Noname manuscript No.
(will be inserted by the editor)

Effects of Dark Pools on Financial Markets' Efficiency and Price-Discovery Function: An Investigation by Multi-Agent Simulations

Takanobu Mizuta · Shintaro Kosugi · Takuya Kusumoto · Wataru Matsumoto · Kiyoshi Izumi

Received: date / Accepted: date

Abstract In financial stock markets, dark pools, in which order books or quotes are not provided, are becoming widely used. However, increasing the use of dark pools would raise regulatory concerns as it may ultimately affect the quality of the price-discovery function in the lit markets, which are normal markets in which all order books are provided to investors. This may destabilize a market and heighten financial systemic risk.

In this study, we investigated effects of a dark pool on financial markets' efficiency and the price-discovery function by using an artificial market model. We found that the markets are made more efficient by raising the share of the trading value amount of the dark pool by a certain level. However, raising the share above the level makes the market significantly inefficient. This indicates that the dark pool has an optimal usage rate for market efficiency.

The smart order routing (SOR) is transmitting market orders to the dark pool, and this leads the depth of limit orders to become thicker. The thicker limit orders absorb market orders, and thus a market price is still stable near a fundamental price. On the other hand, when too many waiting orders are stored in the dark pool, the orders absorb market orders in the lit market by SOR and prevent the market price converging to the fundamental price. This causes the market price to stay very different from the fundamental price and makes the lit market inefficient.

We also discuss mechanisms by which a dark pool makes a market efficient or inefficient by using a simple equation model. The equations suggest that if

Takanobu Mizuta
SPARX Asset Management Co., Ltd., Tokyo, Japan
E-mail: mizutata@gmail.com

Takuya Kusumoto · Wataru Matsumoto
Nomura Securities Co., Ltd., Tokyo, Japan

Kiyoshi Izumi
School of Engineering, The University of Tokyo, Tokyo, Japan
CREST, JST, Tokyo, Japan

the trading value amount is higher in dark pools than in lit markets, markets become inefficient. This suggests that when the usage rate of dark pools is low, dark pools rarely destroy the price-discovery function even though a large buy-sell imbalance occurs. On the other hand, when the usage rate of dark pools is very high, dark pools very easily destroy the price-discovery function by a very slight buy-sell imbalance. We also compared results of the equations with those of simulations and found similar tendencies.

1 Introduction

In financial stock markets, dark pools, in which order books or quotes are not provided, are becoming widely used, especially by institutional investors [SEC(2010)]. In dark pools, investors can trade a large block of stock without market impacts because investors need not show their orders to anyone else, and reducing market impacts by such investors may make markets more efficient [Johnson(2010)].

However, increasing the use of dark pools would raise regulatory concerns as it may ultimately affect the quality of the price-discovery function in the lit markets, which are normal markets in which all order books are provided to investors. This may destabilize a market and heighten financial systemic risk [EC(2010)], [Ye(2012)]. Therefore, for example in Europe, regulators are discussing introducing a volume cap regulation for dark pools, specifically a 8% limit for the trading volume for each stock [Bowley(2014)].

Dark pools are very difficult to discuss by only using results of empirical studies. Because so many factors cause price formation in actual markets, an empirical study cannot isolate the pure contribution of existing new types of markets such as dark pools to price formation. Furthermore, empirical studies cannot investigate situations that have never occurred before in real financial markets.

We usually discuss whether new types of markets should be spread or not on the basis of their effects on price formation. An artificial market, which is a kind of a multi-agent simulation, can isolate the pure contribution of these new types of markets to the price formation and can treat situations that have never occurred [LeBaron(2006)], [Chen et al(2012)], [Cristelli(2014)]. These are strong points of the artificial market simulation study.

Many studies have investigated the effects of several new regulations and effects of new types of markets by using artificial market simulations, for example, investigating effects of short selling regulations [Yagi et al(2010)], transaction taxes [Westerhoff(2008)], financial leverages [Thurner et al(2012)], circuit breakers [Kobayashi and Hashimoto(2011)], price variations [Yeh and Yang(2010)], [Mizuta et al(2015a)], tick sizes [Mizuta et al(2013)], speedup of exchange computer system [Mizuta et al(2015c)], and dark pools [Mo and Yang(2013)], [Mizuta et al(2015b)].

Indeed, [Mo and Yang(2013)] investigated dark pools by using artificial market simulations. However, they have not investigated situations that have never occurred before such as usage rates of dark pools that are much higher than those at present because they also use real historical stock prices. [Mizuta et al(2015b)] investigated whether dark pools reduce market impacts or not by using artificial market simulations. However, they neither discuss market efficiency nor markets' price-discovery function, and their market selection model was not realistic.

Therefore, in this study, we investigated effects of a dark pool on financial markets' efficiency and price-discovery function by using an artificial market model. This is a very important investigation into financial systemic risk because making a market inefficient and losing the price-discovery function may make the market unstable and increase financial systemic risk. In this study, we additionally implemented smart order routing (SOR) to the model of [Mizuta et al(2015b)] to treat actual market selection of investors. We discussed quantitatively how the spreading of dark pools beyond our experience could affect the price-discovery function and aimed to clarify the mechanism of dark pools that makes a market efficient or inefficient.

2 Artificial Market Model

We built a simple artificial market model in which smart order routing (SOR) was additionally implemented to the model of [Mizuta et al(2015b)], which had been built on the basis of the model of [Chiarella and Iori(2002)].

The model of [Mizuta et al(2015b)] succeed at replicating high frequency micro structures such as execution rates, cancel rates, one tick volatility, and so on, which were not replicated by the model of [Chiarella and Iori(2002)]. Their model [Chiarella and Iori(2002)] was very simple but replicated long-term statistical characteristics observed in real financial markets: fat-tail and volatility-clustering.

The simplicity of the model is very important for this study. We explain the basic concept for building our artificial market model in the Appendix.

2.1 Order Process

The model treats only one risk asset and non-risk asset (cash). The number of agents is n. First, at time $t = 1$, agent 1 orders to buy or sell the risk asset; after that at $t = 2$, agent 2 orders; at $t = 3, 4, , , n$, agents $3, 4, , , n$ order respectively. At $t = n + 1$, going back to the first, agent 1 orders, and at $t = n + 2, n + 3, , n + n$, agents $2, 3, , , , n$ order respectively, and this cycle is repeated. Note that time t passes even if no deals are done.

Agents always order only one share and can short-sell freely. The quantity of holding positions is not limited, so agents can take any shares for both long and short positions to infinity.

2.2 Lit Market and Dark Pool

The model has two markets: one lit market, which provides all order books to investors, and one dark pool, which provide no order books.

The lit market adopts a continuous double auction to determine a market price of the risk asset. A continuous double auction is an auction mechanism where multiple buyers and sellers compete to buy and sell some financial assets in the market, and where transactions can occur at any time whenever an offer to buy and an offer to sell match [Friedman(1993)], [TSE(2012)]. The minimum unit of price change is δP. The buy order price is rounded off to the nearest fraction, and the sell order price is rounded up to the nearest fraction. When an agent orders to buy (sell), if there is a lower-price sell order (a higher-price buy order) than the agent's order, dealing is immediately done. We call this a "market order". If there is not, the agent's order remains in the order book. We call this a "limit order". The remaining order is canceled t_c after the order time.

There are many ways to determine trade prices in dark pools. In the model, the dark pool adopts an average of the highest buy order price and the lowest sell order price in the lit market as a trade price. This method is adopted by many dark pools in real financial markets [Johnson(2010)]. Agents do not specify an order price in the dark pool. When the agent orders one unit buy (sell) to the dark pool, trading is done immediately if the dark pool contains opposite waiting sell (buy) orders. If there are no opposite orders, the order remains and waits for opposite orders to come in. In the dark pool, therefore, only either buy or sell orders remain. The same as in a lit market, the remaining order is canceled t_c after the order time.

2.3 Market Selection Model

We investigated two cases: one without and the other with smart order routing (SOR).

In the case without SOR, the agents order to the dark pool with probability d and to the lit market with probability $1-d$.

In the case with SOR, agents select a market in the following way. When the order is a market order and opposite limit orders are waiting in the dark pool, the agents order to the dark pool. When the order is a limit order, or when the order is a market order and opposite limit orders are not waiting in the dark pool, the agents order to the dark pool with probability d, the same as in the case without SOR.

For example, when the agent makes a market buy (sell) order, if limit sell (buy) orders are waiting in the dark pool, the order is always matched in the dark pool. If no waiting limit sell (buy) orders are waiting in the dark pool, the order goes to the dark pool with probability d. Also, when an agent makes a limit order, the order goes to the dark pool with probability d.

Of course, investors prefer to make a market order in a dark pool than in a lit market if there are opposite waiting orders in the dark pool, because a trade price of a market order in a dark pool is always improved by half of a bid ask spread[1] than that in a lit market.

Indeed in actual financial markets, an investor cannot see whether opposite limit orders are waiting in dark pools or not. However, it is said that, in the case of making a market order, at first the investor orders to a dark pool, and then if the order is not matched, s/he cancels the order soon and orders to a lit market as a market order. Therefore, this SOR model is not only reasonable but also essential for a market selection model with a dark pool.

Because [Mizuta et al(2015b)] did not implement such a SOR in their model, they could not discuss effects of the SOR to market efficiency. An important contribution of this study is to discuss these SOR effects.

2.4 Agent Model

An agent j determines an order price and buys or sells by the following process. Agents use a combination of fundamental value and technical rules to form expectations on a risk asset's returns. An expected return of the agent j is

$$r_{e,j}^t = \frac{1}{w_{1,j} + w_{2,j} + u_j} \left(w_{1,j} \log \frac{P_f}{P^t} + w_{2,j} r_{h,j}^t + u_j \epsilon_j^t \right). \qquad (1)$$

where $w_{i,j}$ is the weight of term i of the agent j and is determined by random variables uniformly distributed in the interval $(0, w_{i,max})$ at the start of the simulation independently for each agent. u_j is the weight of the third term of the agent j and is also determined by random variables uniformly distributed in the interval $(0, u_{max})$ at the start of the simulation independently for each agent. P_f is a fundamental value that is constant. P^t is a market price of the risk asset at time t[2]. ϵ_j^t is a noise determined by random variables of normal distribution with an average 0 and a variance σ_ϵ. $r_{h,j}^t$ is a historical price return inside an agent's time interval τ_j, and $r_{h,j}^t = \log{(P^t/P^{t-\tau_j})}$. τ_j is determined by random variables uniformly distributed in the interval $(1, \tau_{max})$ at the start of the simulation independently for each agent.

The first term of Eq. (1) represents a fundamental strategy: an agent expects a positive return when the market price is lower than the fundamental value, and vice versa. The second term of Eq. (1) represents a technical strategy: an agent expects a positive return when historical market return is positive, and vice versa.

After the expected return has been determined, an expected price is

$$P_{e,j}^t = P^t \exp{(r_{e,j}^t)}. \qquad (2)$$

[1] A bid ask spread is defined as the difference between the best bid price (the highest buy limit order price) and best ask price (the lowest sell limit order price).

[2] When the dealing is not done at t, P^t remains at the last market price P^{t-1}, and at $t = 1$, $P^t = P_f$

An order price $P^t_{o,j}$ is determined by random variables normally distributed in an average $P^t_{e,j}$ and a standard deviation P_σ, where P_σ is a constant.

Buy or sell is determined by a magnitude relationship between the expected price $P^t_{e,j}$ and the order price $P^t_{o,j}$, i.e.,

When $P^t_{e,j} > P^t_{o,j}$, the agent orders to buy one share,

When $P^t_{e,j} < P^t_{o,j}$, the agent orders to sell one share.

3 Simulation Results

[Mizuta et al(2015b)] searched for adequate model parameters verified by statistically existing stylized facts and market micro structures to investigate effects of dark pools on market stability. They found parameters to replicate both long-term statistical characteristics and very short-term micro structures of real financial markets[3]. Specifically, we set, $n = 1,000, w_{1,max} = 1, w_{2,max} = 10, u_{max} = 1, \tau_{max} = 10,000, \sigma_\epsilon = 0.06, P_\sigma = 30, t_c = 20,000, \delta P = 0.1, P_f = 10,000$. We ran simulations to $t = 10,000,000$.

3.1 Market Efficiency

In this study, we compared several statistical values of simulation runs for various d = 0, 0.025, 0.05, 0.075, 0.1, 0.125, 0.15, 0.175, 0.2, 0.25, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 under not only other parameters that were fixed but also the same random number table. We simulated these runs 100 times, changing the random number table each time, and we used averaged statistical values of 100 times runs.

We introduced "market inefficiency" M_{ie} that directly measures market efficiency,

$$M_{ie} = \frac{1}{t_e} \sum_{t=1}^{t_e} \frac{|P^t - P_f|}{P_f}, \tag{3}$$

where || means absolute value. M_{ie} is always greater than zero, and $M_{ie} = 0$ means a market is perfectly efficient. The larger the M_{ie}, the less efficient the market[4].

Fig. 1 shows market inefficiency M_{ie} for various shares of a trading value amount of the dark pool (D) in the cases with and without SOR (Smart Order Routing). Note that D is different from the probability of ordering to the dark pool d. We defined D as $D = V_D/(V_D + V_L)$, where V_D is the trading value amount in the dark pool and V_L is the trading value amount in the lit market.

[3] We explain the verification of the model in the Appendix.

[4] This index is sometimes used in experimental financial studies of people, in which this market inefficiency is sometimes called "RAD" (Relative Absolute Deviation) [Stöckl et al(2010)]. Many indications for measuring market efficiency have been proposed [Verheyden et al(2013)]. A feature of M_{ie} is that M_{ie} is calculated by using a fundamental price P_f directly, which is never observed in empirical studies. We can also use M_{ie} in simulation and experimental studies because we can exactly define a fundamental price.

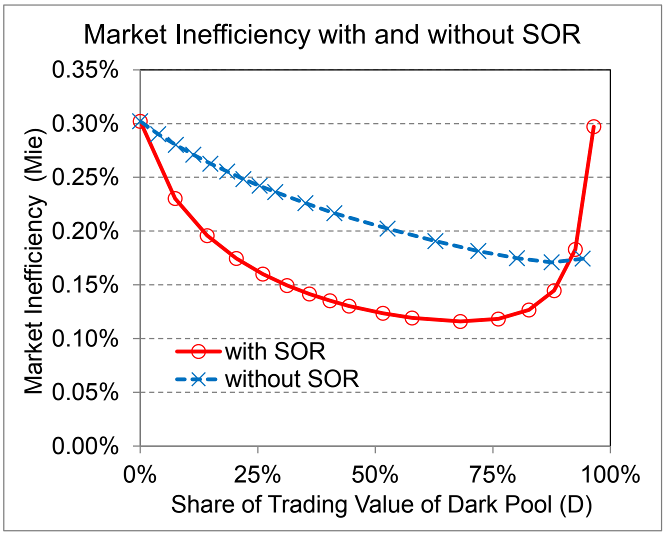

Fig. 1 Market inefficiency M_{ie} for various shares of trading value amount of the dark pool (D).

d is almost the same as the share of numbers of all orders that include limit and market orders. On the other hand, D counts only market orders.

In the case without SOR, M_{ie} monotonically and gradually decreased with D as Fig. 1 shows. On the other hand, in the case with SOR, M_{ie} decreased more sharply in $D \lesssim 70\%$, and M_{ie} increased significantly in $D \gtrsim 70\%$. This indicates that there is an optimal usage rate of the dark pool for the market efficiency. Next, we discuss why M_{ie} decreased in $D \lesssim 70\%$ in section 3.2 and then why M_{ie} increased significantly in $D \gtrsim 70\%$ in section 3.3.

3.2 Becoming Efficient: lower than optimal usage rate of Dark Pool

Fig. 2 shows execution rates in the lit market for various D in the cases with and without SOR. In the case without SOR, the execution rates were stable and moderately decreased. On the other hand, in the case with SOR, the execution rates more rapidly decreased. This indicates that the SOR reduces the execution rates in the lit market.

Fig. 3 shows a mechanism for reducing the execution rates in the case with SOR. As we mentioned in section 2.3, the agents order to the dark pool when the order is a market order and opposite limit orders are waiting. In other cases, the agents order to the dark pool with probability d. Therefore, the number of market orders to the lit market is reduced by the number of market orders sent to the dark pool by the SOR when opposite limit orders are waiting in the dark pool. Thus, the execution rates in the lit market with the SOR are lower than those without the SOR.

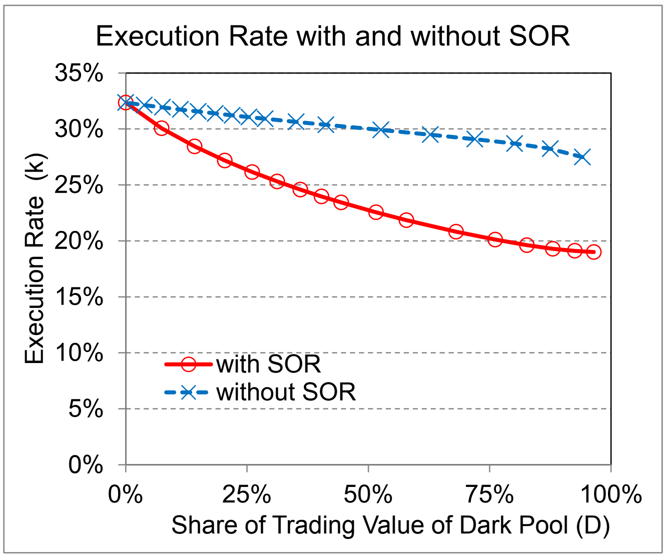

Fig. 2 Execution rates in the lit market for various D.

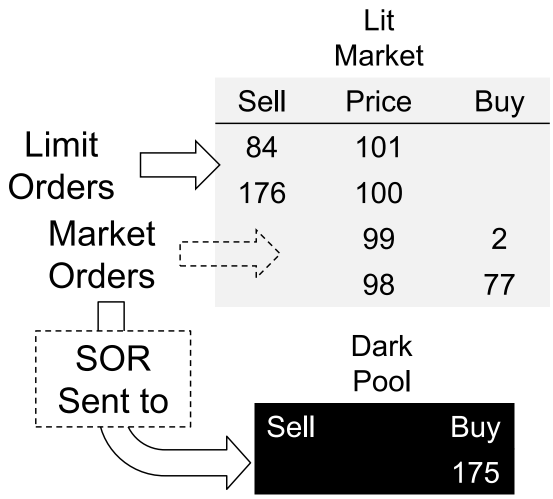

Fig. 3 A mechanism for reducing the execution rates in the case with SOR.

This mechanism raises the limit orders relative to market orders in the lit market. Fig. 4 shows the depth of limit orders and bid ask spreads in the lit market for various D in the case with SOR. Here, depth of limit orders is defined as an average number of buy limit orders from $(1-0.001) \times P^t$(a market price) to P^t per average number of executed trades within one day $(20,000$ time steps), and a bid ask spread is defined as the difference between

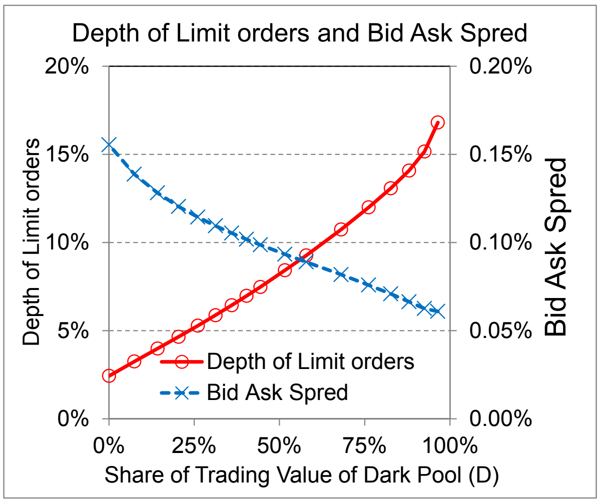

Fig. 4 Depth of limit orders and bid ask spreads in the lit market for various D.

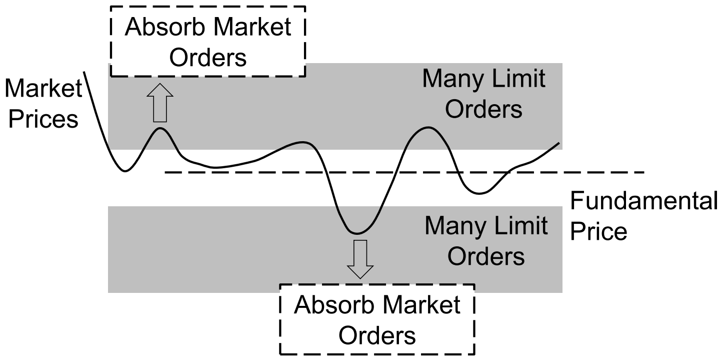

Fig. 5 A mechanism for making the lit market efficient in the case with SOR.

best bid price (the highest buy limit order price) and best ask price (the lowest sell limit order price) per P_f (a fundamental price). Indeed, by increasing D, the depth of limit orders becomes thicker, and this leads to narrowing down the bid ask spread.

Fig. 5 shows a mechanism of how the thicker depth of limit orders makes the lit market efficient. First, P^t is very near P_f. When the depth of limit orders becomes thicker, even though agents make market orders departing from P_f, thicker limit orders absorb these market orders, and thus P^t is still stable near P_f.

We summarize why M_{ie} decreased in $D \lesssim 70\%$ as shown in Figures 3 and 5. Execution rates in the lit market are reduced by more market orders being sent to the dark pool by the SOR than limit orders increasing D. This leads

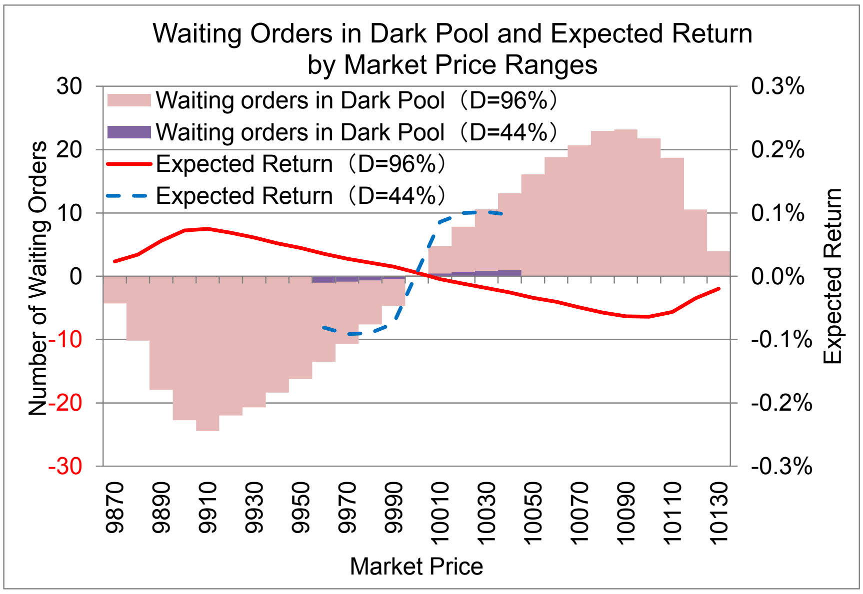

Fig. 6 Number of waiting limit orders in the dark pool and averaged $r^t_{e,j}$ (expected return) for all agents by various P^t ranges.

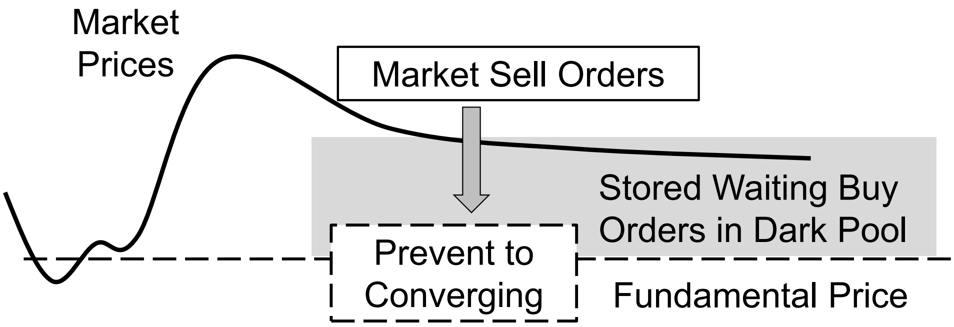

Fig. 7 A mechanism of how the lit market is made inefficient.

to the depth of limit orders becoming thicker, and these thicker limit orders absorb market orders. Thus, a market price is still stable near a fundamental price.

3.3 Becoming Inefficient: Too High usage rate of Dark Pool

We discuss why M_{ie} increased significantly in $D \gtrsim 70\%$. Fig. 6 shows a number of waiting limit orders in the dark pool and $r^t_{e,j}$ (expected return) for all agents averaged by various P^t ranges in the case of $D = 44\%$ and 96% with SOR.

Table 1 Summarized Probabilities for all cases of Orders

		lit $(1-d)$	dark pool (d)
buy (a)	market order (k)	$ak(1-d)$	$ad/2$
	limit order $(1-k)$	$a(1-k)(1-d)$	$ad/2$
sell $(1-a)$	market order (k)	$(1-a)k(1-d)$	$(1-a)d/2$
	limit order $(1-k)$	$(1-a)(1-k)(1-d)$	$(1-a)d/2$

A sign for the number of waiting orders in the dark pool indicates buy or sell orders: plus means buy orders, and minus means sell orders.

In the case of $D = 44\%$, a few waiting orders are stored in the dark pool and averaged $r^t_{e,j}$ are positive when P^t is higher than $P_f(= 10,000)$ and vice versa[5]. On the other hand, in the case of $D = 96\%$, many waiting orders are stored in dark pools and averaged $r^t_{e,j}$ are the opposite to those in the case of $D = 44\%$.

Fig. 7 shows how the lit market is made inefficient. When P^t becomes much higher than P_f, many buy waiting orders are stored in the dark pool and averaged $r^t_{e,j}$ are negative, which means that agents make market sell orders. These market sell orders could have led to P^t reducing and converging to P_f. However, many buy waiting orders stored in the dark pool absorb these market sell orders and prevent P^t converging to P_f. Therefore, P^t is maintained at a much higher price than P_f, and the lit market is made inefficient. When P^t becomes much smaller than P_f, the opposite occurs. This is why M_{ie} increased significantly in $D \gtrsim 70\%$.

4 Theoretical Discussion

4.1 Simple Equation Model

In this section, we discuss mechanisms by which a dark pool makes a market efficient or inefficient by using a simple equation model.

We define a, k, and d as the probabilities of an agent making a buy order, making a market order in the lit market without SOR, and ordering to the dark pool the same as in previous sections, respectively. Obviously, $1-a$, $1-k$ and $1-d$ are the probabilities of an agent making a sell order, making a limit order in the lit market without SOR, and ordering to the lit market is, respectively.

In the simulation results, few limit orders in the dark pool were canceled. This means that all limit orders in the dark pool were executed with market orders and that all numbers of limit orders were the same as all numbers of market orders in the dark pool. Therefore, we defined probabilities of an agent making both limit and market orders in the dark pool as 1/2.

[5] When there were not enough data in a P^t range, we did not plot them in Fig. 6. In the case of $D = 44\%$, data exist in narrower P^t ranges than those in the case $D = 96\%$ because there are no trades outside these P^t ranges.

Table 2 Summarized Probabilities: Summation of buy and sell, $a = 1/2$

	lit $(1-d)$	dark pool (d)
market order (k)	$k(1-d)$	$d/2$
limit order $(1-k)$	$(1-k)(1-d)$	$d/2$

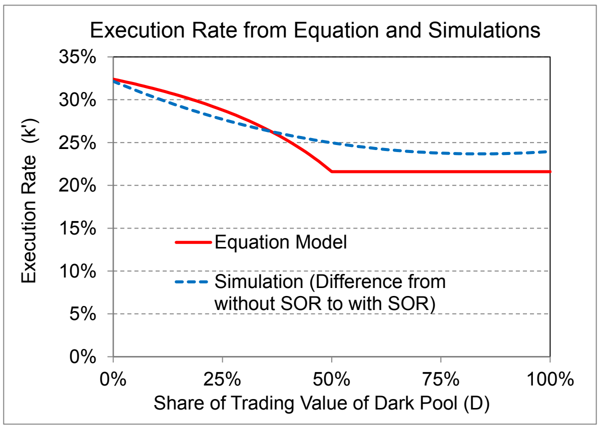

Fig. 8 k' from Eqs.(9) and (10) and the difference between execution rates without and with SOR of simulation results from Fig. 2 for various D.

These definitions derived that, for example, the probability of an agent making a market buy order in the lit order is $ak(1-d)$. Table I summarizes such probabilities for all cases of orders.

Note that from Table I, we can easily derive a share of the trading value amount of the dark pool (D)[6] as defined in the previous section as

$$D = \frac{d/2}{k(1-d) + d/2}. \tag{4}$$

In the following subsections, we investigate market efficiency by the simple equation model from Table I and compare it with simulation results.

4.2 Becoming Efficient: lower than optimal usage rate of Dark Pool

Here, let us discuss an execution rate in the lit market with SOR (k'). For simplicity, we fixed $a = 1/2$ and summed the cases with both a buy and a sell order. Thus, Table I is changed to Table II.

The SOR reduces the probability of market orders to the lit market, $k(1-d)$. When $k(1-d) > d/2$, which means the trading volume in the lit market

[6] Note that in our model agents always order one unit. Therefore, D is exactly the same as the probability of an agent making an market order in the dark pool per the probability of an agent making an market order in both the dark pool and the lit market.

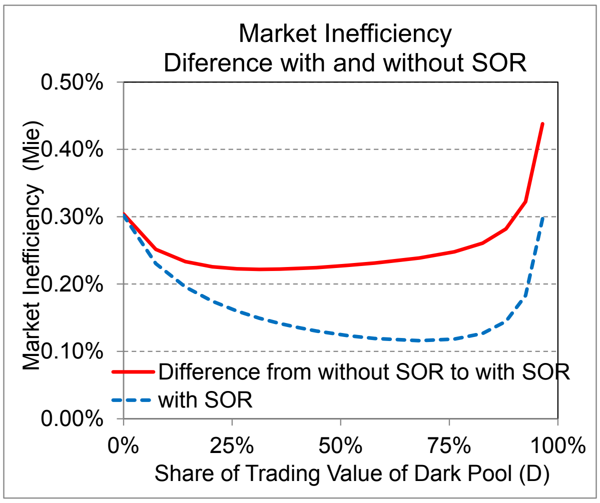

Fig. 9 Market inefficiencies (M_{ie}) with SOR and difference between M_{ie} without SOR and M_{ie} with SOR of simulation results from Fig. 1 for various D.

is greater than that in the dark pool (therefore, this is exactly the same as $D < 1/2$), the number of market orders sent to the dark pool by SOR is limited by the number of limit orders in the dark pool because existing market orders to the lit market outnumber limit waiting orders in the dark pool. In short, reduction of market orders to the lit market by SOR depends on the number of limit orders to the dark pool. Therefore, we can assume the SOR reduces the probability of market orders to the lit market by $\alpha d/2$, where α is constant from zero to 1. Thus, the probability of market orders to the lit market becomes $k(1-d) - \alpha d/2$.

On the other hand, when $k(1-d) < d/2$, the opposite occurs. The number of market orders sent to the dark pool by SOR is limited by the number of market orders to the lit market because existing limit orders in the dark pool outnumber market orders to the lit market. Therefore, we can assume the SOR reduces the probability of market orders to the lit market by $\alpha k(1-d)$, so the probability of market orders to the lit market becomes $k(1-d) - \alpha k(1-d)$.

Therefore, when $k(1-d) > d/2$ (exactly the same as $D < 1/2$), by using Eq.(4), k' becomes,

$$k' = \{k(1-d) - \alpha d/2\}/(1-d) \quad (5)$$

$$= \left(1 - \alpha \frac{D}{1-D}\right) k, \quad (6)$$

On the other hand, when $k(1-d) < d/2$ (exactly the same as $D > 1/2$), by using Eq.(4), k' becomes,

$$k' = \{k(1-d) - \alpha k(1-d)\}/(1-d) \quad (7)$$

$$= k(1-\alpha). \quad (8)$$

Because these two equations should become exactly the same when $k(1-d) = d/2$, we can derive $\alpha = 1/3$, easily[7]. Lastly,

$$k' = \left(1 - \frac{1}{3}\frac{D}{1-D}\right) k \quad (\text{When } D < \frac{1}{2}) \tag{9}$$

$$k' = \frac{2}{3} k \qquad (\text{When } D > \frac{1}{2}). \tag{10}$$

Fig. 8 shows k' from Eqs.(9) and (10) and the difference between execution rates without and with SOR of simulation results from Fig. 2 for various D. We adjusted zero points of them at $D = 0$ to make them exactly the same as the execution rate of simulation results at $D = 0$. We used the difference between execution rates without and with SOR because Eqs.(9) and (10) include the effect only of SOR on reduction of the execution rates and we want to isolate the pure effect of SOR from simulation results.

k' from Eqs.(9) and (10) showed that behaviors of k' are very different between $D < 1/2$ and $D > 1/2$. When $D < 1/2$, the number of market orders sent to the dark pool by SOR depends on the number of limit orders in the dark pool. This means that the more orders made in the dark pool (increasing D), the more orders sent to the dark pool by SOR. Thus, D increases, and k' decreases. On the other hand, when $D > 1/2$, the number of market orders sent to the dark pool by SOR depends on the number of market orders to the lit market. This means that increasing D leads to fewer orders to the lit market and fewer orders sent to the dark pool by SOR. Thus, even though D increases, k' never decreases.

This indicates that, whether markets become efficient or inefficient, the magnitude relationship between the number of market orders to the lit market and the dark pool is intrinsically important; in short, whether $D > 1/2$ or $D < 1/2$. Therefore, this suggests the optimal usage rate of the dark pool for the market efficiency is $D = 1/2$.

The simulation result of k' (the difference between execution rates without and with SOR) was similar to that of Eqs.(9) and (10); when $D < 1/2$, D increases, k' decreases, and when $D > 1/2$, k' is stable. However, decreasing shapes are different; Eqs.(9) and (10) draw a convex upward, the simulation result draws a convex downward, and saturation levels are slightly different. We consider these differences are derived by a mechanism other than SOR and are subjects for a future study.

Fig. 9 shows market inefficiencies (M_{ie}) with SOR and the difference between M_{ie} without SOR and M_{ie} with SOR for simulation results from Fig. 1 for various D. We adjusted the zero point of the difference at $D = 0$ to make it exactly the same as M_{ie} with SOR at $D = 0$. The difference in M_{ie} showed that when $D > 1/2$, D increases, M_{ie} increases; this means that SOR could

[7] We can interpret $\alpha = 1/3$ as follows. There are three cases for waiting limit orders in the dark pool: buy orders waiting, sell orders waiting, and no orders waiting. We assume the three cases have the same probability of occurring. The market order is sent to the dark pool by SOR in the only one case of the three cases, i.e., when opposite orders are waiting in the dark pool. This derives $\alpha = 1/3$.

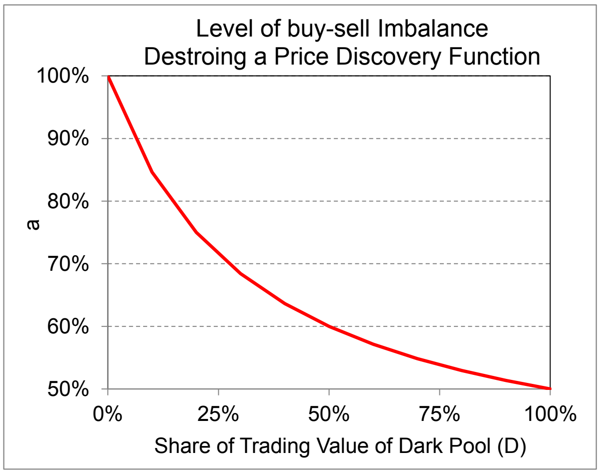

Fig. 10 The boundary of buy-sell order imbalance ($a = (1 + D)/(1 + 3D)$) at which the price-discovery function is destroyed or not for various D.

not make the lit market efficient when $D > 1/2$. This also suggests that a trading volume in dark pools higher than that in lit markets makes markets inefficient.

4.3 Becoming Inefficient: Too High usage rate of Dark Pool

In this subsection, we explain how a too high usage rate of the dark pool makes the lit market significantly inefficient.

We discuss the case in which D is sufficiently higher than 1/2, a large enough number of waiting buy orders are stored in the dark pool, and the market price is higher than the fundamental price. In this case, all market sell orders to the lit market are sent to the dark pool by SOR. Therefore, from Table I, we can derive that the probability an agent makes a sell order in the dark pool is $(1 - a)d + k(1 - a)(1 - d)$. Of course, the probability an agent makes a buy order to the dark pool is ad. If there are more new buy orders to the dark pool than new sell orders to the dark pool, that is

$$ad > (1 - a)d + k(1 - a)(1 - d), \tag{11}$$

waiting buy orders in the dark pool are not reduced and this prevents a market price from converting to the fundamental price. In short, Eq.(11) is the condition in which the price-discovery function of the lit market is destroyed.

Using Eq.(4), we can simplify Eq.(11) to

$$a > \frac{1 + D}{1 + 3D}. \tag{12}$$

Remember that a is the probability of an agent making a buy order. In other words, a is the ratio of buy orders to all orders, and $1 - a$ is the ratio of sell orders to all orders. Thus, a means an imbalance between the numbers of buy and sell orders, $a = 50\%$ means the numbers of buy and sell orders are perfectly balanced, and $a = 100\%$ means no sell orders. Therefore, $a = (1+D)/(1+3D)$ is the boundary of whether the price-discovery function is destroyed or not. Note that this discussion is exactly the same in the opposite case of buy and sell.

Fig. 10 shows $a = (1 + D)/(1 + 3D)$ for various D. The upper side of the line satisfies Eq.(12). When $D = 20\%$, $a = 75\%$; when more than 75% of all orders are buy orders, then less than 25% are sell orders and the number of buy orders is more than three times that of sell orders, so the price-discovery function is destroyed. When buy orders are more than three times as numerous as sell orders, we can say that a very large buy-sell imbalance has occurred and consider this to be a rare case. This suggests that dark pools with $D = 20\%$ rarely destroy the price-discovery function.

On the other hand, when $D = 90\%$, $a = 51\%$ (which is a very slight buy-sell imbalance), and this suggests that dark pools very easily destroy the price-discovery function.

5 Conclusion and Future Studies

In this study, we investigated how a dark pool, in which no order books are provided, affects financial markets' efficiency and price-discovery function by using the artificial market model. This is a very important investigation into financial systemic risk because making a market inefficient and losing the price-discovery function may make the market unstable and increase financial systemic risk. In this study, we additionally implemented a smart order routing (SOR) to the model of Mizuta et al. [Mizuta et al(2015b)] to treat actual market selection of investors. We discussed quantitatively how spreading of dark pools beyond our experience could affect the price-discovery function. We also aimed to clarify the mechanism of a dark pool that makes a market efficient or inefficient.

We found that market inefficiently (M_{ie}) was decreased sharply by raising the share of the trading value amount of the dark pool (D) in $D \lesssim 70\%$. On the other hand, in $D \gtrsim 70\%$, M_{ie} increased significantly. This indicates that there is an optimal usage rate of the dark pool for the market efficiency.

The reason M_{ie} decreased in $D \lesssim 70\%$ is that the execution rates in the lit market are reduced by more market orders being sent to the dark pool by the SOR than limit orders increasing D. This leads the depth of limit orders to become thicker, these thicker limit orders absorb market orders, and thus the market price is still stable near the fundamental price (see also Figures 3 and 5).

The reason M_{ie} increased significantly in $D \gtrsim 70\%$ is as follows. When a market price (P^t) becomes much higher than the fundamental price (P_f), many waiting buy orders are stored in the dark pool and averaged estimated returns ($r^t_{e,j}$) for all agents are negative, which means that agents make market sell orders. These market sell orders could have made P^t converge to P_f, but many waiting buy orders stored in the dark pool absorb these market sell orders and prevent P^t converging to P_f. Therefore, P^t maintains a much higher price than P_f, and the lit market is made inefficient. When P^t becomes much lower than P_f, the opposite occurs (see also Fig. 7).

We also discussed mechanisms by which a dark pool makes a market efficient or inefficient by a simple equation model. The equations about an execution rate we derived indicate that whether $D > 1/2$ or $D < 1/2$ is intrinsically important to whether markets become efficient or inefficient. Therefore, this suggests that the optimal usage rate of the dark pool for the market efficiency is $D = 1/2$ and that a trading volume amount in dark pools higher than that in lit markets makes markets inefficient. We also compared results of the equations with those of simulations and found similar tendencies.

We also derived an equation showing the boundary of a buy-sell imbalance at which dark pools destroy the price-discovery function. We also discussed that when the usage rate of dark pools is low, for example $D = 20\%$, the equation suggests that dark pools rarely destroy the price-discovery function even though a large buy-sell imbalance occurs. On the other hand, when the usage rate of dark pools is very high, for example $D = 90\%$, this equation suggests that dark pools very easily destroy the price-discovery function by a very slight buy-sell imbalance.

A future study is to investigate more details of the optimal usage rate of dark pools for the market efficiency. Our results suggested the optimal usage rate was around $50\% - 70\%$, which is much higher than about 8%, which is the cap level of dark pools that European regulators are discussing [Bowley(2014)]. However, we could not determine the precise level of the optimal usage rate of dark pools for the market efficiency.

Another future study is comparing the simulation results with empirical data. Indeed, we cannot observe M_{ie} of real markets by empirical data because we cannot find fundamental prices in real markets. On the other hand, we can observe an execution rate, depth of limit orders and a bid ask spread of each stock in real lit markets from empirical data. In addition, we can estimate D of each stock from some statistics. D are different from one stock to another. Therefore we can draw figures such as Fig. 2 and Fig. 4 from empirical data plotting each stock having different D, execution rates and so on. To compare these figures with Fig. 2 and Fig. 4, we can compare the simulation results with empirical data.

We also observe buy-sell imbalances in real lit markets from empirical data. Using them and Fig. 10, we can discuss how much D may destroy the price-discovery function in real financial markets.

An artificial market can isolate the pure contribution of these new types of markets to the price formation and can treat such markets on a usage rate higher than we have ever experienced. However, outputs of the artificial market simulation study would not be accurate or credible forecasts of the actual future. The artificial market simulations need to show possible mechanisms affecting price formation by many simulation runs and to gain new knowledge and intelligence; conversely, the artificial market simulations are limited in that their outputs would not certainly but only possibly occur in actual financial markets. Therefore, for more detailed discussions, we should compare the simulation results to results of studies using other methods, e.g. empirical studies.

Appendix

Basic Concept for Building Model

An artificial market, which is a kind of a multi-agent simulation, can isolate the pure contribution of these system changes to the price formation and can treat the changes that have never been employed [LeBaron(2006)], [Chen et al(2012)], [Cristelli(2014)]. These are the strong points of the artificial market simulation study.

However, outputs of the artificial market simulation study would not be accurate or credible forecasts of the actual future. The artificial market simulation needs to show possible mechanisms affecting the price formation by many simulation runs, e.g. searching for parameters, purely comparing before/after the changing, and so on. The possible mechanisms shown by these simulation runs will give us new knowledge and intelligence about effects of the changes to price formation in actual financial markets. Other study methods, e.g. empirical studies, would not show such possible mechanisms. The artificial market simulation studies also need to gain such new knowledge and intelligence.

Indeed, artificial markets should replicate macro phenomena existing generally for any asset and any time. Price variation, which is a kind of macro phenomena, is not explicitly modeled in artificial markets. Only micro processes, agents (general investors), and price determination mechanisms (financial exchanges) are explicitly modeled in artificial markets. Macro phenomena are emerging as the outcome interactions of micro processes. Therefore, the simulation outputs should replicate general macro phenomena at least to show that simulation models are probable in actual markets.

However, it is not a primary purpose for the artificial market to replicate specific macro phenomena only for a specific asset or a specific period. An unnecessary replication of macro phenomena leads to models that are over-fitted and too complex. Such models would prevent our understanding and discovering mechanisms affecting the price formation because of related factors increasing.

Indeed, artificial market models that are too complex are often criticized because they are very difficult to evaluate [Chen et al(2012)]. A too complex model not only would prevent our understanding mechanisms but also could output arbitrary results by over-fitting too many parameters. The simpler the

Table 3 Statistics without Dark Pool

trading	execution rate	32.3%
	cancel rate	26.1%
	number of trades / 1 day	6467
standard deviations	for 1 tick	0.0512%
	for 1 day (20000 ticks)	0.562%
	kurtosis	1.42
autocorrelation coefficient for square return	lag	
	1	0.225
	2	0.138
	3	0.106
	4	0.087
	5	0.075

models, the more difficult arbitrary results are to obtain, and the easier the model is to evaluate.

Therefore, we built an artificial market model that is as simple as possible and do not intentionally implement agents to cover all the investors who would exist in actual financial markets.

Verification of the Model

In many previous artificial market studies, the models were verified to see whether they could explain stylized facts such as a fat-tail, volatility-clustering, and so on [LeBaron(2006)], [Chen et al(2012)], [Cristelli(2014)]. A fat-tail means that the kurtosis of price returns is positive. Volatility-clustering means that the square returns have positive autocorrelation, and the autocorrelation slowly decays as its lag becomes longer. Many empirical studies, e.g. [Sewell(2006)], have shown that both stylized facts (the fat-tail and volatility-clustering) exist statistically in almost all financial markets. Conversely, they also have shown that only the fat-tail and volatility-clustering are stably observed for any asset and in any period because financial markets are generally unstable.

Indeed, the kurtosis of price returns and the autocorrelation of the square returns are stably and significantly positive, but the magnitudes of these values are unstable and very different depending on asset and/or period. The kurtosis of price returns and the autocorrelation of the square returns were observed to have very broad magnitudes of about $1 \sim 100$ and about $0.01 \sim 0.2$, respectively [Sewell(2006)].

For the above reasons, an artificial market model should replicate these values as significantly positive and within a reasonable range as we mentioned. It is not essential for the models to replicate specific values of stylized facts because these stylized facts' values are unstable in actual financial markets.

Table III lists statistics in which there is only one lit market. All statistics are averages of 100 simulation runs, and all the following figures use the average of 100 simulation runs. We define $20,000$ time steps as 1 day because the number of trades within 20,000 time steps is almost the same as that in actual markets per day. All statistics; execution rates, cancel rates[8], standard deviations of returns for one tick and one day[9], kurtosis of price returns, and the autocorrelation coefficient for square returns[10] are of course almost the same as the results of [Mizuta et al(2015b)] because the models do not differ except for the market selection model, which has smart order routing (SOR). [Mizuta et al(2015b)] showed that this mode replicated very short term micro structure, execution rates, cancel rates, and standard deviations of returns for one tick and replicated long-term statistical characteristics, fat tail, and volatility clustering, observed in real financial markets. Therefore, the model was verified to investigate the effect of dark pools on market stability.

Disclaimer & Acknowledgments

Note that the opinions contained herein are solely those of the authors and do not necessarily reflect those of SPARX Asset Management Co., Ltd. and Nomura Securities Co., Ltd. On behalf of all authors, the corresponding author states that there is no conflict of interest. The authors are grateful to Execution Service Department, Nomura Securities Co., Ltd. for their valuable comments from a practical financial perspective on this paper.

References

[Bowley(2014)] Bowley A (2014) Agreement of mifid ii reforms, instinet incorporated. URL http://instinet.com/docs/msr/2014/Agreement_of_MiFID_II_Reforms-Quick_Analysis.pdf

[Chen et al(2012)] Chen SH, Chang CL, Du YR (2012) Agent-based economic models and econometrics. Knowledge Engineering Review 27(2):187–219

[Chiarella and Iori(2002)] Chiarella C, Iori G (2002) A simulation analysis of the microstructure of double auction markets. Quantitative Finance 2(5):346–353

[EC(2010)] European Commission (2010) Public consultation review of the markets in financial instruments directive (mifid). Consultation Report 8

[Cristelli(2014)] Cristelli M (2014) Complexity in Financial Markets, Modeling Psychological Behavior in Agent-Based Models and Order Book Models, Springer

[Friedman(1993)] Friedman D (1993) The double auction market institution: A survey. The Double Auction Market: Institutions, Theories, and Evidence pp 3–25

[Johnson(2010)] Johnson B (2010) Algorithmic Trading & DMA: An introduction to direct access trading strategies. 4Myeloma Press

[Kobayashi and Hashimoto(2011)] Kobayashi S, Hashimoto T (2011) Benefits and limits of circuit breaker: Institutional design using artificial futures market. Evolutionary and Institutional Economics Review 7(2):355–372

[LeBaron(2006)] LeBaron B (2006) Agent-based computational finance. Handbook of computational economics 2:1187–1233

[8] The execution rate is the ratio of the number of trades to that of all orders. The cancel rate is the ratio of the number of cancels to that of all orders + cancels.

[9] In our model, though overnight returns do not exist, the standard deviations of returns for one day correspond to intraday volatility in real financial markets.

[10] We used returns for 10 time units' intervals (corresponding to about 10 seconds) to calculate the statistical values for the stylized facts. In this model, time passes by an agent just ordering even if no dealing is done. Therefore, the returns for one tick (one time) include many zero returns, and they will bias statistical values. This is the reason we use returns for about 10 time units' intervals.

[Mizuta et al(2013)] Mizuta T, Hayakawa S, Izumi K, Yoshimura S (2013) Investigation of relationship between tick size and trading volume of markets using artificial market simulations. In: JPX working paper, Japan Excgange Group, 2, URL http://www.jpx.co.jp/english/corporate/research-study/working-paper/

[Mizuta et al(2015a)] Mizuta T, Izumi K, Yagi I, Yoshimura S (2015a) Investigation of price variation limits, short selling regulation, and uptick rules and their optimal design by artificial market simulations. Electronics and Communications in Japan 98(7):13–21, DOI 10.1002/ecj.11684, URL http://dx.doi.org/10.1002/ecj.11684

[Mizuta et al(2015b)] Mizuta T, Kosugi S, Kusumoto T, Matsumoto W, Izumi K, Yagi I, Yoshimura S (2015b) Effects of price regulations and dark pools on financial market stability: An investigation by multiagent simulations. Intelligent Systems in Accounting, Finance and Management, DOI 10.1002/isaf.1374, URL http://dx.doi.org/10.1002/isaf.1374

[Mizuta et al(2015c)] Mizuta T, Noritake Y, Hayakawa S, Izumi K (2015c) Impacts of speedup of market system on price formations using artificial market simulations. In: JPX working paper, Japan Excgange Group, 9, URL http://www.jpx.co.jp/english/corporate/research-study/working-paper/

[Mo and Yang(2013)] Mo SYK, Yang MPSY (2013) A study of dark pool trading using an agent-based model. In: Computational Intelligence for Financial Engineering Economics (CIFEr), 2013 IEEE Symposium Series on Computational Intelligence on, pp 19–26

[SEC(2010)] Commodity Futures Trading Commission and Securities Exchange Commission (2010) Concept release on equity market structure. Federal Register 75(13):3594–3614

[Sewell(2006)] Sewell M (2006) Characterization of financial time series. http://finance.martinsewell.com/stylized-facts/

[Stöckl et al(2010)] Stöckl T, Huber J, Kirchler M (2010) Bubble measures in experimental asset markets. Experimental Economics 13(3):284–298

[Thurner et al(2012)] Thurner S, Farmer J, Geanakoplos J (2012) Leverage causes fat tails and clustered volatility. Quantitative Finance 12(5):695–707

[TSE(2012)] Tokyo Stock Exchange (2012) Guide to tse trading methodology. URL http://www.jpx.co.jp/learning/tour/books-brochures/detail/tvdivq0000005u66-att/trading_methodology.pdf

[Verheyden et al(2013)] Verheyden T, De Moor L, Van den Bossche F (2013) A tale of market efficiency. Review of Business and Economic Literature 58(2):140–158

[Westerhoff(2008)] Westerhoff F (2008) The use of agent-based financial market models to test the effectiveness of regulatory policies. Jahrbucher Fur Nationalokonomie Und Statistik 228(2):195

[Yagi et al(2010)] Yagi I, Mizuta T, Izumi K (2010) A study on the effectiveness of short-selling regulation using artificial markets. Evolutionary and Institutional Economics Review 7(1):113–132

[Ye(2012)] Ye M (2012) Price manipulation, price discovery and transaction costs in the crossing network. Price Discovery and Transaction Costs in the Crossing Network (March 14, 2012)

[Yeh and Yang(2010)] Yeh C, Yang C (2010) Examining the effectiveness of price limits in an artificial stock market. Journal of Economic Dynamics and Control 34(10):2089–2108