Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

The Option Value in Timing Derivative Trades

Feike Drost, Thijs van der Heijden and Bas Werker

DP 05/2015-085

The option value in timing derivative trades

Feike C. Drost[*], Thijs van der Heijden[†], and Bas J.M. Werker[‡]

May 2, 2015

Abstract

Risk-neutral traders executing derivative trades on behalf of portfolio managers maximize their expected profit compared to trading at pre-determined times by timing trades, using the quickly changing risk exposures of derivative baskets. The optimal order submission strategy is a sequence of stop orders with a time-varying stop price. Timing a straddle trade yields up to 20bps per day in a frictionless world, and up to 72bps per day on the S&P500. A CRRA trader is willing to pay up to 51bps of the value of the derivatives to switch from trading at a fixed time to the optimal timing strategy.

JEL codes : G11, G13.

Key words and phrases : derivatives trading, execution timing, optimal stopping, dynamic programming, straddles, dynamic order strategies.

[*]Econometrics and Finance Group, Tilburg University, P.O. Box 90153, 5000 LE, Tilburg, The Netherlands.

[†]Department of Finance, The University of Melbourne, 198 Berkeley Street, Carlton VIC 3010, Australia, email: thijsv@unimelb.edu.au

[‡]Econometrics and Finance Group, Tilburg University, P.O. Box 90153, 5000 LE, Tilburg, The Netherlands. We thank Ramon van den Akker, Joost Driessen, Bruce Grundy, Frank de Jong, Ralph Koijen, André Lucas, Eric Renault, Leonidas Rompolis, and seminar participants at the University of Newcastle for very helpful comments. Computer code to reproduce the results in this paper is available from the second author upon request.

1 Introduction

In many investment banks and hedge/mutual funds, (derivative) traders who execute trades on behalf of portfolio managers are given a deadline by which to execute certain trades, for example the end of the trading day, reflecting the common industry practice of daily portfolio rebalancing. In addition, the trader will be given an objective against which their performance will be benchmarked, for example immediate execution, or a Time-Weighted Average Price (TWAP). The compensation of the trader, acting as an agent for the portfolio manager, is often directly linked to realized profits/losses relative to the benchmark strategy, providing incentives for the trader to maximize the expected profit relative to the benchmark strategy. Importantly, the expected-profit maximizing strategy for the trader will maximize the expected return on the portfolio from the moment it is initiated to the end of the trading period for the portfolio manager. To a first order approximation, the trader can be seen to be risk-neutral with respect to the price uncertainty associated with delayed execution. We show that when trading (baskets of) derivative securities, timing the trade execution yields economically meaningful benefits compared to benchmark trading strategies that use fixed trading times for a price-taking trader, even in the absence of market frictions.

The frictionless world of the Black-Scholes model is a natural starting point for our analysis of the optimal execution time. It is well-known that the Black-Scholes model cannot match observed option prices (see, e.g., Rubinstein (1985) and Rubinstein (1994)), but the model is still often used in risk management and its single source of risk provides a simple framework to study the implementation problem of a derivative trader.

Rather than letting the trader be completely free in the execution timing, resulting in a potentially large tracking error relative to the benchmark trading strategy, the portfolio manager may choose to impose a risk-management constraint in the spirit of the Value-at-Risk condition used in Danielsson, Shin, and Zigrand (2012). In our set up, such a constraint takes the form of a stop-loss level, which maximizes the purchase price of the desired position to a predetermined value, presumably linked to the value of the position at the start of the trading period. The stop-loss level can also be seen as introducing risk aversion for the trader.

The strategy maximizing expected-profit, or optimally timing trade execution, as followed by the risk-neutral, price-taking derivative trader is the solution to a finite-horizon multiple optimal stopping problem which we formulate and solve using a finite-difference method in Section 2.2. Specifically, we split the problem of trading the portfolio into a sequence of single optimal stopping problems, extending the results of Haggstrom (1967). At each point in time, the trader has to decide how many of the remaining baskets of options in the portfolio, if any, to trade immediately. Executing a trade will turn out to be optimal only when the instantaneous expected excess return on the basket is sufficiently positive.

The optimal timing strategy translates into a dynamic order submission strategy, consisting of a sequence of stop orders with a time-varying stop price, as opposed to a single limit order with a fixed price. When buying a straddle, a basket consisting of a long put and a long call with the same maturity and strike, a sequence of buy-stop orders should be submitted to the market with stop prices that are almost linearly decreasing over time.

In the Black-Scholes world, expected benefits ("gains") from timing the buying/selling of a straddle are economically meaningful – up to several tens of basis points with a one-day trading horizon – compared to either of three benchmark strategies: trading at the opening of the market, at the close of the market, or using a TWAP strategy, i.e., trading an equal fraction of the total position at regularly spaced intervals in time. For a delta-neutral straddle, which has a zero instantaneous expected excess return, the expected gains from trading optimally are unsurprisingly modest at about 2bps colorredof the opening straddle price per day using the default parameters. The cost of imposing the stop-loss when trading the straddle ranges between zero and 4.6bps per day, depending on moneyness, and amount to a bit over 1bp per day for the delta-neutral straddle. In the special case that the value of the portfolio is a monotonic function of the price of the underlying asset, for example when considering a single plain vanilla put or call option, Proposition 2.1 provides sufficient conditions to determine the optimal trading time analytically in the Black-Scholes world. Under the standard assumption that the equity premium is positive, it is optimal to buy a single call and the stock itself immediately, while buying a single put would only occur at the end of the allowed trading period.

In the empirical section, we employ the expected-return-maximizing trading strategy derived in the Black-Scholes world to purchase a near-maturity straddle on the S&P500 index every day using an extensive data set with high-frequency S&P500 option quotes between July 2000 and December 2012. Economically meaningful gains of up to 33bps per day for an at-the-money (ATM) straddle are obtained when compared to the three benchmark strategies. Surprisingly, adding the stop-loss level *increases* the average gain from the optimal timing strategy for an ATM straddle to 72bps (compared to trading at the close of the market), mainly because the stop-loss helps to avoid having to trade at the close on days with substantially negative market movements. By doing so, the stop-loss reduces the left tail of the gain distribution, so the average is raised while simultaneously the variance is reduced. This raises the question how a trader with alternative, risk-averse, preferences would value the two strategies relative to the benchmarks, even though the risk associated with delayed trade execution was not taken into account in the derivation of the optimal timing strategy. A CRRA trader with a coefficient of risk aversion equal to two would be willing to pay up to 32bps of the value of the straddle to switch from one of the benchmark strategies to the optimal timing strategy, and up to 51bps to switch to the optimal timing strategy that includes the stop-loss.

In both set ups, the average gain obtained when purchasing a straddle on the S&P500 index exceed the theoretical expected gain from the Black-Scholes model. We argue that because of i.i.d. returns and no market frictions, the expected gains in the Black-Scholes model are a conservative estimate of the average gains that can be obtained using market data. We conduct extensive simulations using stylized facts in the data to show how the individual assumptions affect the performance of our strategy. In particular, we show how the inclusion of the stop-loss level helps make the strategy robust against serial correlation in intraday index returns, and how priced variance risk mainly affects the performance of the strategy for straddles around the ATM level. We note that even though variance risk carries a negative risk premium (see, e.g., Bollerslev, Tauchen, and Zhou (2009)) and the straddle's variance beta is positive, it is not optimal to always delay purchasing the straddle until the end of the trading period. Like the equity risk beta in the Black-Scholes model, the variance beta will change as either the underlying asset price or the spot volatility changes. Expected future changes in the betas and their relative importance for the value of the straddle generate a trade-off between waiting and trading immediately. The simulations also show that gains are robust to intraday seasonality in trading costs and to potential synchronisation issues between the index and options market, although Muravyev, Pearson, and Broussard (2013) note that latencies, the time it takes to transfer information from one end of a communications network to another, for option markets are just as low as for equity markets nowadays.

Throughout, a portfolio is defined as an (ordered) collection of baskets, and a basket (or "option strategy") as a set of derivatives that have to be traded simultaneously for exogenous reasons. Option strategies can be constructed with ease from the standard traded plain vanilla put and call options available on derivative exchanges. Major exchanges like the Chicago Board Options Exchange (CBOE) allow investors to request a single quote for the whole strategy[1] and trade the basket in a single transaction, even though the actual holdings in the investor's portfolio will show the puts and calls separately once the trade is completed. The single transaction reduces execution risk and (fixed) transaction costs for the investor and facilitates trading in option series that may be illiquid when traded in isolation like far-out-of-the-money puts and calls. Trading in option strategies is an important part of the total market for (equity index) derivatives. Fahlenbrach and Sandås (2010) report that for a sample of the FTSE-100 index option market spanning 2001-2004, 37% of all trades are related to baskets. Those trades represent about 75% of total trading volume in the sample and a total premium revenue that is about three times bigger than that generated by trades in individual options. In the FTSE-100 sample, almost one in five basket trades involves a straddle or strangle[2], with most trading taking place close to the ATM level.

The order in which various baskets have to be traded is often the result of risk management considerations. As an example, a trader may be prohibited from entering into a naked short put position, but may be able to invest in either a bear put spread or add a short put position with a long put position in place already. Exchange margin rules may also favour a certain order of the baskets in the portfolio. For example, CBOE margin requirements for a short ATM put combined with a long out-of-the-money put are lower than for naked short positions[3]. Santa-Clara and Saretto (2009) show that margin require-

[1] See, e.g., CBOE Annual report 2004, page 8 and CBOE Annual report 2007, page 3 (`https://www.cboe.com/AboutCBOE/AnnualReportArchive/AnnualReport2007.pdf`).

[2] A strangle is very similar to a straddle, the only difference being that the put and call option in a strangle have different strike prices.

[3] See `http://www.cboe.com/micro/margin/introduction.aspx` for details.

ments affect option returns in an economically meaningful way, providing an incentive to investors to think about ordering their trades so as to lower margin impacts. Gârleanu and Pedersen (2011) present evidence that margin requirements affect equilibrium holdings and prices.

The starting point of our analysis is a given portfolio of (baskets of) derivative securities, i.e., we take the solution to the asset and security allocation problems as well as the order in which the various baskets have to be traded as given[4]. The option to time trade execution can be seen as a generalized version of a swing option, a derivative contract with multiple exercise rights, with a payoff function that is specific to each exercising time[5]. As discussed in Section 4.8, optimal execution of stock portfolios focuses purely on transaction costs and price impact[6].

The remainder of the paper is structured as follows. In Section 2, we formulate and provide a solution method for the trade execution timing problem of a given portfolio of derivatives. A simple solution exists for portfolios whose value is a monotonic function of the price of the underlying asset, and a dynamic programming algorithm can be used to find the optimal timing strategy for non-monotonic portfolios. In Section 3, we use the algorithm to analyse in depth the optimal timing of purchasing or selling a straddle in the Black-Scholes world, showing the value of being able to choose the trading time, and demonstrating how to formulate the corresponding order submission strategy. In Section 4, we use intraday data on both the S&P500 index and its associated options as traded on the CBOE to time the purchase of a near-maturity straddle. We attribute the difference between the realized gains for the S&P500 straddles and the expected gains under the Black-Scholes assumptions to three main factors: intraday seasonality in option bid-ask spreads, serially correlated intraday index returns and the variance risk premium/leverage effect. Section 5 concludes.

2 Timing value for single options and baskets

The starting point of our analysis is an investor/portfolio manager, seeking to hold a given portfolio (H) of derivative securities on a single underlying asset by a given time (T), for example the end of the day. The elements of H, (h_i), are *baskets* of securities to be traded simultaneously, indexed by i, $i = 1, \ldots, I$. A basket may contain a simple derivative like a plain vanilla call or put option, or a combination of them like a straddle (long put and call with same time to maturity and exercise price).

If the investor has risk aversion equal to the representative agent and executes trades themselves, the discounted derivative price process will be a martingale for the investor in utility terms and hence she will be indifferent between trading now or any time before the end of the trading period. There are

[4]Dert and Oldenkamp (2000), Liu and Pan (2003) and Papahristodoulou (2004) study asset allocation/portfolio choice models using derivatives.

[5]Carmona and Touzi (2008), Carmona and Dayanik (2008) and Benth (2011) among others study the pricing of these kinds of options.

[6]Chan and Lakonishok (1995) is an early example. Harris and Hasbrouck (1996) study the choice between market and limit orders.

at least two ways to deviate from this set up and motivate the analysis below. Either the investor has lower risk aversion than the representative agent or the investor delegates the execution of trades to a third party with a different risk aversion or utility function. We will focus on the latter interpretation in the remainder of the paper, although the results would be the same if the investor themselves were risk neutral.

The portfolio manager (principal) delegates the execution of the portfolio trades to a dedicated risk-neutral trader (agent), who is instructed to trade the baskets h_i in sequence, i.e. h_1 should be traded first, then h_2 and so on. Furthermore, all units of a basket should be traded at once, so order splitting is not allowed. The manager also tells the trader the characteristics of each basket. For example, if basket i consists of a single plain vanilla option, its characteristics are the expiration time $\hat{T}_i$, the strike price K_i, the exercise style (European or American), and whether it is a call or a put option. Throughout, we assume that the characteristics are set at the start of the trading period, either reflecting existing holdings that need to be unwound or being determined as the output of an asset allocation model run with the spot/forward price of the underlying asset at the start of the trading period as an input. The price of basket i, denoted by $f_i(S_t, t)$, is a function of these characteristics as well as calendar time t and the value of the underlying asset S_t.

A straightforward way to obtain the portfolio would be to trade all baskets at $t = 0$, the start of the trading period. This strategy carries no execution risk because all current prices are known, and therefore it constitutes a natural benchmark trading strategy. In this section, we seek to find the trading strategy that will maximize the expected performance relative to this benchmark strategy. Trading at the start of the period will turn out to be optimal only in very specific cases. The problem faced by the trader can be seen as finding the stopping/trading time τ_i for each basket $i = 1, \ldots, I$ that minimizes (maximizes) the expected purchase (sale) price of the total portfolio,

$$V = \inf_{0 \leq \tau_1 \leq \tau_2 \leq \ldots \leq \tau_I \leq T} E\left\{ \sum_{i=1}^{I} h_i \mathrm{e}^{-r\tau_i} f_i(S_{\tau_i}, \tau_i) \right\}, \tag{1}$$

where r is the risk-free rate, which we assume constant. In words, (1) says that at time τ_i, the trader trades h_i units of basket i and adds h_i times the discounted cost $\mathrm{e}^{-r\tau_i} f_i(S_{\tau_i}, \tau_i)$ to the total cost. The value function V in (1) can be interpreted as the amount of money the investor needs to set aside in expectation to obtain the desired portfolio exposure. Trading all securities at the start of the period amounts to choosing $\tau_1 = \tau_2 = \ldots = \tau_I = 0$. The difference between the two strategies is the current value of the "option to time the trade". We write "option" although it contains both a right and an obligation for its holder. For $t \in [0, T)$, the trader has the option to either trade or wait. At the same time, she is also obliged to trade any remaining baskets at the market price at time T.

Other than through combining subsequent baskets, there is no provision for risk aversion with respect to trading at uncertain future prices included in (1). Including risk aversion in a general sense, for example by specifying the agent's utility function, is hard because the non-linearity will introduce dependence on higher-order moments and path-dependence. As discussed in the introduction, as an alternative heuristic to account for risk aversion, we propose the inclusion of a stop-loss level, i.e., a maximum (minimum) price to pay (receive) for the total portfolio. The instant that price is observed in the market (either as the sum of current market prices before the first basket is traded, or as the sum of the prices paid/received for baskets traded previously and the currently observed market prices for baskets not yet traded), trading of all remaining baskets will be triggered. In a model where the value of the underlying asset is the only source of uncertainty, such a stop-loss level can be modelled as a (potentially time-varying) value of the underlying asset itself ($\bar{S}_\tau$). Denoting by τ_i the trading times obtained from (1), and defining τ_{SL} as the hitting time of the stop-loss level,

$$\tau_{SL} = \min\left\{0 \leq t \leq T : S_t = \bar{S}_t\right\}, \tag{2}$$

the trading times with the stop-loss added ($\tau_{i,SL}$) are given by

$$\tau_{i,SL} = \min\left\{\tau_i, \tau_{SL}\right\}, \qquad i = 1, \ldots, I. \tag{3}$$

Two remarks apply to the problem descriptions in (1) and (3). First, as stated in the introduction we assume that portfolio weights are given and think of them being the outcome of a portfolio optimization exercise. Note that the holdings can be negative which means selling/writing the options, in which case (1) maximizes the expected revenue from selling/writing.

Second, as stated above, the portfolio H does not only describe the holdings of the different baskets, but also the order in which the baskets should be traded. This order could be motivated by risk management arguments. Consider, for example, a put spread, consisting of a short position in one put and a long position in another put with the same maturity but a lower strike. Selling the put with the high strike before buying the put with the low strike entails a risk when the underlying asset moves by a large amount in between the sale and purchase. Imposing that the total position should be traded simultaneously reduces this risk. As documented in Fahlenbrach and Sandås (2010) for example, major option exchanges provide infrastructure to trade option baskets in one trade rather than a basket of trades exactly for this reason. In the remainder of this section we first consider the set of problems for which (1) a corner solution is optimal. Subsequently, we present a numerical algorithm to solve the more general problem.

2.1 Optimal trading of a single derivative

Suppose the trader wants to hold by time T the portfolio H consisting of a position h in a single derivative security. The price of the derivative equals $f(S_t, t)$, which we assume to be twice continuously differentiable with respect to S. In addition, assume the risk-free rate r to be constant and suppose the stock price follows an Itô process where the drift $\mu(S_t, t)S_t$ is continuous in both arguments, and where the volatility $\sigma(S_t, t)S_t$ satisfies the usual regularity conditions[7],

$$\mathrm{d}S_t = \mu(S_t, t)S_t\mathrm{d}t + \sigma(S_t, t)S_t\mathrm{d}W_t, \tag{4}$$

[7]See e.g. Theorem 4.5.3 in Kloeden and Platen (1999).

where W_t is a standard Brownian motion. Finally, assume no arbitrage in continuous time. Under these assumptions, Proposition 2.1 extends results in Kukush, Mishura, and Shevchenko (2006) (compare their Lemma 2.2). The optimal time to buy the security depends only on the sign of the product of three factors: the difference between the drift and the risk-free rate, the derivative of the security price with respect to the stock price and the sign of the position.

Proposition 2.1 *Suppose the agent wants to trade h units ($h > 0$ for a buy) of a single security with price function $f(S_t, t)$ on an arbitrage-free market where the dynamics of the stock is given by (4) and the risk-free rate equals r. Then, optimizing (1), the agent optimally trades the asset*

i) at $t = 0$ if $h\frac{\partial f(s,t)}{\partial s}(\mu(s,t) - r)s$ is positive for all s, t,

ii) at $t = T$ if $h\frac{\partial f(s,t)}{\partial s}(\mu(s,t) - r)s$ is negative for all s, t,

iii) The agent is indifferent about the trading time τ if $f(s,t)$ just depends on t or if $\mu(s,t) = r$ for all s, t.

The proof is in Appendix A. Proposition 2.1 essentially states that, for derivatives whose value is a monotonic function of the underlying asset, the timing problem (1) is equivalent to maximizing the exposure to equity risk of the derivative portfolio in a frictionless world.

In the remainder, we focus on a Black-Scholes world, i.e. $\mu(s,t) \equiv \mu$ and $\sigma(s,t) \equiv \sigma$ in (4). Corollary 2.1 describes the optimal trading decisions when the desired portfolio consists of a single call or put or a stock, being equivalent to a call option with zero strike price.

Corollary 2.1 *Suppose that the stock price S follows a geometric Brownian motion with drift $\mu > r$. Then it is optimal to*

i) buy a call option or sell a put option immediately,

ii) buy a put option or sell a call option at the end of the trading horizon.

Translated into an order submission strategy, Corollary 2.1 states that, for plain vanilla options, a market order at either the start or end of the period should be used, a result which differs from the general order submission strategy we discuss below. In the remainder of this paper, we assume a positive equity premium ($\mu > r$).

Proposition 2.1 and Corollary 2.1 both require the value of the derivative portfolio to be a monotonic function of the underlying asset price. Note that market frictions, like intraday seasonality patterns in bid-ask spreads as considered in Section 4.4, will generate an optimal trading time that is different from the corner solution for monotonic securities, independent of whether the agent is the representative agent. However, the point we want to stress in this paper, even without such frictions, for portfolios whose value is not a monotonic function of the underlying asset price the optimal time to trade will in general be different from either the start or end of the trading period. The next section makes explicit the trade-offs that underlie that optimal timing result.

2.2 Dynamic programming solution to the optimal trading of baskets

Suppose we want to trade a *portfolio* of (baskets of) options in a given order. Proposition 2.1 and Corollary 2.1 imply that, for any basket whose value is monotonic in the value of the underlying asset, the expected excess return on the basket until its maturity determines whether we trade at the start or end. A positive (negative) expected excess return leads to trading immediately (at the end). From Corollary 2.1, this idea can be used as long as the ordering of baskets in the portfolio is such that all call purchases and put sales are to be completed before any call sales or put purchases, so that the portfolio can be split into two baskets of options. The first basket will contain all call purchases/put sales and the second basket all call sales/put purchases. In this section we develop the intuition as to why this decision is no longer optimal when trading baskets whose value is not monotonic in the value of the underlying asset, for example a straddle.

We will study trading the straddle in detail in Sections 3 and 4 below, and use it here to illustrate the intuition behind the dynamic programming algorithm. The delta, the first derivative of the straddle with respect to the value of the underlying asset, is negative for low values of the underlying asset and positive for high values. Proposition 2.1 appears to suggest the following optimal strategy for purchasing such a straddle. As long as the delta is negative, delay the purchase. When delta changes sign, buy immediately. However, this logic is not correct as it ignores the value of having the flexibility of choosing the time to buy and the implications of surrendering that flexibility. Even when delta is slightly positive, it may still be optimal to wait because of the possibility that future movements in the underlying asset value make the delta negative again. This reasoning is similar to the exercise decision for real or American options: it is only optimal to exercise directly (trade immediately in our context) if the option is sufficiently in the money (if the delta of the basket is sufficiently positive).

As an alternative way to obtain the intuition for the optimal timing strategy in a general setting, use the stochastic differential equation in (4) to rewrite the agent's optimization problem (1) in terms of the expected return on the portfolio as follows[8]. Include in (1) either the price of the portfolio at the start of the trading period or it's discounted expected value at the end of the trading period as a normalisation. Doing so will change the level of the value function, but will not alter the optimal decision. Then, employ the Black-Scholes partial differential equation (PDE) in the same way as in the proof of Proposition 2.1 to obtain

$$\begin{aligned}\tilde{V}_1 &\equiv V - \sum_{i=1}^{I} h_i f_i(S_0, 0) \\ &= \inf_{0\le\tau_1\le\tau_2\le\ldots\le\tau_I\le T} E\left\{\sum_{i=1}^{I} h_i \int_0^{\tau_i} \left(\mu(S_u, u) - r\right) \mathrm{e}^{-ru} \frac{\partial f_i}{\partial \log S}(S_u, u) \mathrm{d}u\right\}, \end{aligned} \tag{5}$$

[8]Thanks to Leonidas Rompolis for suggesting this alternative interpretation.

or, equivalently,

$$
\begin{aligned}
\tilde{V}_2 &\equiv E\left\{\sum_{i=1}^{I} h_i \mathrm{e}^{-rT} f_i(S_T, T)\right\} - V \\
&= \sup_{0\leq\tau_1\leq\tau_2\leq\ldots\leq\tau_I\leq T} E\left\{\sum_{i=1}^{I} h_i \int_{\tau_i}^{T} (\mu(S_u,u)-r)\,\mathrm{e}^{-ru}\frac{\partial f_i}{\partial \log S}(S_u,u)\mathrm{d}u\right\}, \qquad (6)
\end{aligned}
$$

where the log-delta $\partial f_i/\partial \log S =: f_{i,\log S}$ is a positive transformation of the delta of the derivative $\partial f_i/\partial S =: f_{i,S}$. In words, (5) states that the optimal trading time minimizes the expected risk exposure of the portfolio *before* it is acquired, while (6) maximizes the expected risk exposure of the portfolio from the moment the position is taken until the end of the trading period. Proposition 2.1 as well as Equations (5) and (6) show that the optimal timing decision in the Black-Scholes world is driven by the exposure to the equity risk premium, the compensation for the single risk factor in that model. Importantly, and perhaps counter-intuitively, changes in the time value of the baskets does not influence the optimal timing strategy. Summarizing, for general baskets of options whose deltas may change sign, there is no a priori reason to conclude that always trading at a fixed time will be optimal, but such a strategy can still serve as a benchmark. The delta of the straddle changes from negative to positive around the ATM level.

We use a trinomial tree to solve the optimal stopping problem in (1) following Hull (2012). Rubinstein (2000) and Hull (2012) show the trinomial tree to be equivalent to an explicit finite difference method as used, for example, in Bender and Dokuchaev (2014). The tree has an evenly spaced grid in the time dimension with $N+1$ points, $0 = t_0 < t_1 < \ldots < t_N = T$. Let $\Delta := t_n - t_{n-1}$ denote the time-step. From one time point to the next, the underlying asset price S moves either up by a factor $u \equiv \exp(\sigma\sqrt{3\Delta})$, down by a factor $d \equiv 1/u$, or remains at the same level. The associated probabilities for the up and down move are

$$
p_{\text{up}} = \frac{1}{6} + \sqrt{\frac{\Delta}{12\sigma^2}}\left(\mu - \delta - \frac{\sigma^2}{2}\right), \quad p_{\text{down}} = \frac{1}{6} - \sqrt{\frac{\Delta}{12\sigma^2}}\left(\mu - \delta - \frac{\sigma^2}{2}\right). \qquad (7)
$$

At time step $n \in \{0, 1, \ldots, N\}$ in the tree, there are $2n+1$ nodes indexed by $j = 0, 1, \ldots, 2n$, with the top node corresponding to $j = 0$. In the last period, the option prices in each node are computed using a Black-Scholes formula[9]. Option prices at nodes earlier in time then follow by working backwards in the tree, using the risk-neutral probabilities obtained by replacing μ in (7) with the risk free rate r.

The discrete-time equivalent of problem (1) reads, with $\tau_i \in \Pi_\Delta \equiv \{t_n\}_{n=0,\ldots,N}$,

$$
\hat{V}(0,0,0) = \min_{\tau_1\leq\tau_2\leq\ldots\leq\tau_I} E\left\{\sum_{i=1}^{I} h_i \mathrm{e}^{-r\tau_i} f_i(S_{\tau_i}, \tau_i)\right\}, \qquad (8)
$$

with $\hat{V}(n,j,m)$ the value function of the discrete-time version of the dynamic program in (1), starting at time step n and node j, given that m baskets have been traded so far. We do not impose that basket i needs to be traded *strictly* after basket $i-1$ in (1). Therefore, at any node in the tree we have to

[9]Extending the tree up to the maturity of the option contract ($\hat{T}$) and working backwards to obtain option prices at time T yields identical results but is computationally more cumbersome.

determine the optimal number of remaining baskets to trade immediately given the number of baskets already traded.

Starting from an arbitrary node j at time step n with m baskets already traded, immediate trading of the next k baskets adds an amount $\sum_{i=m+1}^{m+k} h_i f_i(S(n,j),n)$ to the total cost (with $m+k \leq I$). By Bellman's principle of optimality, the value of waiting for the remaining $I-m-k$ baskets equals the expectation of the optimal policy starting next period, given that at that point $m+k$ baskets have been traded already. From node (n,j), the next period node $(n+1,j)$ will be reached with probability p_{up}, node $(n+1,j+2)$ with probability p_{down} and node $(n+1,j+1)$ with probability $1-p_{\text{up}}-p_{\text{down}}$, so the discounted expected value of waiting equals

$$E_{n,j,m+k}\left\{\hat{V}(n+1)\right\} = \Big[p_{\text{up}}\hat{V}(n+1,j,m+k) + p_{\text{down}}\hat{V}(n+1,j+2,m+k) \\ +(1-p_{\text{up}}-p_{\text{down}})\hat{V}(n+1,j+1,m+k)\Big]\,\mathrm{e}^{-r\Delta t}.$$

The algorithm now reads:

1. Start at $T=t_N$ $(n=N)$ and compute the value function of the dynamic programming problem at each node $j=0,\dots,2N$ for each $m=0,\dots,I$,

$$\hat{V}(N,j,m) = \sum_{i=m+1}^{I} h_i f_i(S(N,j),N), \quad j=0,\dots,2N.$$

2. Given $\hat{V}(n,j,m)$ for each j and m, go one step backwards in the tree to time step $n-1$. By Bellman's principle of optimality, the value function at node j when m baskets have already been traded before time step $n-1$ equals

$$\hat{V}(n-1,j,m) = \min_{0\leq k\leq I-m}\Bigg\{\Big[p_{\text{up}}\hat{V}(n,j,m+k) + p_{\text{down}}\hat{V}(n,j+2,m+k) \\ +(1-p_{\text{up}}-p_{\text{down}})\hat{V}(n,j+1,m+k)\Big]\,\mathrm{e}^{-r\Delta t} \\ +\sum_{i=m+1}^{m+k} h_i f_i(S(n,j),n-1)\Bigg\}, \quad j=0,\dots,2(n-1), \quad m=0,\dots,I. \tag{9}$$

3. Repeat Step 2., iterating backwards in time until $t_0=0$ $(n=0)$ and obtain $\hat{V}(0,0,0)$.

The value function $\hat{V}(0,0,0)$ equals the minimum expected discounted value of the portfolio acquired between time 0 and time T. For each node in the tree we determine the optimal decision. Therefore, at each point in time the algorithm also outputs a collection of underlying asset values for which trading at least one basket immediately is optimal (the stopping region). Depending on the number and type of baskets in the option portfolio, there can be multiple stopping regions at any given point in time. The trinomial tree approach is convenient in that we do not need to make assumptions about the shape of the continuation and stopping regions a priori. Hence, alternating stopping and continuation regions in the underlying asset value dimension can be dealt with.

The level of the underlying asset which leaves the agent indifferent between waiting at least one more period and trading immediately is called the stopping boundary. In time-homogeneous optimal stopping problems, such as the valuation of a perpetual American option, the stopping boundary is a constant value S^*. However, in (1) the optimal action to take for a given value of the underlying asset will generally depend on calendar time, so the stopping boundary will be a function of calendar time, S_t^*[10].

The additional complexity from trading I different baskets instead of 1 for a given ordering of the baskets is of order I^2. At each node we have to keep track of I possible values, each of which describes the optimal policy given a number of baskets already traded. For each of the I values, we have to compute at most I numbers to determine the optimal future policy. Hence, solving the optimal trading problem for a portfolio with a reasonably sized number of baskets is feasible.

To conclude this section, we note that our algorithm can be used to find the optimal trading strategy when various sequences of baskets are allowed, as long as the sequence will be fixed before the first basket is traded. This can be achieved by examining every allowed sequence and picking the one with the lowest value function at time zero. A more flexible approach, which we leave for future research, would be to allow switching between sequences after trading has commenced. In such a set up, it may be optimal to trade the first basket in one node but to trade only the second basket in the neighbouring node, introducing an additional path-dependency going forward.

3 Trading a Black-Scholes straddle

The remainder of the paper focuses on trading a straddle: a simple basket of options whose value is a non-monotonic function of the underlying asset value and which can be traded directly in various markets. In this section, we study in detail trading this option strategy in the Black-Scholes world, using the algorithm of Section 2.2. To assess the economic value of optimally timing the trades, we use parameter values in line with the S&P500 data used in Section 4. Specifically, the volatility of the underlying asset is set to $\sigma = 16\%$ annualized, the equity premium 5%, the risk-free rate to $r = 2.4\%$ and dividend rate equal to $\delta = 1.8\%$. The current price of the underlying asset is normalized to 1. The maturity ($\hat{T}$) is set equal to one month and the trading horizon (T) to one trading day. We consider a range of moneyness levels, defined as the ratio of strike price relative to forward price at the start of the trading period, between 0.98 and 1.05, with the focus mostly on the delta-neutral straddle having strike $K = \exp((r - \delta + \sigma^2/2)\hat{T})$.

To make the results in this section comparable to Section 4, we define a normalized *gain* from using the optimal timing strategy as compared to a fixed trading time benchmark strategy as the difference in purchase price, divided by the cost of trading at the start of the trading period (market open). For

[10] In calculations, we extend the tree backwards in time past the current time far enough to ensure that the dimensions of the tree are not binding.

example, the *gain* from trading using the optimal timing strategy relative to trading at the close equals

$$\text{gain} = \frac{\text{cost@close} - \text{cost@optimal timing}}{|\text{cost@open}|}. \tag{10}$$

The absolute value function ensures a correct sign when there is a cash inflow rather than outlay, for example when writing a straddle. In this section, "gain" refers to the expected gain, whereas in the empirical analysis of Section 4 it will refer to a time-series of realized gains. Trading at the open is a natural strategy when portfolio holdings are determined at the open, since this strategy entails no execution risk. The market close is a natural time to trade when the trading horizon equals a single trading day, although it implies exposure to substantial price uncertainty from the point of view of the start of the trading period.

3.1 Buying a straddle

The two main results of purchasing a straddle can be summarized as follows. First, expected gains from using the optimal strategy versus buying at the start or the end are economically meaningful with values up to 24 basis points (bps) for moneyness 1.05 (compared to trading at the open) or 26bps for moneyness 0.95 (compared to the close). Second, the optimal strategy corresponds to submitting a sequence of buy-stop orders[11].

The expected gain compared to the open (close) increases (decreases) with moneyness, as shown in Figure 1. For the short-term delta-neutral straddle, the expected gain from following the optimal timing strategy versus the open or close equals 2bps. As Figure 1 shows, adding the stop-loss reduces the expected gains, leading to negative expected gains for high moneyness straddles when compared to trading at the close. For a delta-neutral straddle, the gain is reduced to about 0.7bps when the stop-loss constraint is added.

Intuition about the pattern of gains as a function of moneyness as well as the level of gains can be obtained from (5) or (6). Gains are increasing in moneyness for the straddle when compared to trading at the open. For low moneyness levels, the straddle delta is positive at the start of the trading period, and, since the expected return on the underlying asset is positive, is expected to increase further over time. For those straddles, there is only a small probability of the delta turning negative before the end of the trading period. Hence, the expected return until the end of the trading period is maximized by trading immediately, obtaining a zero gain when benchmarked against trading at the open and a high expected gain when benchmarked against trading at the close. For moneyness levels above the delta-neutral moneyness, the straddle delta is negative initially, which makes it attractive to delay trading. For moneyness levels close to delta-neutral, neither strategy that trades at a fixed time will maximize the expected return. Either the initial delta is negative, or, if it's positive, the probability of it turning

[11] A stop or stop-loss order is an order to buy or sell a security via a market order once its price reaches a specified price, the stop price. When the stop price is reached, the stop order becomes a market order (`http://www.sec.gov/investor/alerts/trading101basics.pdf`).

negative before the end of the trading horizon is too high, both of which give rise to value of waiting. Turning to the level of gains: for the delta-neutral straddle its log-delta, $f_{\ln S}(S_t, t)$, is zero at $t = 0$ and will be close to zero in expectation for the remainder of the trading period, leading to relatively low expected gains. In Section 4 we argue that the expected gains obtained here are a (very) conservative estimate of what can be achieved when buying a straddle using data on the S&P500 index, for which the Black-Scholes assumptions of i.i.d. returns and no market frictions do not hold.

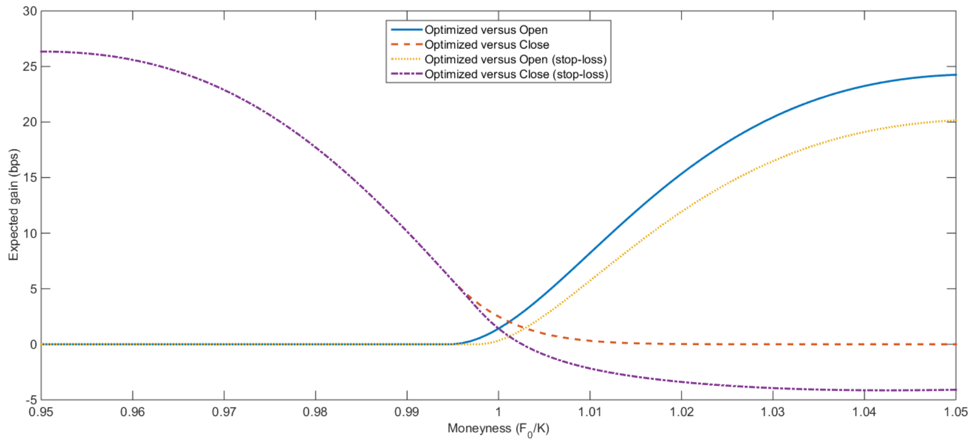


Figure 1: Expected gain of buying straddle using optimal timing strategy (with and without stop-loss), as function of moneyness (forward price at start of trading day/strike) in Black-Scholes world, compared to trading at start or end of trading day. Parameter values as detailed in Section 3.

Turning to the order submission strategy, a buy is triggered when the delta becomes sufficiently positive. This is illustrated in Figure 2(a)[12]. Starting from the current price of the underlying (log-return = 0), the straddle is bought as soon as the right-hand side stopping boundary in Figure 2(a) is hit, which can be implemented as a buy-stop order. Since the straddle price at which trading is triggered is not fixed over time, a sequence of buy-stop orders is needed with stop prices that decrease almost linearly over time. This is in contrast to the results for single options for which Corollary 2.1 established the optimality of a market order at either the start or end of the trading period.

This stopping region takes the form $\{S : S > S_t^*\}$. This means that as long as the cumulative return on the underlying asset since the start of the trading period is negative, it will always be optimal to delay trading even though the straddle price will increase when the underlying asset price decreases. In such cases, one can imagine a trader deciding not to wait but to cut their loss by overruling the strategy and buying the straddle at market price. Such behaviour can be captured by considering the version of the optimal timing problem that includes a stop-loss level $\underline{S} < S(0)$ which triggers immediate purchase of the straddle when hit. Like the optimal stopping boundary itself, this stop-loss level can be implemented as a sequence of buy-stop orders, with the stop price illustrated by the left-hand boundary in Figure 2(a).

Note that such an additional stop-loss adds an (ex-post) constraint to optimization problem (1), so

[12] In both this section as well as Section 4, the time of day refers to New York time.

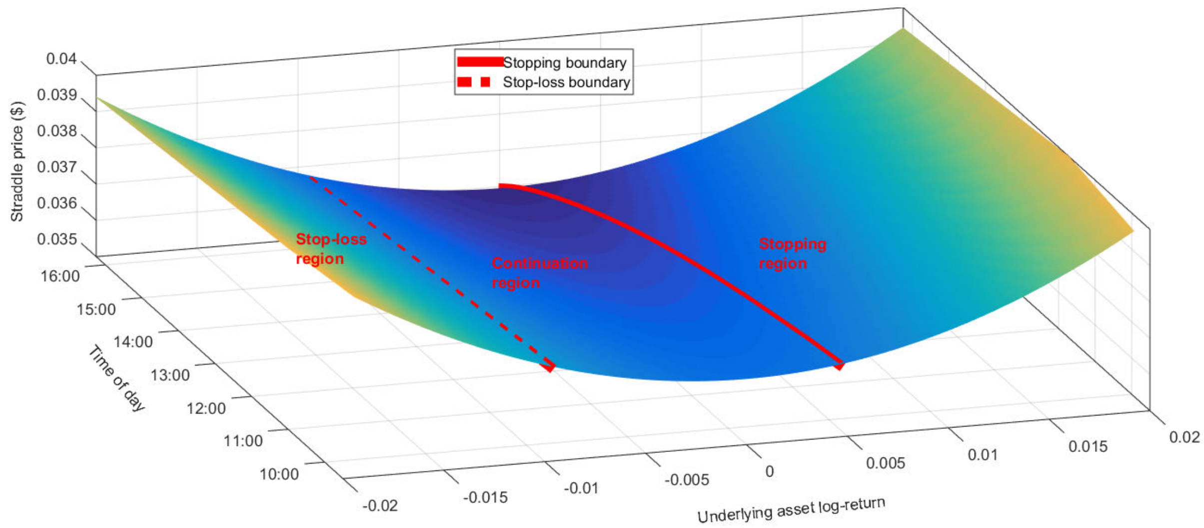
the expected gain of the optimal timing strategy with the stop-loss will be smaller than the expected gain of the optimal timing strategy without stop-loss. As we will argue in Section 4, the motivation to include the stop-loss is that it makes our results more robust against violations of the assumptions of i.i.d. returns and frictionless markets.

In the tree, the expected excess return on the underlying asset is positive, so for a positive delta the straddle value is expected to go up which would suggest that trading immediately would be optimal. Yet, for small delta values the expected increase is more than offset by the benefits derived from keeping alive the option to trade at a later point in time. Figure 2(b) illustrates how it is optimal to wait purchasing the zero-delta straddle by plotting the value of the straddle delta at the optimal stopping boundary. Specifically, close to the start of the trading period it is only optimal to trade the straddle when the delta is about 0.11, corresponding to a value of the underlying asset of about 1.007, i.e. 70bps above the (normalized) value at the start of the trading period, 1. As the end of the trading horizon draws closer the value of the option to delay trading decreases and it becomes optimal to trade for values of delta closer to zero (i.e. lower prices of the underlying asset), as indicated by the decreasing pattern in Figure 2(b).

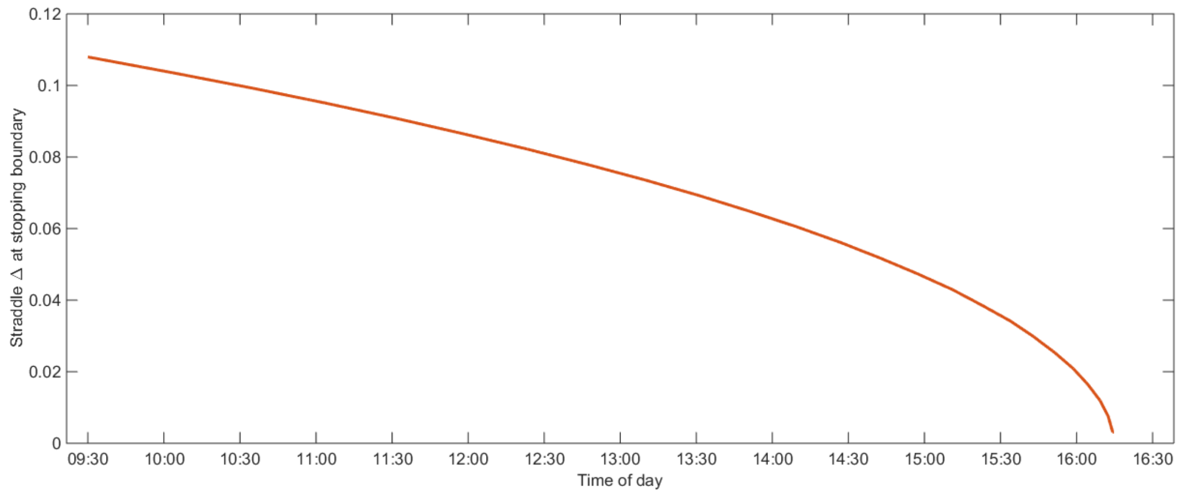
Finally, Figures 2(c) and (d) zoom in on the problem at the start of the trading period. The current price of the straddle, and its delta, as a function of the underlying asset price, are plotted in 2(c). As shown in Figure 2(d), the solution to (1) satisfies the usual value-matching and smooth pasting conditions of free boundary problems where the payoff is a sufficiently smooth function, see Dixit and Pindyck (1994).

3.2 Comparative statics

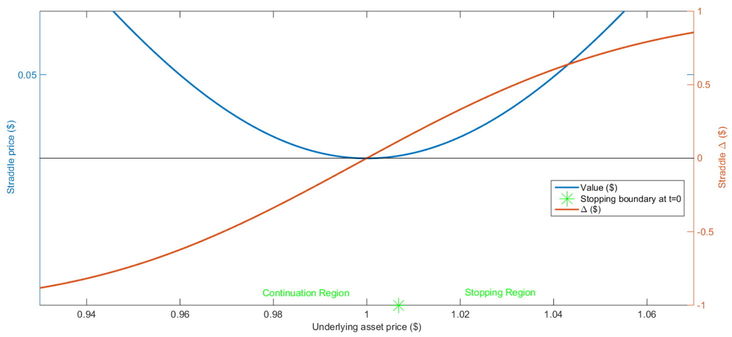
We briefly examine the effect that changing the value of key input parameters moneyness, risk-free rate, and volatility has on the optimal timing decision and expected gain. We once again concentrate on buying the straddle, but include in Figure 3 already the results for selling a straddle that is the subject of the next section. As the moneyness level changes, the stopping boundary shifts almost in a parallel fashion as shown in Figure 3(a). The shift is not completely parallel, which can be seen from (5) or (6), realising that the straddle delta as a function of the underlying asset value will become sharper as time passes (at expiration, the delta will be either -1, 0 or 1). It is optimal to buy the straddle with moneyness 0.99 at the start of the period, as implied by the stopping boundary lying below zero. Although not presented here, we note that the delta at the optimal stopping boundary is independent of the moneyness level and that the length of the time step in the tree has no influence on the expected gain. As shown in Figure 3(c), an increase in the risk-free rate makes trading at a later instant more attractive because the effect of discounting will be larger. Nevertheless, changing the interest rate has hardly any effect on the expected gains compared to trading at the open or close. This is driven by the fact that we keep the ratio of strike price over the initial forward price constant at one, so a change in interest rate implies a



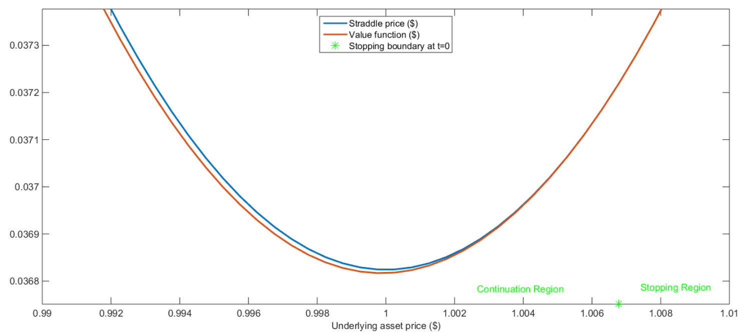
(a) Optimal stopping and stop-loss boundaries



(b) Optimal stopping boundary in terms of straddle delta

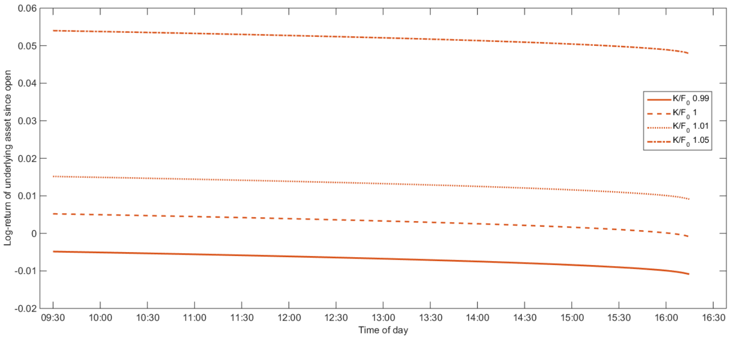


(c) Straddle value ($) and delta at $t = 0$

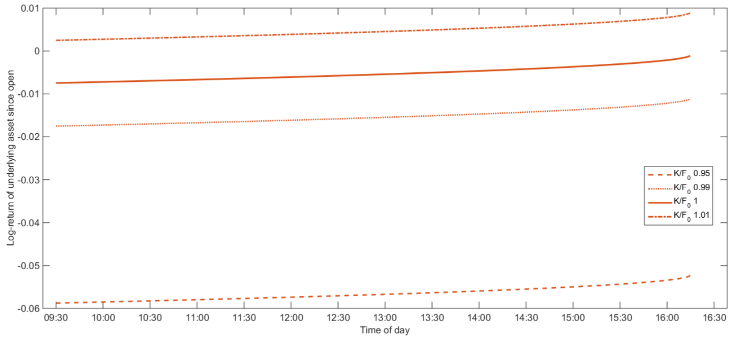


(d) Straddle value ($) and value function at $t = 0$

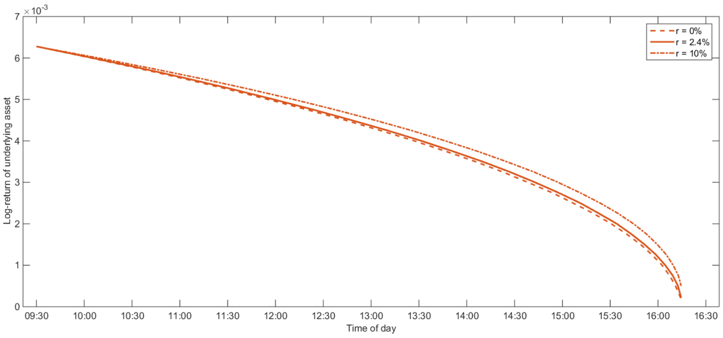
Figure 2: Purchasing a delta-neutral straddle in the trinomial tree ($K = \exp((r - \delta + \sigma^2/2)T)$). Default parameter values: $\mu = 5\% + r$, $r = 2.4\%$, dividend rate 1.8%, $\sigma = 16\%$, time step 15 sec, trading horizon 1 day, option maturity 1 month, initial value of underlying asset 1. The stop-loss level in (a) is set to -1% log-return from the start of the trading period.



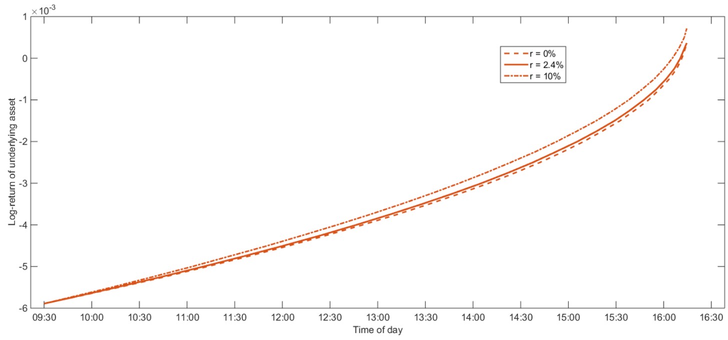
(a) Buy, moneyness K/F_0



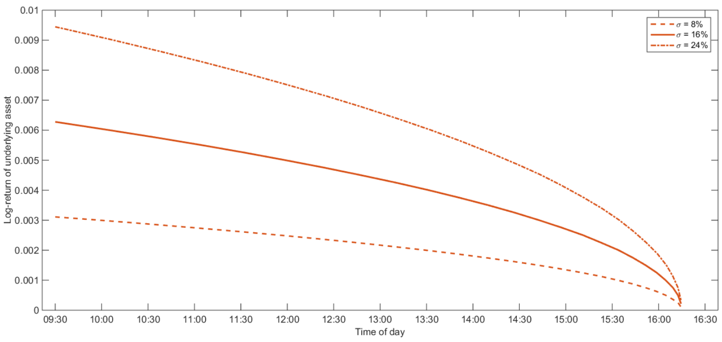
(b) Sell, moneyness K/F_0



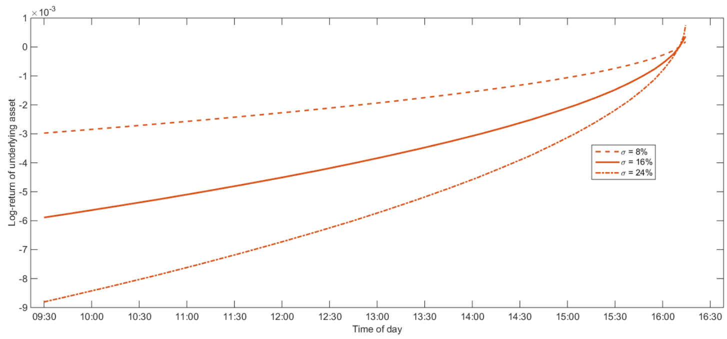
(c) Buy, risk-free rate r



(d) Sell, risk-free rate r



(e) Buy, volatility σ



(f) Sell, volatility σ

Figure 3: Optimal stopping boundary (S_t^*) for buying (left panels) and selling (right panels) a straddle. Default parameter values: moneyness $K/F_0 = 1$, $\mu = 5\% + r$, $r = 2.4\%$, dividend rate 1.8%, $\sigma = 16\%$, time step 15 sec, trading horizon 1 day, option maturity 1 month, initial value of underlying asset 1.

different dollar strike price. If we would keep the strike price fixed instead, the expected gain relative to the open is inversely related to the interest rate with a zero interest rate yielding an expected gain of 3bps. Compared to the close, the expected gain is increasing with an increasing interest rate, yielding 6bps with an interest rate of 10%.

If the volatility of the underlying asset approaches zero, problem (1) becomes deterministic. Assuming that the equity premium is positive, (5) shows that trading is optimal when the (log-)delta becomes positive for the first time, making it a special case for which the logic laid out at the start of Section 2.2 holds. As the volatility increases, the risk associated with delaying trade execution will decrease because the straddle delta as a function of the underlying asset flattens with increasing volatility. Combined with the favourable effect of discounting over a longer period of time, keeping the option to trade alive becomes more attractive. The increased value of waiting shows up as an outward shift of the stopping boundary as shown in Figure 3(e). This result is in line with observations in Criscuolo and Waelbroeck (2014), who study optimal stock portfolio liquidation when the stock price dynamics follows the Heston stochastic volatility model. They note that high volatility makes waiting more interesting, but also increases the execution risk, i.e., uncertainty about future prices. The expected gain for different volatility levels is a combination of the expected gain in the deterministic problem and the value of waiting. Keeping the straddle delta-neutral (which again implies changing the dollar strike price as the volatility changes), halving the volatility from 16% to 8% annualized about doubles the expected gain to 4bps relative to trading at the open or close.

As can be seen from (5), plugging in the Black-Scholes assumption ($\mu(s,t) \equiv \mu$), the first-order effect of an increase in the drift μ is to drive up the expected gain from timing the trade. Since the expectation in (5) is taken over the real-world probability measure, there is a second-order effect as well. A higher μ implies higher values of delta are more likely, and for a given value of delta the probability of it turning negative in the future will be lower. Hence, a higher μ makes trading earlier more attractive so the stopping boundary will be pushed down when μ increases.

Summarizing, although dependent on moneyness, the optimal time to buy a straddle is generally strictly between the begin and the end of the trading horizon. The value of the trade option when purchasing a straddle is determined by the trade-off for positive delta values between trading immediately which allows capturing the equity risk premium, and waiting which allows the trader to avoid the consequences of any future negative delta values.

3.3 Selling a straddle

In (1), selling a straddle is equivalent to putting $h_i = -1$ for the put and call, yielding a globally concave payoff function. Alternatively, keep $h_i = 1$ and replace inf by sup in (1). Both formulations illustrate the key message of this section: optimal selling is not just the mirror image of optimal buying, as displayed in Figure 3(b). As shown in Figures 3(d) and 3(f), changing the risk-free rate or volatility causes the stopping boundaries to shift in a different way when selling the straddle compared to buying it.

The intuition is as follows. In contrast to the case of buying a straddle, a higher risk-free rate hurts when selling a straddle because it will increase the discounting of the option premium to be received at some point in the future. Since the discounting is larger the further in the future trading takes place, the stopping boundary is pushed towards zero rather than away from zero as in the case of buying, compare Figure 3(d) to Figure 3(c).

Similarly, the delta of the short straddle is negative for sufficiently high values of the underlying asset. Note that the price of the straddle is negative in this case, representing the premium received when selling it. Hence, a negative delta implies that a further increase in the underlying asset price increases the premium income received from selling the straddle. Combined with the positive expected return on the underlying asset, the trader optimally defers trading when the current delta is negative.

Trading immediately is only optimal when the stock price drops far enough below the zero-delta stock price, where the short straddle delta is positive and the possibility of a sufficiently positive return on the underlying asset to make the delta change sign again is outweighed by the risk of the delta staying positive or becoming even more positive due to further declines in the underlying asset price. When volatility is high, the straddle delta is a relatively flat function of the underlying asset value, so the potential change in delta is low, which means waiting does not carry a lot of risk. At the same time, waiting means increased discounting of any premium income received from selling the straddle, so volatility and risk-free rate work in opposite directions when selling a straddle.

To conclude, we note that although the trade-offs and, therefore, decisions are different when selling a straddle compared to buying a straddle, the magnitudes of gains are quite similar with the pattern in the moneyness dimension generally reversed. For the delta-neutral straddle using the default parameter values, the expected gain is 2bps compared to trading at either the open or close. The expected gain compared to the close increases with moneyness to obtain an expected gain of 15bps at moneyness 1.02. Compared to the open, the expected gain decreases with moneyness; for moneyness 0.98, an expected gain of 18bps can be obtained in the tree. Similar to the case of buying, changing the risk-free rate but keeping the moneyness fixed has a negligible affect on gains. Keeping the strike fixed instead leads to gains that are increasing with moneyness compared to the open and decreasing compared to the close. Only for the volatility is the pattern the same: a lower volatility means a higher expected gain also when selling the delta-neutral straddle.

4 Buying an S&P500 straddle

Section 3 documents that economically meaningful expected gains are possible when optimally timing trading a straddle in the Black-Scholes world. The question then becomes how the optimal timing strategy derived in the Black-Scholes world performs when applied to market data, where the Black-Scholes assumptions do not hold. In this section, we study the actual gains from using the stopping boundary as derived in Section 2 'as is'. We show that these actual gains are often even larger than those found in Section 3; in Sections 4.4-4.6 we point out that such additional gains are the natural result of well-known deviations from Black-Scholes assumptions, in particular time-varying bid-ask spreads in option markets, serial correlation in high-frequency index returns and the presence of volatility as a priced risk factor.

4.1 Data

We implement the optimal timing trading strategy outlined above to buying a near-maturity straddle on the S&P500 index every day when the CBOE is open[13]. The options quotes and S&P500 index values at the one second frequency are from Thomson Reuters Tick History (TRTH), which started recording S&P500 index option prices on July 7, 2000. Our sample runs until December 31, 2012, so it contains both the 2001 and 2008 market downturns. On net, the S&P500 index value did not change much during our sample period, opening at 1456 on July 7, 2000 and closing at 1426 on December 31, 2012. While the TRTH coverage is generally quite good, all quotes are missing from October 2001 till the end of January 2002, as well as for a large part of February 2002. This leaves 123 days in 2000, 163 in 2001, 213 in 2002 and 250-252 in other years, for a total of 3011 days on which option quotes are recorded.

The S&P500 index options traded on the CBOE are European-style, cash-settled contracts. The expiration date is the Saturday following the third Friday of the month. The options are AM-settled which means that the expiration value of the index is determined by its opening value on the third Friday of the month. The S&P500 index options market is one of the most liquidly traded equity options market, with the CBOE reporting a daily volume of about 700,000 contracts in 2012[14], with each contract representing a notional value of \$100 times the index level[15].

The TRTH data contains both trades and quotes at the one-second level, but for our purposes we retain the quotes only. The following filters apply. Option quotes are only retained if they have

1. both a bid and ask quote,
2. a bid-ask spread that is non-negative,
3. a time-stamp within the regular CBOE trading hours (8:30AM-3:15PM Chicago time, 9:30AM-4:15PM New York time as used in all figures) and after the first observed quote for the underlying S&P500 index itself on that day,
4. satisfy a simple arbitrage bound for calls (and similarly for puts), with δ the index dividend yield,

$$C(K, S, r, T-t, \sigma, \delta) \geq \max\left(S\mathrm{e}^{-\delta(T-t)} - K\mathrm{e}^{-r(T-t)}, 0\right),$$

[13]The CBOE is the exclusive exchange for S&P500 index options, see CBOE's Annual Report 2013, p1.
[14]`http://www.cboe.com/micro/spx/introduction.aspx`
[15]`http://www.cboe.com/products/indexopts/spx_spec.aspx`

5. are arbitrage-free in the strike dimension,

$$C_{bid}(K_1) - C_{ask}(K_2) \leq K_2 - K_1 \quad \forall K_2 > K_1,$$

and similarly for puts.

In order to increase the comparability of strategies between trading days and to reduce expiration effects, we focus on short-term options, being those with the shortest time-to-maturity with a minimum of two weeks. The average time-to-maturity of the options in the sample equals 28 calendar days. Market makers on the CBOE are required to provide continuous markets and firm quotes which can be traded against for a reasonable size. This makes buying at the ask and selling at the bid a feasible strategy. We therefore interpret observed quotes as remaining valid until a new quote is observed.

The current tick size for S&P500 index call and put options trading above \$3.00 is 10 cents, so for a straddle the spread will be 20 cents minimum, when computed as the sum of the call and put spreads. This sum is a conservative estimate of the actual straddle spread, since market makers facing a quote request for a straddle rather than a single option are free to quote a tighter spread on the combination order as long as they stick to the minimum price increment rules. However, we only observe spreads on plain vanilla put and call options in the data, so in the remainder we will infer the straddle spread and price as the sum of the quotes observed for the legs. The first percentile of the distribution of dollar spreads equals \$1.40, with the median spread being \$3.50, so the 10 cent tick size is only occasionally binding. We note that the 2008 financial crisis had a profound effect on S&P500 index option spreads. For an ATM straddle, the minimum spread of 20 cents was observed each year from 2002 to 2007. In 2008 however, the minimum spread for the straddle was \$0.70, and in the remainder of our sample a spread of 20 cents was not observed anymore.

The number of pairs of puts and calls having the same strike varies considerably over time. The median number of pairs per day equals 35, but it ranges between 7 and 56, with more strike pairs available towards the end of the sample. Figure 4 plots the time series. The S&P500 index dividend yield and the risk-free rate are obtained from OptionMetrics. The continuously compounded zero coupon rates are linearly interpolated to the maturity of the options. Together with the index value at the time the first option quotes of the day are observed, these inputs are used to determine the index forward price with the same the maturity as the option contracts. We assume a positive equity premium of five percent annually, noting that in Section 2.1 we established that the magnitude of the equity premium has only a second order effect on the optimal timing decision.

Finally, we use realized volatility data downloaded from the Oxford-Man realized library, see Heber, Lunde, Shephard, and Sheppard (2009). This realized volatility is computed during the time the market is open. In theory, the variance during the time the market is closed should be added to obtain the total variance. For the S&P500 index however, about 99% of the total variance is realized during active trading hours so we choose to use only that part.

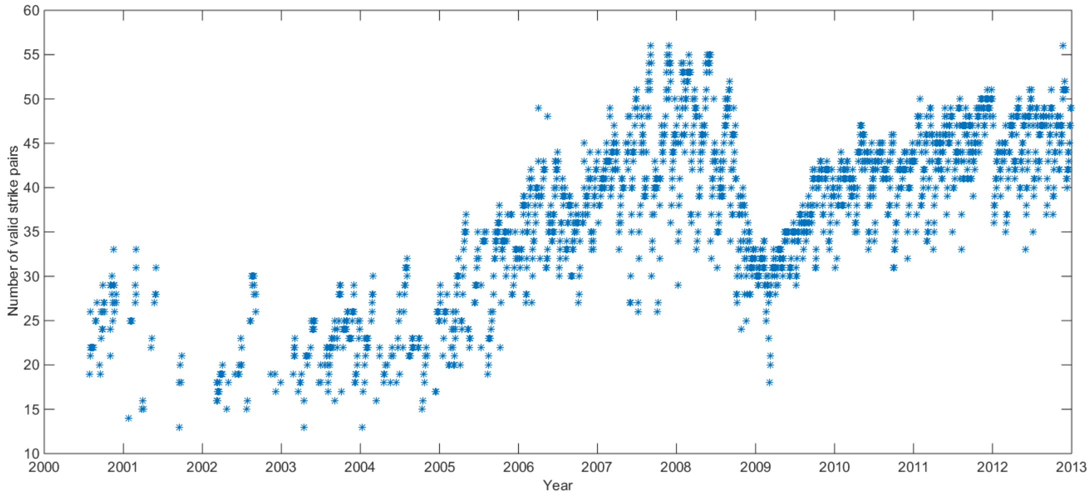


Figure 4: Number of strike pairs available each day for straddle trading in the S&P500 option data.

4.2 Empirical implementation

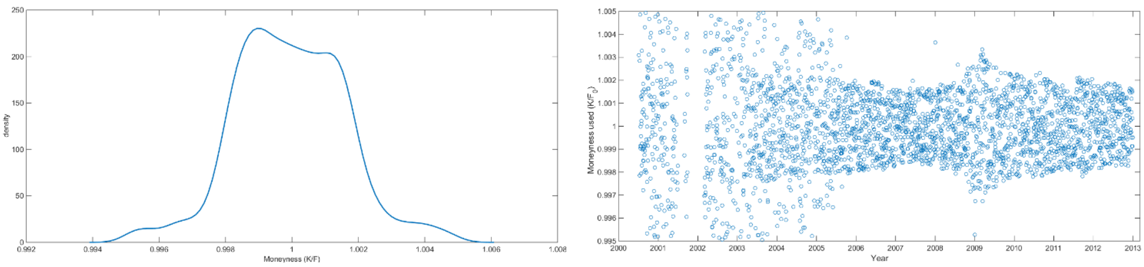
An important input into the option pricing model is the volatility parameter. An extensive literature on volatility has demonstrated that volatility is time-varying, with a high persistence. At the start of the trading period, we need a prediction for the volatility of the underlying asset over the remaining life of the options. Using the Corsi (2009) HAR-RV regression model, we capture this randomness by generating volatility forecasts based on historical realized volatilities from Heber, Lunde, Shephard, and Sheppard (2009). The parsimonious HAR-RV model, which uses historical volatilities averaged over different frequencies to capture both short-term and long-term movements, has been shown to perform well, see e.g. Bollerslev and Todorov (2011) and Shephard and Sheppard (2010). We use daily, weekly (last 5 days), and monthly (last 22 days) realized volatilities as regressors. We generate a volatility forecast for an horizon of h days as follows. On day t, we estimate the HAR-RV volatility model using five years (1250 observations) of historical data up to date $t-1$. The estimated coefficients are used to generate a prediction for the volatility over the period $[t, t+h)$ as the average of the daily forecasts on days $t, t+1, \ldots, t+h-1$. The volatility forecast generated this way has an average annualized value of 15.9% with a standard deviation of 8.5% in the sample, ranging between 7% and 141% (during the 2008 financial crisis).

Using the optimal timing strategy, we buy straddles with different moneyness levels ranging from 0.98 to 1.05, in increments of at least 0.01. Each day and for each moneyness level m, the strike (K_m) is selected from the set of available strikes $\mathcal{K}$ in the data as the closest strike to the product of moneyness and forward price at the market open (F_0) under the condition that the absolute difference does not exceed 50 basis points,

$$K_m = \text{argmin}_{k \in \mathcal{K}} \left(\frac{k}{F_0} - m \right)^2 \text{ s.t. } \left| \frac{k}{F_0} - m \right| \leq 0.0050. \tag{11}$$

The combination of a 100bps increment in moneyness and a 50bps limit on the absolute deviation ensures that any given strike in the data will be associated with at most one moneyness levels. The trade-off is that for some days, an appropriate strike may not be available, which happens mainly in the earlier years of the sample when the strike grid is coarser. In the tables below, for each moneyness level the number of available days is indicated in the column "Days".

The average actual moneyness in the sample equals the desired moneyness. For the ATM straddle, Figure 5 displays the dispersion around that average in two ways. The kernel estimate of the density in Figure 5(a) is almost symmetric with the probability mass concentrated between actual moneyness 0.998 and 1.002. Figure 5(b) shows that the distribution has become more concentrated over time as the S&P500 index options market grew more liquid and the strike grid became denser.



(a) Kernel density estimate (b) Time series

Figure 5: Moneyness level (strike/forward price at market open) as used for the ATM straddle on the S&P500 index.

For each day and moneyness level, we calculate the appropriate stopping region as the output of the dynamic programming algorithm of Section 2.2, using the actual moneyness level of the strike as input and a time step equal to 15 seconds. From the time series of index values we find the first time the index value hits the stopping region or, following the discussion in Section 3, the stop-loss level which is implemented at a -1% log-return. The -1% level corresponds to a one-standard deviation decrease on a daily basis using the average volatility. Denoting the optimal stopping boundary at time t by S_t^*, the stopping time is thus given by

$$\tau^* = \min_{0 \leq \tau \leq T} \left\{ S_\tau \geq S_\tau^* \vee S_\tau \leq S_0 \mathrm{e}^{-0.01} \right\}. \tag{12}$$

The trade is assumed to take place at that time against the prevailing bid quote for a sale or ask quote for a purchase. If the stopping region is not hit before the end of the trading day, then we trade against the closing quotes, determined by the final option quote observed before 3:15PM Chicago local time (4:15PM New York time as used in the figures).

In contrast to the tree, the opening price in dollars of an S&P500 straddle of a given moneyness varies considerably over time. For example, the opening ask price of a straddle with a moneyness (strike/forward price at market open) equal to 1, ranges from \$20.40 to \$184.80, which is due both to changing volatility as well as a changing S&P500 level[16].

[16]The opening price of the straddle is lowest on 29/06/2005 (strike price \$1205) with an implied volatility based on the

For each day t and moneyness level i, we compute the *gain* or *cost saving* from trading using the optimal timing strategy relative to trading using the benchmark strategy following (10), and report the time series average (along with its t-statistic) for each moneyness level in the second column of Table 1. The default analysis takes a conservative approach by buying at the prevailing ask quote. In Section 4.4 we also consider trading at the average of the bid and ask quotes to gauge the effect of intraday seasonality in option bid-ask spreads. In each panel of Table 1, we report the fraction of days on which the two competing strategies coincide and the fraction on which the optimal timing strategy outperforms the alternative in the fourth and fifth column of Table 1.

In addition to trading at the open and the close, we consider as third benchmark strategy a TWAP schedule where an equal part of the straddle is purchased every 15 minutes between 9.45AM and 4PM New York time. This strategy is introduced to reduce the influence of the specific sample that we have at our disposal and can be seen as a buy-side agency algorithm as defined in Hasbrouck and Saar (2013). Such algorithms aim to minimize the execution cost and are heavily used in practice. A survey on 750 institutional investors in Trade Magazine (2012) reports that the majority of them use either VWAP or TWAP-based algorithms to execute their (equity) trades. Brugler (2014) documents clock-time periodicity in the form of regular spikes in trading activity for FTSE100 stocks during 2012 at one and five-minute intervals, and attributes those to the use of Value-Weighted Average Price (VWAP) and Time-Weighted Average Price (TWAP) algorithms being used by investors to execute their trades, in line with results in Schied, Schöneborn, and Tehranchi (2010) that deterministic liquidation strategies for stock portfolios are optimal for mean-variance and exponential utility investors. Hasbrouck and Saar (2013) and Easley, de Prado, and O'Hara (2012) report similar spikes for trading in NASDAQ stocks and S&P500 E-mini futures, respectively. We do not take any fixed trading costs of the trading strategies into account. Hence, insofar as fees on a per-trade basis exist, the total gains of the TWAP strategy are overestimated relative to the alternative strategies that use only a single time point to trade.

4.3 Purchasing using the optimal timing strategy with stop-loss

Table 1 presents our main empirical result: using the trading strategy from Section 3 with a stop-loss yields consistent and economically meaningful cost savings relative to all benchmark strategies for straddles that are close to ATM over a long sample period. Average gains up to 72.6 basis points per day can be obtained for the ATM (moneyness equal to 1) straddle, being the most interesting security for multiple reasons. Firstly, this straddle can be bought on the largest number of days in the sample, which is consistent with the idea that the strike grid is generally finest around the ATM level. For the TWAP strategy, we require quotes to be available on all 15-minute intervals between 9.45AM and 4PM, so this strategy cannot be implemented on days when the market opens late or closes early. Examples

opening midquote of 8.3%, and an opening spread of $2.00. The opening price is highest on 24/10/2008 (strike price $860) when the midquote-based implied volatility reaches 62.2% and the opening spread equals $80, which gradually decreases to about $20 later in the day.

of such days are the Friday following Thanksgiving, and Boxing Day. The open and close strategies automatically adjust in that case to the actual opening/closing time of the exchange.

Secondly, recall that for the parameter settings of Section 3.1 it is not optimal to buy the ATM straddle at the start of the trading period. In the data, the optimal timing strategy buys at the open on 3% of all days, at the close on 13% of all days, at an intermediate time by hitting the stop-loss on 20% of all days and at an intermediate time by hitting the stopping boundary on the remaining 64% of all days.

There are two reasons why the optimal timing strategy sometimes buys at the open in the data. Firstly, the moneyness of the options in the data is not exactly equal to the desired moneyness, as shown in Figure 5. Secondly, moneyness is defined relative to the forward index price at the start of the trading period, while the strike used in the dynamic programming algorithm is defined relative to the spot index value at the start of the trading period. The difference between spot and forward prices is driven by time-variation in dividend yields, interest rates and the time-to-maturity of the forward contract, see Figure 6. A high forward factor happens when the interest rate is high and the dividend yield low, as in the early 2000s. On days where the forward factor is below one, it may be optimal to trade the ATM straddle at the opening of the market.

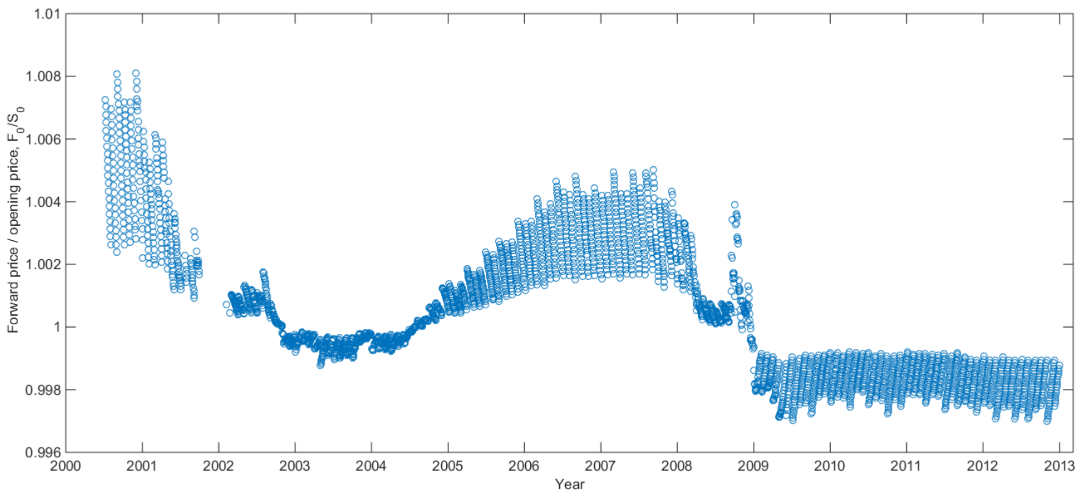


Figure 6: S&P500 index forward price at the market open as a fraction of the spot index value over time. Sample period July 2000 to December 2012. Dividend and zero-coupon yield data are from OptionMetrics.

Section 3 also established that straddles with a low (high) moneyness should optimally be bought at the open (close), which is partially confirmed by columns four and five in Table 1. Buying at the open indeed outperforms buying at the close as can be seen from panel (b) of Table 1. Panel (c) of that same table however suggests that buying the straddle following a TWAP strategy leads to a lower price than buying at the open. In Section 4.4, we will show that this difference is largely driven by intraday seasonality in bid-ask spreads that is not accounted for in the Black-Scholes world of Section 3.

While the optimal timing strategy with stop-loss outperforms its competitors on average, the time series of cost savings is skewed and fat-tailed, with profits and losses of up to several tens of percentage points on some days. The skewness for the gains of the ATM straddle compared to the open (close) equals -0.4 (4.6) and the kurtosis 11.6 (66.7). We note that the inference about the average remains valid even with the non-normality of the raw distribution.

Trading using the TWAP strategy outperforms the optimal timing strategy for the low moneyness straddles. This is largely driven by the optimal timing strategy buying directly at the opening of the market, when spreads are generally higher than during the rest of the day, see Figure 8 in Section 4.4. Figure 7 shows that for moneyness levels equal to or slightly larger than one, the outperformance of the optimal timing strategy with stop-loss over the TWAP strategy is consistent over the trading day.

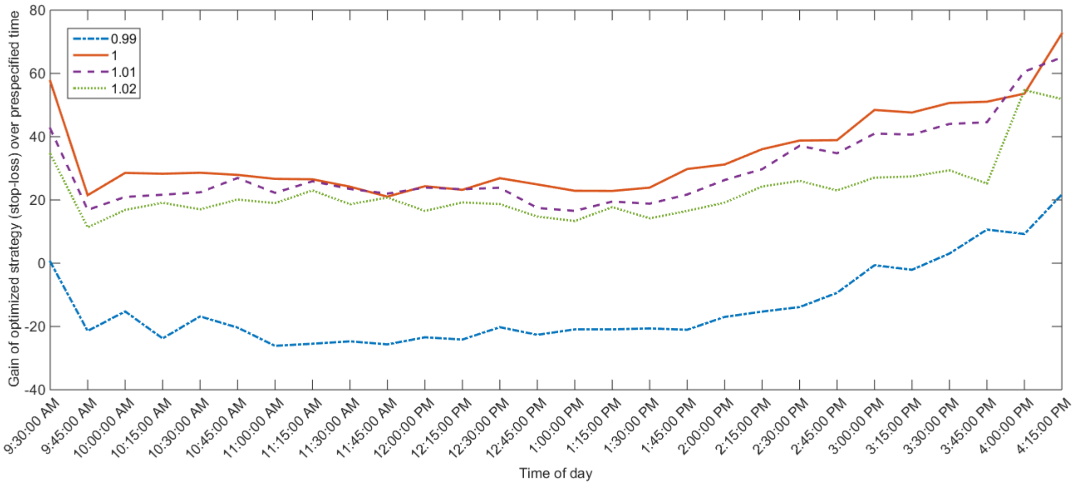


Figure 7: Buying a short-maturity (minimum two weeks) straddle with different moneyness levels (strike/forward index value at market open). The average gain of the optimal timing strategy with a stop-loss level at -1% intraday log-return relative to buying at a fixed point in time is plotted. Sample period July 2000 to December 2012 (3011 trading days). All purchases are conducted at the prevailing ask quotes.

In Table 2, we redo the analysis without the stop-loss. Following the arguments in Section 3 and the simulation results in Appendix B, removing the stop-loss level should increase the average gain. In contrast, in the data the average cost saving from using the optimal timing strategy decreases. For the ATM straddle for example, the average gain of the optimal timing strategy slips by about 39-40bps. Without the stop-loss, the cost of the TWAP strategy is slightly lower than that of the optimal timing strategy, albeit not statistically significant.

Comparing the results in Tables 1 and 2, the main effect of introducing the stop-loss level is to cut off a large part of the left tail of the gain distribution. In Section 4.5 we will argue that the most extreme negative gains are driven by strings of negative intraday index returns. As an example of the effect of this positive serial correlation, consider the worst day in the sample in terms of performance of the optimal

Moneyness	Opt vs open gain *(t-stat)*	Days	Opt equals open fraction of days	Opt better than open
0.98	0.0 *(0.0)*	2773	1.00	0.00
0.99	0.5 *(1.42)*	2821	0.98	0.01
1	57.6 *(9.90)*	2830	0.03	0.62
1.01	42.6 *(3.89)*	2817	0.01	0.59
1.02	34.6 *(2.27)*	2788	0.00	0.56
1.05	36.0 *(1.93)*	2551	0.00	0.52

(a) Optimal timing versus opening

Moneyness	Opt vs close gain *(t-stat)*	Days	Opt equals close fraction of days	Opt better than close
0.98	14.7 *(1.22)*	2773	0.00	0.44
0.99	21.7 *(2.10)*	2821	0.00	0.42
1	72.6 *(7.77)*	2830	0.13	0.43
1.01	65.1 *(6.85)*	2817	0.58	0.23
1.02	51.9 *(4.95)*	2788	0.70	0.16
1.05	38.7 *(3.22)*	2551	0.73	0.15

(b) Optimal timing versus close

Moneyness	Opt vs TWAP gain *(t-stat)*	Days	Opt equals TWAP fraction of days	Opt better than TWAP
0.98	-24.4 *(-3.22)*	2722	0.00	0.42
0.99	-16.8 *(-2.68)*	2766	0.00	0.41
1	32.4 *(6.19)*	2777	0.00	0.53
1.01	29.1 *(3.95)*	2755	0.00	0.56
1.02	22.4 *(2.31)*	2730	0.00	0.55
1.05	22.4 *(1.91)*	2490	0.00	0.54

(c) Optimal timing versus TWAP

Table 1: Buying a short-maturity (minimum two weeks) straddle with different moneyness levels (strike/forward index value at market open), with a stop-loss level at -1% intraday log-return. Sample period July 2000 to December 2012 (3011 trading days). All purchases are conducted at the prevailing ask quotes. Column 2 in each panel contains the realized average gain in basis points of following the optimal timing strategy instead of a fixed trading time benchmark (market open, market close or TWAP), with t-statistics in parentheses. Column 3 contains the number of days on which the strategy could be implemented. Columns 4 and 5 show the fraction of days on which the optimal timing strategy equals the competing strategy or performs strictly better, respectively.

timing strategy without the stop-loss relative to trading at the opening, February 27, 2007. During this so-called "Chinese correction", the S&P500 index gradually lost more than 3% of its value. The stopping boundary was never hit, but option prices had spiked up along with volatility. The straddle was only bought at the close of the market, at a price 110% higher than its opening value (contributing about -3bps to the average gain over the whole sample). The stop-loss was hit at 12:10PM New York time, when the straddle price had only gone up by 14% compared to the open price.

As suggested above, the model as postulated in Section 2 is only a crude approximation to the true data-generating process of the S&P500 index and its options. In the following three sub-sections, we will demonstrate how the Black-Scholes world in which we solve for the optimal trading strategy can be seen as a conservative setting for the potential gains when purchasing a straddle. In particular, we argue that the data contains at least three sources of predictability, driving the difference between the gains from the model and the gains documented in Tables 1 and 2 and adding value to the stop-loss level: intraday seasonality in bid-ask spreads, serial correlation in intraday index returns and the leverage effect.

4.4 Intraday seasonality in option bid-ask spreads

There model of Section 2 abstracts from bid-ask spreads, while in the data put and call option spreads are time-varying and generally highest near the opening of the market. Figure 8 illustrates the bid-ask spread intraday seasonality pattern by plotting the distribution of the relative bid-ask spread during the day for an ATM straddle on the S&P500 index[17]. This pattern is consistent with results in Duarte, Lou, and Sadka (2006), who find that, for a sample of intraday S&P500 index option quotes during 2001-2002, the average intraday spread is very close to the closing spread. The simulations in Appendix B show how the intraday seasonality pattern has a positive effect on the gains compared to trading at the open in the Black-Scholes world. For the S&P500 data, we assess the effect of intraday seasonality by buying at the midpoint of the bid and ask quotes ("midquote") instead of the ask, while noting that, being price takers, small traders may not be able to trade at the midquote in practice. It is also appropriate at this stage to recall that we compute the straddle spread as the sum of the call and put spread that make up the straddle. Our results remain valid if actual straddle spreads when quoted as a single trade are lower than the sum of the call and put spread, as long as the straddle spread shows the same intraday seasonality pattern as displayed in Figure 8.

Comparing Table 3 to Table 1 confirms the intuition. Consider Panel (a) first, where the benchmark is trading at the open. In that case, the trade decision of the open and optimal timing strategies coincide for the 0.98 and 0.99 moneyness straddles, so the gain is the same whether trading at the midquote or

[17]It is not clear what generates the spikes around 10AM and 4PM New York time. The 10AM spike could be due to the release of the daily position and daily margin reports by the Options Clearing Corporation at 10AM. Upon closer inspection, the 10AM spike is only present in the data from 2009 onwards. The drift lower towards 4PM and subsequent spike up could be due to the NYSE and NASDAQ markets closing, although the S&P500 index futures contract, which is the default hedge product for S&P500 index options, continues to trade until 4:15PM without interruption.

Moneyness	**Opt vs Open** **gain *(t-stat)***	**Days**	**Opt equals Open**	**Opt better than Open**
			fraction of days	
0.98	0.0 *(0.0)*	2773	1.00	0.00
0.99	-1.5 *(-1.52)*	2821	0.98	0.01
1	18.4 *(1.93)*	2830	0.03	0.63
1.01	-2.7 *(-0.18)*	2817	0.00	0.61
1.02	-12.4 *(-0.65)*	2788	0.00	0.60
1.05	-1.9 *(-0.09)*	2551	0.00	0.56

(a) Optimal timing versus opening

Moneyness	**Opt vs Close** **gain *(t-stat)***	**Days**	**Opt equals Close**	**Opt better than Close**
			fraction of days	
0.98	14.7 *(1.22)*	2773	0.00	0.44
0.99	19.8 *(1.91)*	2821	0.00	0.42
1	33.4 *(5.42)*	2830	0.29	0.34
1.01	19.7 *(6.66)*	2817	0.81	0.11
1.02	4.9 *(3.45)*	2788	0.94	0.03
1.05	0.7 *(0.85)*	2551	0.99	0.00

(b) Optimal timing versus close

Moneyness	**Opt vs TWAP** **gain *(t-stat)***	**Days**	**Opt equals TWAP**	**Opt better than TWAP**
			fraction of days	
0.98	-24.4 *(-3.22)*	2722	0.00	0.42
0.99	-18.8 *(-3.00)*	2766	0.00	0.41
1	-7.9 *(-1.19)*	2777	0.00	0.52
1.01	-17.8 *(-1.94)*	2755	0.00	0.57
1.02	-26.0 *(-2.27)*	2730	0.00	0.56
1.05	-17.2 *(-1.28)*	2490	0.00	0.55

(c) Optimal timing versus TWAP

Table 2: Buying a short-maturity (minimum two weeks) straddle with different moneyness levels (strike/forward index value at market open). Sample period July 2000 to December 2012 (3011 trading days). All purchases are conducted at the prevailing ask quotes. Column 2 in each panel contains the realized average gain in basis points of following the optimal timing strategy instead of a fixed trading time benchmark (market open, market close or TWAP), with t-statistics in parentheses. Column 3 contains the number of days on which the strategy could be implemented. Columns 4 and 5 show the fraction of days on which the optimal timing strategy equals the competing strategy or performs strictly better, respectively.

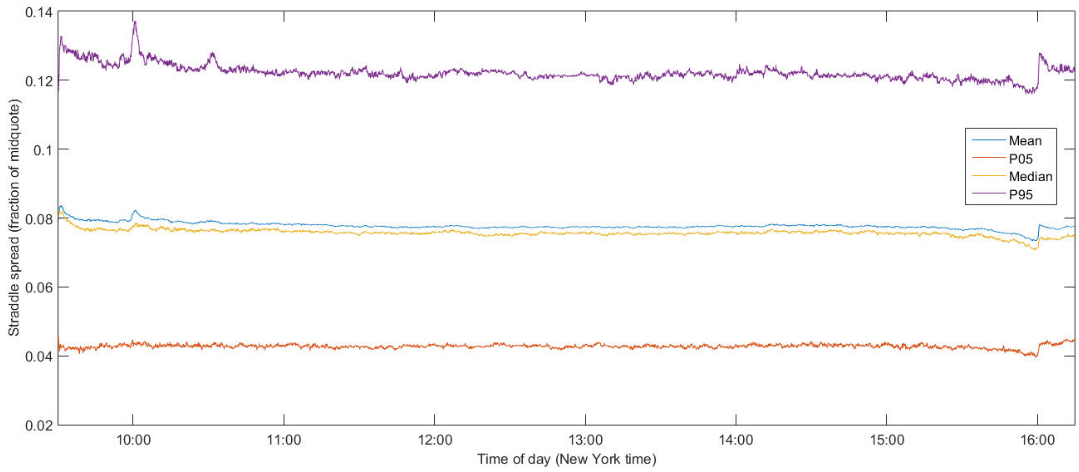


Figure 8: Intraday pattern of dollar bid-ask spreads as a fraction of the midquote for an S&P500 index ATM straddle at the open of the market (Strike/Forward price = 1). Straddle dollar spreads are computed as the sum of the dollar spread on the put and the call. Sample period July 2000 to December 2012. "P05" and "P95" denote the 5th and 95th percentiles, respectively.

the ask price. For moneyness 1 and higher the gain decreases when trading at the midquote instead of at the ask, partially because the optimal timing strategy trades after the open, avoiding the higher opening spread. Not all of the gain can be explained this way however. For the ATM straddle, the median bid-ask spread starts at around 8.1% of the midquote at the open, and decreases steadily during the day to about 7.2% of the open midquote at 4PM. Using those values, the *maximum* gain from not trading at the open would be $90/2 = 45$bps, still smaller than the 57bps in Table 1. For most of the day, the bid-ask spread hovers around 7.6% of the straddle midquote, so a more reasonable estimate would be that about $50/2 = 25$bps of the 57bps average gain could possibly be attributed to intraday seasonality in option bid-ask spreads.

Figure 7 shows the same point in a different way. Focusing on the 0.99 moneyness again, the average gain increases with calendar time during the second half of the trading day, even though the relative spread decreases with calendar time, at least up to 4PM. If intraday seasonality in spreads were the only explanation for the performance of the optimal timing strategy (which buys at the high spread at the market open), then the average gain should be decreasing with calendar time. This result holds true for all moneyness levels, but for moneyness equal to 1 and up, the comparison is not as clean as the trading epoch of the optimal timing strategy is not a fixed point in time.

If the benchmark is TWAP, the gains for the low moneyness straddles are higher when the trades are executed at the midquote than at the ask price. The optimal timing strategy still buys at the open, but the price improvement from switching to midquote is bigger at the open than at later points in time. The improvement ranges between 8-16bps, slightly smaller than the difference in the half or effective spread.

Moneyness	Opt vs open gain *(t-stat)*	Days	Opt equals open	Opt better than open
			fraction of days	
0.98	0.0 *(0.0)*	2773	1.00	0.00
0.99	0.5 *(1.55)*	2821	0.98	0.01
1	31.2 *(5.95)*	2830	0.03	0.60
1.01	20.8 *(1.87)*	2817	0.01	0.58
1.02	26.0 *(1.65)*	2788	0.00	0.55
1.05	30.3 *(1.59)*	2551	0.00	0.52

(a) Optimal timing versus opening

Moneyness	Opt vs close gain *(t-stat)*	Days	Opt equals close	Opt better than close
			fraction of days	
0.98	23.3 *(1.89)*	2773	0.00	0.45
0.99	35.9 *(3.45)*	2821	0.00	0.43
1	67.0 *(7.36)*	2830	0.13	0.42
1.01	62.6 *(6.47)*	2817	0.58	0.22
1.02	52.4 *(4.86)*	2788	0.70	0.16
1.05	39.6 *(3.23)*	2551	0.73	0.15

(b) Optimal timing versus close

Moneyness	Opt vs TWAP gain *(t-stat)*	Days	Opt equals TWAP	Opt better than TWAP
			fraction of days	
0.98	-14.5 *(-1.89)*	2722	0.00	0.43
0.99	-2.7 *(-0.43)*	2766	0.00	0.42
1	28.1 *(5.79)*	2777	0.00	0.52
1.01	24.5 *(3.28)*	2755	0.00	0.56
1.02	22.6 *(2.25)*	2730	0.00	0.55
1.05	26.6 *(2.22)*	2490	0.00	0.54

(c) Optimal timing versus TWAP

Table 3: Buying a short-maturity (minimum two weeks) straddle with different moneyness levels (strike/forward index value at market open), with a stop-loss level at -1% intraday log-return. Sample period July 2000 to December 2012 (3011 trading days). All purchases are conducted at the mid price of the prevailing bid and ask quotes. Column 2 in each panel contains the realized average gain in basis points of following the optimal timing strategy instead of a fixed trading time benchmark (market open, market close or TWAP), with t-statistics in parentheses. Column 3 contains the number of days on which the strategy could be implemented. Columns 4 and 5 show the fraction of days on which the optimal timing strategy equals the competing strategy or performs strictly better, respectively.

4.5 Serial correlation in intraday returns

The Black-Scholes model used in Section 2 assumes returns are i.i.d. over any time period. Prior research, for example Fisher (1966) and Froot and Perold (1995), suggests that this is a reasonable assumption for low-frequency portfolio returns, but may break down when considering higher frequencies, with non-synchronous trading of portfolio constituents generating positive autocorrelation in portfolio returns. Fleming, Ostdiek, and Whaley (1996) show this for an early sample of five minute returns on the S&P500 index. They find that for their sample, which spans January 1988 to March 1991, the first-order autocorrelation of S&P500 index returns measured on a five-minute horizon is 31%.

As Figure 9 shows, increased trading activity and market liquidity have led to the i.i.d. assumption being reasonable for S&P500 index returns measured at a frequency of at least 5 minutes during our more recent sample. Five minute intervals are often used in realized volatility calculations for that reason (see e.g. Zhang, Mykland, and Aït-Sahalia (2005) for a discussion). However, when returns are measured at the 15-second frequency as used in the trinomial tree, Figure 9 shows that there is substantial positive serial correlation in returns for low lags with a first order autocorrelation of about 28%.

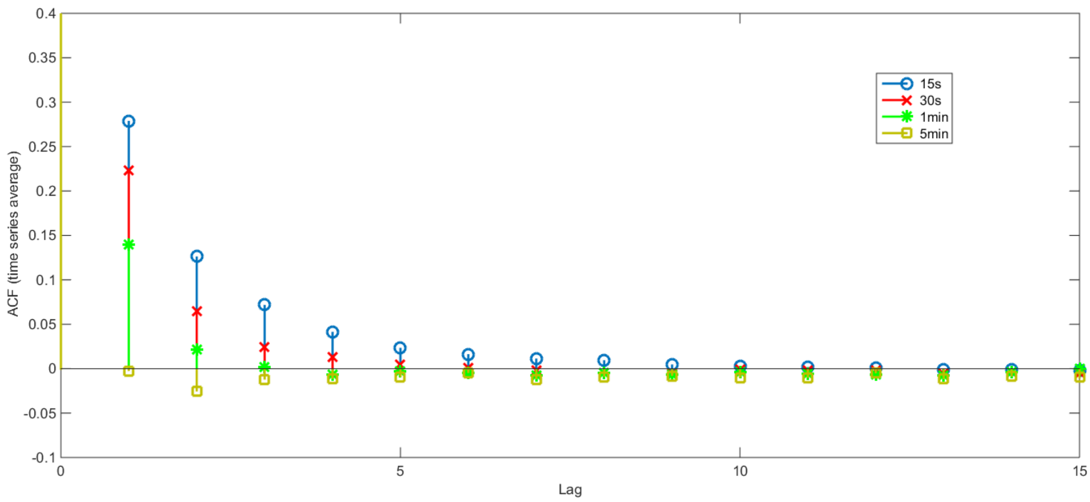


Figure 9: Autocorrelation functions of intraday S&P500 index returns, measured on different frequencies as indicated in the figure. The autocorrelation function is computed per day, and the time-series average is shown.

The positive autocorrelation can negatively affect the gains from our optimal timing trading strategy when buying the (ATM) straddle, because trading is only triggered if the cumulative intraday return hits a certain positive threshold. Strings of negative returns in particular, for example during the "Chinese correction" as discussed in Section 4.3, will delay the purchase until the end of the trading period at a potentially much higher option price. A stop-loss level provides a simple way to mitigate the effect of serially correlated intraday returns, without reverting to a much more complicated data generating process when deriving the optimal timing strategy. In this case, the stop-loss can be interpreted as a way to deal with (aversion to) model risk. To show this, we conduct the following experiment. Assume that the only deviation from geometric Brownian motion in the underlying asset price process is due to serial correlation that only affects the drift. Option prices, which do not depend on the drift, can then still be computed using the Black-Scholes formula. We take the trading epochs from the S&P500 data as derived in Section 4.3 for the optimal timing strategy both with and without stop-loss. At each trading epoch, rather than trading at the observed option price in the data, we buy the option at the Black-Scholes price, using the actual strike, the observed S&P500 index value at the time of trading and the volatility forecast from the HAR-RV model for that day as inputs, with risk-free rate, dividend yield and time-to-maturity directly taken from the OptionMetrics database. Figure 10 shows that the stop-loss consistently adds value, up to 31bps for the straddle with moneyness 1.02. In contrast, when the data-generating process is the geometric Brownian motion of the Black-Scholes model, adding the stop-loss decreases the average gain by up to 3bps, see Figure 1 and the simulation results in Appendix B.

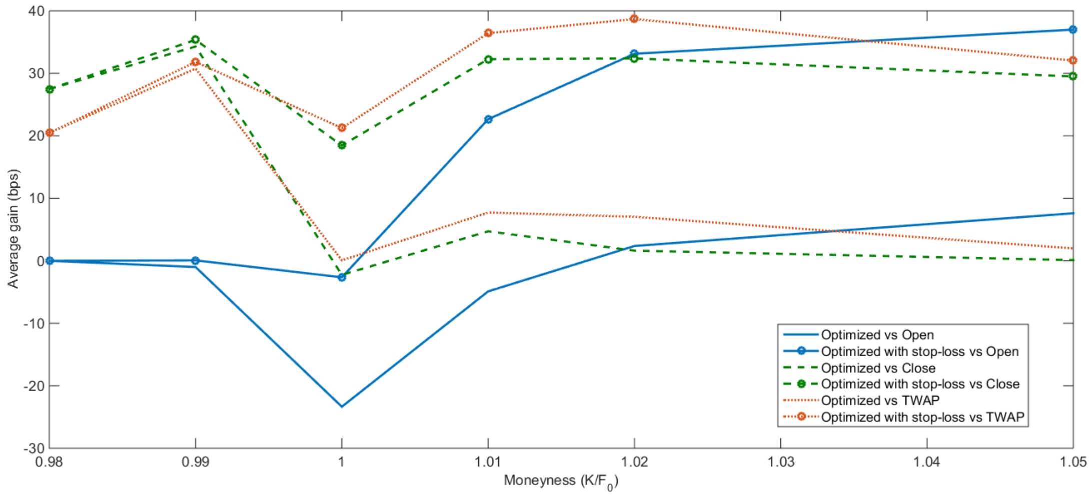


Figure 10: Average gains from trading using optimal timing strategy with or without stop-loss, compared to trading at either the open or close of the market or a TWAP strategy. Trading epochs are taken from the analysis in Section 4.3. Option prices are computed using Black-Scholes formula using a volatility forecast from the HAR-RV model.

We note that even though trading activity in general and high-frequency trading in particular have increased during our sample, the serial correlation of high-frequency index returns remains high. At the 15 second frequency, the time-series average of first order autocorrelation computed per day decreases from about 0.4 in 2001 to around 0.25 in 2012. If the serial correlation in returns would be the only factor driving the cost savings from using the optimal timing strategy, then this should be visible as a downward trend in average gains over time. Figure 11 does not show such trend.

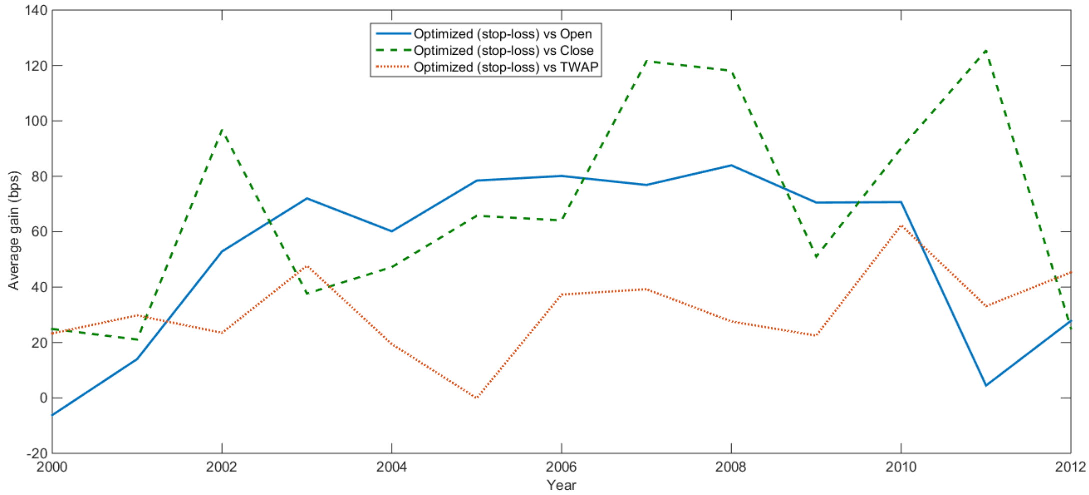


Figure 11: Buying a short-maturity (minimum two weeks) S&P500 straddle with different moneyness levels (strike/forward index value at market open), with a stop-loss level at -1% intraday log-return. Average gains per year from trading using optimal timing strategy with or without stop-loss, compared to trading at either the open or close of the market or a TWAP strategy. Sample period July 2000 to December 2012 (3011 trading days).

4.6 Stochastic volatility, the variance risk premium and the leverage effect

In the previous section, by computing option prices using the Black-Scholes formula, we shut down the effect of stochastic volatility and the variance risk premium on option prices and hence the performance of our optimal timing trading strategy. For stock market indices, there is plenty of evidence that index return volatility is stochastic and a priced risk factor, see for example Bollerslev, Tauchen, and Zhou (2009) and Bollerslev, Gibson, and Zhou (2011). Option pricing models have long incorporated stochastic volatility, see, e.g., the continuous-time Heston (1993) model and discrete-time Heston and Nandi (2000) model. Under the physical measure, the dynamics of the underlying asset S and its spot variance V in the Heston (1993) model are given by

$$\mathrm{d}S_t = (\mu - \delta)S_t\mathrm{d}t + \sqrt{V_t}S_t\mathrm{d}Z_{1,t},$$
$$\mathrm{d}V_t = \kappa\left(\theta - V_t\right)\mathrm{d}t + \omega\sqrt{V_t}\mathrm{d}Z_{2,t},$$

where $Z_{1,t}$ and $Z_{1,t}$ are standard Brownian motions under the physical measure with $\mathrm{d} < Z_{1,t}, Z_{2,t} >= \rho\mathrm{d}t$. Following Heston (1993) and specifying the market price of variance risk as $\lambda\sqrt{V_t}/\omega$, the dynamics under the risk-neutral measure can be written as

$$\mathrm{d}S_t = (r - \delta)S_t\mathrm{d}t + \sqrt{V_t}S_t\mathrm{d}\tilde{Z}_{1,t}, \tag{13}$$
$$\mathrm{d}V_t = \tilde{\kappa}\left(\tilde{\theta} - V_t\right)\mathrm{d}t + \omega\sqrt{V_t}\mathrm{d}\tilde{Z}_{2,t}, \tag{14}$$

where $\tilde{Z}_{1,t}$ and $\tilde{Z}_{1,t}$ are standard Brownian motions under the risk-neutral distribution with $\mathrm{d} < \tilde{Z}_{1,t}, \tilde{Z}_{2,t} >= \rho\mathrm{d}t$, $\tilde{\kappa} = \kappa + \lambda$ and $\kappa\theta = \tilde{\kappa}\tilde{\theta}$.

Empirically, priced stochastic volatility has been shown to be an important factor in explaining option prices; see, e.g., Bakshi, Cao, and Chen (1997). If variance risk is indeed priced, then the exposure of the option portfolio to variance risk over time will be a determinant of the trading strategy that seeks to minimize the expected purchase price of the straddle. Using the same PDE-based argument as employed in the proof of Proposition 2.1, and letting the variance risk premium be a constant λ times the spot variance V_t, we can express (6), the expected return from the moment of trading τ until the end of the trading period, in the Heston (1993) stochastic volatility model as

$$E\left\{ \mathrm{e}^{-rT} f(S_T, V_T, T) - \mathrm{e}^{-r\tau} f(S_\tau, V_\tau, \tau) \right\} = (\mu - r)\, E \int_\tau^T \mathrm{e}^{-ru} f_{\ln S}(S_u, V_u, u)\mathrm{d}u + \lambda E \int_\tau^T \mathrm{e}^{-ru} f_{\ln V}(S_u, V_u, u)\mathrm{d}u. \qquad (15)$$

Empirically, λ has been found to be negative, while for a long straddle $f_{\ln V} = V f_V$, with f_V the vega of the straddle, is positive. Therefore, an optimal timing strategy in the Heston model will delay trading relative to the Black-Scholes model. In other words, the optimal stopping boundary derived in a Heston model would generally lie above that of a Black-Scholes model with comparable parameter values.

Simulation results documented in Appendix B confirm that using a stopping boundary derived in a Black-Scholes world when the underlying asset price process follows Heston dynamics with a negative variance risk premium and negative correlation between return and volatility innovations generates substantial positive gains compared to trading at the open. The largest gains are obtained for straddles that are closest to being delta- and vega-neutral, corresponding to a moneyness close to 1.01 for the chosen parameters as shown in Figure 12. At the same time, compared to the close the optimal timing strategy realizes small negative gains. As suggested by (15), the variance risk premium is the main driver of these results. Switching off the leverage effect leads to small increases in the magnitude of gains and losses. Gains and losses are smaller in the presence of a leverage effect because the joint occurrence of a negative return and positive spot variance innovation will reduce both the log-delta $f_{\ln S}$ and the log-vega $f_{\ln V}$ at the same time for moneyness levels below about 1.01, so that the reduction in the exposure to the positive equity premium is (partly) offset by a reduction in the exposure to the negative variance risk premium. The stop-loss does not add value in this set up; the intuition is that hitting the stop-loss means trading earlier than otherwise, and hence being exposed to the negative variance risk premium for a longer period of time, decreasing the expected return on the straddle position. We conclude that the presence of a priced volatility risk factor can explain some of the difference between theoretical and empirical gains, mainly when comparing against trading at the open.

Finally, we consider the combination of serial correlation and leverage effect/variance risk premium. In Section 4.5 we analysed the value of the stop-loss in relation to serial correlation, without taking into account a leverage effect or changes in the variance risk premium. If innovations in returns and volatility are negatively correlated, then the value added of the stop-loss is potentially even higher on days when strings of negative intraday index returns are realized. This is the case when the variance risk premium

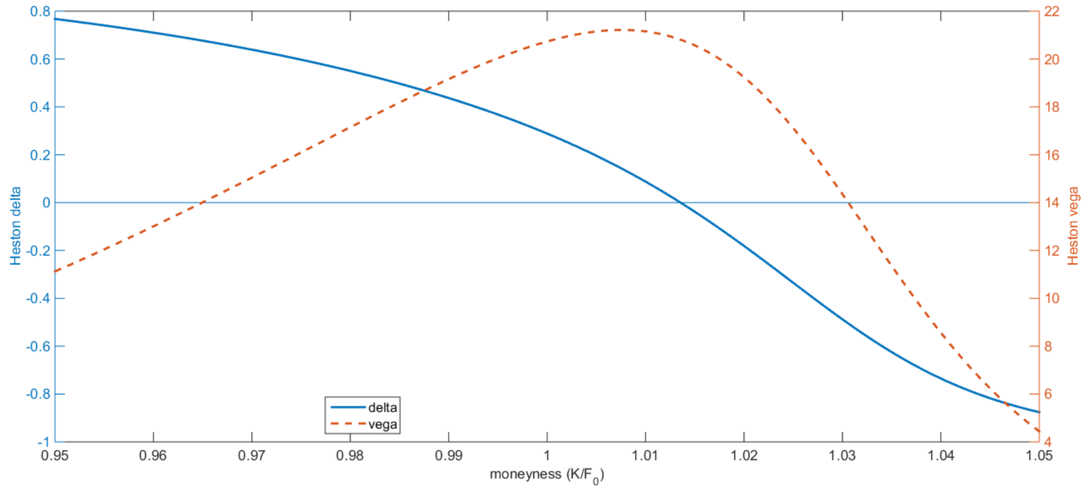


Figure 12: Delta and vega of a straddle in the Heston model as a function of moneyness (K/F_0). The parameters in (13) and (14) are set to $r = 0.024$, $\delta = 0.018$, $\tilde{\kappa} = 3.2$, $\tilde{\theta} = 0.04$, $V_0 = 0.0256$, $\omega = 1$, $\rho = -0.8$.

is time-varying and co-moves with the level of (implied) volatility, as documented empirically by, for example, Bollerslev, Tauchen, and Zhou (2009) and Carr and Wu (2009), because then the negative returns tend to coincide with an increase in volatility and the variance risk premium, which in turn pushes option prices up.

The prediction for the performance of the optimal timing strategy when faced with the combination of strings of positively correlated positive intraday index returns and the leverage effect/variance risk premium is more subtle. On the one hand, positively serially correlated returns increase the attractiveness of trading earlier, suggesting the boundary should be lowered relative to the i.i.d. returns case. At the same time, the leverage effect in this case implies a positive return will be associated with a decrease in (implied) volatility and hence decreasing option prices. This will make delaying the purchase more attractive, which can be achieved by shifting the boundary upward. We examine the net effect by using an alternative definition of the stopping boundary, which lies slightly above the stopping boundary used in the remainder of the paper. If the serial correlation is more important, then the gains from the optimal strategy using the alternative boundary should be lower, whereas the gains will be higher if the leverage effect dominates.

Using the alternative stopping boundary ("UB" in Table 4) for the optimal timing strategy generates a slight improvement in performance for most moneyness levels with the largest gain for the ATM straddle at 2.8 basis points, which is statistically significant at the 10% level. In addition, for the 0.99 moneyness straddle, the improvement of 1 bps is statistically significant at the 5% level. As far as the time at which the trade is triggered is concerned, the effect of switching between the upper and lower boundary is very small for both low and high moneyness levels. Trading times are the same for the vast majority of days for low and high moneyness straddles, as indicated in the fourth column of Table 4. However, for the

ATM straddle the trading time when using the upper boundary will be different from the trading time using the lower boundary on about two out of three days. On 40% of days, the strategy using the upper bound outperforms the strategy using the lower bound. On these days, the median trade for the ATM straddle takes place 92 seconds later when using the upper bound. While the economic significance of the performance improvement from using the alternative stopping boundary is limited, the results in Table 4 suggest that as far as positively correlated strings of positive index returns are concerned, the performance of the strategy is more sensitive to a decrease in (implied) volatility stemming from the leverage effect than to the serial correlation itself. For future reference, we also include the metrics analyzed in Section 4.7 in the table.

4.7 Alternative risk preferences

The stop-loss level has been introduced to capture, in a simple way, risk-aversion of agents. In the previous section we studied deviations from the assumed Black-Scholes model, i.e., the stop-loss level can be interpreted as an aversion to model risk. We now analyse the risk-return properties of the optimal timing strategy in more detail. We consider three metrics in particular.

The first metric we consider is the risk premium ψ for a CRRA investor with risk aversion coefficient $\gamma = 2$, solving

$$(1-\psi)^{1-\gamma} = E\left\{(1+\text{gain})^{1-\gamma}\right\}, \tag{16}$$

where the expectation is estimated by the sample average of gains as defined in (10). A positive value of ψ means that the investor would have to be paid this fraction of the portfolio value in order to prefer the optimal timing strategy over the competitor. Hence, a negative amount means that the investor is willing to pay to get exposure to the optimal timing strategy.

The second metric is the maximum value of the risk aversion coefficient A such that a mean-variance investor would still prefer the optimal timing strategy over a specific competitor,

$$A = 2\frac{E\{\text{gain}\}}{\text{Var}\{\text{gain}\}}. \tag{17}$$

A negative value means that only a risk-seeking mean-variance investor would be willing to use the optimal timing strategy versus the competing strategy. Once again, expectations and variances are estimated by sample equivalents.

The third metric is the annualized Sharpe ratio SR, implicitly assuming that the gain is realized at the end of the day,

$$SR = \sqrt{252}\frac{E\{\text{gain}\}}{\sqrt{\text{Var}\{\text{gain}\}}}, \tag{18}$$

which is conservative because the actual gain is often realized on a horizon much shorter than one day.

Given the earlier discussion, we focus on the ATM straddle. Table 5 shows that the optimal timing strategy with stop-loss scores well in risk-reward terms, with the expection of the comparison with the TWAP strategy. For the open strategy for example, the CRRA investor would be willing to pay 48

Moneyness	**UB vs LB gain** ***(t-stat)***	**Days**	**UB = LB**	**UB cheaper**	ψ **(bps)**	A	SR
			fraction of days				
0.98	0.0 *(0.0)*	2763	1.00	0.00	0.00	NA	NA
0.99	1.0 *(2.30)*	2811	0.97	0.02	-0.91	39.41	0.69
1	2.8 *(1.65)*	2820	0.40	0.40	-1.99	6.82	0.49
1.01	-0.6 *(-0.77)*	2805	0.84	0.10	0.83	-6.64	-0.23
1.02	-0.0 *(-0.05)*	2778	0.96	0.02	0.06	-0.97	-0.02
1.05	0.1 *(0.27)*	2541	1.00	0.00	-0.05	8.82	0.09

(a) Optimal timing strategy with stop-loss

Moneyness	**UB vs LB gain** ***(t-stat)***	**Days**	**UB = LB**	**UB cheaper**	ψ **(bps)**	A	SR
			fraction of days				
0.98	0.0 *(0.0)*	2773	1.00	0.00	0.00	NA	NA
0.99	-0.3 *(-0.39)*	2821	0.97	0.02	0.47	-3.65	-0.12
1	1.5 *(0.73)*	2830	0.36	0.43	-0.29	2.60	0.22
1.01	-0.1 *(-0.12)*	2817	0.83	0.11	0.32	-0.96	-0.04
1.02	0.1 *(0.15)*	2788	0.95	0.03	0.01	1.93	0.04
1.05	0.2 *(0.76)*	2551	0.99	0.00	-0.19	21.60	0.24

(b) Optimal timing strategy

Table 4: Buying a short-maturity (minimum two weeks) straddle with different moneyness levels (strike/forward index value at market open), with a stop-loss level at -1% intraday log-return. Sample period July 2000 to December 2012 (3011 trading days). The comparison is between the optimal timing strategy using the default stopping boundary ("LB") and the optimal timing strategy using an alternative stopping boundary ("UB") that provides an upperbound on the continuous-time, continuous-state stopping boundary. All purchases are conducted at the prevailing ask quotes. Column 2 in each panel contains the realized gain in basis points of following the optimal timing strategy instead of trading at the opening or close, respectively (with t-statistics in parentheses). Column 3 contains the number of days on which the strategy could be implemented. Columns 4 and 5 show the fraction of days on which the optimal timing strategy equals the competing strategy or performs strictly better, respectively. The last three columns refer to the metrics discussed in Section 4.7. ψ is the risk premium of a CRRA investor with $\gamma = 2$. A is the maximum coefficient of risk aversion for a mean-variance investor such that she prefers the optimal timing strategy. SR is the annualized Sharpe ratio of the daily gains distribution.

basis points of the value of the derivative portfolio to get access to the optimal timing strategy. Any mean-variance investor with a coefficient of risk aversion smaller than 12.0 would prefer the optimal timing strategy with the stop-loss over trading at the open.

Comparing panel 5a with panel 5b demonstrates the interpretation of the stop-loss level as a means to incorporate risk aversion into the problem. The stop-loss strategy is consistently preferred on all risk-return metrics to the strategy without the stop-loss, mainly because adding the stop-loss substantially reduces the variance of the gain distribution. This is in contrast with the results of the simulations of the Black-Scholes model in Appendix B, where the strategy without stop-loss is preferred on all metrics.

Competing Strategy	ψ **(bps)**	A	SR
Open	-48.02	12.02	2.96
Close	-50.77	5.88	2.31
TWAP	-23.39	8.01	1.77

(a) Optimal timing strategy with stop-loss

Competing Strategy	ψ **(bps)**	A	SR
Open	-32.06	1.43	0.58
Close	-23.07	6.22	1.62
TWAP	29.94	-1.33	-0.37

(b) Optimal timing strategy

Table 5: Risk-return analysis of gains from trading using optimal timing (stop-loss) strategy versus fixed time benchmark strategies for an ATM straddle (Strike/Forward price = 1), with all trades completed at the prevailing ask quote. ψ is the risk premium of a CRRA investor with $\gamma = 2$. A is the maximum coefficient of risk aversion for a mean-variance investor such that she prefers the optimal timing strategy. SR is the annualized Sharpe ratio of the daily gains distribution.

4.8 Trading with a lag

Trading of S&P500 index options was historically done in a trading pit, and only moved to the CBOE's hybrid trading system that combines electronic and open out-cry trading in 2007[18]. This may potentially introduce a lag in reporting option quotes back to the OPRA feed, leading to the time series of the S&P500 index and its options being out of sync. In addition, Ding, Hanna, and Hendershott (2014) show that consolidated data feeds like the Thomson Reuters Tick History data base we use here are frequently lagging direct exchange data feeds when multiple exchanges are involved. To address these issues, we rerun the analysis using a trading lag of 60 seconds for the optimal timing and stop-loss strategy. That is, the moment the stopping boundary or stop-loss level is hit, we wait 60 seconds before trading at the then current quotes. Table 6 shows that while cost savings decrease somewhat compared to Table 1, they remain both economically and statistically significant for straddle around the ATM level. The decrease is about 15bps across the board, independent of the benchmark strategy used.

The analysis conducted so far suggests several avenues for further research. The focus in this paper

[18] See `http://www.cboe.com/AboutCBOE/AnnualReportArchive/AnnualReport2007.pdf`.

Moneyness	**Opt vs Open**	**Days**	**Opt equals Open**	**Opt better than Open**
	gain *(t-stat)*		**fraction of days**	
0.98	0.0 *(0.0)*	2773	1.00	0.00
0.99	0.9 *(2.21)*	2821	0.98	0.01
1	42.0 *(7.18)*	2830	0.03	0.60
1.01	29.7 *(2.66)*	2817	0.00	0.59
1.02	21.6 *(1.39)*	2788	0.00	0.56
1.05	19.9 *(1.05)*	2551	0.00	0.52

(a) Optimal timing versus opening

Moneyness	**Opt vs Close**	**Days**	**Opt equals Close**	**Opt better than Close**
	gain *(t-stat)*		**fraction of days**	
0.98	14.7 *(1.22)*	2773	0.00	0.44
0.99	22.2 *(2.14)*	2821	0.00	0.42
1	57.0 *(6.26)*	2830	0.13	0.41
1.01	52.1 *(5.59)*	2817	0.59	0.22
1.02	38.9 *(3.73)*	2788	0.70	0.15
1.05	22.6 *(1.90)*	2551	0.73	0.14

(b) Optimal timing versus close

Moneyness	**Opt vs TWAP**	**Days**	**Opt equals TWAP**	**Opt better than TWAP**
	gain *(t-stat)*		**fraction of days**	
0.98	-24.4 *(-3.22)*	2722	0.00	0.42
0.99	-16.4 *(-2.62)*	2766	0.00	0.41
1	16.6 *(3.31)*	2777	0.00	0.50
1.01	16.2 *(2.20)*	2755	0.00	0.55
1.02	9.4 *(0.96)*	2730	0.00	0.55
1.05	6.2 *(0.52)*	2490	0.00	0.53

(c) Optimal timing versus TWAP

Table 6: Buying a short-maturity (minimum two weeks) straddle with different moneyness levels (strike/forward index value at market open), with a stop-loss level at -1% intraday log-return. Sample period July 2000 to December 2012 (3011 trading days). All purchases are conducted at the prevailing ask quotes, 60 seconds after the trade signal was generated. Column 2 in each panel contains the realized average gain in basis points of following the optimal timing strategy instead of a fixed trading time benchmark (market open, market close or TWAP), with t-statistics in parentheses. Column 3 contains the number of days on which the strategy could be implemented. Columns 4 and 5 show the fraction of days on which the optimal timing strategy equals the competing strategy or performs strictly better, respectively.

has been on index options because of the higher liquidity in these markets, but the analysis can be used when trading options on individual stocks as well. Duan and Wei (2009) and Christoffersen, Fournier, and Jacobs (2013) show the importance of the systematic risk exposure of the underlying stocks for the dynamics of individual equity option prices, while Driessen, Maenhout, and Vilkov (2009) argue that individual stock options do not carry a variance risk premium. Muravyev and Pearson (2015) argue that sophisticated option traders are able to time trade execution, using recent stock price changes to predict the direction of option quote updates and limit the effective spread paid.

The trinomial tree model could be extended to account for (some of) the stylized facts that we have identified as creating a wedge between the gains from trading in the S&P500 data vis-á-vis the Black-Scholes world. In addition to the factors analysed here, recent work in derivative pricing has shown that the value of derivatives is also affected by volatility-of-volatility risk. In particular, Huang and Shaliastovich (2014) argue that delta-hedged equity index option strategies are exposed to volatility *and* volatility-of-volatility risk, both of which carry negative risk premiums, and therefore earn negative returns on average. In determining whether extending the trinomial model is worthwhile, one has to balance the improved fit to stylized facts against the increased complexity and loss of tractability.

In an extended model where the price of the underlying asset is no longer the sole source of risk, the definition of the stop-loss level may have to be modified too as the equivalence between a maximum portfolio price and an underlying asset value will cease to exist. A definition of the stop-loss in terms of the price of the derivative basket itself will have to be used, for example a deviation of 10% relative to the price at the start of the trading period. In the current empirical analysis, the stop-loss level is independent of the spot volatility of the index, i.e., on high-volatility days it is more likely to be hit than on low-volatility days. Alternatively, the stop-loss could be defined as a fixed number of daily return standard deviations.

The results so far highlight the high-frequency changes in option deltas and therefore also their betas, which in the Black-Scholes world equal the product of the underlying asset beta and the option elasticity,

$$\beta_{\text{Option}} = \beta_{\text{asset}} \Delta_{\text{option}} \frac{\text{Asset price}}{\text{Option price}}. \tag{19}$$

This may have consequences for return measurement of delta-hedged strategies as in Bertsimas, Kogan, and Lo (2000) and Bakshi and Kapadia (2003), and for dynamic factor loadings as studied by Engle (2014).

Throughout, we have assumed no price impact from trading to focus solely on the benefits from trade timing in view of high-frequency changes in option risk exposures for trade timing. In stock markets, the optimal liquidation of large positions with the aim of minimizing the price impact has been studied extensively, see for example Bertsimas and Lo (1998), Almgren and Chriss (2000), Gatheral (2009), Huberman and Stanzl (2004), Huberman and Stanzl (2005) and Schied, Schöneborn, and Tehranchi (2010). Hasbrouck (1991) develops an empirical measure to quantify price impact in the stock market.

For derivatives markets, much less work has been done to date. While the market for S&P500 options is quite liquid and the contract size is large, implying that positions of reasonable size can be traded without affecting the market price, the supply of S&P500 index options is not perfectly elastic. A few papers that take a look at a specific form of price impact in options markets are Cho and Engle (1999) and Muravyev (2014). Bollen and Whaley (2004), Gârleanu, Pedersen, and Poteshman (2009) and Ni, Pan, and Poteshman (2008) show that order flow imbalances affect daily movements in the implied volatility surface.

5 Conclusion

We study the trade execution part of the portfolio delegation problem where a portfolio manager hires a trader to do the trading. The trader is a) given a certain amount of time to complete trading a derivatives portfolio, and b) evaluated based on profits/losses relative to an execution strategy that trades at a time during the day. In this setting, the trader holds an option to delay trading whenever the value of a basket of options to be traded simultaneously is a non-monotonic function of the value of the underlying asset. The optimal trading strategy, defined as maximizing the expected profit for the trader, uses a random execution time, which is the solution of an optimal stopping problem. Equivalently, the optimal trading strategy corresponds to an order submission strategy that consists of a series of stop-orders with time-varying stop prices, rather than a fixed-price limit order.

In a Black-Scholes world, we solve the optimal stopping problem using an explicit finite difference method and demonstrate the option to trade has substantial value when trading a straddle of up to several tens of basis points. Using the trading strategy derived in the Black-Scholes model to trade a straddle on the S&P500 index, average gains up to 72bps per day are achieved for short-maturity, at-the-money straddles.

A risk management constraint in the form of a stop-loss level can be added to the optimal timing strategy. This stop-loss level can be interpreted as a way to reduce model risk since it enhances the robustness of the optimal timing strategy when the Black-Scholes assumptions are violated. When trading the S&P500 index straddle, the addition of the stop-loss leads to a substantial improvement in the value of the trading option, even though it lowers the value of the trading option in the Black-Scholes world.

We identify and analyse three ways in which the data-generating process of S&P500 index options deviates from the Black-Scholes assumptions of i.i.d. returns and no transaction costs. Firstly, S&P500 straddle bid-ask spreads display intraday seasonality: they are high at the market open and slowly decrease during the day until the moment the market for the S&P500 stocks closes at 4PM New York time. The performance of the optimal timing strategy is qualitatively similar when trading at the midquote instead of the bid/ask.

Secondly, high-frequency returns on the S&P500 index display considerable positive serial correlation when measured at sub five-minute intervals. The autocorrelation gives value to the stop-loss, which we demonstrate by taking the observed time-series of S&P500 returns and the trading strategy generated in the Black-Scholes model, but, rather than trading at market prices, trade the options at Black-Scholes prices with a constant volatility.

Thirdly, the presence of volatility as a priced risk factor with a negative price of risk makes delaying purchasing the straddle more attractive, because it has a positive exposure to volatility risk. A simulation study in which we apply the trading strategy derived under the Black-Scholes assumptions on a sample of option prices and underlying asset returns generated from the Heston (1993) stochastic volatility model confirms this. The negative correlation between innovations in the asset price and its volatility has a small dampening effect on the magnitude of gains and losses.

The optimal timing strategy was derived under the assumption of risk-neutrality, so for any risk-averse agent the utility improvement gained by using the optimal timing strategy relative to some fixed-time trading strategy provides a lower bound on the potential utility improvement. Buying the S&P500 straddle using the optimal timing strategy, a CRRA-trader with risk aversion coefficient equal to two would be willing to pay up to 50bps of the value of the straddle to switch from a fixed-time strategy to the optimal timing strategy.

References

Almgren, R., and N. Chriss (2000): "Optimal Execution of Portfolio Transactions," *Journal of Risk*, 3, 5–39.

Bakshi, G., C. Cao, and Z. Chen (1997): "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance*, 52(5), 2003–2049.

Bakshi, G., and N. Kapadia (2003): "Delta-Hedged Gains and the Negative Market Volatility Risk Premium," *Review of Financial Studies*, 16(2), 527–566.

Bender, C., and N. Dokuchaev (2014): "A first-order BSPDE for swing option pricing," *Mathematical Finance*, pp. n/a–n/a.

Benth, Fred Espen; Lempa, J. K. N. T. (2011): "On the optimal exercise of swing options in electricity markets," *Journal of Energy Markets*, 4(4), 3–28.

Bertsimas, D., L. Kogan, and A. W. Lo (2000): "When is time continuous," *Journal of Financial Economics*, 55(2), 173–204.

Bertsimas, D., and A. W. Lo (1998): "Optimal Control of Execution Costs," *Journal of Financial Markets*, 1, 1–50.

Bollen, N. P., and R. E. Whaley (2004): "Does Net buying Pressure Affect the Shape of Implied Volatility Functions," *Journal of Finance*, 59, 711–752.

Bollerslev, T., M. Gibson, and H. Zhou (2011): "Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities," *Journal of Econometrics*, 160(1), 235–245.

Bollerslev, T., G. Tauchen, and H. Zhou (2009): "Expected stock returns and variance risk premia," *Review of Financial Studies*, 22(11), 44634492.

Bollerslev, T., and V. Todorov (2011): "Tails, fears, and risk premia," *The Journal of Finance*, 66(6), 2165–2211.

Brugler, J. (2014): "Dark Trading and Intraday Market Quality," University of Cambridge working paper.

Carmona, R., and S. Dayanik (2008): "Optimal Multiple Stopping of Linear Diffusions," *Mathematics of Operations Research*, 33(2), 446–460.

Carmona, R., and N. Touzi (2008): "Optimal Multiple Stopping and Valuation of Swing Options," *Mathematical Finance*, 18(2), 239–268.

Carr, P., and L. Wu (2009): "Variance Risk Premiums," *Review of Financial Studies*, 22(3), 1311–1341.

Chan, L. K., and J. Lakonishok (1995): "The Behavior of Stock Prices Around Institutional Trades," *The Journal of Finance*, 50(4), 1147–1174.

Cho, Y.-H., and R. F. Engle (1999): "Modeling The Impacts of Market Activity On Bid-Ask Spreads In The Option Market," working paper, University of California San Diego.

Christoffersen, P., M. Fournier, and K. Jacobs (2013): "The factor structure in equity options," University of Toronto working paper.

Corsi, F. (2009): "A Simple Approximate Long-Memory Model of Realized Volatility," *Journal of Financial Econometrics*, 7(2), 174–196.

Criscuolo, A. M., and H. Waelbroeck (2014): "Effect of Volatility Fluctuations on Optimal Execution Schedules," Portware LLC working paper.

Danielsson, J., H. S. Shin, and J.-P. Zigrand (2012): "Procyclial leverage and endogenous risk," London School of Economics working paper.

Dert, C., and B. Oldenkamp (2000): "Optimal Guaranteed Return Portfolios and the Casino Effect," *Operations Research*, 48(5), 768–775.

Ding, S., J. Hanna, and T. Hendershott (2014): "How slow is the NBBO? A comparison with direct exchange feeds," *The Financial Review*, 49(2), 313–332.

Dixit, A. K., and R. S. Pindyck (1994): *Investment Under Uncertainty*. Princeton University Press, Princeton, New Jersey.

Driessen, J., P. J. Maenhout, and G. Vilkov (2009): "The price of correlation risk: Evidence from equity options," *The Journal of Finance*, 64(3), 1377–1406.

Duan, J.-C., and J. Wei (2009): "Systematic risk and the price structure of individual equity options," *Review of Financial Studies*, 22(5), 1981–2006.

Duarte, J., X. Lou, and R. Sadka (2006): "Can Liquidity Events explain the low-short-interest puzzle? Implications from the options market," University of Washington Working Paper.

Easley, D., M. M. L. de Prado, and M. O'Hara (2012): "The Volume Clock: Insights into the High-Frequency Paradigm," *Journal of Portfolio Management*, 39(1), 19–29.

Engle, R. F. (2014): "Dynamic Conditional Beta," NYU Stern Working Paper.

Fahlenbrach, R., and P. Sandås (2010): "Does information drive trading in option strategies," *Journal of Banking & Finance*, 34, 2370–2385.

Fisher, L. (1966): "Some New Stock-Market Indexes," *Journal of Business*, 39(1), 191–225.

Fleming, J., B. Ostdiek, and R. E. Whaley (1996): "Trading costs and the relative rates of price discovery in stock, futures, and option markets," *Journal of Futures Markets*, 16(4), 353–387.

Froot, K. A., and A. F. Perold (1995): "New Trading Practices and Short-Run market efficiency," *Journal of Futures Markets*, 15, 731–765.

Gârleanu, N., and L. H. Pedersen (2011): "Margin-based Asset Pricing and Deviations from the Law of One Price," *Review of Financial Studies*, 24(6), 1980–2022.

Gârleanu, N., L. H. Pedersen, and A. M. Poteshman (2009): "Demand-Based Option Pricing," *Review of Financial Studies*, 22(10), 4259–4299.

Gatheral, J. (2009): "No-Dynamic-Arbitrage and Market Impact," Working Paper, New York University.

Haggstrom, G. W. (1967): "Optimal Sequential Procedures when More Than one Stop is Required," *The Annals of Mathematical Statistics*, 38(6), 1618–1626.

Harris, L., and J. Hasbrouck (1996): "Market vs. Limit Orders: The SuperDOT Evidence on Order Submission Strategy," *The Journal of Financial and Quantitative Analysis*, 31(2), 213–231.

Hasbrouck, J. (1991): "Measuring The Information Content of Stock Trades," *Journal of Finance*, 46(1), 179–207.

Hasbrouck, J., and G. Saar (2013): "Low-latency trading," *Journal of Financial Markets*, 16(4), 646–679.

Heber, G., A. Lunde, N. Shephard, and K. Sheppard (2009): "Oxford-Man Institute's realized library," version 0.1.

Heston, S. L. (1993): "A Closed-Form Solution For Options With Stochastic Volatility With Applications To Bond and Currency Options," *Review of Financial Studies*, 6(2), 327–343.

Heston, S. L., and S. U. Nandi (2000): "A closed-form GARCH option valuation model," *Review of Financial Studies*, 13(3), 585–625.

Huang, D., and I. Shaliastovich (2014): "Volatility-of-volatility risk," University of Pennsylvania Working Paper.

Huberman, G., and W. Stanzl (2004): "Price Manipulation and Quasi-Arbitrage," *Econometrica*, 72(4), 1247–1275.

——— (2005): "Optimal Liquidity Trading," *Review of Finance*, 9, 165–200.

Hull, J. (2012): *Options, futures and other derivatives*. Pearson, 8th edn.

Kloeden, P., and E. Platen (1999): *Numerical solution of stochastic differential equations*. Springer, 3rd edn.

Kukush, A. G., Y. S. Mishura, and G. M. Shevchenko (2006): "On Reselling of European Option," *Theory of Stochastic Processes*, 12(28)(3-4), 80–93.

Liu, J., and J. Pan (2003): "Dynamic Derivative Strategies," *Journal of Financial Economics*, 69, 401–430.

Muravyev, D. (2014): "Order Flow and Expected Option Returns," Boston College Working Paper.

Muravyev, D., and N. D. Pearson (2015): "Option trading costs are lower than you think," Boston College Working paper.

Muravyev, D., N. D. Pearson, and J. P. Broussard (2013): "Is There Price Discovery in Equity Options?," *Journal of Financial Economics*, 107(2), 259–283.

Ni, S. X., J. Pan, and A. M. Poteshman (2008): "Volatility Information Trading in the Option Market," *Journal of Finance*, 63(3), 1059–1091.

Papahristodoulou, C. (2004): "Option strategies with linear programming," *European Journal of Operational Research*, 157, 246–256.

Peskir, G., and A. Shiryaev (2006): *Optimal Stopping and Free-Boundary Problems*. Birkhäuser, Basel.

Rouah, F. D. (2013): *The Heston model and its extensions in MATLAB and C#*. Wiley Finance Series.

Rubinstein, M. (1985): "Nonparametric Tests of Alternative Option Pricing Models Using All Reported Trades and Quotes on the 30 Most Active CBOE Option Classes from August 23, 1976 Through August 31, 1978," *The Journal of Finance*, 40(2), 455–480.

——— (1994): "Implied Binomial Trees," *The Journal of Finance*, 49(3), 771–818.

——— (2000): "On the Relation Between Binomial and Trinomial Option Pricing Models," *The Journal of Derivatives*, 8(2), 47–50.

Santa-Clara, P., and A. Saretto (2009): "Option strategies: Good deals and margin calls," *Journal of Financial Markets*, 12, 391–417.

Schied, A., T. Schöneborn, and M. Tehranchi (2010): "Optimal Basket Liquidation for CARA Investors is Deterministic," *Applied Mathematical Finance*, 17(6), 471–489.

Shephard, N., and K. Sheppard (2010): "Realising the future: forecasting with high-frequency-based volatility (HEAVY) models," *Journal of Applied Econometrics*, 25(2), 197–231.

Trade Magazine (2012): "The 2012 Algorithmic Trading Survey," Online.

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005): "A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100(472), 1394–1411.

A Proofs

Proof of Proposition 2.1: We follow the reasoning of Dixit and Pindyck (1994). The agent's optimization problem is given by

$$V(S_0, 0) = \min_{0 \leq \tau \leq T} E\left\{h \exp(-r\tau) f(S_\tau, \tau)\right\}, \tag{20}$$

where the minimum is taken over all stopping times $\tau \in [0, T]$. In the following, let f_x be the partial derivative of $f(.,.)$ with respect to x. By Itô's Lemma, the dynamics of the discounted portfolio value $h\exp(-rt)f(S_t, t)$ is given by

$$he^{-rt}\left\{\left[-rf(S_t,t) + f_s\mu(S_t,t)S_t + \frac{1}{2}f_{ss}\sigma^2(S_t,t)S_t^2 + f_t\right]\mathrm{d}t + f_s\sigma(S_t,t)S_t\mathrm{d}W_t\right\}. \tag{21}$$

We employ the Black-Scholes partial differential equation (PDE) to show that the bracketed term only contains the product of f_s and $(\mu(s,t) - r)s$. This no-arbitrage argument yields

$$f_t = rf(s,t) - f_s rs - \frac{1}{2}f_{ss}\sigma^2(s,t). \tag{22}$$

Plugging in (22) into (21), the dynamics of the discounted portfolio value $h\exp(-rt)f(S_t,t)$ becomes

$$he^{-rt}\left\{f_s\left[\mu(S_t,t)-r\right]S_t\mathrm{d}t+f_s\sigma(S_t,t)S_t\mathrm{d}W_t\right\}. \tag{23}$$

To prove the first statement we need to show that the expectation of the discounted portfolio value at any stopping time is at least equal to the current value.

Assume $hf_s\left[\mu(s,t)-rs\right]>0\ \forall(s,t)$ and the standard regularity conditions on $\mu(s,t)$ and $\sigma(s,t)$, so that the discounted portfolio value is a submartingale. The optional stopping theorem (see for example Peskir and Shiryaev (2006), Theorem A2 on p60) for a submartingale states that

$$E\left\{he^{-r\tau}f(S_\tau,\tau)\mid\mathcal{F}_0\right\}\geq hf(S_0,0)\ \forall\tau\in[0,T],$$

which proves the first statement. The second and third statement follow similarly by showing that the discounted portfolio value process is a supermartingale (second statement) or a martingale (third statement). □

B Simulation results

We conduct a series of Monte Carlo simulations to examine the effects of including a stop-loss level, study the impact of relaxing assumptions about bid-ask spreads and the leverage effect/variance risk premium. In the default setting, we simulate 100,000 paths of an underlying asset that follows a Geometric Brownian Motion, simulated on a time step of 15 seconds with 16% annualized return volatility, a continuously-compounded risk-free rate of 2.4%, an equity premium of 5% and a dividend rate of 1.8%. In the Heston model, the correlation between the two Brownian motions (ρ) is set to either 0 or -0.8, and the constant (λ) in the variance risk premium is set to either 0 or -1.8, which is similar to the average value found in Bollerslev, Gibson, and Zhou (2011). The volatility-of-volatility is set to 1, the long-run average variance under the physical distribution and the current spot variance are set to $0.16^2 = 0.0256$, and the mean reversion parameter under the physical distribution is set to 5. We use the same seed for each simulation, so the Brownian motion driving the asset price takes the same values in each model.

The straddles to be purchased have a maturity of 1 month, and the trading horizon equals 1 day. The moneyness is defined as the ratio of the strike over the forward price at the start of the trading period; we analyse the moneyness levels used in Section 4: $0.98, 0.99, 1, 1.01, 1.02$ and 1.05. In the default set up and when studying the impact of intraday seasonality in bid-ask spreads, the Black-Scholes formula is used to compute the price of the straddle. Option prices in the Heston model are computed using the code for direct integration of the Heston characteristic function by Rouah (2013). For all simulations, the stopping boundary that determines the trading epoch is derived in the frictionless market underlying the dynamic programming algorithm of Section 2.2, using a trinomial tree with a time-step equal to 15 seconds. Where applicable, the stop-loss level is set to a constant -1% intraday log-return measured from the start of the trading period. The gain is computed using (10).

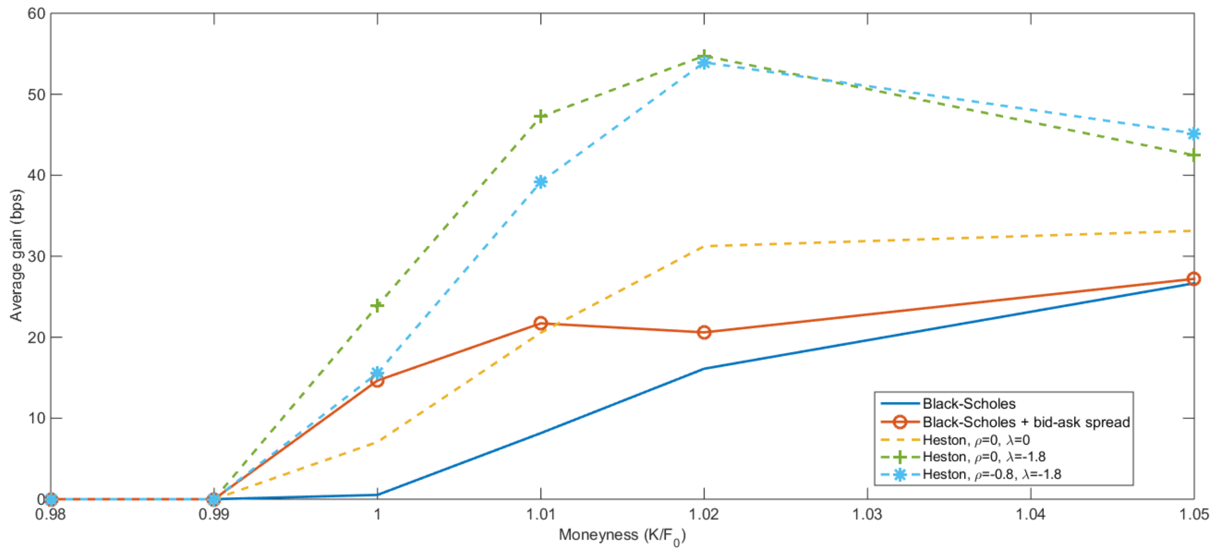
For the different simulation models, Figure 13 plots the average gains of the optimal timing strategy against the usual benchmark strategies of trading at either the open or close of the market, or a TWAP strategy. The Black-Scholes model without any modifications (solid blue line) is included as a comparison; the pattern of gains as a function of moneyness is very similar to the one displayed in Figure 1, albeit less smooth because fewer moneyness levels are analysed here. As expected, adding the stop-loss does not add value to our strategy in the Black-Scholes world as shown in Figure 13(d).

Adding a transaction cost in the form of a time-varying average bid-ask spread as displayed in Figure 8 to the Black-Scholes option prices increases the gain of the optimal timing strategy compared to buying at the open by an economically meaningful 15-20 bps for ATM straddles. The spread at the close is lower than during most of the rest of the day, which explains why the gains of the optimal timing strategy are low when comparing to trading at the close. The exception is the 0.98 moneyness straddle, which the optimal timing strategy always buys at the open and hence it has a fixed price in the simulations where the starting values are the same for each run. Trading at the close will incur a lower transaction cost, but it turns out that the variance of the close price from the perspective of the open is so large that it more than offsets the effect of the spread. Adding a stop-loss to the optimal timing strategy in this setting affects performance adversely, albeit by an almost negligible single basis point at most.

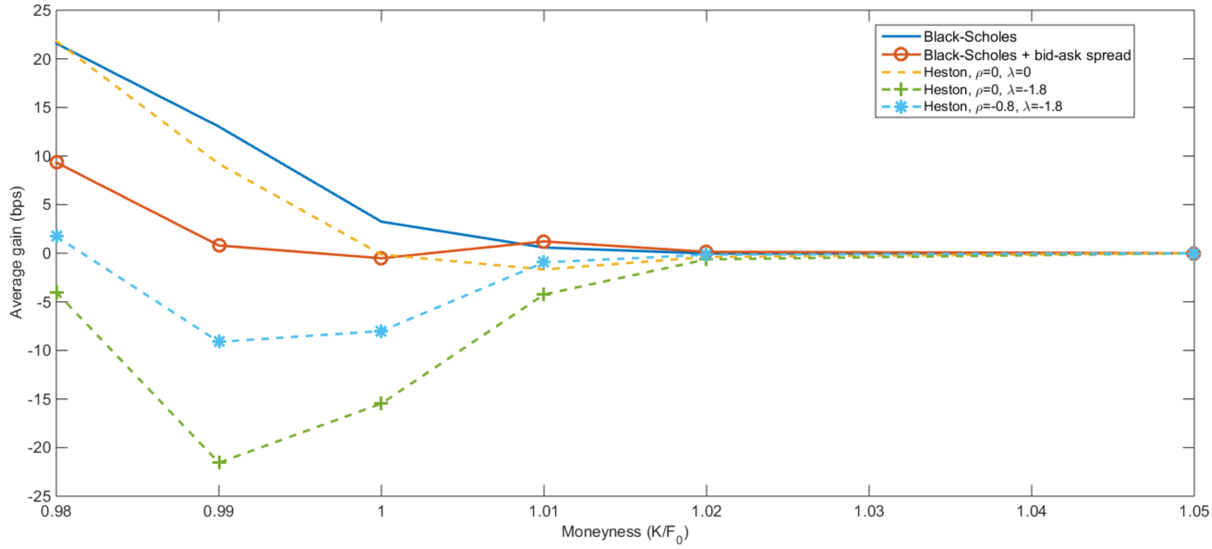
The pattern of gains from using the optimal timing strategy in the Heston model simulation is similar to that of the Black-Scholes with transaction costs, but it is driven by an additional factor. Without a leverage effect or variance risk premium, the gains are slightly higher than for Black-Scholes when compared to the open, which is caused by the combination of the specific simulation run, the relatively high volatility-of-volatility and the relatively high persistence of volatility shocks, causing the average spot variance to decrease slightly over the day, thereby lowering option prices later in time. Repeating the simulation with a mean reversion parameter equal to 100 will yield gains that are almost indistinguishable from those in the Black-Scholes model.

Adding a (negative) variance risk premium leads to higher option prices and, through (15), additional value of waiting and thus higher gains. As argued in Section 4.6, the optimal stopping boundary for a straddle in the Heston model lies above the Black-Scholes boundary, so the Black-Scholes-based boundary used here will trigger trading too early, leading to lower, even negative, gains when the benchmark strategy is buying at the close. The optimal timing strategy outperforms the TWAP strategy for most moneyness levels, so timing the transaction still has value, even when the trading trigger is derived in a misspecified model. The stop-loss does not add value in the Heston model, the reason being that adding the stop-loss implies that trading is potentially triggered earlier in time, implying a larger exposure to the negative variance risk premium. Extending the Heston model to also incorporate the leverage effect results in a reduction of both gains and losses compared to all benchmark strategies, but also further lowering the value of adding a stop-loss. In unreported results, we verify that when the persistence of the volatility shocks is very high ($\kappa = 0.156$ and $\tilde{\kappa} = 0.1$), then the stop-loss does add value for straddles with a moneyness around 1.0.

Comparing the patterns of gains in Figure 13 to those in Table 1, we conclude that the different simulation models, in combination with the results in Figure 10, go a long way in explaining the difference between the gains in the Black-Scholes world and those in the S&P500 data.

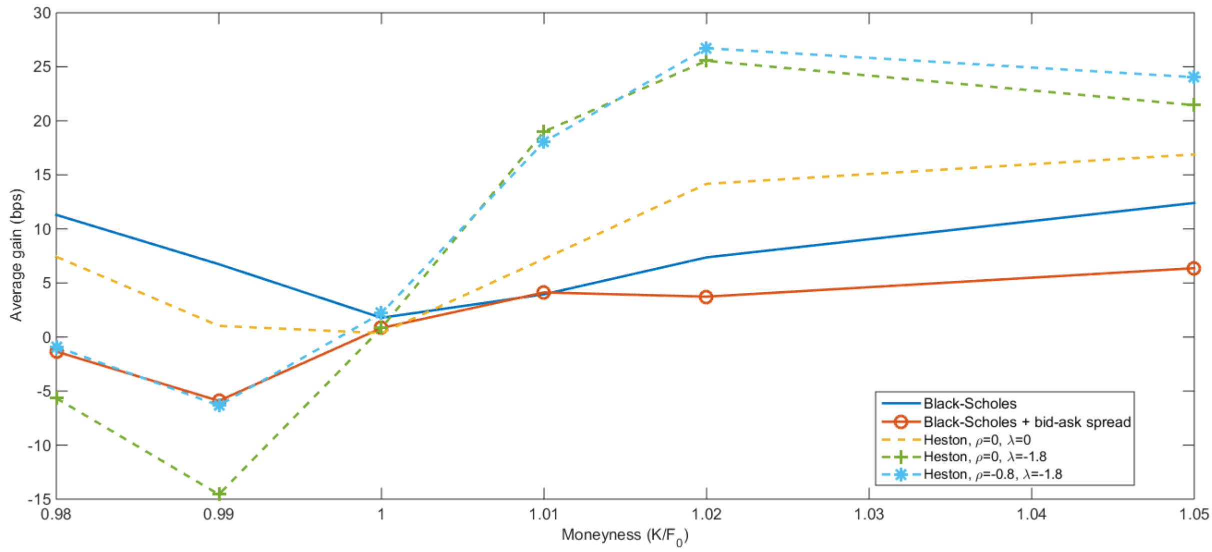


(a) Optimal timing versus Open

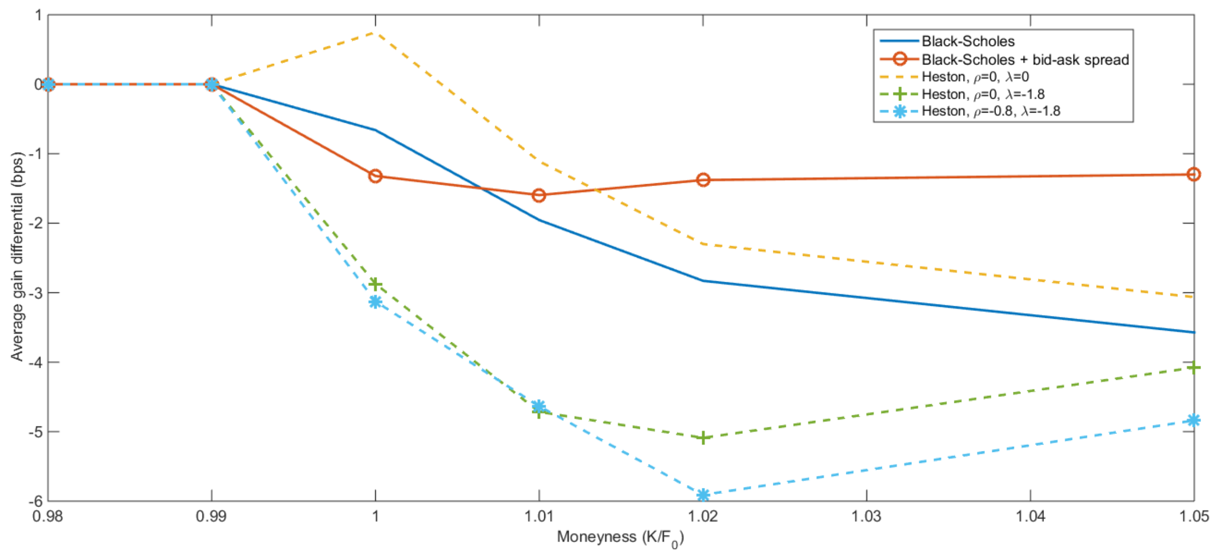


(b) Optimal timing versus Close

Figure 13: Realized gains for purchasing a straddle in simulation studies with parameter values as outlined in Appendix B. "Black-Scholes" refers to the standard Black-Scholes model analysed in Section 3. "Black-Scholes + bid-ask spread" assumes Black-Scholes dynamics and theoretical option prices, but the price paid includes a spread that is time-varying using the average value in Figure 8. The three versions of the Heston model differ in the assumption about the correlation between return and volatility innovations (ρ), and the presence of a variance risk premium (λ).



(c) Optimal timing versus TWAP



(d) Differential gain of including stop-loss

Figure 13: Realized gains for purchasing a straddle in simulation studies with parameter values as outlined in Appendix B (cont'd).