Received: 22 December 2022 | Accepted: 25 March 2023

DOI: 10.1002/fut.22419

RESEARCH ARTICLE

FUTURES MARKETS WILEY

American strangle options with arbitrary strikes

Tsvetelin S. Zaevski

Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, Bulgaria

Correspondence
Tsvetelin S. Zaevski, Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Acad. Georgi Bonchev Str., Block 8, 1113 Sofia, Bulgaria.
Email: t_s_zaevski@abv.bg and t_s_zaevski@math.bas.bg

Funding information
Science and Education for Smart Growth Operational Program (2014–2020) and cofinanced by the European Union through the European Structural and Investment funds, Grant/Award Number: BG05M2OP001-1.001-0003; Bulgarian National Science Fund, Grant/Award Number: KP-06-N32/8

Abstract

The so-called American strangle options are examined in this paper. Their main characteristic is the combined put and call feature. The holder has the right to exercise prematurely choosing the option's style—put or call. We abandon the traditional assumption that the put strike is below the call one considering arbitrary values. We also assume that the put and call weights are different. The equations for the early exercise boundaries are derived in the perpetual case. After that we approximate numerically these boundaries for the finite maturity options maximizing the option holder's utility. On the basis of them we apply a Crank–Nicolson finite difference method to the corresponding Black–Scholes-style partial differential equation to obtain the fair option price.

KEYWORDS
American strangle options, Crank–Nicolson method, optimal regions, pricing

JEL CLASSIFICATION
C41, G12, G13

1 | INTRODUCTION

The strangle options appear as a financial instrument which preserves against the financial risk when the investor expects some large deviations but he is not sure for the direction of these movements. This happens in the high volatility periods. In fact this is a very often situation supported by the volatility clustering financial market phenomena. Roughly said, the strangle options are a combination between an American put and call. Its holder has the right how to exercise the option—as a put or as a call. This way the strangles keep the investor's interest from both upward and downward shocks. Obviously, these options lead to two-sided optimal stopping problem. The owner would exercise the option as a call if the underlying asset reaches some large enough value. On the contrary, if the asset falls enough, then the holder will exercise the option as a put.

Usually, the options without maturity restrictions are easier to research due to the absence of the coercive exercise at the maturity. As a consequence, the optimal boundaries are flat. First results for these perpetual options can be found in Beibel and Lerche (1997) and Shiryaev (1999), see also Gapeev and Lerche (2011) and Qiu (2020). Later, several authors turn to the strangle pricing problem under a finite maturity horizon. A major subclass of such options is the so-called straddles. Their main feature is the matching put and call strikes. A Laplace transform method is applied to the American straddle options in Alobaidi and Mallier (2002). The same authors examine the behavior of these options near the maturity—see Alobaidi and Mallier (2006). Chiarella and Ziogas (2005) overcome some problems with the Laplace transform using the Fourier one. Alternatively, Kang et al. (2017) use the Laplace–Carlson transform. An approach based on deriving the limits for the boundaries by the use of capping is presented by Ma et al. (2018). The variational inequalities method is presented by Jeon and Oh (2019). This method is closely related to the corresponding two-sided free boundary differential problem which describes the strangle pricing. Another such approach is presented by Qiu (2020)—the related integral equations are derived via the so-called early exercise premium (EEP). Also, Abdou and Moraux (2016) use an EEP method for pricing and hedging. Recently, Jeon and Kim (2022) derive an analytic valuation formula under mean-reversion assumptions.

In all these works, the assumption is that the put strike is less than the call strike (only for the straddles they are equal). We remove this restriction considering arbitrary strikes. We also assume that the put and call features are presented with different weights.

We establish our model in terms of Shiryaev et al. (1995). We assume that the asset does not pay dividends, but we introduce an extra discount factor. It turns out that a dividend model can be translated into an additionally discounted nondividend one. Also, this way we introduce a time structure in the option payoff. Our first step is to prove several propositions which give the shape of the optimal regions. It turns out that the put-optimal region consists of all points below some function—this function is the put exercise boundary. Analogously all points above another function—the call-optimal boundary—are call-optimal.

The importance of the asymptotic case motivates us to consider first the perpetual strangles. As we mentioned above the optimal boundaries are flat. We derive the equations which they have to solve and prove that these equations have unique solutions. Something more, we show that the obtained roots lead to correct boundaries—the call one is above the boundary value at the maturity, whereas the put one is below. The approach we use is based on some exit properties of a Brownian motion from a strip. This way we consider the strangle option price as a two-dimensional function w.r.t. its boundaries and search for its maximum. A similar method is applied in Zaevski (2020a, 2020b), but for another class of American-style derivatives, namely, the game options—put and call. They lead again to a two-dimensional optimization problem. However, there exists a significant difference—while the game options lead to a sup–inf (min–max) problem, a max–max problem arises for the strangles. This way we are looking for a saddle point for the game options, but for the maximum of the strangle price w.r.t. both variables.

As for the ordinary American calls, the undiscounted case is specific—it is never optimal for a strangle's holder to exercise earlier the option as a call. This allows us to derive close-form formulas for the put boundary and for the price. It is interesting fact that the existing call right influences the option although this right never will be used. Something more this call impact appears only through the corresponding call weight, but not through the call strike.

Having in mind the results for the perpetual options, we turn to the finite maturity horizon. It turns out that the put boundary is a decreasing function w.r.t. the time to maturity, whereas the call one increases. Using this and the derived terminal values (at the maturity and at the infinity), we construct an algorithm to approximate the whole boundaries. It is based on maximizing the financial outcome of the option holder. Once we estimate the optimal boundaries, the free boundary problem for strangle pricing turns to a boundary value problem in a given region. We use the Crank–Nicolson finite difference approach to solve numerically this equation.

The paper is organized as follows. In Section 2 we establish our model. The shape of the optimal regions is obtained in Section 3. The perpetual options are considered in Section 4, whereas the finite maturity case is examined in Section 5. Some numerical results are provided in Section 6. We conclude in Section 7.

2 | PRELIMINARIES

Let the asset price be presented by the log-normal process

$$dS_t = rS_t\,dt + \sigma S_t\,dB_t \tag{1}$$

under the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, Q)$, where B_t is a Brownian motion and the measure Q is risk neutral. Assume that the risk-free interest rate and the additional discount factor are the constants r and λ, respectively. We do not impose a positiveness for the risk-free rate, but we require $\lambda \geq 0$ and $r + \lambda > 0$. Let $T \leq \infty$ is the maturity date—note that we allow its absence when $T = \infty$. We shall denote by t the current time and by $\tau = T - t$ the time to maturity—we use t and τ without writing explicitly their meaning hereafter. As we mentioned above, the option's holder may choose the option feature when he exercises. If he prefers a put characteristic, he receives $C_1 > 0$ shares of a put option with the strike K_1. Analogously, if the holder chooses a call feature, he receives the payoff of $C_2 > 0$ call options with the strike K_2. Hence, the payoff of the strangle option can be written as

$$N(t, x) = e^{-\lambda t}\max\{C_1(K_1 - x)^+, C_2(x - K_2)^+\}. \tag{2}$$

This means that the option's holder would receive the amount of $N(t,x)$ if he exercises the option in moment t at value $S_t = x$. The following simple proposition gives the relation between the additional discount factor and the dividends. Its proof can be found in Zaevski (2020a), Proposition 2.3.

Proposition 2.1. *If in addition the underlying asset pays continuously dividends at rate δ, then this model is equivalent to a nondividend model with risk-free rate $r - \delta$ and discount factor $\lambda + \delta$.*

This parametrization is used in Shiryaev et al. (1995). Note that the risk-free rate is not necessarily positive. Proposition 2.1 allows us to examine dividend models in our nondividend framework.

Let D_0 be the value which makes the put and call payoffs equal

$$D_0 := \frac{C_1K_1 + C_2K_2}{C_1 + C_2}. \tag{3}$$

We shall name it put–call barrier. We shall denote the strangle option price by the function $F(t,x)$ assuming that in the moment t the underlying asset value is $S_t = x$. Obviously, if the immediate exercise is optimal in the point (t,x) then $F(t,x) \equiv N(t,x)$. We denote by $\mathcal{T}_{[t,T]}, t < T$, the set of all stopping times with values in the interval $[t,T]$. We view them as the possible strategies for the option's holder, that is, the exercise happens when the stopping time occurs. We are interested in the optimal strategy, that is, the stopping time which maximizes the financial result of the option. In the following definition we formalize this.

Definition 2.1. Suppose that the option is alive at the moment t and the current asset value is $S_t = x$—we denote by $E^{t,x}$ the expectation just under the assumption $S_t = x$. The strategy ζ is optimal if it maximizes the expected future option flows

$$E^{t,x}[e^{-r\zeta}N(\zeta, S_\zeta)]. \tag{4}$$

The function $N(t,x)$ is given in Equation (2).

The next step is to define the so-called early exercise (or optimal) and continuation regions and the corresponding boundaries. We shall denote them by Υ and $\overline{\Upsilon}$, respectively.

Definition 2.2.

1. A point $(t,x) \in \Upsilon$ if for every $\zeta \in \mathcal{T}_{[t,T]}$

$$N(t,x) \geq E^{t,x}[e^{-r(\zeta - t)}N(\zeta, S_\zeta)]. \tag{5}$$

2. The region in which the option's holder exercises the option as a put/call shall be named put/call-optimal region and shall be denoted by Υ^p and Υ^c, respectively. Note that $\Upsilon = \Upsilon^p \cup \Upsilon^c$. Let us state for convenience that if the call and put features lead to one and the same result, the option is exercised as a call. This is possible only when $x = D_0$. Under this assumption we can think that $\Upsilon^p \cap \Upsilon^c = \varnothing$.
3. The continuation region is defined as $\overline{\Upsilon} = \{[t,T] \times (0,\infty) \setminus \Upsilon\}$.
4. We have two early exercise boundaries—between the continuation region and the put- or call-optimal region. We shall denote these boundaries by $A(t)$ and $B(t)$, respectively.

The following proposition gives the time dependence of the option price.

Proposition 2.2. *If the asset price at moment t is x, $S_t = x$, then the option price is given by $F(t,x) = e^{-\lambda t}F(0,x)$.*

Proof. See the proof of Proposition 2.2 from Zaevski (2020a)—although this paper is devoted to a different class of the American-style options, namely, the game options, the present proposition can be proven in a very similar way. □

We assume hereafter that the initial moment is zero. Also, we impose the following natural assumption.

Condition 2.1. Let us mark the initial value of the process S_t by a superscript, that is, S_t^x means the value of the process at time t conditioned on $S_0 = x$.

1. If $x < y$, then $S_t^x < S_t^y$ for every t.
2. For a fixed sample path, the asset price tends to infinity, if its initial value tends to infinity.
3. For a fixed sample path, the asset price tends to zero when its initial value tends to zero.

3 | OPTIMAL REGIONS

We shall provide a series of propositions for the form of the early exercise regions. Remind that we denote by $E^{t,x}$ the expectation under the assumption $S_t = x$. The first one states that call early exercising is never optimal when $\lambda = 0$, a fact appearing for many different derivatives with American call features.

Proposition 3.1. *We have $\Upsilon^c \equiv \varnothing$ when $\lambda = 0$.*

Proof. We have that $r > 0$ since $r + \lambda > 0$ and $\lambda = 0$. Suppose that the set Υ^c is not empty and it contains the point (t, x). Note that the value x has to be above the call strike, $x > K_2$. Let ζ be a stopping time from the set $\mathcal{T}_{[t,T]}$. We use that $e^{-rt}S_t$ is a martingale to obtain

$$\begin{aligned} E^{t,x}[e^{-r\zeta}N(\zeta, S_\zeta)] &\le e^{-rt}C_2(x - K_2) \\ &= E^{t,x}[e^{-r\zeta}C_2S_\zeta] - C_2K_2e^{-rt} \\ &< E^{t,x}[e^{-r\zeta}C_2(S_\zeta - K_2)] \\ &\le E^{t,x}[e^{-r\zeta}C_2(S_\zeta - K_2)^+] \\ &\le E^{t,x}[e^{-r\zeta}\max\{C_1(K_1 - S_\zeta)^+, C_2(S_\zeta - K_2)^+\}] \\ &\equiv E^{t,x}[e^{-r\zeta}N(\zeta, S_\zeta)]. \end{aligned} \tag{6}$$

The contradiction leads to $\Upsilon^c \equiv \varnothing$. □

The next proposition shows that in the opposite case, that is, $\lambda > 0$, the call-optimal region Υ^c is not empty.

Proposition 3.2. *If $\lambda > 0$, then the call-optimal region Υ^c is not empty. Also if $(t, x) \in \Upsilon^c$ and $y > x$, then $(t, y) \in \Upsilon^c$.*

Proof. Assume that $x > \max\{K_1, K_2\}$ and therefore $(t, x) \notin \Upsilon^p$. Hence, the immediate exercise has the result $e^{-\lambda t}(x - K_2)$. Let ζ be some stopping time from the set $\mathcal{T}_{[t,T]}$ and let us compare it with the immediate exercise. We shall denote by $f(t, x; \zeta)$ the difference between the outcomes of the strategy ζ and the immediate exercise. Having in mind that the process $e^{-rt}S_t$ is a martingale and thus $x = E^{t,x}[e^{-r(\zeta - t)}S_\zeta]$ we present the function $f(t, x; \zeta)$ as

$$\begin{aligned} f(t, x; \zeta) &:= E^{t,x}[e^{-r\zeta}N(\zeta, S_\zeta)] - e^{-(r+\lambda)t}C_2(x - K_2) \\ &= E^{t,x}\left[e^{-(r+\lambda)\zeta}\max\begin{pmatrix} C_2K_2e^{(r+\lambda)(\zeta-t)} + C_1K_1 - S_\zeta(C_1 + C_2e^{\lambda(\zeta-t)}), \\ -e^{\lambda(\zeta-t)}C_2S_\zeta + e^{(r+\lambda)(\zeta-t)}C_2K_2, \\ C_2K_2(e^{(r+\lambda)(\zeta-t)} - 1) - S_\zeta C_2(e^{\lambda(\zeta-t)} - 1), \\ -e^{\lambda(\zeta-t)}C_2S_\zeta + e^{(r+\lambda)(\zeta-t)}C_2K_2 \end{pmatrix}\right]. \end{aligned} \tag{7}$$

Due to the second statement of Condition 2.1 we conclude that for every stopping time ζ function (7) is negative for a large enough initial value x. Therefore immediate exercising is optimal for these initial conditions. Hence, the call-optimal region Υ^c is not empty.

Suppose now that $x \in \Upsilon^c$ and $y > x$. Therefore $f(t, x; \zeta) \leq 0$ for every stopping time $\zeta \in \mathcal{T}_{[t,T]}$. The function $f(t, x; \zeta)$ is defined in Equation (7). Let us observe that this function is decreasing w.r.t. the variable x due to the first statement of Condition 2.1 and therefore $f(t, y; \zeta) \leq 0$ too. Since this is true for an arbitrary $\zeta \in \mathcal{T}_{[t,T]}$, the point (t, y) is call-optimal too. □

Below we present the analogue of Proposition 3.2 for the put region.

Proposition 3.3. *The put-optimal region Υ^p is not empty. Something more, if $(t, x) \in \Upsilon^p$ and $y < x$, then $(t, y) \in \Upsilon^p$ too.*

Proof. We shall proceed in a similar way. Below we define the comparison function for an arbitrary stopping time $\zeta \in \mathcal{T}_{[t,T]}$ taking into attention that $N(t, x) = e^{-\lambda t} C_1(K_1 - x)$ for small enough initial conditions x. Let us denote again by $f(t, x; \zeta)$ the difference between the outcomes of the strategy ζ and the immediate exercise. Using the martingality of the process $e^{-rt}S_t$ we obtain the function $f(t, x; \zeta)$ as

$$
\begin{aligned}
f(t, x; \zeta) &:= E^{t,x}[e^{-r\zeta} N(\zeta, S_\zeta)] - e^{-(r+\lambda)t} C_1(K_1 - x) \\
&= E^{t,x}\left[e^{-(r+\lambda)\zeta} \max\begin{pmatrix} -C_1K_1(e^{(r+\lambda)(\zeta-t)} - 1) + C_1 S_\zeta(e^{\lambda(\zeta-t)} - 1), \\ e^{\lambda(\zeta-t)} C_1 S_\zeta - e^{(r+\lambda)(\zeta-t)} C_1 K_1, \\ -C_2K_2 - e^{(r+\lambda)(\zeta-t)} C_1K_1 + S_\zeta(C_1 e^{\lambda(\zeta-t)} + C_2), \\ -e^{(r+\lambda)(\zeta-t)} C_1 K_1 + e^{\lambda(\zeta-t)} C_1 S_\zeta \end{pmatrix} \right].
\end{aligned}
\tag{8}
$$

The third statement of Condition 2.1 shows that for small enough initial values x and every stopping time ζ function (8) is negative. Therefore the immediate exercise is optimal for such initial conditions and thus the put-optimal region Υ^p is not empty.

Let x be in the put-optimal region, $x \in \Upsilon^p$, and $y < x$. Hence, $f(t, x; \zeta) \leq 0$ for every stopping time $\zeta \in \mathcal{T}_{[t,T]}$. The function (8) is increasing w.r.t. the variable x since the first statement of Condition 2.1 holds. Therefore $f(t, y; \zeta) \leq 0$ too, which leads to the put-optimality of the point (t, y). □

The following proposition establishes the behavior of the optimal boundaries $A(\tau)$ and $B(\tau)$.

Proposition 3.4. *The put boundary $A(\tau)$ is nonincreasing w.r.t. the time to maturity, whereas the call one $B(\tau)$ is nondecreasing.*

Proof. On the basis of Propositions 3.2 and 3.3 we can use similar arguments to Jacka (1991), Proposition 2.2. □

The next step is to obtain the values of these boundaries at the maturity.

Proposition 3.5. *The value of the put boundary at the maturity is*

$$
D_1 \equiv A(0) = \min\left\{ K_1, \frac{C_1K_1 + C_2K_2}{C_1 + C_2}, \frac{r + \lambda}{\lambda} K_1 \right\}. \tag{9}
$$

Proof. First, note that D_1 cannot be above the put strike K_1 because the option's holder will receive nothing. Also, it cannot be above the put–call barrier (3), because in the opposite case the holder will prefer to exercise the option as a call.

Suppose now that a point near to maturity (t, x) is put-optimal, $(t, x) \in \Upsilon^p$. In this region the option price function $F(t, x)$ has to satisfy the variational inequality

$$
F_t(t, x) + \mathcal{A}F(t, x) - rF(t, x) \leq 0. \tag{10}
$$

We shall find the largest value of x for which inequality (10) holds for the limit function $N(t,x)$, which in this case turns to the put payoff $N(t,x) = e^{-\lambda t}C_1(K_1 - x)$. We have

$$e^{-\lambda t}C_1[-\lambda(K_1 - x) - \mathcal{A}\{x\} - r(K_1 - x)] = C_1[\lambda x - (r+\lambda)K_1]. \tag{11}$$

The largest value of x for which formula (11) is not positive is

$$d_1 = \frac{r+\lambda}{\lambda}K_1. \tag{12}$$

Note that Equation (9) can be written as $D_1 = \min\{K_1, D_0, d_1\}$. Let us examine first the case $r < 0$ which turns D_1 to $D_1 = \min\{D_0, d_1\}$. We shall prove that all points below D_1 are put-optimal near the maturity. Suppose that this is not true for some value $x < D_1$. Note that this point cannot be call-optimal since the option's holder will prefer to exercise the option as a put. Hence the point (t,x) is in the continuation region, $(t,x) \in \overline{\Upsilon}$, and therefore the statement (10) turns to equality in this point. Hence,

$$\begin{aligned}
0 &< \lim_{t\to T}\frac{F(t,x) - N(t,x)}{T-t} \\
&= -\lim_{t\to T}\frac{F(T,x) - F(t,x)}{T-t} + \lim_{t\to T}\frac{N(T,x) - N(t,x)}{T-t} \\
&= \mathcal{A}F(T,x) - rF(T,x) + N_t(T,x) \\
&= C_1[-rxe^{-\lambda T} - re^{-\lambda T}(K_1 - x) - \lambda e^{-\lambda T}(K_1 - x)] \\
&= e^{-\lambda T}C_1[-(r+\lambda)K_1 + \lambda x] < 0.
\end{aligned} \tag{13}$$

The contradiction confirms that $(t,x) \in \Upsilon^p$. Suppose now that $r \ge 0$. Let us mention that in this case $d_1 \ge K_1 > x$ and therefore using the analogous arguments as above we reach to the same contradiction (13). □

Proposition 3.6. *The value of the call boundary at the maturity is*

$$D_2 \equiv B(0) = \max\left\{K_2, \frac{C_1K_1 + C_2K_2}{C_1 + C_2}, \frac{r+\lambda}{\lambda}K_2\right\}. \tag{14}$$

Equation (14) can be written also as $D_2 \equiv B(T) = \max\{K_2, D_0, d_2\}$ *for*

$$d_2 = \frac{r+\lambda}{\lambda}K_2. \tag{15}$$

Proof. The proof is analogous to the proof of Proposition 3.5. The main difference is that we search for the lowest value of x for which inequality (10) holds having in mind that the payoff function turns to $N(t,x) = e^{-\lambda t}C_2(x - K_2)$. □

Remark 3.1. We can see that $D_2 = \infty$ when $\lambda = 0$ which is in accordance with Proposition 3.1.

Corollary 3.1. *We have that* $D_1 \le D_2$. *The equality holds in the following cases:*

1. $r \ge 0$ *and*

$$K_2 \le \frac{\lambda C_1 K_1}{(r+\lambda)C_1 + rC_2}. \tag{16}$$

2. $r < 0$ *and*

$$K_2 \le \frac{(r+\lambda)C_2 + rC_1}{\lambda C_2}K_1. \tag{17}$$

Proof. If $K_1 < K_2$, then $D_1 < D_2$. Hence, the equality may hold only when $K_2 \leq K_1$. Suppose first that $r \geq 0$. Therefore $D_1 = D_0$ and

$$D_2 = \max\left\{D_0, \frac{r+\lambda}{\lambda}K_2\right\}. \tag{18}$$

Hence, $D_2 = D_0$ when $\frac{r+\lambda}{\lambda}K_2 \leq D_0$ which is equivalent to inequality (16). Analogously, if $r < 0$ we can prove that $D_1 = D_2 = D_0$ when inequality (17) holds. □

4 | PERPETUAL OPTIONS

The next step in our study is to obtain both boundaries of a strangle option assuming that $T = \infty$. These boundaries have to be flat since (A) the option's holder has no time horizon, (B) the asset price is a Markov process, and (C) the specific of the payoff function $N(t, x)$.

4.1 | Existence of discounting

Suppose first that $\lambda > 0$. The undiscounted case is considered later. Propositions 3.2 and 3.3 show that the exercise regions have to be of the form $\Upsilon^p = (0, \overline{A}\,]$ and $\Upsilon^c = (\overline{B}, \infty]$ for some constants $\overline{A} < \overline{B}$. They have to satisfy the following conditions:

$$\overline{A} < K_1, \quad \overline{A} < D_0 \equiv \frac{C_1K_1 + C_2K_2}{C_1 + C_2}, \quad \overline{B} > K_2, \quad \overline{B} \geq D_0 \equiv \frac{C_1K_1 + C_2K_2}{C_1 + C_2}, \tag{19}$$

because (A) if $\overline{A} \geq K_1$ or $\overline{B} \leq K_2$, the option's holder will receives nothing, (B) if $\overline{A} \geq D_0$ the holder will exercise the option as a call, and (C) if $\overline{B} < D_0$ the holder will prefer the put feature.

Let us sketch first the approach we use to derive the optimal boundaries. We examine financial derivatives related to the first exit of the underlying asset from a strip (A, B), $A < B$. The constants A and B are chosen in a way to satisfy conditions (19). We examine the price of such derivative as a function of its boundaries A and B. This function is obtained through some exit properties of a Brownian motion from a strip. We have to derive the maximum of this function. We do this by fixing one of the boundaries and finding the value for another which maximizes the price function. We prove the existence and uniqueness of these maxima. Note that we divide all prices (strikes, initial price, and boundaries) to the fixed boundary—we do that to unify the optimization problems to the intervals $(0, 1)$ or $(1, \infty)$. Thus we obtain a two-dimensional system for the boundaries. We extract from this system one polynomial-style equation and prove that it has a unique solution. This way we derive the optimal boundaries and using them we find the fair option price.

Let us denote by ζ^A and ζ^B the first hitting moments to the levels A and B, and let $\zeta = \zeta^A \wedge \zeta^B$. Due to the log-normality of the asset price process, we can view ζ^A and ζ^B as the first hitting times of a Brownian motion with drift

$$\psi = \frac{r}{\sigma} - \frac{\sigma}{2} \tag{20}$$

to the values

$$\begin{aligned} \tilde{A} &= \frac{\ln A - \ln S_0}{\sigma} < 0, \\ \tilde{B} &= \frac{\ln B - \ln S_0}{\sigma} > 0. \end{aligned} \tag{21}$$

If the exit happens from the lower boundary, then the derivative's holder receives the put payoff $e^{-\lambda t}C_1(K_1 - A)$. Analogously, if the asset exits from the upper boundary the derivative pays an amount of $e^{-\lambda t}C_2(B - K_2)$. Hence, if the asset starts from a point x, $A < x < B$, then the price of such derivative has to be

$$\begin{aligned} f(A, B, x) = C_2(B - K_2)E^x\left[e^{-(r+\lambda)\zeta^B}I_{\zeta^B \le \zeta^A}\right] \\ + C_1(K_1 - A)E^x\left[e^{-(r+\lambda)\zeta^A}I_{\zeta^A < \zeta^B}\right]. \end{aligned} \tag{22}$$

Let us define the constants p and q as

$$\begin{aligned} p &:= 2\sqrt{\left(\frac{r}{\sigma^2} - \frac{1}{2}\right)^2 + 2\frac{r+\lambda}{\sigma^2}}, \\ q &:= \sqrt{\left(\frac{r}{\sigma^2} - \frac{1}{2}\right)^2 + 2\frac{r+\lambda}{\sigma^2}} + \frac{r}{\sigma^2} - \frac{1}{2}. \end{aligned} \tag{23}$$

Note that $p \ge q + 1$ and the equality stands when $\lambda = 0$. The expectations in formula (22) can be obtained through the following lemma reported in Borodin and Salminen (2015) as Equations (3.0.5a and b).

Lemma 4.1. *Let ζ be the first exit of a Brownian motion with drift μ from a strip (a, b). Then we have*

$$\begin{aligned} E\left[e^{-y\zeta}I_{\zeta=\zeta^a}\right] &= e^{\mu a}\frac{\sinh(b\sqrt{2y+\mu^2})}{\sinh((b-a)\sqrt{2y+\mu^2})}, \\ E\left[e^{-y\zeta}I_{\zeta=\zeta^b}\right] &= e^{\mu b}\frac{\sinh(-a\sqrt{2y+\mu^2})}{\sinh((b-a)\sqrt{2y+\mu^2})}. \end{aligned} \tag{24}$$

Estimating expectations (24), written for $y = r + \lambda$, in formula (22) we derive

$$f(A, B, x) = C_1(K_1 - A)\left(\frac{A}{x}\right)^q \frac{B^p - x^p}{B^p - A^p} + C_2(B - K_2)\left(\frac{B}{x}\right)^q \frac{x^p - A^p}{B^p - A^p}. \tag{25}$$

Let us fix the value B and make the following change of variables $a = \frac{A}{B}$, $k_1 = \frac{K_1}{B}$, $k_2 = \frac{K_2}{B}$, and $y = \frac{x}{B}$. Note that $k_2 \le 1$ and $C_1 + C_2 - C_1k_1 - C_2k_2 \ge 0$ due to restrictions (19). Price function (25) turns to

$$\begin{aligned} f(a) &= \frac{B}{y^q}\frac{C_1(k_1 - a)a^q(1 - y^p) + C_2(1 - k_2)(y^p - a^p)}{1 - a^p} \\ &= \frac{B}{y^q}\left(\frac{-a^pC_2(1 - k_2) - a^{q+1}C_1(1 - y^p) + a^qC_1k_1(1 - y^p) + C_2(1 - k_2)y^p}{1 - a^p}\right). \end{aligned} \tag{26}$$

The next step is to recognize which value of a maximizes the function

$$g(a) = \frac{-a^pC_2(1 - k_2) - a^{q+1}C_1(1 - y^p) + a^qC_1k_1(1 - y^p) + C_2(1 - k_2)y^p}{1 - a^p} \tag{27}$$

in the interval $(0, 1]$. Its derivative is

$$g_a(a) = \frac{1 - y^p}{(1 - a^p)^2}a^{q-1}\left[\begin{matrix} -a^{p+1}C_1(p - q - 1) + a^pC_1k_1(p - q) \\ -a^{p-q}pC_2(1 - k_2) - aC_1(q + 1) + qC_1k_1 \end{matrix}\right]. \tag{28}$$

Let us define the function $h(\cdot)$ as

$$h(a) = -a^{p+1}C_1(p-q-1) + a^pC_1k_1(p-q) - a^{p-q}pC_2(1-k_2) - aC_1(q+1) + qC_1k_1. \tag{29}$$

We show in Appendix A.1 that derivative (28) has a unique root in the interval $(0, 1]$. Note that this root is just $a = 1$ when $C_1 + C_2 - C_1k_1 - C_2k_2 = 0$. Assume now $C_1 + C_2 - C_1k_1 - C_2k_2 > 0$ and let us consider the open interval $(0, 1)$. The mentioned above root leads to a maximum for function (27) since

$$\begin{aligned} h(0) &= qC_1k_1 > 0, \\ h(1) &= -p(C_1 + C_2 - C_1k_1 - C_2k_2) < 0. \end{aligned} \tag{30}$$

We shall parametrize now w.r.t. the variable a. We can rewrite the equation $h(a) = 0$ as

$$-a^{p+1}C_1(p-q-1) + a^pC_1\frac{K_1}{B}(p-q) - a^{p-q}pC_2\left(1 - \frac{K_2}{B}\right) - aC_1(q+1) + qC_1\frac{K_1}{B}. \tag{31}$$

Hence Equation (31) leads to

$$B(a) = \frac{a^pC_1K_1(p-q) + a^{p-q}pC_2K_2 + qC_1K_1}{a^{p+1}C_1(p-q-1) + a^{p-q}pC_2 + aC_1(q+1)}. \tag{32}$$

Let us denote by $A_1(a)$ the corresponding put boundary. Using $A_1(a) = aB(a)$ we derive

$$\begin{aligned} A_1(a) &= \frac{a^pC_1K_1(p-q) + a^{p-q}pC_2K_2 + qC_1K_1}{a^pC_1(p-q-1) + a^{p-q-1}pC_2 + C_1(q+1)} \\ &= \frac{p-q}{p-q-1}a\frac{a^qC_1K_1 + C_2K_2 + \dfrac{q}{p-q}\left(C_2K_2 + \dfrac{C_1K_1}{a^{p-q}}\right)}{a^{q+1}C_1 + C_2 + \dfrac{q+1}{p-q-1}\left(C_2 + \dfrac{C_1}{a^{p-q-1}}\right)}. \end{aligned} \tag{33}$$

Let us define the following functions:

$$\begin{aligned} X_1(a) &= a(a^qC_1K_1 + C_2K_2), \\ X_2(a) &= a\frac{q}{p-q}\left(C_2K_2 + \frac{C_1K_1}{a^{p-q}}\right), \\ X_3(a) &= a^{q+1}C_1 + C_2, \\ X_4(a) &= \frac{q+1}{p-q-1}\left(C_2 + \frac{C_1}{a^{p-q-1}}\right). \end{aligned} \tag{34}$$

In such a way the function (33) can be rewritten as

$$A_1(a) = \frac{p-q}{p-q-1}\frac{X_1(a) + X_2(a)}{X_3(a) + X_4(a)}. \tag{35}$$

Let us fix now the boundary A and change the variables for price function (25) as $b = \frac{B}{A}$, $k_1 = \frac{K_1}{A}$, $k_2 = \frac{K_2}{A}$, and $y = \frac{x}{A}$. We have $k_1 > 1$ and $C_1k_1 + C_2k_2 - C_1 - C_2 > 0$ due to limitations (19). Now price function (25) turns to

$$\begin{aligned} f(b) &= \frac{A}{y^q}\frac{C_1(k_1-1)(b^p - y^p) + C_2(b - k_2)b^q(y^p - 1)}{b^p - 1} \\ &= \frac{A}{y^q}\left[\frac{b^pC_1(k_1-1) + b^{q+1}C_2(y^p - 1) - b^qC_2k_2(y^p - 1) - C_1(k_1-1)y^p}{b^p - 1}\right]. \end{aligned} \tag{36}$$

We have to find the value of b which maximizes the function

$$g(b) = \frac{b^pC_1(k_1-1) + b^{q+1}C_2(y^p-1) - b^qC_2k_2(y^p-1) - C_1(k_1-1)y^p}{b^p-1}. \tag{37}$$

Its derivative can be written as

$$\begin{aligned} g_b(b) &= \frac{y^p-1}{(b^p-1)^2}b^{q-1}\begin{bmatrix} -b^{p+1}C_2(p-q-1) + b^pC_2k_2(p-q) \\ +b^{p-q}pC_1(k_1-1) - b(q+1)C_2 + qC_2k_2 \end{bmatrix} \\ &= \frac{y^p-1}{(b^p-1)^2}b^{q-1}h(b) \end{aligned} \tag{38}$$

for

$$h(b) = -b^{p+1}C_2(p-q-1) + b^pC_2k_2(p-q) + b^{p-q}pC_1(k_1-1) - b(q+1)C_2 + qC_2k_2. \tag{39}$$

We prove in Appendix A.2 that function (39) has a unique root larger than one. It leads to the maximum for function (38) since

$$\begin{aligned} h(1) &= p(C_1k_1 + C_2k_2 - C_1 - C_2) > 0, \\ h(\infty) &= -\infty. \end{aligned} \tag{40}$$

The root of the function (39) leads to

$$A(b) = \frac{b^pC_2K_2(p-q) + b^{p-q}pC_1K_1 + qC_2k_2}{b^{p+1}C_2(p-q-1) + b^{p-q}pC_1 + b(q+1)C_2}. \tag{41}$$

Using $b = \frac{1}{a}$ and denoting by $A_2(a)$ the put boundary we obtain

$$\begin{aligned} A_2(a) &= a\frac{a^pqC_2K_2 + a^qpC_1K_1 + (p-q)C_2K_2}{a^p(q+1)C_2 + a^{q+1}pC_1 + (p-q-1)C_2} \\ &= \frac{q}{q+1}\frac{a^{p-q}C_2K_2 + C_1K_1 + \dfrac{p-q}{q}\left(C_1K_1 + \dfrac{C_2K_2}{a^q}\right)}{a^{p-q-1}C_2 + C_1 + \dfrac{p-q-1}{q+1}\left(C_1 + \dfrac{C_2}{a^{q+1}}\right)} \\ &= \frac{p-q}{p-q-1}\frac{X_2(a)a^p + X_1(a)}{X_4(a)a^p + X_3(a)}. \end{aligned} \tag{42}$$

Having in mind Equations (33) and (42) we conclude that the equation $A_1(a) = A_2(a)$ has to be solved. It leads to

$$(1-a^p)(X_1(a)X_4(a) - X_2(a)X_3(a)) = 0. \tag{43}$$

The inequality $a < 1$ leads to

$$X_1(a)X_4(a) - X_2(a)X_3(a) = 0. \tag{44}$$

In such a way we derive the value of a as the solution of Equation (44) which also can be written as

$$\begin{aligned} H(a) &:= a^{p+1}C_1C_2K_2\alpha - a^pC_1C_2K_1\beta - a^{p-q}C_2^2K_2(\beta-\alpha) \\ &\quad - a^{q+1}C_1^2K_1(\beta-\alpha) - aC_1C_2K_2\beta + C_1C_2K_1\alpha = 0, \end{aligned} \tag{45}$$

where the constants α and β are

$$\begin{aligned} \alpha &= \frac{q}{q+1}, \\ \beta &= \frac{p-q}{p-q-1}. \end{aligned} \tag{46}$$

We prove in Appendix B that Equation (45) has a unique solution in the interval $a \in (0, 1)$. Something more, this solution leads to boundary values $\overline{A}$ and $\overline{B}$ such that $\overline{A} \leq D_1$ and $\overline{B} \geq D_2$, which validate their consistency.

We can formulate the results above in the following theorem.

Theorem 4.1. *Suppose that $\lambda > 0$ and let $\overline{a}$ be the unique solution of Equation (45). Then the optimal boundaries can be derived as $\overline{A} = A_1(\overline{a}) \equiv A_2(\overline{a})$ and $\overline{B} = \frac{\overline{A}}{\overline{a}}$, where the functions $A_1(a)$ and $A_2(a)$ are given in Equations (33) and (42). These boundaries lead to option price (25).*

4.2 | Nondiscounted model

Now we shall derive the closed-form formulas when the additional discount factor is zero, $\lambda = 0$. As a consequence we have $r > 0$. It is proven in Appendix A.2 that function (39) is positive for $b > 1$. This means that price function (25) is increasing. This corresponds to Proposition 3.1, which says that early exercising as a call is never optimal in the absence of discounting. We shall use an approach similar to the one presented in Section 4.1 method having in mind that the call boundary does not exist. We need to consider a one-sided exit problem related to exercising the option as a put. This way we have a one-dimensional price function and we are looking for the boundary value that maximizes it.

Suppose now that the option's holder exercises when the underlying asset hits the value $A \in (0, \min\{K_1, D_0\})$—we shall denote this moment by ζ^A. Thus the option price can be written as

$$\begin{aligned} Y(A) = {} & E^x\left[e^{-r\zeta^A} C_1 (K_1 - S_{\zeta^A})^+ I_{\zeta^A < \infty}\right] \\ & + \lim_{T\to\infty} E^x\left[e^{-rT} \max\{C_1(K_1 - S_T)^+, C_2(S_T - K_2)^+\} I_{T < \zeta^A}\right]. \end{aligned} \tag{47}$$

Note that the function above depends on the variable A through the stopping time ζ^A. First, suppose that $K_1 \leq K_2$—this leads to the domain $A \in (0, K_1)$. We have

$$\max\{C_1(K_1 - x)^+, C_2(x - K_2)^+\} = C_1(K_1 - x)^+ + C_2(x - K_2)^+ \tag{48}$$

and therefore the option price (47) turns to

$$\begin{aligned} Y(A) = {} & C_1(K_1 - A) E^x\left[e^{-r\zeta^A} I_{\zeta^A < \infty}\right] \\ & + C_1 \lim_{T\to\infty} E^x\left[e^{-rT}(K_1 - S_T)^+ I_{T<\zeta^A}\right] + C_2 \lim_{T\to\infty} E^x\left[e^{-rT}(S_T - K_2)^+ I_{T<\zeta^A}\right] \\ = {} & C_1(K_1 - A) E^x\left[e^{-r\zeta^A} I_{\zeta^A < \infty}\right] + C_1 \lim_{T\to\infty} P_{DO}(x, A, K_1) + C_2 \lim_{T\to\infty} C_{DO}(x, A, K_2). \end{aligned} \tag{49}$$

We denote above by $P_{DO}(x, A, K)$ and $C_{DO}(x, A, K)$ the prices of the down-and-out barrier options (put and call, respectively) with strike K and barrier A. Taking the limit $T \to \infty$ for these prices—see, for example, formulas (10.45) and (10.48) from Zhang (1997)—we obtain

$$\begin{aligned} & \lim_{T\to\infty} P_{DO}(x, A, K_1) = 0, \\ & \lim_{T\to\infty} C_{DO}(x, A, K_2) = x\left(1 - \left(\frac{A}{x}\right)^{1+\frac{2r}{\sigma^2}}\right). \end{aligned} \tag{50}$$

We shall use the following lemma for the Laplace transform of the Brownian motion's first hitting time.

Lemma 4.2. *If ζ is the first hitting moment of a Brownian motion with drift μ to a negative level a, then*

$$E[e^{-r\zeta}I_{\zeta<\infty}] = e^{(\sqrt{\mu^2+2r}+\mu)a}. \tag{51}$$

Having in mind that the stopping time ζ^A can be viewed as the first hitting of a Brownian motion with drift $\mu = \frac{r}{\sigma} - \frac{\sigma}{2}$ to the value $\frac{1}{\sigma}\ln\left(\frac{A}{x}\right)$ and using Lemma 4.2 we obtain

$$E^x[e^{-r\zeta^A}I_{\zeta^A<\infty}] = \left(\frac{A}{x}\right)^{\frac{2r}{\sigma^2}}. \tag{52}$$

Combining Equations (50) and (52) we transform the option price (49) to

$$Y(A) = C_2x + \left(\frac{A}{x}\right)^{\frac{2r}{\sigma^2}}[-A(C_1+C_2) + C_1K_1]. \tag{53}$$

Its A-derivative is

$$Y'(A) = \left(\frac{A}{x}\right)^{\frac{2r}{\sigma^2}}\left[\frac{2rC_1K_1}{\sigma^2A} - (C_1+C_2)\left(\frac{2r}{\sigma^2}+1\right)\right]. \tag{54}$$

Considering the function

$$h(A) = \frac{2rC_1K_1}{\sigma^2A} - (C_1+C_2)\left(\frac{2r}{\sigma^2}+1\right), \tag{55}$$

we see that it decreases starting from $h(0) = +\infty$ and finishes at the negative value $h(K_1) < 0$. Hence it has a unique root

$$\overline{A} = \frac{2rC_1K_1}{(C_1+C_2)(2r+\sigma^2)}, \tag{56}$$

which leads to the maximum of the option price. Hence, $\overline{A}$ is just the optimal boundary.

Suppose now that $K_2 < K_1$ and thus $A \in (0, D_0)$. We have

$$\max\{C_1(K_1-x)^+, C_2(x-K_2)^+\} = C_1(D_0-x)^+ + C_2(x-D_0)^+ + C_1C_2\frac{K_1-K_2}{C_1+C_2}. \tag{57}$$

Hence, the second term of option price (47) turns to a sum of down-and-out call and put options and a bond. When we take the limit $T \to \infty$, the put and bond prices vanish due to the first equation of (50). Something more the second equation of (50) shows that the limit of the call does not depend on the strike K_2 and therefore formula (53) still holds. We have to check that $\overline{A} < D_0$ to conclude that the optimal boundary is again $\overline{A}$.

We can summarize the derived results in the following theorem.

Theorem 4.2. *If $\lambda = 0$ then the early exercising of a perpetual strangle option is never optimal as a call. On the contrary, the option holder exercises as a put when the underlying reaches the level $\overline{A}$, where $\overline{A}$ is given by Equation (56). If the starting point is below this value $x \le \overline{A}$, then the option price is $C_1(K_1-x)$. Otherwise, if $x > \overline{A}$, then the price is $Y(\overline{A})$, where the function $Y(\cdot)$ is given by Equation (53)*

Remark 4.1. Regardless that exercising as a call is never optimal, the put boundary and the option price depend on the call features by the number of shares C_2, but not on the strike K_2.

Remark 4.2. Let us see what changes if $C_2 = 0$ in light of Remark 4.1. This way boundary value (56) turns to $\frac{2r}{2r+\sigma^2}K_1 = \frac{q}{q+1}K_1$, because $q = \frac{2r}{\sigma^2}$ when $\lambda = 0$. In fact, when $C_2 = 0$ we have a classical put option and this value is namely its optimal boundary—see, for example, formula (3.17) from Zaevski (2020b).

5 | FINITE MATURITY HORIZON

Suppose now that the maturity is finite, $T < \infty$.

5.1 | Deriving the exercise boundaries

For the functions $A(t)$ and $B(t)$, $A(t) < B(t)$, we define a European-style derivative, which expires as a put if the asset falls below $A(t)$ and as a call if it rises above $B(t)$. We shall name these instruments $(A(t), B(t))$-European options. The corresponding stopping times shall be denoted by ζ^A and ζ^B and the lower one by ζ, $\zeta = \zeta^A \wedge \zeta^B$.

Suppose that $0 \equiv t_0 < t_1 < t_2 < \dots < t_n \equiv T$ be an increasing time sequence and $a(t)$ and $b(t)$ be two continuous piecewise linear functions w.r.t. it

$$\begin{aligned} a(t) &= \sum_{i=1}^{n} a_i(t) I_{t\in[t_{i-1},t_i]} \equiv \sum_{i=1}^{n} (a_{1,i}t + a_{2,i}) I_{t\in[t_{i-1},t_i]}, \\ b(t) &= \sum_{i=1}^{n} b_i(t) I_{t\in[t_{i-1},t_i]} \equiv \sum_{i=1}^{n} (b_{1,i}t + b_{2,i}) I_{t\in[t_{i-1},t_i]}, \end{aligned} \tag{58}$$

$a_i(t_i) = a_{i+1}(t_i)$ and $b_i(t_i) = b_{i+1}(t_i)$, $i = 1, 2, \dots, n-1$. We impose the condition $a(t) < b(t)$ and additionally $a(0) < 0 < b(t)$. As we have mentioned above, we shall approximate the optimal boundaries as an exponent of piecewise linear functions—$A(t) = \exp(a(t))$ for the put boundary and $B(t) = \exp(b(t))$ for the call one. The values of these functions at the grid nodes shall be denoted by $\alpha_i = a(t_i)$, $\beta_i = b(t_i)$, $A_i = A(t_i)$, and $B_i = B(t_i)$, $i = 0, 1, \dots, n$. Let us introduce the functions

$$\begin{aligned} c(t) &= \sum_{i=1}^{n} c_i(t) I_{t\in(t_{i-1},t_i]} \equiv \sum_{i=1}^{n} (c_{1,i}t + c_{2,i}) I_{t\in(t_{i-1},t_i]}, \\ d(t) &= \sum_{i=1}^{p} d_i(t) I_{t\in(t_{i-1},t_i]} \equiv \sum_{i=1}^{p} (d_{1,i}t + d_{2,i}) I_{t\in(t_{i-1},t_i]} \end{aligned} \tag{59}$$

for

$$\begin{aligned} c_{1,i} &= \frac{a_{1,i} - r}{\sigma} + \frac{\sigma}{2}, \quad i = 1, \dots, n, \\ c_{2,i} &= \frac{a_{2,i} - \ln(x)}{\sigma}, \quad i = 1, \dots, n, \\ d_{1,i} &= \frac{b_{1,i} - r}{\sigma} + \frac{\sigma}{2}, \quad i = 1, \dots, n, \\ d_{2,i} &= \frac{b_{2,i} - \ln(x)}{\sigma}, \quad i = 1, \dots, n. \end{aligned} \tag{60}$$

In fact the stopping times ζ^A and ζ^B can be viewed as the first hitting moments of the Brownian motion to the functions $c(t)$ and $d(t)$, respectively. Using the notations above, we present the price of an $(A(t), B(t))$-European option as

$$
\begin{aligned}
G(x,T;A(t),B(t)) &= E^x\left[C_1 e^{-(r+\lambda)\zeta^A}(K_1 - S_{\zeta^A})^+ I_{\zeta^A=\zeta,\zeta<T}\right] \\
&\quad + E^x\left[C_2 e^{-(r+\lambda)\zeta^B}(S_{\zeta^B} - K_2)^+ I_{\zeta^B=\zeta,\zeta<T}\right] \\
&\quad + E^x\left[C_1 e^{-(r+\lambda)T}(K_1 - S_T)^+ I_{T\le\zeta,S_T\in(D_1,D_0)}\right] + E^x\left[C_2 e^{-(r+\lambda)T}(S_T - K_2)^+ I_{T\le\zeta,S_T\in(D_0,D_2)}\right] \\
&= C_1K_1\sum_{i=1}^{n} E\left[e^{-(r+\lambda)\zeta^A} I_{\zeta\in(t_{i-1},t_i],\zeta=\zeta^A}\right] - C_1 x\sum_{i=1}^{n} e^{\sigma c_{2,i}} E\left[e^{-\psi_{1,i}\zeta^A} I_{\zeta\in(t_{i-1},t_i],\zeta=\zeta^A}\right] \\
&\quad + C_2 x\sum_{i=1}^{n} e^{\sigma d_{2,i}} E\left[e^{-\psi_{2,i}\zeta^B} I_{\zeta\in(t_{i-1},t_i],\zeta=\zeta^B}\right] - C_2K_2\sum_{i=1}^{n} E\left[e^{-(r+\lambda)\zeta^B} I_{\zeta\in(t_{i-1},t_i],\zeta=\zeta^B}\right] \\
&\quad + C_1K_1 e^{-(r+\lambda)T} P(v_1 < B_T < l, T \le \zeta) - C_1 x e^{-\psi_3 T} E\left[e^{\sigma B_T} I_{v_1<B_T<l,T\le\zeta}\right] \\
&\quad + C_2 x e^{-\psi_3 T} E\left[e^{\sigma B_T} I_{l<B_T<v_2,T\le\zeta}\right] - C_2K_2 e^{-(r+\lambda)T} P(l < B_T < v_2, T \le \zeta),
\end{aligned}
\tag{61}
$$

where

$$
\begin{aligned}
\psi_{1,i} &= (r+\lambda) - \left(r - \frac{\sigma^2}{2}\right) - \sigma c_{1,i} = \lambda + \frac{\sigma^2}{2} - \sigma c_{1,i}, \\
\psi_{2,i} &= (r+\lambda) - \left(r - \frac{\sigma^2}{2}\right) - \sigma d_{1,i} = \lambda + \frac{\sigma^2}{2} - \sigma d_{1,i}, \\
\psi_3 &= \lambda + \frac{\sigma^2}{2}, \\
l &= \frac{1}{\sigma}\ln\left(\frac{D_0}{x}\right) - \left(\frac{r}{\sigma} - \frac{\sigma}{2}\right)T, \\
v_1 &= \frac{1}{\sigma}\ln\left(\frac{D_1}{x}\right) - \left(\frac{r}{\sigma} - \frac{\sigma}{2}\right)T, \\
v_2 &= \frac{1}{\sigma}\ln\left(\frac{D_2}{x}\right) - \left(\frac{r}{\sigma} - \frac{\sigma}{2}\right)T.
\end{aligned}
\tag{62}
$$

We shall use the following proposition to derive the expectations in formula (61).

Proposition 5.1. *Assume that $y > 0$, whereas we do not impose restrictions for z. If the functions $a(\cdot)$ and $b(\cdot)$ are linear, then*

$$
\begin{aligned}
L_a(t,y;a_1,a_2,b_1,b_2) &= E\left[e^{-y\zeta} I_{\zeta\le T,\zeta=\zeta^A}\right] \\
&= e^{a_2\left(\sqrt{a_1^2+2y}-a_1\right)} P_1^l\left(T; \sqrt{a_1^2+2y}, a_2, b_1 + \sqrt{a_1^2+2y} - a_1, b_2\right), \\
L_b(t,y;a_1,a_2,b_1,b_2) &= E\left[e^{-y\zeta} I_{\zeta\le T,\zeta=\zeta^B}\right] \\
&= e^{b_2\left(\sqrt{b_1^2+2y}-b_1\right)} P_2^l\left(T; a_1 + \sqrt{b_1^2+2y} - b_1, a_2, \sqrt{b_1^2+2y}, b_2\right), \\
V(z,z,T,a_1,a_2,b_1,b_2) &= E\left[e^{zB_T} I_{\zeta>T,B_T<z}\right] \\
&= \exp\left(\frac{z^2T}{2}\right)\left\{\begin{array}{l}\left(N\left(\dfrac{z - zT}{\sqrt{T}}\right) - N\left(\dfrac{a(T) - zT}{\sqrt{T}}\right)\right) \\ + \displaystyle\sum_{j=1}^{\infty}\sum_{k=1}^{4}\left[\begin{array}{l} s_k \exp\left(\lambda_{j,k} z + \dfrac{\lambda_{j,k}^2 + 2\eta_{j,k}}{2T}\right) \\ \times\left(N\left(\dfrac{z - (zT + \lambda_{j,k})}{\sqrt{T}}\right) - N\left(\dfrac{a(T) - (zT + \lambda_{j,k})}{\sqrt{T}}\right)\right)\end{array}\right]\end{array}\right\}.
\end{aligned}
\tag{63}
$$

If the functions $a(\cdot)$ and $b(\cdot)$ are piecewise linear, then

$$E\left[e^{-y\zeta} I_{\zeta\in(t_{m-1},t_m)}, \zeta=\zeta_{a,b}\right] = \int_{\alpha_1,\ldots,\alpha_{m-1}}^{\beta_1,\ldots,\beta_{m-1}} \left(\begin{array}{l} \prod_{i=1}^{m-1}\left(1-\sum_{j=1}^{\infty} q_{i,j}(x_{i-1},x_i)\right) \dfrac{\exp\left(-\dfrac{(x_i-x_{i-1})^2}{2(t_i-t_{i-1})}\right)}{\sqrt{2\pi(t_i-t_{i-1})}} \\ e^{-yt_{m-1}} L_{a,b}(t_m-t_{m-1}, y; a_{m,1}, \alpha_{m-1}-x_{m-1}, b_{m,1}, \beta_{m-1}-x_{m-1}) \end{array} \right) dx_1 \ldots dx_{m-1},$$

$$E\left[e^{zB_T} I_{B_T<z,\zeta>T}\right] = \int_{\alpha_1,\ldots,\alpha_{n-1}}^{\beta_1,\ldots,\beta_{n-1}} \left(\begin{array}{l} \prod_{i=1}^{n-1}\left(1-\sum_{j=1}^{\infty} q_{i,j}(x_{i-1},x_i)\right) \dfrac{\left(\exp\left(-\dfrac{(x_i-x_{i-1})^2}{2(t_i-t_{i-1})}\right)\right)}{\sqrt{2\pi(t_i-t_{i-1})}} \\ e^{zx_{n-1}} V(z, z-x_{n-1}, t_n-t_{n-1}; a_{n,1}, \alpha_{n-1}-x_{n-1}, b_{n,1}, \beta_{n-1}-x_{n-1}) \end{array} \right) dx_1 \ldots dx_{n-1}, \tag{64}$$

where the functions $L_a(\cdot)$, $L_b(\cdot)$, and $V(\cdot)$ are given by formulas (63).

Proof. The proofs can be found in Zaevski (2021). □

Note that the probabilities in formula (61) can be obtained using the last statement of (64) for $z=0$. We shall use an alternative parametrization for the option price—$G(x; t_0, t_1, \ldots t_n; A_0, A_1, \ldots, A_n; B_1, B_2, \ldots, B_n)$ instead $G(x, T; A(t), B(t))$. Once we have a formula for $(A(t), B(t))$-European option, we can use the following backward algorithm.

1. The boundaries at the maturities are given in Propositions 3.5 and 3.6—$A_n = D_1$ and $B_n = D_2$.
2. Suppose that we have derived all values $A_m, A_{m+1}, \ldots, A_n$ and $B_m, B_{m+1}, \ldots, B_n$ for some $m<n$.
3. We derive the put boundary in the following way. For the constants $A<x$ let $B(x,A)$ be this value which maximizes

$$G(x; 0, t_m-t_{m-1}, \ldots t_n-t_{m-1}; A, A_m, \ldots, A_n; B, B_m, \ldots, B_n) \tag{65}$$

amongst all $B>x$. Let us view Equation (65) as a function of A and let $A(x)$ be this argument which maximizes

$$G(x; 0, t_m-t_{m-1}, \ldots t_n-t_{m-1}; A, A_m, \ldots, A_n; B(x,A), B_m, \ldots, B_n). \tag{66}$$

Our approximation for the put boundary value A_{m-1} is the largest x for which $x=A(x)$. In fact, this is the largest initial value of the underlying asset for which immediate exercising as a put is optimal.

4. Analogously we obtain the call boundary. Let for a fixed $x<B$, $A(x,B)$ be this value which maximizes function (65) w.r.t. the variable A. Also, let $B(x)$ maximizes

$$G(x; 0, t_m-t_{m-1}, \ldots t_n-t_{m-1}; A(x,B), A_m, \ldots, A_n; B, B_m, \ldots, B_n) \tag{67}$$

amongst all $B>x$.

5. Thus we approximate the call boundary B_{m-1} as the smallest x for which $x=B(x)$. We illustrate the way we derive the boundaries in Figure 1a—we present there the difference $B(x)-x$. Our approximate value is the smallest x for which this difference is zero and it is marked by a circle. The used parameters are reported in Section 6.

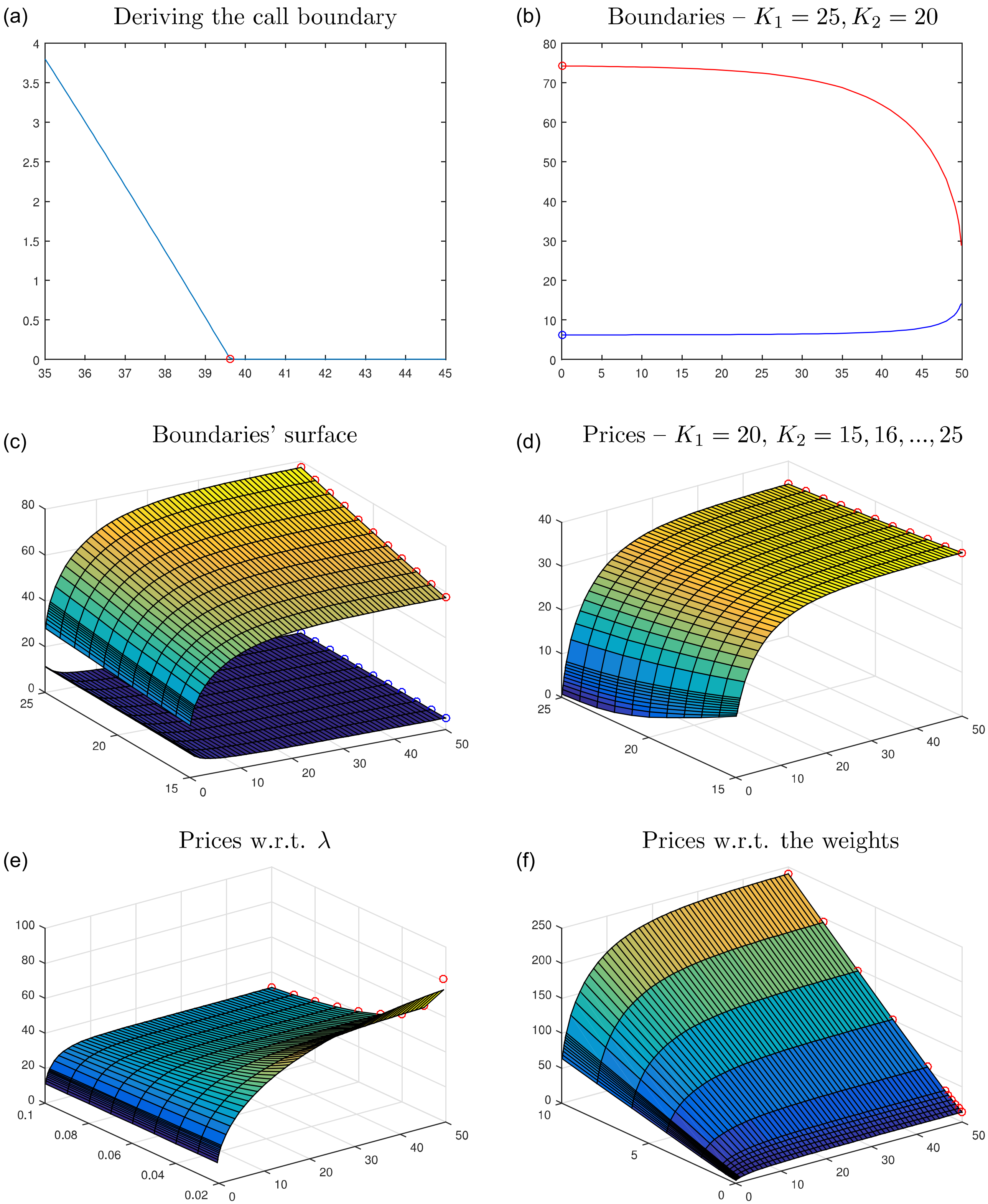

FIGURE 1 Optimal boundaries and option prices. (a) Deriving the call boundary, (b) boundaries—$K_1 = 25$, $K_2 = 20$, (c) boundaries' surface, (d) prices—$K_1 = 20$, $K_2 = 15, 16, ..., 25$, (e) prices w.r.t. λ, and (f) prices w.r.t. the weights. [Color figure can be viewed at wileyonlinelibrary.com]

5.2 | Finite difference approach for option pricing

Once we approximate the exercise boundaries, we can view option pricing as the boundary value problem

$$\begin{aligned} &F_t(t,x) + rxF_x(t,x) + \frac{1}{2}\sigma^2x^2F_{xx}(t,x) - rF(t,x) = 0,\\ &F(t,A(t)) = e^{-\lambda t}C_1(K_1 - x), \quad t \in (0,T),\\ &F(t,B(t)) = e^{-\lambda t}C_2(x - K_2), \quad t \in (0,T),\\ &F(T,x) = e^{-\lambda T}\max\{C_1(K_1 - x), C_2(x - K_2)\}, \quad x \in (D_1, D_2). \end{aligned} \tag{68}$$

The differential equation holds in the region $(t,x) \in \{(0,T) \times (A(t),B(t))\}$ and the boundary constraints are imposed on the lower, upper, and right boundaries. We shall use the Crank–Nicolson finite difference scheme to solve numerically Equation (68). Our algorithm is as follows:

1. We divide the time and state intervals uniformly into M and N subintervals $T \equiv t_1 > t_2 > \cdots > t_M \equiv 0$ and $A(0) \equiv x_1 < x_2 < \cdots < x_N \equiv B(0)$, respectively. We shall denote by $F(m,n)$ the option prices at the nodes. Note that we work backward—in such a way the values $F(1,n)$ correspond to the prices at the maturity and $F(M,n)$ are the initial option prices.
2. We use the algorithm presented in Section 5.1 to approximate the boundaries $A(t)$ and $B(t)$ at $\overline{M} < M$ points. After this we use a cubic spline interpolation to find the boundary values on the whole time grid—$A_1, A_2, ..., A_M$ and $B_1, B_2, ..., B_M$. Note again that A_1 and B_1 are the values at the maturity, respectively, D_1 and D_2, whereas A_M and B_M are the initial ones.
3. We incorporate the terminal condition by

$$F(1,n) = e^{-\lambda T}\max\{C_1(K_1 - x_n), C_2(x_n - K_2)\}. \tag{69}$$

4. For every $m = 1, 2, ..., M$ we denote by l_m the highest $n \in \{0, 1, 2, ..., N\}$ for which $x_{l_m} < A_m$. Also let k_m be the lowest n such that $x_{k_m} > B_m$.
5. The lower and upper boundary conditions are integrated by

$$\begin{aligned} F(m,n) &= e^{-\lambda t_m}C_1(K_1 - x_n) \quad \forall m \ \ and \ \ n \le l_m\\ F(m,n) &= e^{-\lambda t_m}C_2(x_n - K_2) \quad \forall m \ \ and \ \ n \ge k_m. \end{aligned} \tag{70}$$

6. Suppose that we have derived the values of $F(i,n)$ for all n and $i < m$. We derive the values $F(m,n)$ by the Crank–Nicolson scheme. The derivatives are approximated by Equations (C1) from Appendix C. Thus the equation from the boundary value problem (68) acquires the form (C2). Rearranging these equations for $n \in \{l_m + 1, l_m + 2, ..., k_m - 1\}$ we obtain the linear system for $F(m,n)$ composed of Equations (C3), (C4), and (C5).

6 | NUMERICAL RESULTS

We present some numerical experiments in this section. The main values we use are as follows—the initial asset value is $S_0 = 22$, the risk-free rate is $r = -0.02$, the additional discount factor is $\lambda = 0.05$, the volatility is $\sigma = 0.3$, the put and call strikes are $K_1 = \$25$ and $K_2 = \$20$, the corresponding weights are $C_1 = 3$ and $C_1 = 2$, the time to maturity is in the interval $\tau \in (0, 50)$. We shall mention expressly the used values if we vary some of them. The parameters we use for the Crank–Nicolson method are—$\overline{M} = 16$ points for the optimal boundaries obtained by the two steps algorithm, $M = 64$ points for the time grid, and $N = 1000$ steps for the state space. Note that these points are enough due to the quadratic convergence.

In Figure 1b we present the put and call boundaries w.r.t the actual time—the put boundary is the lower one. We mark by circles the perpetual values derived via Theorem 4.1. We can see that these values are the limits when the time to maturity increases.

TABLE 1 Option prices and optimal boundaries.

K_1	20	22	24	25
Weight ratio $\frac{C_1}{C_2} = 0.5$				
$K_2 = 20$	{4.3621;52.5720}	{4.7448;53.6082}	{5.1220;54.6609}	{5.3057;55.2218}
	$12.4648	$13.4766	$14.5628	$15.1438
$K_2 = 22$	{4.4103;56.8144}	{4.7985;57.8294}	{5.1814;58.8657}	{5.3747;59.3756}
	$10.8370	$11.9038	$13.1632	$13.7671
$K_2 = 24$	{4.4537;61.0725}	{4.8466;62.0707}	{5.2356;63.0865}	{5.4255;63.6108}
	$9.5939	$10.6888	$11.9389	$12.5846
$K_2 = 25$	{4.4738;63.2076}	{4.8693;64.1964}	{5.2603;65.2015}	{5.4532;65.7156}
	$9.0320	$10.1333	$11.3953	$12.0483
Weight ratio $\frac{C_1}{C_2} = 1$				
$K_2 = 20$	{4.8290;60.1743}	{5.2781;62.1678}	{5.7315;64.1334}	{5.9502;65.1782}
	$15.8613	$18.0208	$20.1833	$21.5037
$K_2 = 22$	{4.8556;64.2656}	{5.3119;66.1944}	{5.7627;68.1735}	{5.9877;69.1654}
	$14.3319	$16.4861	$18.9870	$20.2409
$K_2 = 24$	{4.8801;68.3783}	{5.3386;70.2827}	{5.7948;72.2092}	{6.0219;73.1878}
	$13.0994	$15.2889	$17.8092	$19.1157
$K_2 = 25$	{4.8913;70.4431}	{5.3514;72.3342}	{5.8062;74.2744}	{6.0362;75.2179}
	$12.5413	$14.7396	$17.2781	$18.5920
Weight ratio $\frac{C_1}{C_2} = 2$				
$K_2 = 20$	{5.1233;73.8414}	{5.6225;77.3750}	{6.1217;80.9273}	{6.3690;82.7279}
	$22.8514	$27.2151	$31.9439	$34.4314
$K_2 = 22$	{5.1346;77.7337}	{5.6360;81.2233}	{6.1345;84.7604}	{6.3847;86.5286}
	$21.3369	$25.6815	$30.7194	$33.2898
$K_2 = 24$	{5.1454;1.6612}	{5.6470;85.1202}	{6.1482;88.6067}	{6.3979;90.3696}
	$20.1202	$24.4991	$29.5812	$32.2168
$K_2 = 25$	{5.1507;83.6339}	{5.6533;87.0708}	{6.1543;90.5508}	{6.4042;92.2996}
	$19.5654	$23.9636	$29.0628	$31.7074
Weight ratio $\frac{C_1}{C_2} = 3$				
$K_2 = 20$	{5.2285;86.1083}	{5.7437;91.0003}	{6.2594;95.9010}	{6.5167;98.3662}
	$29.8512	$36.4821	$43.6647	$47.4866
$K_2 = 22$	{5.2351;89.8781}	{5.7513;94.7181}	{6.2671;99.5959}	{6.5247;102.0336}
	$28.3389	$34.9378	$42.4953	$46.3864
$K_2 = 24$	{5.2419;93.6623}	{5.7576;98.5050}	{6.2741;103.3287}	{6.5322;105.7617}
	$27.1472	$33.7310	$41.3697	$45.3392
$K_2 = 25$	{5.2433;95.6113}	{5.7595;100.4213}	{6.2786;105.1839}	{6.5350;107.6390}
	$26.5958	$33.2461	$40.8602	$44.8374

The rest of the experiments we provide is parametrized just w.r.t. the time to maturity instead of the current time. We present in Figure 1c,d the surfaces of the boundaries and the option prices fixing the put strike as $K_1 = \$20$ and varying the call one among $K_2 = \$15, \$16, ..., \$25$. Figure 1e shows the behavior of the option prices varying the discount factor as $\lambda \in (0.02, 0.1)$, whereas Figure 1f presents the behavior w.r.t. the put and call weights. We vary the ratio $c = \frac{C_1}{C_2}$ as $c \in \{0.01, 0.2, 0.4, 0.6, 0.9, 1, 2, 4, 6, 8, 10\}$. The circles show again the perpetual values.

Some options' prices together with the corresponding boundaries are presented in Table 1. The varied parameters are both strikes—$K_{1,2} \in \{20, 22, 24, ,25\}$—as well as the weight's ratio $\frac{C_1}{C_2} \in \{0.5, 1, 2, 3\}$. The time to maturity is fixed at $T = 3$. We report in the first row the optimal boundaries—the first is the put one, and the second is the call one. The second row gives the option price.

7 | CONCLUSIONS

Specific financial instruments, namely, American strangle options, are examined in this paper. They exhibit jointly put and call features giving to the holder the right to choose how to exercise—as a call or as a put. Arbitrary values for the strikes have been considered—it is possible for the call strike to be less than the put one. Also, the call and put options in the strangle derivative are presented by different weights. It turns out that the pricing of such derivatives leads to two-sided optimal stopping problems. Closed-form formulas for the optimal boundaries as well as for the fair price have been obtained for the perpetual versions of these instruments. It turns out that early exercising as a call is never optimal if the discount (dividend) rate is missing. However, the call right has its impact—it appears via the call weight, but not through the call strike. On the other hand, a numerical approach is constructed to approximate the optimal boundaries when the maturity is finite. On the basis of them, the Crank–Nicolson finite difference scheme is adapted to the raising boundary value problem related to option pricing. Several numerical experiments are provided.

ACKNOWLEDGMENTS

The author would like to express sincere gratitude to the editor Prof. Dr. Bart Frijns and to the anonymous reviewer for the helpful and constructive comments which substantially improve the quality of this paper. This research has been partially supported by Grant No. BG05M2OP001-1.001-0003, financed by the Science and Education for Smart Growth Operational Program (2014–2020) and cofinanced by the European Union through the European Structural and Investment funds and by the project KP-06-N32/8 with the Bulgarian National Science Fund.

ORCID

Tsvetelin S. Zaevski http://orcid.org/0000-0002-1118-4189

REFERENCES

Abdou, S. L., & Moraux, F. (2016). Pricing and hedging American and hybrid strangles with finite maturity. *Journal of Banking & Finance*, *62*, 112–125.

Alobaidi, G., & Mallier, R. (2002). Laplace transforms and the American straddle. *Journal of Applied Mathematics*, *2*(3), 121–129.

Alobaidi, G., & Mallier, R. (2006). The American straddle close to expiry. *Boundary Value Problems*, *2006*, 1–14.

Beibel, M., & Lerche, H. R. (1997). A new look at optimal stopping problems related to mathematical finance. *Statistica Sinica*, *7*(1), 93–108.

Borodin, A., & Salminen, P. (2015). Handbook of Brownian motion—Facts and formulae. In *Probability and its applications*. Birkhäuser Basel.

Chiarella, C., & Ziogas, A. (2005). Evaluation of American strangles. *Journal of Economic Dynamics and Control*, *29*(1–2), 31–62.

Gapeev, P. V., & Lerche, H. R. (2011). On the structure of discounted optimal stopping problems for one-dimensional diffusions. *Stochastics an International Journal of Probability and Stochastic Processes*, *83*(4–6), 537–554.

Jacka, S. D. (1991). Optimal stopping and the American put. *Mathematical Finance*, *1*(2), 1–14.

Jeon, J., & Kim, G. (2022). Analytic valuation formula for American strangle option in the mean-reversion environment. *Mathematics*, *10*(15), 2688.

Jeon, J., & Oh, J. (2019). Valuation of American strangle option: Variational inequality approach. *Discrete & Continuous Dynamical Systems—B*, *24*(2), 755.

Kang, M., Jeon, J., Han, H., & Lee, S. (2017). Analytic solution for American strangle options using Laplace–Carson transforms. *Communications in Nonlinear Science and Numerical Simulation*, *47*, 292–307.

Ma, J.-T., Li, W.-Y., & Cui, Z.-Y. (2018). Valuation of American strangles through an optimized lower-upper bound approach. *Journal of the Operations Research Society of China*, *6*(1), 25–47.

Qiu, S. (2020). American strangle options. *Applied Mathematical Finance*, *27*(3), 228–263.
Shiryaev, A. N., Kabanov, Y., Kramkov, D., & Mel'nikov, A. (1995). Toward the theory of pricing of options of both European and American types. II. Continuous time. *Theory of Probability & Its Applications*, *39*(1), 61–102.
Shiryaev, A. N. (1999). *Essentials of stochastic finance: Facts, models, theory* (Vol. 3). World Scientific.
Zaevski, T. (2020a). Discounted perpetual game call options. *Chaos, Solitons & Fractals*, *131*, 109503.
Zaevski, T. (2020b). Discounted perpetual game put options. *Chaos, Solitons & Fractals*, *137*, 109858.
Zaevski, T. (2021). Laplace transforms of the Brownian motion's first exit from a strip. *Comptes rendus de l'Académie bulgare des Sciences*, *74*(5), 669–676.
Zhang, P. (1997). *Exotic options*. World Scientific.

How to cite this article: Zaevski, T. S. (2023). American strangle options with arbitrary strikes. *The Journal of Futures Markets*, 1–24. https://doi.org/10.1002/fut.22419

APPENDIX A: UNIQUENESS OF THE SOLUTIONS

A.1 | Put boundary

We shall prove now that the function $h(\cdot)$ defined in Equation (29) has a unique root in the interval $(0, 1)$. First we consider the undiscounted case, that is, $\lambda = 0$. The function $h(\cdot)$ turns to

$$h(a) = a^{q+1}C_1k_1 - a(q+1)(C_1 + C_2 - C_2k_2) + qC_1k_1 \tag{A1}$$

and its derivative is

$$h_a(a) = (q+1)[a^qC_1k_1 - C_1 - C_2 + C_2k_2]. \tag{A2}$$

We have $h_a(1) = -(q+1)[C_1 + C_2 - C_1k_1 - C_2k_2] < 0$. Therefore $h_a(a)$ is negative in the interval $(0, 1)$ since it is an increasing function. Hence, the function $h(\cdot)$ is decreasing and therefore it has a unique root due to inequalities (30).

Suppose now $\lambda > 0$ or equivalently $p > q + 1$. We can present function (29) as

$$h(a) = C_1\left[-\bar{h}(a) - a^{p-q}\frac{p}{C_1}(C_1 + C_2 - C_1k_1 - C_2k_2)\right] \tag{A3}$$

for

$$\bar{h}(a) = a^{p+1}(p-q-1) - a^pk_1(p-q) - a^{p-q}p(1-k_1) + a(q+1) - qk_1. \tag{A4}$$

The function $\bar{h}(a)$ is examined in Zaevski (2020a), Appendix B.1—there it is denoted by $h(a; 0)$. It may behave in two ways—(A) starts from a negative value and increases to zero or (B) starts from a negative value, increases to a positive maximum, and decreases to zero. Hence, function (A3) may be increasing only when it is negative (only in case [B]). Having in mind boundary values (30) we conclude that function (29) has just one root in the interval $(0, 1)$.

A.2 | Call boundary

First we consider the case $\lambda = 0$ or equivalently $p = q + 1$. Function (39) turns to

$$h(b) = b^{q+1}C_2k_2 + b(q+1)(C_1k_1 - C_1 - C_2) + qC_2k_2 \tag{A5}$$

and its derivative is

$$h_b(b) = (q+1)(b^qC_2k_2 + C_1k_1 - C_1 - C_2). \tag{A6}$$

This function starts from a positive value for $b = 1$ and increases, which makes it positive for all $b > 1$. Hence, the function $h(b)$ is increasing and therefore it is positive for $b > 1$ too since $g(1) = p(C_1k_1 + C_2k_2 - C_1 - C_2) > 0$.

Suppose now that $\lambda > 0$. Function (39) can be presented as

$$h(b) = C_2\left[-\bar{h}(b) + b^{p-q}\frac{p}{C_2}(C_1k_1 + C_2k_2 - C_1 - C_2)\right] \tag{A7}$$

for

$$\bar{h}(b) = b^{p+1}(p-q-1) - b^p k_2(p-q) + b^{p-q}p(k_2-1) + b(q+1) - qk_2. \tag{A8}$$

Let us change the variable b to $d = \frac{1}{b}$ turning in this way the interval $(1, \infty)$ to $(0, 1)$. The function $\bar{h}(d)$ written as $\bar{h}(d; 0)$ is examined in Appendix B.1 from Zaevski (2020b). There are two cases for its behavior—starts from a positive value and decreases to zero or (B) starts from a positive value decreases to a negative minimum and increases to zero. We can conclude that the function $h(d)$ starts from a negative value for $d = 0$ and finishes at a positive value for $d = 1$. Also it may decrease only if it is positive—this is possible only when the case (B) holds. Hence, it has a unique root for $d \in (0, 1)$ and therefore the same is true for $b > 1$.

APPENDIX B: EXISTING AND UNIQUENESS OF THE OPTIMAL BOUNDARIES

First we shall prove that Equation (45) has a unique solution in the interval $(0, 1)$. Note that $\beta > \alpha$. Hence, $H(0) = \alpha C_1C_2K_1 > 0$ and $H(1) = -(\beta-\alpha)(C_1C_2K_1 + C_1C_2K_2 + C_1^2K_1 + C_2^2K_2) < 0$. Also, function (45) can be decomposed to $H(a) = C_1C_2\alpha H_1(a) + H_2(a)$, where

$$\begin{aligned} H_1(a) &= (K_1 - aK_2)(1 - a^p), \\ H_2(a) &= -(\beta-\alpha)a\left(a^{p-1}C_1C_2K_1 + a^{p-q-1}C_2^2K_2 + a^qC_1^2K_1 + C_1C_2K_2\right). \end{aligned} \tag{B1}$$

If $K_1 \geq K_2$, then both functions $H_1(a)$ and $H_2(a)$ are decreasing which leads to the existence and uniqueness of the solution. Suppose now that $K_1 < K_2$. The first and second derivatives of the function $H_1(a)$ are

$$\begin{aligned} H_1'(a) &= a^pK_2(p+1) - a^{p-1}K_1p - K_2, \\ H_1''(a) &= a^{p-2}p[aK_2(p+1) - K_1(p-1)]. \end{aligned} \tag{B2}$$

Hence, the derivative $H_1'(a)$ starts from the negative value $H_1'(0) = -K_2$, decreases to a minimum, and increases to the positive value $H_1'(1) = p(K_2 - K_1)$. Therefore, the function $H_1(a)$ starts from the positive value $H_1(0) = K_1$, decreases to a negative minimum, and increases to zero. Hence, the solution of Equation (45) exists and it is unique since the function $H_2(a)$ is decreasing and negative. We shall denote this solution by $\bar{a}$. Note that $\bar{a} < \frac{K_1}{K_2}$ since $H_1(\bar{a}) \leq 0$ for $a \geq \frac{K_1}{K_2}$.

To continue our discussion we need to introduce the following two functions.

$$\begin{aligned} G_1(a) &= \frac{p-q}{p-q-1}a\frac{a^qC_1K_1 + C_2K_2}{a^{q+1}C_1 + C_2}, \\ G_2(a) &= \frac{q}{q+1}\frac{a^{p-q}C_2K_2 + C_1K_1}{a^{p-q-1}C_2 + C_1}. \end{aligned} \tag{B3}$$

Equations (35), (42), and (44) show that $\bar{A} = A_1(\bar{a}) = A_2(\bar{a}) = G_1(\bar{a}) = G_2(\bar{a})$. Note that the equations $G_1(a) = G_2(a)$ and $H(a) = 0$ are equivalent and therefore the first one has the same unique solution. Let us examine now the behavior of functions (B3). The derivative of the function $G_1(a)$ is

$$G_1'(a) = C_2\frac{p-q}{p-q-1}\frac{-a^{q+1}C_1K_2q + a^q(q+1)C_1K_1 + C_2K_2}{(a^{q+1}C_1 + C_2)^2}. \tag{B4}$$

If we denote by $l(\cdot)$ the function

$$l(a) = -a^{q+1}C_1K_2q + a^q(q+1)C_1K_1 + C_2K_2, \tag{B5}$$

then its derivative is

$$l'(a) = -a^{q-1}q(q+1)C_1(aK_2 - K_1). \tag{B6}$$

It is always positive when $K_1 \geq K_2$ and it is positive for $a < \frac{K_1}{K_2}$ and negative, otherwise. We conclude that the function $G_1(a)$ is increasing when $K_1 \geq K_2$. Otherwise, if $K_1 < K_2$, then the function $G_1(a)$ is increasing or first increases and then decreases. The second case holds when $l(1) = qC_1(K_1 - K_2) + C_1K_1 + C_2K_2 < 0$.

Let us turn to the function $G_2(a)$. We have

$$G_2'(a) = C_2\frac{q}{q+1}a^{p-q-2}\frac{a^{p-q}C_2K_2 + a(p-q)C_1K_2 - (p-q-1)C_1K_1}{(a^{p-q-1}C_2 + C)^2}. \tag{B7}$$

We see that the function $l(a) = a^{p-q}C_2K_2 + a(p-q)C_1K_2 - (p-q-1)C_1K_1$ starts from a negative value and increases. Hence, the function $G_2(a)$ is decreasing or first decreases and then increases. The second case is actual when $l(1) = (p-q)C_1(K_2 - K_1) + C_1K_1 + C_2K_2 < 0$.

We shall prove now that $\overline{A} < D_1$. First we shall show that $\min\{K_1, \overline{K}\} > \overline{A}$, where $\overline{K} := K_1\frac{r+\lambda}{\lambda}$. Formulas (23) lead to

$$\frac{r+\lambda}{\lambda} = \frac{q}{q+1}\frac{p-q}{p-q-1} \tag{B8}$$

and therefore

$$\overline{K} = K_1\frac{q}{q+1}\frac{p-q}{p-q-1}. \tag{B9}$$

If $K_2 < K_1$ or equivalently $D_0 < K_1$, then we can easily check that $G_2(0) < \min\{K_1, \overline{K}\}$ and $G_2(1) < \min\{K_1, \overline{K}\}$ and therefore $\overline{A} = G_2(\overline{a}) < \overline{K}$. If $K_1 \leq K_2$, we have $G_2\left(\frac{K_1}{K_2}\right) = K_1\frac{q}{q+1} < \min\{K_1, \overline{K}\}$. The fact that $\overline{a} < \frac{K_1}{K_2}$ establishes the result.

It remains to be proven that $\overline{A} < D_0$. We need to consider only the case $K_2 < K_1$ due to the result above. Let us examine the dependence on the discount factor λ. Its domain is $[\max\{0, -r\}, \infty)$. Hence the variable q increases in the interval $\left[\max\left\{0, \frac{2r}{\sigma^2}\right\}, \infty\right)$. We shall parametrize w.r.t. the variable q hereafter. Equation (23) gives

$$p(q) = 2q - 2\frac{r}{\sigma^2} + 1. \tag{B10}$$

The functions (B3) can be rewritten as

$$\begin{aligned} G_1(a,q) &= \left(1 + \frac{1}{p(q) - q - 1}\right)\left(K_1 + C_2\frac{aK_2 - K_1}{a^{q+1}C_1 + C_2}\right),\\ G_2(a,q) &= \left(1 - \frac{1}{q+1}\right)\left(aK_2 - C_1\frac{aK_2 - K_1}{a^{p(q)-q-1}C_2 + C_1}\right). \end{aligned} \tag{B11}$$

Having in mind that the function $p(q) - q - 1$ is increasing, we conclude that $G_1(a,q)$ decreases w.r.t. q, whereas $G_2(a,q)$ is a q-increasing function. Let us investigate the behavior of functions $G_1(a,q)$ and $G_2(a,q)$ when $q \to \infty$. We have that $\lim_{q\to\infty} G_1(a,q) = aK_2$ for $a \in (0,1)$ and $\lim_{q\to\infty} G_1(1,q) = D_0$. Analogously, $\lim_{q\to\infty} G_2(a,q) = K_1$ for

$a \in (0,1)$ and $\lim_{q\to\infty} G_2(1,q) = D_0$. We observe that $G_2^{-1}(D_0, q)$ is well defined since $G_2(1,q) = \frac{q}{q+1}D_0 < D_0$. Also, $G_1^{-1}(D_0, q)$ is well defined because it is an a-increasing function. The function $L(q) = G_2^{-1}(D_0,q) - G_1^{-1}(D_0,q)$ is increasing. Let us consider the asymptotic case $q = \infty$. If we suppose that $\overline{a}(\infty) < 1$, then the equation $G_1(\overline{a},\infty) = G_1(\overline{a},\infty)$ leads to $\overline{a} = \frac{K_1}{K_2}$ which is impossible due to $K_1 < K_2$. Hence $\overline{a}(\infty) = 1$ and therefore $\lim_{q\to\infty} G_1^{-1}(D_0,q) = \lim_{q\to\infty} G_2^{-1}(D_0,q) = 1$ due to the observed above shape of the functions $G_1(a,q)$ and $G_2(a,q)$ for $q \to \infty$. This means that $L(\infty) = 0$ and thus $G_1^{-1}(D_0,q) > G_2^{-1}(D_0,q)$ for $q < \infty$ because the function $L(q)$ is increasing. Therefore $\overline{A} = G_1(\overline{a}) < D_0$ due to the a-behavior of the functions $G_1(a)$ and $G_2(a)$.

The inequality $\overline{B} > D_2$ can be proven analogously.

APPENDIX C: FINITE DIFFERENCE TERMS

$$
\begin{aligned}
F_t &= \frac{F(m-1,n) - F(m,n)}{\Delta t},\\
F &= \frac{F(m-1,n) + F(m,n)}{2},\\
F_x &= \frac{F(m-1,n) - F(m-1,n-1) + F(m,n) - F(m,n-1)}{2\Delta x},\\
F_{xx} &= \frac{F(m-1,n+1) - 2F(m-1,n) + F(m-1,n-1)}{2(\Delta x)^2},\\
&\quad + \frac{F(m,n+1) - 2F(m,n) + F(m,n-1)}{2(\Delta x)^2}.
\end{aligned}
\tag{C1}
$$

$$
\begin{aligned}
0 = &\frac{F(m-1,n) - F(m,n)}{\Delta t}\\
&+ \frac{1}{2} r x_n \frac{F(m-1,n) - F(m-1,n-1) + F(m,n) - F(m,n-1)}{\Delta x}\\
&+ \frac{1}{4}\sigma^2 x_n^2 \left(\frac{F(m-1,n+1) - 2F(m-1,n) + F(m-1,n-1)}{(\Delta x)^2} + \frac{F(m,n+1) - 2F(m,n) + F(m,n-1)}{(\Delta x)^2} \right)\\
&- \frac{1}{2} r (F(m-1,n) + F(m,n)).
\end{aligned}
\tag{C2}
$$

- If $n = l_m + 1$, then

$$
\begin{aligned}
&F(m,n)\left(\frac{1}{\Delta t} - \frac{1}{2}\frac{rx_n}{\Delta x} + \frac{1}{2}\frac{\sigma^2 x_n^2}{(\Delta x)^2} + \frac{1}{2}r\right) - F(m,n+1)\frac{1}{4}\frac{\sigma^2 x_n^2}{(\Delta x)^2}\\
= \;& F(m-1,n-1)\left(-\frac{1}{2}\frac{rx_n}{\Delta x} + \frac{1}{4}\frac{\sigma^2 x_n^2}{(\Delta x)^2}\right) + F(m-1,n)\left(\frac{1}{\Delta t} + \frac{1}{2}\frac{rx_n}{\Delta x} - \frac{1}{2}\frac{\sigma^2 x_n^2}{(\Delta x)^2} - \frac{1}{2}r\right)\\
&+ F(m-1,n+1)\frac{1}{4}\frac{\sigma^2 x_n^2}{(\Delta x)^2} - F(m,l_m)\left(\frac{1}{2}\frac{rx_n}{\Delta x} - \frac{1}{4}\frac{\sigma^2 x_n^2}{(\Delta x)^2}\right).
\end{aligned}
\tag{C3}
$$

- If $l_m + 1 < n < k_m - 1$, then

$$
\begin{aligned}
&F(m,n-1)\left(\frac{1}{2}\frac{rx_n}{\Delta x} - \frac{1}{4}\frac{\sigma^2 x_n^2}{(\Delta x)^2}\right) + F(m,n)\left(\frac{1}{\Delta t} - \frac{1}{2}\frac{rx_n}{\Delta x} + \frac{1}{2}\frac{\sigma^2 x_n^2}{(\Delta x)^2} + \frac{1}{2}r\right)\\
&-F(m,n+1)\frac{1}{4}\frac{\sigma^2 x_n^2}{(\Delta x)^2} = F(m-1,n-1)\left(-\frac{1}{2}\frac{rx_n}{\Delta x} + \frac{1}{4}\frac{\sigma^2 x_n^2}{(\Delta x)2}\right)\\
&+ F(m-1,n)\left(\frac{1}{\Delta t} + \frac{1}{2}\frac{rx_n}{\Delta x} - \frac{1}{2}\frac{\sigma^2 x_n^2}{(\Delta x)^2} - \frac{1}{2}r\right) + F(m-1,n+1)\frac{1}{4}\frac{\sigma^2 x_n^2}{(\Delta x)^2}.
\end{aligned}
\tag{C4}
$$

- If $n = k_m - 1$, then

$$\begin{aligned}
&F(m, n-1)\left(\frac{1}{2}\frac{rx_n}{\Delta x} - \frac{1}{4}\frac{\sigma^2 x_n^2}{(\Delta x)^2}\right) + F(m, n)\left(\frac{1}{\Delta t} - \frac{1}{2}\frac{rx_n}{\Delta x} + \frac{1}{2}\frac{\sigma^2 x_n^2}{(\Delta x)^2} + \frac{1}{2}r\right) \\
&= \ F(m-1, n-1)\left(-\frac{1}{2}\frac{rx_n}{\Delta x} + \frac{1}{4}\frac{\sigma^2 x_n^2}{(\Delta x)^2}\right) + \ F(m-1, n)\left(\frac{1}{\Delta t} + \frac{1}{2}\frac{rx_n}{\Delta x} - \frac{1}{2}\frac{\sigma^2 x_n^2}{(\Delta x)^2} - \frac{1}{2}r\right) \\
&+F(m-1, n+1)\frac{1}{4}\frac{\sigma^2 x_n^2}{(\Delta x)^2} + \ F(m, n+1)\frac{1}{4}\frac{\sigma^2 x_n^2}{(\Delta x)^2}.
\end{aligned} \tag{C5}$$