i^B \rangle} \right) + \sum_{i=1}^{23} g_i D_{i,t} + u_t.
\]

(24)

A comparison of the coefficients of equations (24) and (21) shows that $d_i^{ln} = a_0^{pl} + \sum_{i=1}^{L} (c_i^{pl} + d_i^{pl})$, $a_i^{ln} = a_i^{pl}$, $b_i^{ln} = b_i^{pl}$, $e_i^{ln} = e_i^{pl}$, $f_i^{ln} = f_i^{pl}$, $c_i = c_i^{pl} \beta$ and $d_i = d_i^{pl} \beta$.

These equations are validated according to tables 5 and 4, except for the first one about $a_0$. In particular, figure 5 is to verify equation (25). It is found in figure 5(a) that the points are located around the dashed diagonal line $y = x$. A linear regression shows that

\[
y = 1.074x + 0.003,
\]

(26)
In this work, we have analyzed the mechanism of immediate price impact by investigating the trade size, the bid–ask spread, the standing volumes of the first few price levels at the bid and ask sides of the limit order book and the price level gaps. We have proposed two regression

Table 5. Cross-sectional regression results of the logarithmic model in (21) with $L = 5$ for all stocks. We obtain that $\alpha = 0.15$ for PB trades, $\alpha = 0.20$ for PS trades, $\alpha = 0.55$ for FB trades and $\alpha = 0.50$ for FS trades. The numbers in parentheses are the $p$-values. The estimates of the coefficients in red indicate that the associated variables are significant at the 5% level.

| Variable | PB | PS | FB | FS |
|----------|----|----|----|----|
| $R^2$-adj | 0.425 | 0.216 | 0.401 | 0.396 |
| $p$-value | 0.000 | 0.000 | 0.000 | 0.000 |
| $d_0$ | -1.60 (0.000) | -1.16 (0.000) | 0.80 (0.000) | -1.28 (0.000) |
| $a$ for $\omega$ | 2.06 (0.000) | 1.70 (0.000) | 1.78 (0.000) | 2.48 (0.000) |
| $b$ for $S$ | 20.55 (0.000) | 20.89 (0.000) | 6.42 (0.000) | -1.58 (0.101) |
| $c_1$ for $V_1^A$ | -0.254 (0.000) | 0.022 (0.000) | -2.617 (0.000) | 0.064 (0.000) |
| $c_2$ for $V_2^A$ | -0.031 (0.000) | 0.015 (0.000) | -0.595 (0.000) | 0.092 (0.000) |
| $c_3$ for $V_3^A$ | 0.005 (0.000) | 0.003 (0.245) | 0.204 (0.000) | 0.080 (0.000) |
| $c_4$ for $V_4^A$ | 0.011 (0.000) | 0.006 (0.004) | 0.302 (0.000) | 0.067 (0.000) |
| $c_5$ for $V_5^A$ | 0.005 (0.000) | -0.001 (0.661) | 0.260 (0.000) | 0.080 (0.000) |
| $d_1$ for $V_1^B$ | 0.024 (0.000) | -0.266 (0.000) | 0.173 (0.000) | -1.676 (0.000) |
| $d_2$ for $V_2^B$ | 0.014 (0.000) | -0.047 (0.000) | 0.079 (0.000) | -0.279 (0.000) |
| $d_3$ for $V_3^B$ | 0.011 (0.000) | -0.005 (0.035) | 0.104 (0.000) | 0.110 (0.000) |
| $d_4$ for $V_4^B$ | 0.009 (0.000) | 0.004 (0.077) | 0.074 (0.000) | 0.196 (0.000) |
| $d_5$ for $V_5^B$ | 0.005 (0.000) | -0.001 (0.783) | 0.164 (0.000) | 0.213 (0.000) |
| $10^2e_1$ for $G_1^A$ | 3.775 (0.000) | -0.352 (0.000) | 0.936 (0.000) | 0.050 (0.000) |
| $10^2e_2$ for $G_2^A$ | 0.552 (0.000) | -0.349 (0.000) | 0.352 (0.000) | 0.042 (0.009) |
| $10^2e_3$ for $G_3^A$ | -0.188 (0.000) | -0.394 (0.000) | 0.371 (0.000) | 0.038 (0.036) |
| $10^2e_4$ for $G_4^A$ | -0.434 (0.000) | -0.331 (0.000) | -0.043 (0.376) | 0.004 (0.809) |
| $10^2e_5$ for $G_5^A$ | -0.405 (0.000) | -0.386 (0.000) | 0.039 (0.417) | 0.042 (0.017) |
| $10^2f_1$ for $G_1^B$ | -0.231 (0.000) | 3.448 (0.000) | 0.136 (0.000) | 0.716 (0.000) |
| $10^2f_2$ for $G_2^B$ | -0.348 (0.000) | 0.485 (0.000) | 0.052 (0.128) | 0.211 (0.000) |
| $10^2f_3$ for $G_3^B$ | -0.479 (0.000) | -0.296 (0.000) | -0.077 (0.040) | 0.024 (0.230) |
| $10^2f_4$ for $G_4^B$ | -0.241 (0.000) | -0.564 (0.000) | -0.027 (0.481) | -0.036 (0.069) |
| $10^2f_5$ for $G_5^B$ | -0.393 (0.000) | -0.407 (0.000) | -0.026 (0.462) | -0.039 (0.012) |

where $y$ contains the data of $c_i^{ln}$ and $d_i^{ln}$ and $x$ contains the data of $c_i^{pl}$ and $d_i^{pl}$. For better visibility of the points with small values, we plot the absolute values on the log–log scale in figure 5(b). We find that several data points for the PS trades deviate the most from $y = x$, where the most deviated point is for $d_3$. The observation of these deviations is consistent with the facts that, for the PS trades, the adjusted $R$-square is the smallest and that the associated value of $\beta = 0.2$ is not very close to zero.

With the above analysis, we can say that the power-law model and the logarithmic model are in essence equivalent in explaining the immediate price impact.

6. Conclusion

In this work, we have analyzed the mechanism of immediate price impact by investigating the trade size, the bid–ask spread, the standing volumes of the first few price levels at the bid and ask sides of the limit order book and the price level gaps. We have proposed two regression
models, which have quantitatively similar explanatory powers. We found that large trade sizes, wide bid–ask spreads, high liquidity at the same side and low liquidity at the opposite side will cause large price impacts. For most individual stocks, together with other determinants, the liquidity information at the first two price levels on the limit order book is sufficient to have a high explanatory power. When a cross-sectional regression is performed, it is better to include more price levels. Asymmetrical impacts have been unveiled for several investigated measures. In particular, the coefficient of spread for FB orders is significant, while that for FS orders is insignificant.

More interestingly, we have investigated separately filled orders and partially filled orders. For instance, consider PB orders and FB orders. We found that the coefficient of spread for PB orders is larger than that for FB orders ($b_{PB} > b_{FB}$), the coefficient magnitude of $V_A$ for PB orders is smaller than for FB orders ($|c_{1,PB}| < |c_{1,FB}|$), and the coefficient of $G_A$ for PB orders is larger than for FB orders ($e_{1,PB} > e_{1,FB}$). These observations are consistent with the fact that filled orders are more aggressive than partially filled orders, so that liquidity changes have a greater impact on FB orders than on PB orders. The results in this work thus extend our previous work on the relationship between the immediate price impact and the order size [60].

We note that the results of this work seem to differ from those of some previous works [61–63]. This is not surprising since we are analyzing all price impacts, large or small, whereas those previous works focus on large price changes [61–63], where the liquidity is found to play a major role and the trade size a minor role. In contrast, we showed that both trade size and liquidity (characterized by a measure of market width and two measures of market depth) affect the magnitudes of immediate price impacts. The importance of trade sizes is consistent with earlier works [54, 55, 60]. Furthermore, our analysis here uncovers more determinants of immediate price impacts.

**Acknowledgment**

This work was partially supported by the National Natural Science Foundation (grant nos 11075054 and 10905023), the Humanities and Social Sciences Fund sponsored by the Ministry of Education of China (no. 09YJCZH042) and the Fundamental Research Funds for the Central Universities.
References

[1] Mantegna R N and Stanley H E 2000 An Introduction to Econophysics: Correlations and Complexity in Finance (Cambridge: Cambridge University Press)

[2] Fama E F 1965 J. Bus. 38 34–105

[3] Bachelier L 1900 Théorie de la Spéculaison PhD Thesis University of Paris, Paris

[4] Osborne M P M 1959 Oper. Res. 7 145–73

[5] Black F and Scholes M 1973 J. Polit. Econ. 81 637–54

[6] Mandelbrot B B 1963 J. Bus. 36 394–419

[7] Mandelbrot B B 1960 Int. Econ. Rev. 1 79–106

[8] Mandelbrot B B 1961 Econometrica 29 517–43

[9] Mandelbrot B B 1962 J. Polit. Econ. 71 421–40

[10] Mantegna R N and Stanley H E 1995 Nature 376 46–9

[11] Gopikrishnan P, Meyer M, Amaral L A N and Stanley H E 1998 Eur. Phys. J. B 3 139–40

[12] Liu Y H, Gopikrishnan P, Cizeau P, Meyer M, Peng C K and Stanley H E 1999 Phys. Rev. E 60 1390–400

[13] Gopikrishnan P, Plerou V, Amaral L A N, Meyer M and Stanley H E 1999 Phys. Rev. E 60 5305–16

[14] Plerou V, Gopikrishnan P, Amaral L A N, Meyer M and Stanley H E 1999 Phys. Rev. E 60 6519–29

[15] Makowiec D and Gnaciński P 2001 Acta Phys. Pol. B 32 1487–500

[16] Bertram W K 2004 Physica A 341 533–46

[17] Yan C, Zhang J W, Zhang Y and Tang Y N 2005 Physica A 353 425–32

[18] Qiu T, Zheng B, Ren F and Triemer S 2007 Physica A 378 387–98

[19] Drożdż S, Forczek M, Kwapień J, Oświęcimka P and Rak R 2007 Physica A 383 59–64

[20] Pan R K and Sinha S 2007 Europhys. Lett. 77 58004

[21] Pan R K and Sinha S 2008 Physica A 387 2055–65

[22] Tabak B M, Takami M Y, Cajueiro D O and Petitiniga A 2009 Physica A 388 59–62

[23] Eryiğit M, Çukur S and Eryiğit R 2009 Physica A 388 1879–86

[24] Queiroz S M D and Queiroz S M D, Moyano L G, de Souza J and Tsallis C 2007 Eur. Phys. J. B 55 161–7

[25] Zhang J W, Zhang Y and Kleinert H 2007 Physica A 377 166–72

[26] Gu G F, Chen W and Zhou W X 2008 Physica A 387 495–502

[27] Fuentes M A, Geric A and Vicente J 2009 PLoS One 4 e8243

[28] Geric A, Vicente J and Fuentes M A 2009 Phys. Rev. E 80 065102

[29] Castaing B, Gagne Y and Hopfinger E J 1990 Physica D 46 177–200

[30] Ghăshghaie S, Breymann W, Peinke J, Talkner P and Dodge Y 1996 Nature 381 767–70

[31] Plerou V and Stanley H E 2008 Phys. Rev. E 77 037101

[32] Stanley H E, Plerou V and Gabaix X 2008 Physica A 387 3967–81

[33] Mu G H and Zhou W X 2010 Phys. Rev. E 82 066103

[34] Karppoff J M 1987 J. Finance Quant. Anal. 22 109–26

[35] Wood R A, McNish T H and Ord J K 1985 J. Finance 40 723–39

[36] Jain P C and Joh G H 1988 J. Finance Quant. Anal. 23 269–83

[37] Ying C C 1966 Econometrica 34 676–85

[38] Epp T W 1977 J. Finance Quant. Anal. 12 141–6

[39] Harris L 1987 J. Finance Quant. Anal. 22 127–41

[40] Gallant A R, Rossi P E and Tauchen G 1992 Rev. Finance Stud. 5 199–242

[41] Richardson G, Sefcik S and Thompson R 1986 J. Finance Econ. 17 313–33

[42] Rogalski R J 1978 Rev. Econ. Stat. 60 268–74

[43] Saatcioglu K and Starks L T 1998 Int. J. Forecast. 14 215–25
[47] Loeb T F 1983 Finance Analysts J. 39 39–44
[48] Perold A F and Salomon R S J 1991 Finance Anal. J. 47 31–9
[49] Zhang Y C 1999 Physica A 269 30–44
[50] Farmer J D 2002 Ind. Corp. Change 11 895–953
[51] Almgren R F 2003 Appl. Math. Finance 10 1–18
[52] Gabaix X, Gopikrishnan P, Plerou V and Stanley H E 2003 Nature 423 267–70
[53] Gabaix X, Gopikrishnan P, Plerou V and Stanley H E 2003 Physica A 324 1–5
[54] Lillo F, Farmer J D and Mantegna R 2003 Nature 421 129–30
[55] Lim M and Coggins R 2005 Quant. Finance 5 365–77
[56] Farmer J D and Lillo F 2004 Quant. Finance 4 C7–11
[57] Plerou V, Gopikrishnan P, Gabaix X and Stanley H E 2004 Quant. Finance 4 C11–5
[58] Gabaix X, Gopikrishnan P, Plerou V and Stanley H E 2006 Q. J. Econ. 121 461–504
[59] Gabaix X, Gopikrishnan P, Plerou V and Stanley H E 2007 J. Eur. Econ. Assoc. 5 564–73
[60] Zhou W X 2012 Quant. Finance 12 at press
[61] Farmer J D, Gillemot L, Lillo F, Mike S and Sen A 2004 Quant. Finance 4 383–97
[62] Weber P and Rosenow B 2006 Quant. Finance 6 7–14
[63] Joulin A, Lefevre A, Grunberg D and Bouchaud J P 2008 Wilmott Mag. Sept/Oct 1–7
[64] Næs R and Skjeltorp J A 2006 J. Finance Markets 9 408–32
[65] Biais B, Hillion P and Spatt C 1995 J. Finance 50 1655–89
[66] Gu G F, Chen W and Zhou W X 2008 Physica A 387 5182–8
[67] Press W, Teukolsky S, Vetterling W and Flannery B 1996 Numerical Recipes in FORTRAN: The Art of Scientific Computing (Cambridge: Cambridge University Press)
[68] Bouchaud J P, Gefen Y, Potters M and Wyart M 2004 Quant. Finance 4 176–90
[69] Lillo F and Farmer J D 2004 Stud. Nonlinear Dyn. Econometr. 8 1–33
[70] Gu G F and Zhou W X 2009 Europhys. Lett. 86 48002
[71] Bouchaud J P 2005 Chaos 15 026104
[72] Farmer J D, Gerig A, Lillo F and Mike S 2006 Quant. Finance 6 107–12
[73] Bouchaud J P, Kockelkoren J and Potters M 2006 Quant. Finance 6 115–23
[74] Saar G 2001 Rev. Finance Stud. 14 1153–81