The Navigation of an Iceberg: The Optimal Use of Hidden Orders *

Angelika Esser⋆ and Burkart Mönch⋆⋆

⋆ *DekaBank, Hahnstrae 55, 60528 Frankfurt am Main, Germany*
e-mail: Angelika.Esser@dekabank.de

⋆⋆ *Department of Finance, Faculty of Economics and Business Administration, Uni-PF 77, Goethe University, 60054 Frankfurt am Main, Germany*
e-mail: moench@wiwi.uni-frankfurt.de

January 14, 2005

Abstract. Market participants with large orders to execute are often reluctant to expose these to an open order book in their entirety in order to avoid a potential adverse market impact. Therefore, investors often split large orders into smaller tranches. Iceberg orders facilitate these trading practices by executing such business automatically in the order book. This paper analyzes the rationale for the use of iceberg orders in continuous trading by assessing the costs and benefits of this trading instrument. We present a parsimonious framework that allows the determination of the optimal limit and the optimal peak size of an iceberg order for a static liquidation strategy. Examples with real world order book data demonstrate how the setup can be implemented numerically and provide a deeper insight into relevant properties of the model.

Keywords: optimal liquidation, order book imbalance, limit order, iceberg order, liquidity

JEL classification: G12

1. Introduction

The rapid development in technology over the last couple of years has permitted many stock exchanges to transfer trading from open outcry markets, where market makers or specialists act as intermediaries, to screen-based electronic markets. Typically, electronic trading platforms provide market participants with information on an anonymous open order book during continuous trading in real time. Usually the limits, the accumulated order volumes of each limit, and the number of orders

* This research was partially supported by the Deutsche Forschungsgemeinschaft. We are grateful to the Trading Surveillance Office of the Deutsche Börse AG for providing us a limit order book dataset. We are indebted to Nicole Branger and Christian Schlag for inspiration and insightful discussions. Special thanks for helpful comments go to Matthias Birkner, Götz Kersting, Michael Melvin, Roderich Tumulka, Anton Wakolbinger, and our colleagues Christoph Benkert and Micong Klimes.

in the book at each limit are displayed, so that traders can assess the altering order flow and the market liquidity.

What does the existence of an open order book imply for investment firms who want to submit limit orders, the total volume of which is large relative to others in the market? No doubt, exposing large limit orders in an open order book may reveal the investor's motives for trading and may raise suspicion that the originator of the large order has access to private information about the true value of the security under consideration. Consequently, other market participants change their own order submission strategy, which in turn lowers the probability that the large order will be executed at the prespecified limit. The investment firm then has to choose a less favorable limit if it wants to increase the probability of execution and thus suffers losses from the *indirect* adverse price impact of its large exposure in the order book. A possible solution is not to submit one large limit order but to split the order into several smaller limit orders, which are submitted over time. For this reason many electronic trading platforms introduced so-called iceberg orders. Euronext, the Toronto Stock Exchange, the London Stock Exchange (with its order driven services SETS, SETSmm, and IOB), and XETRA are just some prominent examples. Iceberg orders allow market participants to submit orders with only a certain portion of the order publicly disclosed. The metaphor alludes to the fact that in nature an iceberg's biggest part floats unobservable under the water. Only one-ninth of the mass of ice is seen above the water surface.

An iceberg order is specified by its mandatory limit, its overall volume, and a peak volume. The peak is the visible part of the iceberg order and is introduced into the order book with the original time stamp of the iceberg order according to price/time priority. As soon as the disclosed volume of an iceberg order has received a complete fill and a hidden volume is still available, a new peak is entered into the book with a new time stamp. The new peak behaves in an identical manner to a conventional limit order. From this point of view a pure limit order is basically a special case of an iceberg order where the peak volume coincides with the total order volume.

However, it is important to note that iceberg orders exhibit a less favorable time priority compared with pure limit orders. After the peak of an iceberg order is completely matched, all visible limit orders at the same limit that were entered before the new peak is displayed take priority, i.e. they must have received a complete fill before the new peak comes to execution.

When submitting an iceberg order to the market, several issues have to be considered. Imagine, for example, that the management of a mutual fund has to close a large position in a single stock within one trading day. Using an iceberg order with only a small peak size allows it to minimize the adverse informational impact of disclosing the actual order volume. However, the smaller the peak size the less favorable the time priority of the overall order. Thus, choosing a peak size that is too small seems suboptimal. Such a strategy would significantly lengthen the time to complete execution or would make a complete fill unlikely. Moreover, the fund managers have to choose a reasonable limit for the order. If the limit is too low, one may miss some trade opportunities, i.e. one would give away the chance to participate in raising stock prices. Otherwise, if the limit is too ambitious, the order is unlikely to receive a complete fill.

In the present paper this tradeoff is modeled analytically in a continuous time setup where a large position in a single stock is to be liquidated within a finite trading window.[1] We assume that the investor uses an iceberg order and follows a static strategy, i.e. once the limit and the peak size of the iceberg order are chosen, the trader sticks to this strategy over a fixed period. We then determine the optimal peak size and the optimal order limit by maximizing the expected payoff of the liquidation strategy under certain assumptions concerning the execution risk of the iceberg order. Note that a pure limit order would be also an admissible solution to our optimization problem.

Unless an iceberg sell order is immediately executable, i.e. the limit is so low that it is actually a market sell order, the probability of receiving a complete fill within a finite time horizon is strictly smaller than one. In principle at least two alternative approaches would be able to incorporate execution risk into a liquidation model.

First, one may assume that the investor is forced to trade the remaining shares with a market order if the iceberg order fails to receive complete execution. We call this setup the *self-contained approach.* Market orders are executed immediately. They use liquidity from the book until they are completely filled. Consequently the investor has to bear a liquidity discount, so that he or she gets penalized for every share that could not be sold via the iceberg order. However, in our opinion such a rigorous assumption may not always be justified in practice, especially if the remaining order volume under consideration cannot be absorbed by the market without a significant price change.

In this case, investors typically follow an adaptive strategy, i.e. they review their orders frequently and adjust them if the market moves away from the prespecified order limit. For this reason we also propose a different approach that considers the execution probability as a boundary condition, i.e. only those combinations of peak size and limit

[1] The analysis for a purchasing strategy is symmetric.

are admissible that assure a certain execution probability within a prespecified time horizon. We call this model the *open approach.* Compared with the first one the latter framework is rather flexible and does not require any assumption concerning the liquidation of the unexecuted part of the iceberg order. To get a flavor of the concept, imagine, for example an investor who wants to liquidate a large position, say, within one week. At the end of each trading day the investor inspects the state of the iceberg order and, if necessary, adjusts the limit or the peak size to reach the target.[2] The *open approach* can assist the investor in this procedure. It deals with the optimal combination of order limit and peak size that maximizes the expected liquidation revenues in the case of complete execution, given that the probability to receive a complete fill exceeds a certain level, for example 40% within one trading day. If the order remains partially or completely unexecuted by the end of the first day, the investor may wish to rerun the optimization at the second day and thereby increase the execution probability, let's say, to 60% and so on. If a substantial part of the order is still unexecuted on the last day of the week, the investor will probably choose a minimum execution probability that is close to one. In principle one can also specify a utility function for the investor to model the trade-off between expected payoffs and execution risk. However, in order to keep the problem tractable for exposition we will not address this issue in this paper.

We present a theoretical framework for both the *open* and the *self-contained approach.* Although the underlying assumptions of the latter model are certainly questionable from an empirical point of view we believe that its basic structure may serve as a guideline to build more sophisticated models, for example by implementing an individual penalty function for the unexecuted part of the iceberg order that meets the specific requirements of the investor under consideration. The numerical analysis that illustrates the theoretical part will focus on the *open approach.*

The technical design of the model can be summarized as follows: During continuous trading a transaction takes place if an order becomes executable against orders on the other side of the book. Thus, for an iceberg sell order that is stored on the ask side of the book the dynamics of the best bid price are of special interest. We model the best bid price as a stochastic process in continuous time and assume a constant best bid size. If the stochastic process hits the limit of the iceberg order a

[2] Note that at some exchanges unexecuted iceberg orders are deleted automatically by the system on the end of each trading day and must be resubmitted if desired by the investor at the next trading day. In this case a daily adjustment of order limit and peak size seems very plausible.

transaction is executed and the stochastic process jumps back to the next lower limit. Whether the peak of the iceberg order or another sell order at the same limit is processed at this event depends on the relative time priority of the orders. If new orders with the same limit as the iceberg order are submitted continuously to the book the time priority of the iceberg order deteriorates compared with a pure limit order. The smaller the peak size of the iceberg order, the more often the limit must be hit such that the iceberg order receives a complete fill.

On the other hand, a smaller peak size lowers the adverse informational impact of showing the actual order volume in an open book. We model the drift of the stock price process as a function of the visible order imbalance. When the peak size of an iceberg order enters the book the visible order imbalance changes. We define the order imbalance as the total visible order volume (in number of shares) stored on the bid side of the order book divided by the total visible order volume stored on the bid side and on the ask side of the order book. We exemplify empirically, using order book data, that current variations in the visible order book imbalance are positively correlated with future returns. Thus, the higher the peak size of an iceberg sell order, the smaller the order imbalance and the smaller the expected returns in the next time intervals. Consequently, a higher peak size results in a smaller probability that the stock price process will reach the prespecified limit within the given time horizon.

In total, one can observe two opposite effects if the peak size of an iceberg sell order is reduced in our model:

- The drift of the stochastic process is reduced to a smaller extent when the order enters the book.

- The number of times the limit must be hit in order to process the iceberg order completely increases.

While the first effect is beneficial for the originator of the iceberg order, the latter is not. The proposed framework weights these effects and identifies the optimal combination of peak size and order limit.

The rest of the paper is organized as follows: Section 2 briefly reviews the related literature. The dataset used to exemplify the theoretical ideas throughout this paper is described in Section 3. Section 4 introduces the theoretical setup for both the *self-contained* and the *open approach*. In Section 5 we explicitly model the drift as a function of the order imbalance. The open approach to determine the optimal combination of order limit and peak size is calibrated with a clinical order book data sample in Section 6 so that one can get an impression of the optimal strategies for different scenarios. The paper concludes in Section 7 with a brief summary and a discussion of issues for further research.

2. Related Literature

A number of empirical studies shed light upon the use of hidden orders[3] and the associated motives of traders.

Aitken, Brown, and Walter (1996) state that approximately 6% of orders accounting for 28% of shares traded at the Australian Stock Automated Trading System (SEATS) were undisclosed in 1993. Aitken, Berkman, and Mak (2001) find that undisclosed limit orders are used to reduce the option value of limit orders. This follows the intuition that limit orders can become mispriced when new public information arrives. Some authors, for example Copeland and Galai (1983), therefore characterize limit buy (sell) orders as free put (call) options to other market participants. Pardo and Pascual (2003) use six months of limit order book and transaction data on 36 stocks from the Spanish Stock Exchange (SSE) and report that 26% of all trades (20% of all non-aggressive trades and 42% of all aggressive trades)[4] involve hidden volume. They provide evidence that liquidity suppliers use iceberg orders to mitigate adverse selection costs if new information is released to the market, and that hidden orders temporarily increase the aggressiveness of traders when revealed to the public. D'Hondt, De Winne, and François-Heude (2003) investigate data for six CAC40 stocks traded at Euronext and show that 30% of the depth is hidden in the whole book. The authors highlight that hidden orders are more frequently canceled than usual orders, that iceberg orders are less likely to be totally filled and that the limit of hidden orders is modified more often than that of pure limit orders.

The modeling of optimal liquidation strategies attracts more and more attention by researchers. Bertsimas and Lo (1998), Almgren and Chriss (2000), Hisata and Yamai (2000), Dubil (2002), and Mönch (2003) investigate liquidation strategies for large security positions if market orders are employed as trading instruments. The papers differ from each other mainly in the definition of the stock price dynamics,

[3] Note that at some exchanges the expression "hidden order" is reserved exclusively for orders that are *completely* invisible to other market participants. However, as it is common practice in the literature that is related to our paper we use it as a synonym for an iceberg order.

[4] A buyer- or seller-initiated trade is defined as *aggressive* if it consumes, at least, the best quote on the opposite side of the book.

the modeling of the price impact function, and whether the final time horizon is given exogenously or modeled endogenously.

Compared with market orders, the analysis of optimal liquidation strategies for limit orders is more complex. While the former are matched immediately (provided a sufficient market depth) the execution probability of the latter order type depends critically on the respective limit. Wald and Horrigan (2001) estimate a probit model that characterizes the execution probability depending on a number of variables as the order limit, subsequent realized returns, the bid–ask spread and so on. Lo, MacKinlay, and Zhang (2002) compare empirically three different approaches to determine the execution probability of limit orders using order book data for the 100 largest stocks in the S&P 500 from August 1994 to August 1995. First, they model the execution of a limit order as the first passage time of a geometric Brownian motion to the limit price and find that the predictive power of this setup is only moderate. The first passage time model suffers from important shortcomings. It neither considers the time priority, the order size, a potential adverse impact of revealing large limit orders in the book nor does it distinguish between time-to-first-fill and time-to-completion. As mentioned above, these limitations are eased in the framework proposed in our paper. Second, the authors consider first-passage times determined by historical time series of transaction data. As this approach also ignores the time priority and current market conditions it is not able to represent actual limit order execution times adequately. Finally, the article proposes an econometric model of limit order execution times based on survival analysis and actual limit order data. This empirical approach is a *reduced form* model as it leaves open which mechanism actually causes the execution of a limit order. The model uses eight explanatory variables that are updated in real time to capture current market conditions and three explanatory variables that are updated monthly to model differences across stocks. The authors make some assumptions that may not always be justified empirically to keep the framework tractable. For example the authors assume time-independent covariates. Nevertheless, the empirical framework is able to predict actual limit order executions very well. However, the model requires a continuous update with order book data. Thus, it may be only of limited use if such data are not available in real time.

The intention of our framework is closely related to that of Lo, MacKinlay, and Zhang (2002) although we use a *structural* approach to model explicitly the functionality of the order book. Both approaches focus on the modeling of the execution probability of limit or, in our case, more generally, iceberg orders. In our approach this probability is obtained endogenously and in the parametric model of Lo, MacKinlay, and Zhang (2002) the form of the probability density function is specified exogenously. While Lo, MacKinlay, and Zhang (2002) just present a framework to estimate the time-to-first-fill and the time-to-complete execution of a limit order, our approach is more flexible as it is able to capture every state of partial execution. Furthermore, we extend the analysis to the identification of optimal liquidation strategies for different scenarios.

Cho and Nelling (2000) estimate the execution probability of a limit order conditional on order-specific variables and other variables that capture general market conditions using quote data for 144 NYSE stocks from November 1990 to January 1991. The authors observe that the longer a limit order is outstanding, the less likely it is to receive a complete fill. Furthermore, they find that the execution probability is low when the limit price is far away from the current quote, when the order volume is high, when spreads are narrow, and when volatility is low.

The hypothesis that the order imbalance is a proxy for the execution probability of limit orders and influences the order submission strategy of investors is supported by many authors. Chordia and Subrahmanyam (2002) analyze time series of daily order imbalances and individual stock returns for the period 1988–98 using a comprehensive sample of NYSE stocks. They find that lagged imbalances bear a positive predictive relation to current day returns. Furthermore, they observe that daily imbalances are positively autocorrelated. Ranaldo (2004) uses data of 15 stocks quoted on the Swiss Stock Exchange. He finds that orders are submitted more (less) aggressively when the outstanding order volume on the same (opposite) side of the book is large. Furthermore, he observes that buyers are more concerned about the opposite side of the book, while sellers are more concerned about their own side. Parlour (1998) presents a dynamic setup of an order book, where investors anticipate that the own order placement strategy influences the following order flow and where the execution probability of limit orders is modeled endogenously.

3. Description of the Market and Dataset

The functionality of the liquidation model proposed in this paper will be illustrated by using a clinical order book data sample from the German automated trading system XETRA.

XETRA is an order driven market where investors, by placing limit orders, establish prices at which other participants can buy or sell shares. A trade takes place whenever a counterpart order hits the quotes. The system was introduced in 1997 by the German Stock Exchange. At the time the data were collected, XETRA was open for trading from 9:00 a.m. to 8:00 p.m. The trading day starts with an opening auction, followed by continuous trading, which can be interrupted by one or several intraday auction(s). At the end of the day there is either a closing or an end-of-day auction. On XETRA, all of approximately 6,000 equities listed on the Frankfurt Stock Exchange are tradable. The minimum trading volume is one share. Market participants can see all non-hidden entries on each side of the order book, but trading in XETRA is anonymous, i.e. market participants cannot identify the counterparts. On XETRA there are no dedicated providers of liquidity for blue chips stocks. For small and mid cap stocks, designated sponsors (banks and security firms) are given incentives to provide sufficient liquidity by responding to a quote request within a fixed period of time. Floor trading with market makers on the Frankfurt Stock Exchange still takes place but loses more and more market share. In the blue chip segment merely every tenth share is still traded on the floor.

Based on event histories for 61 trading days (January 03 – March 28, 2002), which were provided by the Trading Surveillance Office of the Deutsche Börse AG, order book sequences were reconstructed. By starting from an initial state, each change in the order book depth caused by entry, filling, cancelation or expiration of orders was considered as prescribed by the market model of XETRA. Due to the huge amount of data only the blue chip share MAN is considered as a representative example of the stocks that are traded in XETRA. MAN is one of Europes leading suppliers of capital goods and systems in the fields of commercial vehicle construction, and mechanical and plant engineering. Over the sample period the daily turnover in MAN on XETRA ranged from 500,000 to 1,000,000 shares, and the order book depth from 80,000 to 150,000 shares. In our dataset we counted 786 iceberg sell orders, 140,948 pure limit sell orders, and 4,130 market sell orders. At first sight the hidden part of the order book seems tiny. However, analyzing the average volume of each trading instrument more deeply changes this impression slightly. Iceberg orders exhibit an average volume of 16,037 shares, whereas pure limit and market orders just have an average order volume of 964 and 1,069 shares. Due to this fact, hidden orders represent a remarkable proportion of 8.24% of the overall volume on the ask side of our sample order book. Pure limit sell orders and market sell orders provide 88.87% and 2.89% of the liquidity on the ask side. Figure 1 shows the spectrum of the observed initial volumes of all hidden sell orders. In Figures 2 and 3 we address the issue of how market participants choose the limit and the peak size in

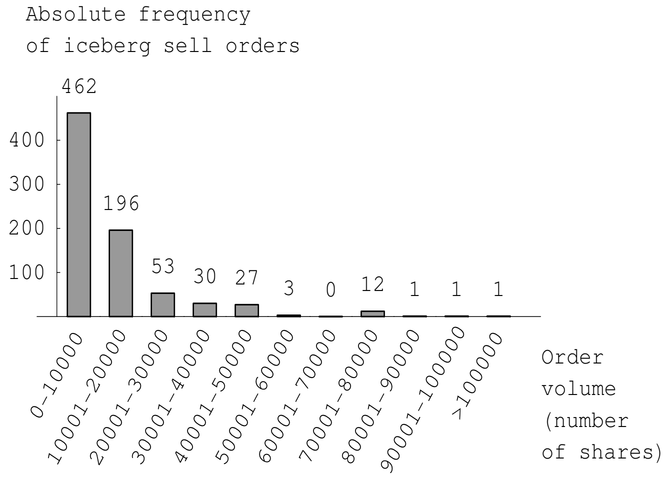

Figure 1. Absolute frequencies of order volumes of all observed iceberg sell orders in the sample.

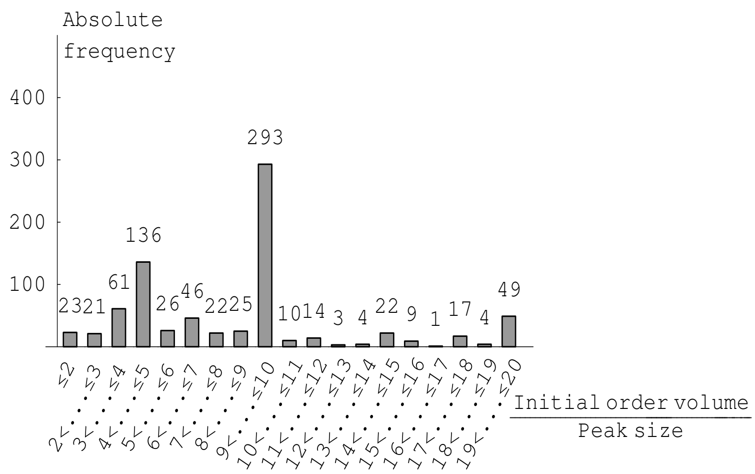

Figure 2. Ratio of initial order volume to peak size of all different iceberg sell orders in the sample.

practice. Obviously there exists a strong preference to specify a peak size that corresponds to a tenth of the overall order volume, as one can

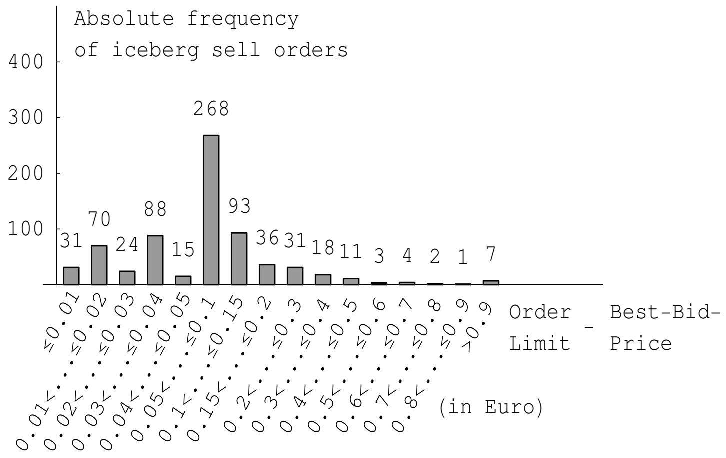

Figure 3. Absolute price difference between the chosen limit and the current best bid price at time of entry of all observed iceberg sell orders in the sample entered during continuous trading.

see in Figure 2. Roughly 37% of all market participants follow such a strategy. To investigate the difference between the chosen limit and the current best bid price at the time of submission we consider only the 702 iceberg sell orders that entered the book during continuous trading. Figure 3 shows that the majority of market participants set the limit between 5 and 15 cents above the best bid price. Note that the average bid–ask spread in our sample is 7 cents and the average midprice €26.91. With respect to the success of the observed trading strategies, Table I delivers an insight into the empirical execution probability of hidden orders. Less than 18% of all iceberg sell orders were executed completely. Almost 30% of all iceberg orders received a partial fill before expiry or cancelation by the investor. The majority (52%) of all hidden sell orders were canceled or expired completely unexecuted. Looking at the median of the observed time between entry and complete execution or deletion one can state that market participants check the state of their orders frequently and cancel them if prices move away from the limit.

Table I. Execution or deletion of iceberg sell orders in the order book sample.

	Number of iceberg sell orders	Average ratio of executed to initial volume; median in parentheses	Average time between entry and complete execution or deletion (hh:mm:ss); median in parentheses
Completely executed	139	100.00% (100.00%)	00:40:48 (00:10:46)
Partially executed	231	32.44% (28.00%)	00:50:04 (00:09:29)
Completely unexecuted	416	0.00% (0.00%)	03:32:43 (00:09:45)

4. The Basic Setup

4.1. General Idea and Dynamics

This section introduces the basic concepts and provides the motivation for the assumptions that have been made. Assume that the large investor holds ϕ_0 shares that should be liquidated before time T. For this purpose the trader submits an iceberg sell order that is stored on the ask side of the order book. The investor assigns a peak size ϕ_p and a limit $\bar{S}$ to the iceberg order. The latter is strictly higher than the initial best bid price S_0 such that the first proportion of the order is not immediately executable.

The best bid price S_t is modeled by a kind of jump-diffusion process. For $S_t < \bar{S}$ it follows a geometric Brownian motion:

$$\mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t \qquad \text{with } S_0 < \bar{S}. \tag{1}$$

Throughout this section the drift μ is assumed to be a constant. In Section 5 we ease this restriction and model the drift as a function of the chosen peak size.

When the process hits the limit of the iceberg order, i.e. $S_{t-} = \bar{S}$, a small downward jump to the next order book entry on the bid side occurs such that $S_t = (1 - \varepsilon)\,\bar{S}$. For sake of simplicity the jump size is modeled as a constant throughout the paper. Figure 4 illustrates the general setting described so far.

Each time the limit $\bar{S}$ is hit by the best bid price a transaction is executed. The transaction size ϕ_s is assumed to be constant over time. Furthermore, we assume that whenever a new tranche of the iceberg order enters the book a fixed volume ϕ_a of other sell orders that exhibit a better time priority is already stored at the same limit.

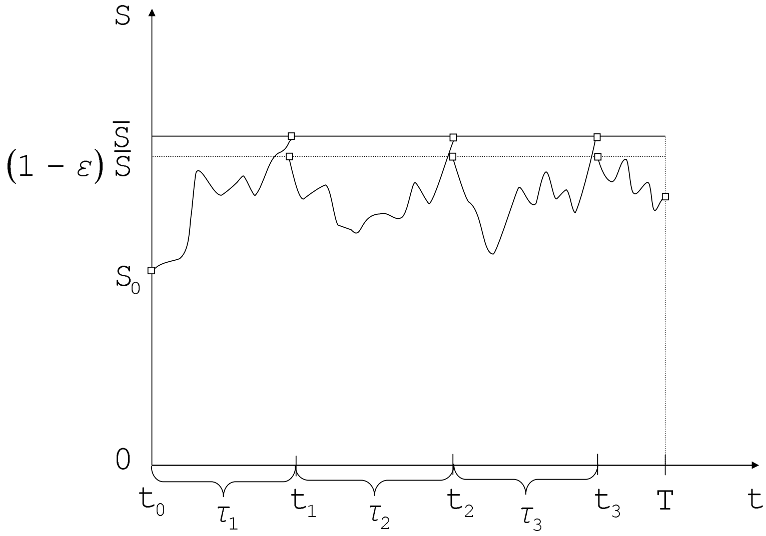

Figure 4. If the best bid price S_t hits the limit of the iceberg order $\bar{S}$, a small downward jump to the next limit on the bid side, i.e. to $(1-\varepsilon)\,\bar{S}$, occurs.

These orders must be matched before the current peak of the iceberg order becomes executable. Thus, one can observe the following sequence of newly displayed order quantities over time: $\phi_a, \phi_p, \phi_a, \phi_p$, and so on. Table II summarizes the notation that is used throughout the rest of the paper.

Table II. Notation used in the proposed liquidation model.

Variable	Meaning
ϕ_0	Total number of shares that have to be liquidated by the large investor before time T
$\phi_p(\leq \phi_0)$	Assigned peak size to the iceberg order
ϕ_a	Volume of other sell orders at limit $\bar{S}$ that is already stored on the ask side when a fresh peak enters the order book
ϕ_s	Total transaction volume that is processed each time the limit is hit

The objective of the investor is the identification of the optimal combination of peak size ϕ_p and limit $\bar{S}$. The time horizon is T, and the investor is interested in the expected payoff of his or her liquidation strategy. To deal with execution risk, we consider two alternative approaches. In the *open approach*, the investor chooses a lower bound P^* for the probability that the submitted iceberg order receives a complete fill up to time T. For the *self-contained approach* we assume that if the iceberg order is not completely executed before time T, the large investor submits a market order to sell the remaining part of the shares. Consequently, the trader has to bear a significant liquidity discount, denoted by $\Psi \cdot S_T$, where $\Psi \gg \varepsilon$. Thus, the investor will receive $(1 - \Psi)\, S_T$ for each of the remaining shares. It seems reasonable to model Ψ as a function of the number of shares that are sold by submitting a market order at T.

The remaining part of this section is dedicated to the derivation of the formulas necessary to implement the *open* and the *self-contained approach*. Obviously, the liquidation value depends on the actual number of times the best bid price hits the prespecified limit of the iceberg order, which in turn is a random variable in the proposed setup. Thus, in Subsection 4.2 we introduce formulas to compute the executed volume of the iceberg order conditional on the event that the limit is hit a certain number of times. In Subsection 4.3 we calculate the probability that the limit is hit a certain number of times. Furthermore, in this subsection we present the objective function of the investor for both the *open* and the *self-contained approach*.

4.2. Execution of the Iceberg Order

The number of times n^* the limit $\bar{S}$ must be hit by the best bid price such that the iceberg order is completely satisfied is given by

$$n^* = \left\lceil \frac{(\phi_0/\phi_p)\phi_a + \phi_0}{\phi_s} \right\rceil = \left\lceil \frac{\phi_0(1 + \phi_a/\phi_p)}{\phi_s} \right\rceil .$$

The brackets $\lceil . \rceil$ are called upper Gaussian brackets, with $\lceil x \rceil = \min\{z \in \mathbb{Z} : \; z \geq x\}$ and $\mathbb{Z}$ as the set of integers. Note that $n^* \cdot \phi_s$ corresponds to the total order volume that is matched after n^* transactions, whereas ϕ_0 shares originate from the iceberg order and $(\phi_0/\phi_p)\phi_a$ shares from other orders.

After the limit is hit n times the number of executed shares of the iceberg order is given by

$$h(n) := \min \left\{ \max \left[\quad \phi_p \left\lfloor \frac{n\phi_S}{\phi_a + \phi_p} \right\rfloor , \right. \right.$$
$$\left. \left. n\phi_S - \phi_a \left(1 + \left\lfloor \frac{n\phi_S}{\phi_a + \phi_p} \right\rfloor \right) \right] , \phi_0 \right\},$$

where $\lfloor x \rfloor = \max\{z \in \mathbb{Z} : \; z \leq x\}$. As long as $n \leq n^*$ the first element of the max-expression is larger than the second if other sell orders at the same limit exhibit a better time priority than the current peak of the iceberg order. If the current peak of the iceberg order takes time priority over other sell orders at the same limit, the second term in the max-expression is larger than the first.

The outstanding order volume of the iceberg order after the n-th hit of the limit is then equal to

$$\phi_0 - h(n).$$

4.3. Liquidation Value

Armed with the results of the previous section one can now calculate the liquidation value conditional on the event that the limit is hit a certain number of times. Let M denote the number of times the limit is hit *before* time T. If $M \geq n^*$, the liquidation value G is given by $G = \phi_0 \bar{S}$, since the iceberg order is completely executed at time T. The *open approach* simply maximizes this expression by solving the following optimization problem:

$$\begin{array}{lll} \max\limits_{\{\phi_p, \bar{S}\}} & \phi_0 \, \bar{S} & \\ s.t. & \mathrm{P}^* & \leq \; \mathrm{P}(M \geq n^*) \\ & S_0 & < \; \bar{S} \\ & \phi_p & \leq \; \phi_0. \end{array}$$

In contrast to the *open approach*, which focuses on the case of full execution, the *self-contained approach* considers also those states of the world where $M < n^*$. In this setup, if $M = n$, the trader receives $h(n)\,\bar{S}$ for the executed part of the iceberg order and $[\phi_0 - h(n)]\, S_T \left[1 - \Psi(\phi_0 - h(n))\right]$ for the remaining part that is liquidated using a market order at time T. Given the realizations of M and S_T the liquidation value can be calculated by

$$G_{M=n} = h(n)\,\bar{S} + [\phi_0 - h(n)]\,[1 - \Psi(\phi_0 - h(n))]\, S_T.$$

However, at time t_0, both M and S_T are random variables. Thus, in order to derive the expected liquidation value one has to weight all possible realizations of the liquidation value G_M by their probabilities, whereas S_T depends on the realization of M. Thus, the expected liquidation value can be written as

$$\mathbb{E}G \;=\; \sum_{n=0}^{\infty} \mathrm{P}(M = n) \times$$

$$
\begin{aligned}
&\left\{ h\,(n)\,\bar{S} + [\phi_0 - h\,(n)]\,[1 - \Psi\,(\phi_0 - h\,(n))] \times \right. \\
&\left. \mathbb{E}\,(S_T|M = n) \right\} \\
= \;& \mathrm{P}(M = 0)\,\phi_0 \cdot [1 - \Psi(\phi_0)]\,\mathbb{E}\,(S_T|\,M = 0) \\
&+ \sum_{n=1}^{n^*-1} \mathrm{P}(M = n) \left\{ h\,(n)\,\bar{S} + [\phi_0 - h\,(n)] \times \right. \\
&\qquad \left. [1 - \Psi\,(\phi_0 - h\,(n))]\,\mathbb{E}\,(S_T|\,M = n) \right\} \\
&+ \mathrm{P}(M \geq n^*)\,\left(\phi_0\,\bar{S}\right).
\end{aligned}
$$

For the *self-contained approach* we need to solve the following optimization problem:

$$
\begin{array}{lll}
\max\limits_{\{\phi_p, \bar{S}\}} & \mathbb{E}G & \\
s.t. & S_0 & < \bar{S} \\
 & \phi_p & \leq \phi_0.
\end{array}
$$

It remains to calculate the following quantities:

- $\mathrm{P}(M = 0)$
- $\mathrm{P}(M = n)$, for $n = 1, \ldots, n^* - 1$
- $\mathrm{P}(M \geq n^*)$
- $\mathbb{E}\,(S_T|\,M = 0)$
- $\mathbb{E}\,(S_T|\,M = n)$, for $n = 1, \ldots, n^* - 1$.

For this purpose we must compute the distributions of the hitting times, denoted by t_i, $i = 1, \ldots, n$. The time periods between two successive hitting times will be denoted by $\tau_i := t_i - t_{i-1},\; i = 2, \ldots, n$ and we will let $\tau_1 = t_1$.

The distribution of τ_1 can be calculated as follows: Since the process for the best bid price S_t up to τ_1 follows a geometric Brownian motion, see equation (1), the logarithm of the process is an arithmetic Brownian motion

$$
\mathrm{d}\,(\ln S_t) = \left(\mu - \sigma^2\big/2\right)\mathrm{d}t + \sigma \mathrm{d}W_t.
$$

Note that if an arithmetic Brownian motion has a negative drift, i.e. when $\mu < \sigma^2/2$ in our model, then τ_1 has a *defective* density function whose integral over $[0, \infty)$ is less than one, see, e.g., Karlin and Taylor (1975), p. 362. Thus, the probability that the limit will *never*(!) be reached is positive. Specifically, the probability for the event $\{\tau_1 < \infty\}$ is given by

$$\mathrm{P}\left(\tau_1 < \infty\right) = \begin{cases} 1 & \text{for } \mu \geq \sigma^2/2 \\ \exp\left[-2\ln\left(\bar{S}/S_0\right)\left|\mu - \sigma^2/2\right|/\sigma^2\right] & \text{for } \mu < \sigma^2/2. \end{cases} \tag{2}$$

The distribution of the first hitting time of $\ln \bar{S}$, starting at $\ln S_0 < \ln \bar{S}$, conditional on the event $\{\tau_1 < \infty\}$, is given by

$$\tilde{f}_{0,1}\left(t\right) = \frac{\ln\left(\bar{S}/S_0\right)}{\sigma\sqrt{2\pi t^3}} \exp\left\{-\frac{\left[\ln\left(\bar{S}/S_0\right) - \left|\mu - \sigma^2/2\right| t\right]^2}{2\sigma^2 t}\right\}. \tag{3}$$

Taking the product of (2) and (3) we can write the unconditional (defective) density of τ_1 as

$$\begin{aligned} f_{0,1}\left(t\right) &= \tilde{f}_{0,1}\left(t\right)\mathrm{P}\left(\tau_1 < \infty\right) \\ &= \frac{\ln\left(\bar{S}/S_0\right)}{\sigma\sqrt{2\pi t^3}} \exp\left\{-\frac{\left[\ln\left(\bar{S}/S_0\right) - \left(\mu - \sigma^2/2\right) t\right]^2}{2\sigma^2 t}\right\}. \end{aligned} \tag{4}$$

At τ_1 the process independently restarts at $\bar{S}(1-\varepsilon)$, following again a geometric Brownian motion. To derive the (defective) density of the first hitting time *after the restart* (denoted by $\tau_2 = t_2 - t_1$) we just need to replace $\ln S_0$ by $\ln[\bar{S}(1-\varepsilon)]$ in equation (4) if we assume a constant drift μ. Thus, for $n^* \geq 2$ one can write

$$f_{n-1,n}\left(t\right) = \frac{-\ln\left(1-\varepsilon\right)}{\sigma\sqrt{2\pi t^3}} \exp\left\{\frac{-\left[-\ln\left(1-\varepsilon\right) - \left(\mu - \sigma^2/2\right) t\right]^2}{2\sigma^2 t}\right\}. \tag{5}$$

Since t_2 can be decomposed into the sum of the two independent random variables τ_1 and τ_2, i.e. $t_2 = \sum_{i=1}^{2} \tau_i$, the (defective) density $f_{0,2}$ of t_2 is simply the convolution of the corresponding (defective) densities, given by

$$f_{0,2}(t) \equiv (f_{0,1} \star f_{1,2})(t) := \int_0^t f_{0,1}(t-u) f_{1,2}(u) \mathrm{d}u.$$

One can proceed by iterating this methodology: At $\tau_i,\ i \geq 2$ the process independently restarts at $\bar{S}(1-\varepsilon)$ following a geometric Brownian motion. Thus, the distribution of $t_n = \sum_{i=1}^{n} \tau_i$ is given by

$$f_{0,n}(t) \equiv (f_{0,n-1} \star f_{n-1,n})(t) = \left(f_{0,1} \star f_{1,2} \star \ldots \star f_{n-1,n}\right)(t). \tag{6}$$

Now, we are able to derive the corresponding probabilities by rewriting the number of hits in terms of hitting times. Since the events

$$\{M = n\} \text{ and } ((t_n < T) \wedge (t_{n+1} \geq T))$$

are identical, the desired probability for $(M \geq 1)$ can be written as

$$\begin{aligned} \mathrm{P}(M = n) &\equiv \mathrm{P}\Big((t_n < T) \wedge (t_{n+1} \geq T)\Big) \\ &= \mathrm{P}(t_n < T) - \mathrm{P}(t_{n+1} < T) \\ &= \mathrm{P}\left(\sum_{i=1}^{n} \tau_i < T\right) - \mathrm{P}\left(\sum_{i=1}^{n+1} \tau_i < T\right) \\ &= \int_0^T f_{0,n}(t)\,\mathrm{d}t - \int_0^T f_{0,n+1}(t)\,\mathrm{d}t. \end{aligned}$$

The probability for the event that the limit is not hit before T is given by

$$\begin{aligned} \mathrm{P}(M = 0) &\equiv \mathrm{P}(t_1 > T) \\ &= 1 - \int_0^T f_{0,1}(t)\,\mathrm{d}t. \end{aligned}$$

The probability for a complete fill of the iceberg order before T can be computed via

$$\begin{aligned} \mathrm{P}(M \geq n^*) &\equiv \mathrm{P}(t_{n^*} \leq T) \\ &= \int_0^T f_{0,n^*}(t)\,\mathrm{d}t. \end{aligned}$$

Now the expected liquidation value, conditional on the event that the limit is hit n times before time T, can be calculated. To simplify the explanation, assume for a moment that the hitting times are deterministic. This assumption will be relaxed later. In this case the expression

$$\mathbb{E}(S_T | \, M = n)$$

is equal to

$$\begin{aligned} &\mathbb{E}_{t_n}\left(S_T \,\middle|\, \max_{t_n < u < T}(S_u) < \bar{S}\right) \\ =\ &\mathbb{E}_{t_n}\left(\exp\left[\ln(S_T/S_{t_n})\right] S_{t_n} \,\middle|\, \max_{t_n < u < T}\left[\ln(S_u/S_{t_n})\right] < \ln(\bar{S}/S_{t_n})\right) \\ =\ &\int_{-\infty}^{\ln(\bar{S}/S_{t_n})} \exp(s)\, S_{t_n} g\left(s \,\middle|\, \max_{t_n \leq u \leq T}\left[\ln(S_u/S_{t_n})\right] < \ln(\bar{S}/S_{t_n}),\ n\right) \mathrm{d}s, \end{aligned}$$

since the n-th hit occurs at t_n and the process independently restarts at t_n following a geometric Brownian motion conditional on the event that the threshold $\bar{S}$ is not hit within the time interval from t_n to T.

One may notice that this scenario is similar to the evaluation of knock-out barrier options (see, e.g., Zhang (1998), pp. 203–259). The formula for the conditional density g of $\ln(S_T/S_{t_n})$ is given in Appendix A.

However, for $M = n \geq 1$, t_n is in fact a random variable. Thus, we need to consider the distribution of t_n, conditional on the event $\{t_n \leq T \bigwedge t_{n+1} > T\}$. Due to the independence and identical distributions of τ_i for $i \geq 2$ this conditional density is given by

$$
\begin{aligned}
f^{cond}(t)\mathrm{d}t &:= \mathrm{P}(t_n \in (t, t+\mathrm{d}t) | \tau_{n+1} > T - t) \\
&= \frac{\mathrm{P}\Big((t_n \in (t, t+\mathrm{d}t)) \wedge (\tau_{n+1} > T - t) \Big)}{\mathrm{P}(\tau_{n+1} > T - t)} \\
&= \frac{f_{0,n}(t)\mathrm{P}(\tau_{n+1} > T - t)\mathrm{d}t}{\mathrm{P}\Big((t_n \leq T) \wedge (t_{n+1} > T) \Big)} \\
&= \frac{f_{0,n}(t)\left(1 - \int_0^{T-t} f_{n,n+1}(s)\mathrm{d}s\right)\mathrm{d}t}{\int_0^T f_{0,n}(u)\,\mathrm{d}u - \int_0^T f_{0,n+1}(u)\,\mathrm{d}u}.
\end{aligned}
$$

Armed with this result we are able to write the conditional expectation of S_T as

$$
\begin{aligned}
&\mathbb{E}\left(S_T | \, M = n\right) \\
&= \int_0^T f^{cond}(t)\, \mathbb{E}_t \left(S_T | \max_{t_n < u < T} (S_u) < \bar{S} \right) \mathrm{d}t \\
&= \int_0^T \frac{f_{0,n}(t)\left(1 - \int_0^{T-t} f_{n,n+1}(s)\mathrm{d}s\right)}{\int_0^T f_{0,n}(u)\,\mathrm{d}u - \int_0^T f_{0,n+1}(u)\,\mathrm{d}u} \times \\
&\quad \Bigg[\int_{-\infty}^{\ln(\bar{S}/S_{t_n})} \exp(s)\, S_t \times \\
&\quad g\left(s | \max_{t \leq u \leq T} [\ln(S_u/S_t)] < \ln(\bar{S}/S_t),\ n \right) \mathrm{d}s \Bigg] \mathrm{d}t. \qquad (7)
\end{aligned}
$$

Note that the integral with respect to t in equation (7) has a singularity at the upper end point of the integration range. Thus, for numerical integration one should use a quadrature routine that can handle functions with end-point singularities.[5]

[5] For example, imsl_d_int_fcn_sing from the IMSL C-Library is such a routine.

Conditional on the event that the limit is not hit before T, the conditional expectation of S_T simplifies to

$$\begin{aligned}
&\mathbb{E}\left(S_T \mid M=0\right) \\
&= \int_{-\infty}^{\ln(\bar{S}/S_0)} \exp(s)\, S_0\, g\left(s \middle| \max_{0\le u\le T}[\ln(S_u/S_0)] < \ln(\bar{S}/S_0),\ n=0\right) \mathrm{d}s.
\end{aligned}$$

The general setup of the alternative approaches is summarized in the following two propositions:

PROPOSITION 1. *The* ***open approach*** *to determine the optimal combination of the peak size and the limit of an iceberg order can be represented by the following optimization problem:*

$$\begin{array}{lll}
\max\limits_{\{\phi_p,\bar{S}\}} & \phi_0\, \bar{S} & \\
s.t. & \mathrm{P}^* & \le \int_0^T f_{0,n^*}(t)\,\mathrm{d}t \\
 & S_0 & < \bar{S} \\
 & \phi_p & \le \phi_0,
\end{array}$$

where P^* *is given exogenously.*

PROPOSITION 2. *The* ***self-contained approach*** *to determine the optimal combination of the peak size and the limit of an iceberg order can be represented by the following optimization problem:*

$$\begin{array}{lll}
\max\limits_{\{\phi_p,\bar{S}\}} & \mathbb{E}G & \\
s.t. & S_0 & < \bar{S} \\
 & \phi_p & \le \phi_0,
\end{array}$$

where $\mathbb{E}G$ *is given by*

$$\begin{aligned}
\mathbb{E}G &= \sum_{n=0}^{\infty} \mathrm{P}(\mathrm{M}=\mathrm{n}) \times \\
&\quad \left\{ h(n)\,\bar{S} + [\phi_0 - h(n)]\,[1-\Psi(\phi_0 - h(n))]\, \mathbb{E}\left(S_T \mid M=n\right) \right\} \\
&= \left[1 - \int_0^T f_{0,1}(t)\,\mathrm{d}t\right] \phi_0 \cdot [1-\Psi(\phi_0)]\, \mathbb{E}\left(S_T \mid M=0\right) \\
&\quad + \sum_{n=1}^{n^*-1} \left[\int_0^T f_{0,n}(t)\mathrm{d}t - \int_0^T f_{0,n+1}(t)\,\mathrm{d}t\right] \times
\end{aligned}$$

$$\left\{h(n)\,\bar{S} + [\phi_0 - h(n)]\,\{1 - \Psi\,[\phi_0 - h(n)]\}\,\mathbf{E}\,(S_T|\, M = n)\right\}$$

$$+ \left[\int_0^T f_{0,n^*}(t)\,\mathrm{d}t\right] \phi_0\,\bar{S},$$

and $\mathbf{E}\,(S_T|\, M = 0)$ *and* $\mathbf{E}\,(S_T|\, M = n)$ *for* $n \geq 1$ *are given by*

$$\begin{aligned}
&\mathbf{E}\,(S_T|\, M = 0) \\
&= \int_{-\infty}^{\ln(\bar{S}/S_0)} \exp(s)\, S_0 \times \\
&\quad g\left(s|\max_{0\leq u\leq T}[\ln(S_u/S_0)] < \ln(\bar{S}/S_0),\ n = 0\right) \mathrm{d}s
\end{aligned}$$

$$\begin{aligned}
&\mathbf{E}\,(S_T|\, M = n) \\
&= \int_0^T \frac{f_{0,n}(t)\left(1 - \int_0^{T-t} f_{n,n+1}(s)\mathrm{d}s\right)}{\int_0^T f_{0,n}(u)\,\mathrm{d}u - \int_0^T f_{0,n+1}(u)\,\mathrm{d}u} \times \\
&\quad \left[\int_{-\infty}^{\ln(\bar{S}/S_{t_n})} \exp(s)\, S_t \times \right. \\
&\quad \left. g\left(s|\max_{t\leq u\leq T}[\ln(S_u/S_t)] < \ln(\bar{S}/S_t),\ n\right) \mathrm{d}s\right] \mathrm{d}t.
\end{aligned}$$

5. Modeling of the Drift Component

Up to now the drift of the best bid price has been assumed to be a constant. This section completes the theoretical framework by modeling explicitly the impact of the peak size on the drift following the intuition that the disclosure of large order volumes has an adverse market impact. For this purpose we will model the drift μ_t as a function of the order imbalance B_t.

Similar to Brown (1997) we define the imbalance B_t of the order book as the number of shares displayed on the bid side divided by the sum of shares displayed on the bid side and the ask side. The imbalance coefficient is bounded by 1 (if no orders are stored on the ask side of the book) and by 0 (if the bid side is empty). The parameter is 0.5 if the ask volume equals the bid volume. Whenever a new peak shows up in the order book the displayed ask volume increases, which in turn reduces B_t.

To keep the setup tractable for exposition, we assume the following simplified scenario: The best bid price exhibits a zero drift $\mu_t \equiv \bar{\mu} = 0$

prior to the submission of the iceberg order ($t < t_0$). Furthermore, suppose that $B_t \equiv \bar{B}$ for $t < t_0$.

As soon as the iceberg order is submitted to the market the symmetry of the order book starts varying. Suppose that each variation in the displayed volume of the iceberg order ϕ_{dp} influences the order book symmetry. The displayed volume of other orders remains constant over time. Thus, we are able to model the imbalance as a function of the displayed volume of the iceberg order only:

$$B_t\left(\phi_{dp}\right) = \frac{c}{d+\phi_{dp}}, \tag{8}$$

where the parameters c (and d) denote the number of shares displayed on the bid side (on the bid side *and* the ask side) before the submission of the iceberg order.

The displayed volume ϕ_{dp} is equal to ϕ_p whenever a new peak is submitted to the order book. When the peak of the iceberg order receives a complete or partial fill the parameter ϕ_{dp} will be reduced. In our setup the displayed volume ϕ_{dp} depends on the number of times the limit was already hit. It can be calculated as

$$\phi_{dp}(n)$$

$$= \begin{cases} \min\Big[\phi_0 - h(n), & \\ \phi_p - \Big\{\max\Big[\Big(\frac{n\phi_s}{\phi_a+\phi_p} - \Big\lfloor\frac{n\phi_s}{\phi_a+\phi_p}\Big\rfloor\Big)(\phi_a+\phi_p) - \phi_a, 0\Big]\Big\}\Big] & \text{if } n < n^* \\ 0 & \text{else.} \end{cases}$$

If the displayed volume of submitted orders has some information content to the market, one would expect a positive relationship between past levels of B_t and future returns. We define μ_n by the recursion

$$\mu_{n+1} \;=\; \mu_n + \beta \cdot \left(B_{t_{n+1}} - B_{t_n}\right) \text{ for } n \geq 0$$

with the initial value

$$\mu_0 \;=\; \bar{\mu} + \beta \cdot \left(B_{t_0} - \bar{B}\right),$$

such that

$$\Delta\mu_n = \beta \Delta B_{t_n}. \tag{9}$$

Figure 5 illustrates this idea. Before time t_0 the drift μ_t is equal to the long-term mean $\bar{\mu}$. At times t_0 the first peak of the iceberg order appears in the book and the drift μ_t is reduced. At times t_3, t_4, t_5, and t_6 the first peak of the iceberg order receives partial fills, which goes along with small upward jumps in the drift μ_t. At time t_7 the first peak becomes completely filled and the second peak appears in the order book, which again causes a downward jump in the drift μ_t. At time t_n^* the iceberg order is completely filled and the drift μ_t reverts towards its long-term mean $\bar{\mu}$.

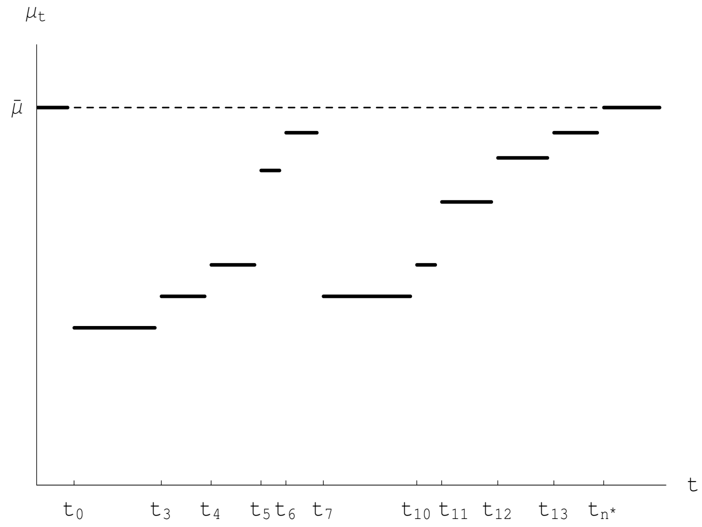

Figure 5. Example for the alternating drift component in our model.

Note that μ_n is a deterministic function of the random variable n in our framework. Thus, we can rewrite equation (4) for the case where the drift depends on the displayed peak size:

$$\begin{aligned} f_{0,1}(t) &= \frac{\ln(\bar{S}/S_0)}{\sigma\sqrt{2\pi t^3}} \times \\ &\quad \exp\left\{\frac{-\left[\ln(\bar{S}/S_0) - (\mu_0 - \sigma^2/2)\,t\right]^2}{2\sigma^2 t}\right\}. \end{aligned} \tag{10}$$

For the subsequent hitting times after the restart we get

$$\begin{aligned} f_{n-1,n}(t) &= \frac{-\ln(1-\varepsilon)}{\sigma\sqrt{2\pi t^3}} \times \\ &\quad \exp\left\{\frac{-\left[-\ln(1-\varepsilon) - (\mu_{n-1} - \sigma^2/2)\,t\right]^2}{2\sigma^2 t}\right\}. \end{aligned} \tag{11}$$

This completes the introduction of the theoretical framework. We now turn to the numerical implementation of the open approach.

6. Numerical Results

To exemplify the formal analysis of the previous section the *open approach* is implemented using the MAN dataset for 61 trading days. The results are presented in the following.

6.1. Parameter Specification

Table III. Estimated parameters for the MAN dataset for 61 trading days.

		No. of obs.	Mean	Std. dev.
ϕ_s	Average volume of all transactions	40,888	868.2	1,257.4
ϕ_a	Average displayed volume of all best ask quotes	158,607	1,554.4	1,752.0
ε	Average relative price difference between the best and the second best bid price	57,290	0.00090724	0.00105471
c	Average number of shares displayed on the bid side of the book	36,661	166,465.87	57,495.87
d	Average number of shares displayed on the bid side and on the ask side of the book	36,661	291,657.18	84,722.34

To implement the model a number of parameters need to be calibrated with order book data. Table III summarizes the results for our clinical sample. To estimate the parameters ϕ_s, ϕ_a, and ε we consider all observed transactions, and best and second best ask quotes with equal weights. For the calibration of the parameters c and d we use the order book data collected at intervals of 1 minute from 9:30 a.m. to 19:30 p.m. To estimate the parameter β we regress 60-minutes-ahead returns on changes of the order imbalance during the past 60 minutes, minute by minute. For this purpose we use best bid quotes and order book data collected at intervals of 1 minute from 9:30 a.m. to 19:30 p.m. We do not consider overnight returns for our analysis.

Table IV reports the results. For the calibration of the volatility parameter we use best bid quotes collected at intervals of 15 minutes from 9:00 a.m. to 20:00 p.m. and do not consider overnight price changes. The estimation for the volatility parameter yields $\sigma = 0.7$. Furthermore, we set $S_0 = €\,28.55$, which is the closing price at March 28, 2002, the last day of our sample period.

Table IV. Estimated regression coefficient for equation (9).

		No. of obs.	Estimate	t-statistic
β	60-minute forecast intervals	29,341	116.647882	22.45

6.2. Numerical Implementation

The computation of $f_{0,n}(t)$ requires the calculation of an n-th iterated convolution given by equation (6). In order to obtain $f_{0,n}(t)$ one needs to calculate $(n-1)$-dimensional integrals. To the best of our knowledge, closed form expressions are not available. Thus, we apply numerical approximations to these integrals. Employing conventional quadrature algorithms or Monte Carlo methods to compute high-dimensional integrals is very time consuming and thus not suitable for the dimensions under consideration in our framework. Therefore, we use interpolating cubic splines $s_{0,n}$ for $n \geq 2$ to approximate the convolutions in the following way:

$$f_{0,n}(t) \approx s_{0,n}(t),$$

such that

$$s_{0,n}(\hat{t}_k) = \int_0^{\hat{t}_k} s_{0,n-1}(\hat{t}_k - u) f_{n-1,n}(u)\,\mathrm{d}u \approx f_{0,n}(\hat{t}_k),$$

where $\hat{t}_k$ denotes the equally spaced spline knots.

Alternatively, one can also invert the Laplace transform of the density function.[6] The Laplace transform for the (defective) density func-

[6] Numerical routines that do this job pretty fast are, for example, imsl_d_inverse_laplace from the IMSL C-Library or C06LAF/C06LBF from the Nag Fortran-Library.

tion of the first hitting time τ_1 is well known (see, e.g., Karlin and Taylor, p. 362) and is given by

$$\mathbb{E}\exp(-\lambda\tau_1) = \exp\left\{ -\frac{\ln(\bar{S}/S_0)}{\sigma^2} \times \left[\sqrt{\left(\mu_0 - \frac{\sigma^2}{2}\right)^2 + 2\sigma^2\lambda} - \left(\mu_0 - \frac{\sigma^2}{2}\right)\right]\right\}.$$

As the following sequences of hitting times τ_i for $i \geq 2$ are identically distributed their Laplace transform is given by

$$\mathbb{E}\exp(-\lambda\tau_i) = \exp\left\{ -\frac{\ln\left[\bar{S}/\bar{S}(1-\varepsilon)\right]}{\sigma^2} \times \left[\sqrt{\left(\mu_{i-1} - \frac{\sigma^2}{2}\right)^2 + 2\sigma^2\lambda} - \left(\mu_{i-1} - \frac{\sigma^2}{2}\right)\right]\right\}.$$

The Laplace transform of the sum of the independent hitting times $t_n = \sum_{i=1}^{n} \tau_i$ is equal to the product of the corresponding exponential functions:

$$\begin{aligned}\mathbb{E}\exp(-\lambda t_n) = \exp\Bigg\{ & -\frac{\ln(\bar{S}/S_0)}{\sigma^2} \times \\ & \left[\sqrt{\left(\mu_0 - \sigma^2/2\right)^2 + 2\sigma^2\lambda} - \left(\mu_0 - \sigma^2/2\right)\right] \\ & + \sum_{i=2}^{n} - \left(\frac{-\ln(1-\varepsilon)}{\sigma^2}\right) \times \\ & \left[\sqrt{\left(\mu_{i-1} - \sigma^2/2\right)^2 + 2\sigma^2\lambda} - \left(\mu_{i-1} - \sigma^2/2\right)\right]\Bigg\}.\end{aligned}$$

6.3. Numerical Examples

For the first example, suppose that the investor wants to liquidate 10,000 MAN shares (approximately 1–2% of daily turnover) within 10 hours. Assume, furthermore, that ϕ_p has to be a multiple of 1,000 shares. Figure 6 represents the optimal limit $\bar{S}$ as a function of the probability P^* that the iceberg order receives a complete fill. For $\mathrm{P}^* \leq 59\%$, the optimal limit is a monotonic decreasing function of P^*. The optimal peak size remains at a constant level of 8,000 shares and is thus insensitive to changes of P^*. Smaller or higher peak sizes reduce the

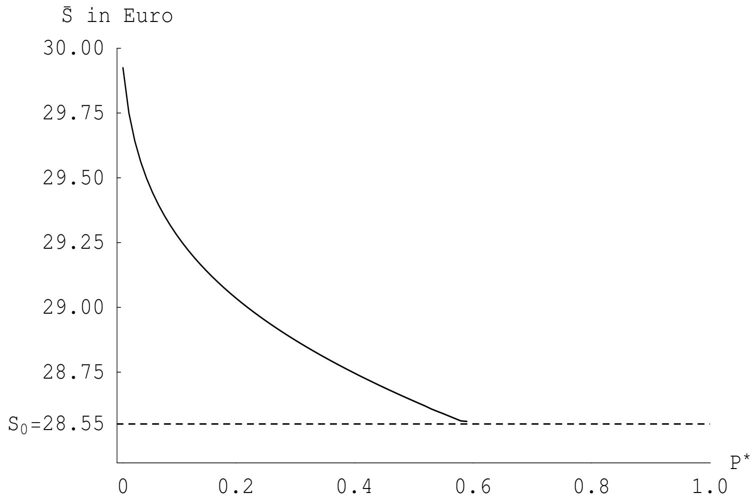

Figure 6. Optimal limit $\bar{S}$ as a function of the probability P^* that the iceberg order receives a complete fill. Other parameters: $T = 10$ hours, $\phi_0 = 10{,}000$ MAN shares.

value of the objective function, for example by approx. 1% if $\mathrm{P}^* = 30\%$ and $\phi_p = 1{,}000$ shares. If the limit is set to €28.56, i.e. the smallest possible value in this example, the probability to observe a complete execution is still less than 60%.

The optimal peak size is significantly higher than peak sizes that were observed empirically in Section 3. Two reasons may explain the difference. First, one may argue that the model systematically underestimates the negative price impact of displaying a large order volume in the book. There are good reasons to believe that a variation of the order imbalance within entries close to the best quotes has a stronger impact on future returns than changes of the order imbalance caused by an entry of an order that possesses a more unfavorable price priority than the majority of other orders already stored in the book. A redefinition of the order imbalance by weighting order book entries differently, depending on their price priority, might solve this problem. Second, one may argue that market participants overestimate the informational impact of revealing large orders in an open book. The empirical exploration of these issues is left for further research.

In the next example we investigate the relationship between the final time horizon T and the optimal combination of limit and peak size. We set $\mathrm{P}^* = 50\%$ and $\phi_0 = 10{,}000$ shares. Figures 7 and 8 report the results. If $T < 6$ hours, then the probability of receiving a complete

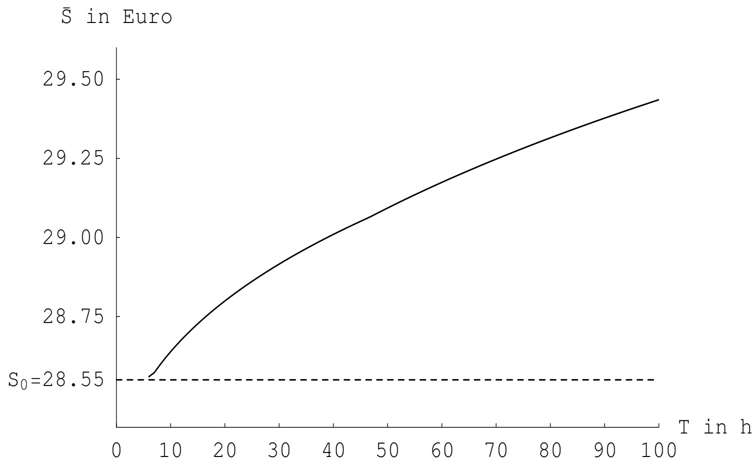

Figure 7. Optimal limit $\bar{S}$ as a function of the final time horizon T. Other parameters: $P^* = 50\%$, $\phi_0 = 10{,}000$ MAN shares.

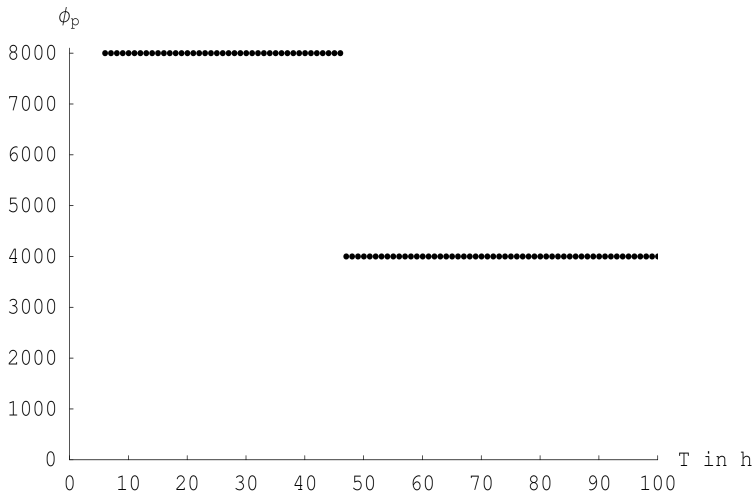

Figure 8. Optimal peak size ϕ_p as a function of the final time horizon T. Other parameters: $P^* = 50\%$, $\phi_0 = 10{,}000$ MAN shares.

fill is less than 50%, no matter which limit is assigned to the order. If $T \geq 6$ hours we can observe two beneficial effects for the originator of the iceberg order. First, as the final time horizon increases, the optimal order limit increases as well. Second, a longer time horizon allows for a reduction of the peak size. However, ϕ_p is not strictly monotonic decreasing in T. Instead we observe a step function. For $T \leq 46$ hours a peak size of 8,000 shares is optimal, for $T > 46$ the optimal peak size is 4,000 shares.

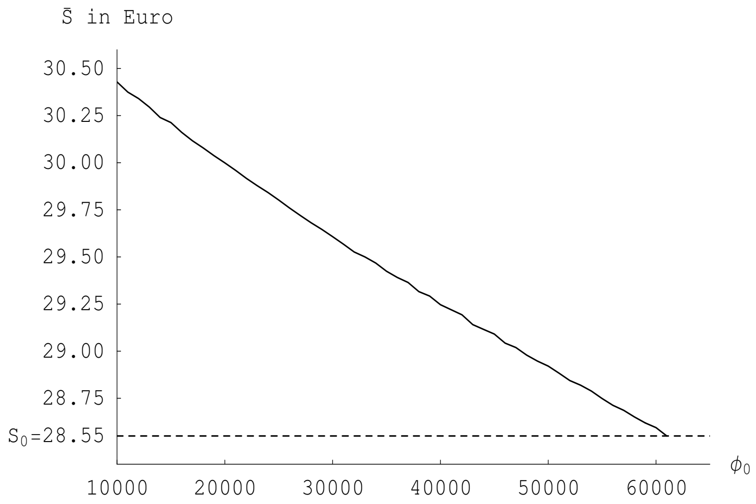

Figure 9. Optimal limit $\bar{S}$ as a function of ϕ_0. Other parameters: $\mathrm{P}^* = 25\%$, $T = 100$ hours.

In the last example (see Figures 9 and 10) we analyze the relationship between the initial position ϕ_0 and the optimal pairs of ϕ_p and $\bar{S}$. We set $\mathrm{P}^* = 25\%$ and $T = 100$ hours. Figure 9 corroborates the hypothesis that if more shares have to be liquidated within the same period of time the limit has to be lowered to keep the execution probability at the same level. Furthermore, an increase in ϕ_0 tends to result in higher peak sizes, as we can observe in Figure 10. However, in some cases the optimal peak size decreases if the initial position is raised. At the first moment this may seem somehow counterintuitive. The main reason for this phenomenon can be found in the discrete setup of the order execution process. Whenever the limit $\bar{S}$ is hit, a fixed transaction size ϕ_s is processed. At the n^*-th hit the last part of the iceberg order, which is given by $\phi_0 - h(n^* - 1)$, becomes executable. However, if $\phi_0 - h(n^* - 1) \ll \phi_s$ a small reduction in ϕ_p would not change n^* but would increase the drift component μ_t and thus the probability that $t_n < T$.

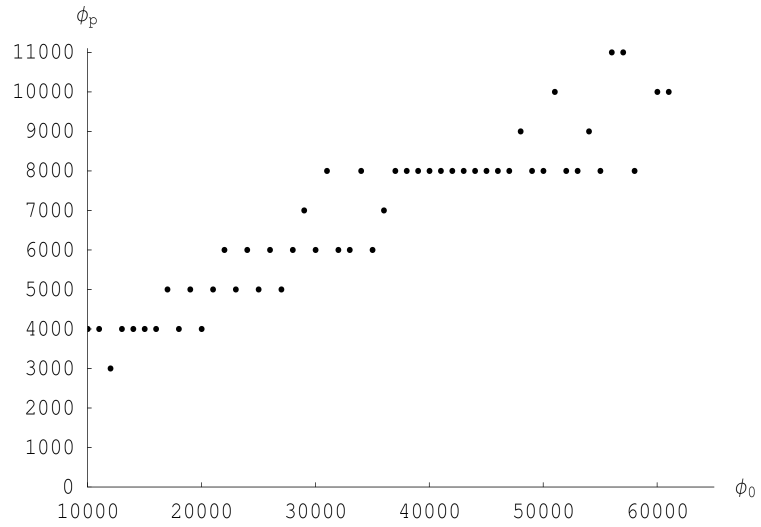

Figure 10. Optimal peak size ϕ_p as a function of ϕ_0. Other parameters: $P^* = 25\%$, $T = 100$ hours.

7. Summary and Conclusion

This paper introduces a setup that allows the determination of the optimal combination of limit and peak size of an iceberg order, given a large position in a security that should be liquidated within a finite time horizon. The framework balances the direct advantage of a large peak size that leads to a better time priority of an iceberg order and the adverse informational impact of revealing large order volumes in an open order book. Furthermore, it assesses the tradeoff between the order limit and the execution probability of the iceberg order. We have presented two approaches to incorporate the execution risk of an iceberg order. The so-called *self-contained approach* assumes that the unexecuted part is liquidated by a market order. The *open approach* is far more flexible as it does not require any assumption concerning the liquidation of the unexecuted part. It identifies the optimal combination of limit and peak size, given a minimum probability of complete order execution. Using real-world order book data we illustrate how the *open approach* can be implemented and explore major properties of the model by modifying input parameters.

To our knowledge, this framework is the first analytical approach that investigates the tradeoff between limit and peak size of an iceberg order, on the one hand, and the resulting execution probability, on the other. This paper is written in search of a stylized model that is able to illustrate the interaction between observable market variables and order specific parameters that are important to analyze iceberg orders as a trading instrument.

The modeling of the best bid price by a Brownian motion or the assumption of constant parameters for order imbalance, transaction size, order flow and the price difference between the best and the second best price are, of course, approximations as the standard deviations in Table III clearly indicate. These simplifications allow us to keep the number of stochastic variables to the minimum required to illustrate the discussed trade-off in a simple way. Further research may focus on introducing more freedom from determinism by modeling more sources of risk, for example, in a simulation-based approach and comparing the empirical performance of the different models. Furthermore, although certainly challenging from a technical point of view, the investigation of dynamic approaches seems highly relevant from an empirical perspective, since many market participants pursue dynamic instead of static limit-setting strategies as shown in Table I in Section 3.

References

Aitken, M. J., P. Brown, and T. Walter (1996): Infrequent trading and firm size as explanation for the intra-day patterns in returns on SEATS, Working Paper, Securities Industry Research Centre of Asia-Pacific.

Aitken, M. J., H. Berkman, and D. Mak (2001): The use of undisclosed limit orders on the Australian Stock Exchange, *Journal of Banking & Finance* 25, 1589–603.

Almgren, R. and N. Chriss (2000): Optimal execution of portfolio transactions, *Journal of Risk* 3 (Winter 2000/2001), 5–39.

Bertsimas, D. and A. W. Lo (1998): Optimal control of execution, *Journal of Financial Markets* 1, 1–50.

Brown, P., D. Walsh, and A. Yuen (1997): The interaction between order imbalance and stock price, *Pacific-Basin Finance Journal* 5, 539–57.

Cho, J. W. and E. Nelling (2000): The probability of limit-order execution, *Association for Investment Management and Research*, September/October, 28–33.

Chordia, T. and A. Subrahmanyam (2002): Order imbalance and individual stock returns, Working Paper, Emory University, Atlanta, USA.

Copeland, T. E. and D. Galai (1983): Information effects on the bid–ask spread, *Journal of Finance* 38, 1457–69.

D'Hondt, C., R. De Winne, and A. François-Heude (2003): Hidden orders on Euronext: Nothing is quite as it seems . . . , Working Paper, FUCaM-Catholic University of Mons, Belgium.

Dubil, R. (2002): Optimal liquidation of large security holdings in thin markets, Working Paper, University of Connecticut, Storrs, USA.

Harrison, M. J. (1990): *Brownian Motion and Stochastic Flow Systems*, Krieger Publishing Company, Malabar, Florida, USA.

Hisata, Y. and Y. Yamai (2000): Research toward the practical application of liquidity risk evaluation methods, *Monetary and Economic Studies*, December, 83–128.
Karlin, S. and H. M. Taylor (1975): *A First Course in Stochastic Processes*, Academic Press, San Diego.
Lo, A. W., A. C. MacKinlay, and J. Zhang (2002): Econometric models of limit-order execution, *Journal of Financial Economics* 65, 31–71.
Mönch, B. (2003): Optimal liquidation strategies, Working Paper, Goethe University, Frankfurt am Main, Germany.
Pardo, Á. and R. Pascual (2003): On the hidden side of liquidity, Working Paper, Universidad de las Islas Baleares, Palma de Malorca, Spain.
Parlour, C. A. (1998): Price dynamics in limit order markets, *Review of Financial Studies* 11, 789–816.
Ranaldo, A. (2004): Order aggressiveness in limit order markets, *Journal of Financial Markets* 7, 53–74.
Wald, J. K. and H. T. Horrigan (2001): Optimal limit order choice, Working Paper, forthcoming 2006 in *Journal of Business* 79.
Zhang, P. G. (1998): *Exotic Options*, World Scientific Publishing, Singapore.

Appendix

A. Conditional Density g of $\ln(S_T/S_{t_n})$

The conditional density g of $\ln(S_T/S_{t_n})$ is given by

$$g\left(s| \max_{t_n \le u \le T} \ln(S_u/S_{t_n}) \; < \; \ln(\bar{S}/S_{t_n}),\, n \right) = \frac{\psi_t\left(s, \ln\left(\bar{S}/S_{t_n}\right)\right)}{\mathrm{P}\left(\ln\left(\max_{t_n \le u \le T} S_u \Big/ S_{t_n}\right) < \ln\left(\bar{S}/S_{t_n}\right)\right)},$$

where

$$\begin{aligned}
\psi_t(x,y) &= (1/\sigma)\exp\Big[\left(\mu^* - \sigma^2\big/2\right) x\big/\sigma^2 \\
&\quad - \left(\mu^* - \sigma^2\big/2\right)^2 (T-t_n)\big/2\sigma^2\Big]\,\delta\left(x/\sigma, y/\sigma\right), \\
\delta(x,y) &= \Big[\varphi\left(x\,(T-t_n)^{-1/2}\right) \\
&\quad -\varphi\left((x-2y)\,(T-t_n)^{-1/2}\right)\Big]\,(T-t_n)^{-1/2},
\end{aligned}$$

and

$$\mathrm{P}\left(\ln\left(\max_{t_n \le u \le T} S_u \Big/ S_{t_n}\right) \; < \; \ln\left(\bar{S}/S_{t_n}\right)\right)$$

$$= \Phi\left(\frac{\ln\left(\bar{S}/S_{t_n}\right) - \left(\mu^* - \sigma^2/2\right)\left(T - t_n\right)}{\sigma\sqrt{\left(T - t_n\right)}}\right)$$
$$- \exp\left[2\left(\mu^* - \sigma^2\Big/2\right)\ln\left(\bar{S}/S_{t_n}\right)\Big/\sigma^2\right] \times$$
$$\Phi\left(\frac{-\ln\left(\bar{S}/S_{t_n}\right) - \left(\mu^* - \sigma^2/2\right)\left(T - t_n\right)}{\sigma\sqrt{\left(T - t_n\right)}}\right),$$

where $\varphi(z)$ denotes the standard normal density function and $\Phi(z)$ the standard normal cumulative distribution function. If the drift is a constant, as assumed in Section 4, set $\mu^* = \mu$. If the drift is modeled as a time-dependent variable, as proposed in Section 5, replace μ^* by μ_{t_n}. For the derivation of the respective formulas , see, e.g., Harrison (1990), pp. 1–16.