國立臺灣大學管理學院財務金融學系

碩士論文

Department of Finance

College of Management

National Taiwan University

Master Thesis

美式賣權的提前履約曲線之凸性探討-使用Kim方法

Discussion of Nonconvexity of Early Exercise Boundary for an American Put Option by Applying Kim's Method

葉力嘉

LI-CHIA YEH

指導教授：張森林 博士、繆維中 博士

Advisor: San-Lin Chung, Ph.D. and Daniel Wei-Chung Miao, Ph.D.

中華民國 112 年 7 月

July, 2023

doi:10.6342/NTU202302497

致謝

首先要感謝張森林教授與繆維中教授這兩年來對我的教導，在定期會議中不論是論文相關的知識或是做研究的態度都讓我學到不少。另外要特別感謝王之彥教授，即使我不是您的學生，依舊願意多次與我討論論文的程式細節。也要感謝崔茂培教授，即使已經在數學系畢業，依舊願意與我討論論文中的數學與數值結果。

另外要特別感謝黃維坪教我將程式改得更有效率，省了我好多時間。也要感謝盧德倫、何世宇、王建安，在我快放棄的時候，跟我討論論文的內容、教我怎麼寫論文。

最後要感謝我的父母，在這兩年的碩士生活中，讓我無憂無慮地生活，有不順心的事，也常聽我訴苦。有你們我才能完成這篇論文。

葉力嘉 謹誌於臺灣大學

民國 112 年 7 月

doi:10.6342/NTU202302497

摘要

我們使用 Kim （1990）論文中的方程式產生出美式賣權之提前履約曲線的數值結果。另外，我們利用此方程式推導出提前履約曲線對時間與殖利率導數的公式。透過對提前履約曲線對殖利率導數的數值結果，我們提出了爲何當殖利率高於無風險利率時，提前履約曲線會出現非凸性的原因。我們也解釋了當殖利率高於無風險利率很多時 (大約高於 1.5 倍)，提前履約曲線會恢復凸性。

關鍵字: 美式選擇權、賣權、提前履約曲線、凸性。

doi:10.6342/NTU202302497

Abstract

We use the equation in Kim (1990) to generate the numerical results of early exercise boundary for the American put options. Then, we evaluate the formula of the derivative of the early exercise boundary with respect to time and dividend yield. By the numerical results of the derivative of the boundary with respect to the dividend yield, we propose reasoning to explain the nonconvexity of the early exercise boundary when the dividend yield is higher than the risk-free rate. And we also explain why the early exercise boundary regains convex when the dividend yield is much higher than the risk-free rate (for approximately $q/r > 1.5$)

Keyword: American option · Put option · Early exercise boundary · Convexity.

doi:10.6342/NTU202302497

Contents

doi:10.6342/NTU202302497

doi:10.6342/NTU202302497

List of Figures

doi:10.6342/NTU202302497

List of Tables

doi:10.6342/NTU202302497

Chapter 1

Introduction

Nowadays, the Black-Scholes model is widely used to value the option price. It has an analytic form for the European option, but no analytic solution has been found for the American option. The main reason is that the early exercise boundary, which represents the optimal price for early exercise before maturity, does not have an exact formula. However, we still know some characteristics of the boundary, such as the monotonicity and the limit value at maturity. In particular, convexity is an interesting topic to discuss.

Note that the regularity of the boundary was proved in Chen et al. (2013). It proved the boundary is C^{∞}[1] in the log normal context, so the derivative of the

[1] The boundary is differentiable for all degrees of differentiation.

doi:10.6342/NTU202302497

boundary makes sense.

The convexity of the American put early exercise boundary with respect to time is still an open question currently. In Chen et al. (2008), it proves that the early exercise boundary of an American put is convex when the dividend yield is zero. In Chen et al. (2013), it proves that nonconvex behavior appears when the dividend yield is higher than the risk-free rate. It gives us the theoretical upper bound of nonconvexity and numerically evaluates the nonconvex range. In Chen's paper, we observe that the upper and lower bound of the nonconvex range increase as the dividend yield increases. The early exercise boundary regains convexity as the dividend yield is significantly higher than the risk-free rate. The cause of this phenomenon is still unclear.

In this paper, we will first use the integral equation provided by Kim (1990) to numerically evaluate the early exercise boundary. Then, we derive the formula of the derivative of the early exercise boundary with respect to time to evaluate the nonconvex range more explicitly. Though the numerical result is not good, we still conclude some possible reasons to explain the numerical error causing the inaccurate results. Finally, we derive the formula of the derivative of the early exercise boundary with respect to the dividend yield. By the numerical results, we will explain how the nonconvex range change as the dividend yield rises.

doi:10.6342/NTU202302497

Chapter 2

Early Exercise Boundary Evaluation

We first introduce the notation in this paper.

Notation:

K: Exercise price of the option

τ: Time to maturity

r: Risk-free rate in the market

q: Dividend yield of the security

σ: Volatility of the return of the security

$B(\tau)$: Early exercise boundary (or optimal exercise price) of the American put at time τ

Δt_1: $\{B(u) \mid 0 \leq u \leq \Delta t_1\}$ is the early exercise boundary range using finite

doi:10.6342/NTU202302497

difference method data

Δt_2: Subinterval length of the numerical integration in Kim's method

Δt_3: Interval length of early exercise boundary recursively evaluating in Kim's method

First, we notice that

$$B(0) = \lim_{\tau \to 0^+} B(\tau) = \begin{cases} \frac{rK}{q} & \text{if } q > r \\ K & \text{if } q \le r \end{cases} \tag{2.1}$$

This is already proved in Kim (1990), and Wilmott et al. (1993). It means that the boundary has a different price at maturity with different conditions of dividend yield. Observe that $B(0) < K$ as $q > r$, we can view it as the one who has stock will earn expectedly $e^{(q-r)\tau}(> 1)$ times if he/she does not early exercise it with τ remaining time. Therefore, the stock price requirement of early exercise is lower than the exercise price K even near the expiry date.

By the equation (12) in Kim (1990), we know that the *live* American put price has

doi:10.6342/NTU202302497

an integral equation representation at the stock price $S(\tau)$:

$$P(S(\tau),\tau;B(\cdot)) = p(S(\tau),\tau) \ + \int_0^{\tau} \Big[rKe^{-r(\tau-u)}N(-d_2(S(\tau),\tau-u;B(u))) \\ - qB(\tau)e^{-q(\tau-u)}N\Big(-d_1\big(S(\tau),\tau-u;B(u)\big)\Big)\Big] du \tag{2.2}$$

where $B(\cdot)$ means $\{B(u) \mid 0 \leq u \leq \tau\}$.

$$p(S,\tau) = Ke^{-r\tau}N\Big(-d_2(S,\tau;K)\Big) - Se^{-q\tau}N\Big(-d_1(S,\tau;K)\Big) \tag{2.3}$$

represents European put option at the time to maturity τ.

$$d_1(S,\tau,K) = \frac{\ln\left(\frac{S}{K}\right) + \left(r - q + \frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}}, \quad d_2(S,\tau;K) = d_1(S,\tau;K) - \sigma\sqrt{\tau} \tag{2.4}$$

$$N(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \tag{2.5}$$

This equation explains the American put option can be separated as the summation of the European put price and the premium value of the opportunity to early exercise.

doi:10.6342/NTU202302497

With taking $S(\tau) = B(\tau)$, American put is early exercised, so we have

$$K - B(\tau) = p\Big(B(\tau), \tau\Big) + \int_0^\tau f\Big(B(\tau), \tau; B(u), u\Big)\, du \tag{2.6}$$

where

$$\begin{aligned} f\Big(B(\tau), \tau; B(u), u\Big) = rKe^{-r(\tau-u)} N\Bigg(-d_2\Big(B(\tau), \tau-u; B(u)\Big)\Bigg) \\ - qB(\tau)e^{-q(\tau-u)} N\Bigg(-d_1\Big(B(\tau), \tau-u; B(u)\Big)\Bigg) \end{aligned} \tag{2.7}$$

By this equation, we can evaluate the early exercise boundary recursively. Below is the methodology.

2.1 Methodology

First, we will evaluate $B(\tau)$ from $\tau = 0$ to the time far from maturity, so $B(\tau)$ value near expiry is important. We use the finite difference method to evaluate $B(\tau)$ for $\tau = 0 \sim \Delta t_1$, with Δt_2 time gap and ΔS price gap. This is to ensure that $B(\tau)$ would not be misestimated too much from the beginning. However, we are concerned that the stepwise function of the finite difference method $B(\tau)$ may not

doi:10.6342/NTU202302497

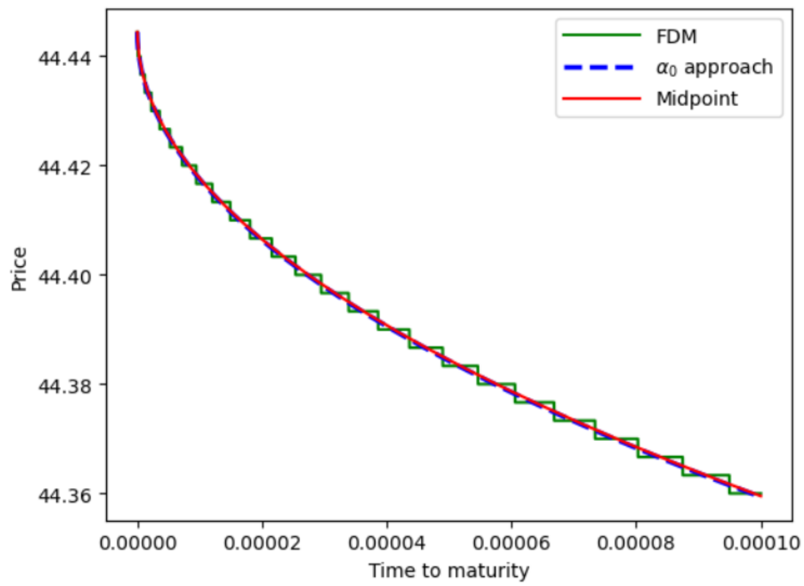


Figure 2.1: $B(\tau)$ comparison of finite difference method

Evans et al. (2002) approximation (α_0 approach) and the modified function (Midpoint) with $K = 50,\ r = 0.04,\ q = 0.045,\ \sigma = 0.3,\ \Delta S = \frac{1}{300},\ \Delta t_2 = 10^{-8}$

be smooth enough. Therefore, we modify it by linearly connecting the midpoint of each price gap, thus overcoming the issue of smoothness. Note that the modified function is almost the same as the approximation function in equation (3.8) in Evans et al. (2002) [2]. See the comparison in Figure 2.1.

RMSE of the function in Evans et al. (2002) (called it $\boldsymbol{\alpha_0}$ **approach**) and the modified function (called it **Midpoint**) is 1.642×10^{-4}, and the differences are shown in Table 2.1.

[2] The early exercise boundary $B(\tau) \approx \frac{r}{q} K e^{-\alpha_0 \sigma \sqrt{2\tau}}$ as τ near 0 where $\alpha_0 \approx 0.4517$.

doi:10.6342/NTU202302497

τ	10^{-5}	2×10^{-5}	3×10^{-5}	4×10^{-5}	5×10^{-5}	6×10^{-5}	7×10^{-5}	8×10^{-5}	9×10^{-5}	10^{-4}
Difference	0.0001538	0.0001322	0.0001637	0.0001579	0.0001677	0.0001723	0.0001553	0.0001781	0.0001683	0.0001700

Table 2.1: Differences of α_0 approach and Midpoint

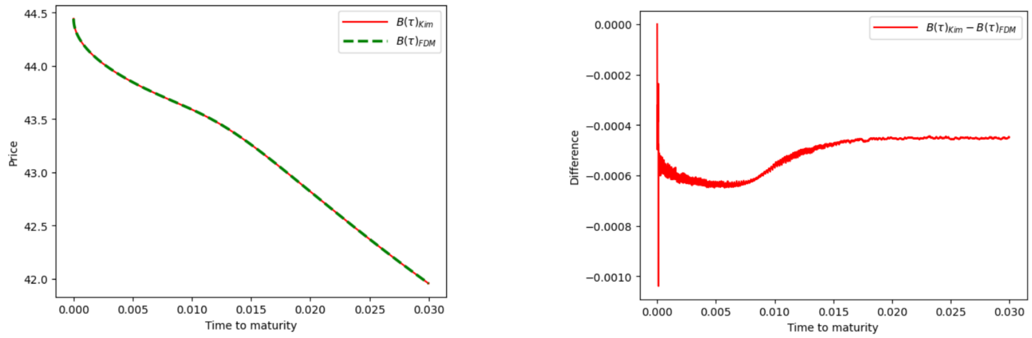
Second, since we find the value from $B(0)$ to $B(\Delta t_1)$, we want to evaluate the next point $B(\tau)$ for $\tau = \Delta t_1 + \Delta t_3$. According to (2.6), suppose $F\Big(S, \tau, B(\cdot)\Big) = -K + S + p(S,\tau) + \int_0^\tau f(S,\tau; B(u), u)\, du$ we take in $S = B(\tau - \Delta t_3)$ and $B(\tau - \Delta t_3) - \delta_1$ as the upper and lower value. By the bisection method, we can figure out $S = B(\tau)$ when $S = \Big|F\Big(S, \tau, B(\cdot)\Big)\Big| < \epsilon_1$. Note that we also define $\{B(u) \mid \Delta t_1 \le u \le \Delta t_1 + \Delta t_3\}$ by the linear function between $B(\Delta t_1)$ & $B(\Delta t_1 + \Delta t_3)$.

Third, we use the trapezoid method to approximate the integration with time gap Δt_2 for each time point in the integration. For $0 < u \le \Delta t_1$, we take into the value of finite difference method data, and for $u > \Delta t_1$ we take into the value we get from our method and connect them by piece-wise linear function.

Fourth, we can recursively evaluate $B(\Delta t_1 + \Delta t_3), B(\Delta t_1 + 2\Delta t_3), B(\Delta t_1 + 3\Delta t_3)$, and so on.

Finally, we can get $B(\tau)$ function from our method, called it **Kim's** method.

doi:10.6342/NTU202302497



(a) $B(\tau)$ comparison between Kim method and finite difference method $\Delta S = \frac{1}{100}, \Delta t = 10^{-7}$

(b) $B(\tau)$ difference between Kim method and finite difference method

Figure 2.2: $B(\tau)$ comparison
With parameter $K = 50,\ r = 0.04,\ q = 0.045,\ \sigma = 0.3$

2.2 Numerical results

In the Figure 2.2, we use $\Delta S = \frac{1}{300},\ \Delta t_1 = 10^{-4},\ \Delta t_2 = 10^{-8},\ \Delta t_3 = 10^{-5},\ \delta_1 = 0.1,\ \epsilon_1 = 10^{-11}$ as the parameter of **Kim's** method (See the Appendix for the choice of ΔS and Δt_2) to compare with finite difference method with $\Delta S = \frac{1}{100}$ and $\Delta t = 10^{-7}$ (called it **FDM** method). We find that $B(\tau)_{Kim}$ is slightly lower than $B(\tau)_{FDM}$. RMSE of them is 5.22×10^{-4}. It verifies the accuracy of $B(\tau)_{Kim}$.

doi:10.6342/NTU202302497

Chapter 3

Derivative of Early Exercise Boundary with respect to time

For the purpose of finding the nonconvex range of the early exercise boundary more explicitly, we want to evaluate the derivative of $B(\tau)$. Then, by finding the local maximum and local minimum of $B_\tau(\tau)$ we can find the nonconvex range of $B(\tau)$.

doi:10.6342/NTU202302497

3.1 Derivation of the formula

Proposition 1. *Given a function of time* $x(u), u \geq 0$ *and a function* $f\big(x(\tau), \tau; x(u), u\big)$ *with* $\tau > u$. *Define*

$$y(\tau) = \int_0^\tau f(x(\tau), \tau; x(u), u)du$$

then

$$\frac{dy}{d\tau} = \lim_{u \to \tau^-} f(x(\tau), \tau, x(u), u) + \int_0^\tau \left[\frac{\partial f(\cdot)}{\partial \tau} + \frac{\partial f(\cdot)}{\partial x(\tau)}\frac{\partial x(\tau)}{\partial \tau}\right] du$$

Proof. An implication of Leibniz's integral rule. Note that $u \to \tau^-$ is required when there is no value at $u = \tau$ □

Theorem 1. *For* $\tau > 0$, *the derivative of the early exercise boundary with respect to time has the following form.*

$$B_\tau(\tau) = \frac{dB(\tau)}{d\tau} = -\frac{p_\tau\Big(B(\tau), \tau\Big) + \frac{1}{2}\Big(rK - qB(\tau)\Big) + \int_0^\tau h\Big(B(\tau), \tau; B(u), u\Big)\, du}{1 + p_{B(\tau)}\Big(B(\tau), \tau\Big) + \int_0^\tau g\Big(B(\tau), \tau; B(u), u\Big)\, du} \tag{3.1}$$

where

$$d_1' = d_1(B(\tau), \tau, K), \qquad d_2' = d_1' - \sigma\sqrt{\tau} \tag{3.2}$$

doi:10.6342/NTU202302497

$$d_1^\star = d_1(B(\tau), \tau - u, B(u)), \qquad d_2^\star = d_1^\star - \sigma\sqrt{\tau - u} \tag{3.3}$$

$$n(x) = N'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \tag{3.4}$$

$$\begin{aligned} g\Big(B(\tau), \tau; B(u), u\Big) = \frac{\partial f(\cdot)}{\partial B(\tau)} =& -\frac{rKe^{-r(\tau-u)}n(d_2^\star)}{B(\tau)\sigma\sqrt{\tau-u}} + \frac{qe^{-q(\tau-u)}n(d_1^\star)}{\sigma\sqrt{\tau-u}} \\ & - qe^{-q(\tau-u)}N(-d_1^\star) \end{aligned} \tag{3.5}$$

$$\begin{aligned} h\Big(B(\tau), \tau; B(u), u\Big) = \frac{\partial f(\cdot)}{\partial \tau} = -rKe^{-r(\tau-u)} & \left[rN(-d_2^\star) + \frac{n(d_2^\star)}{\tau-u}\left(\frac{d_2^\star}{2} - \frac{\ln(\frac{B(\tau)}{B(u)})}{\sigma\sqrt{\tau-u}}\right)\right] \\ + qB(\tau)e^{-q(\tau-u)} & \left[qN(-d_1^\star) + \frac{n(d_1^\star)}{\tau-u}\left(\frac{d_1^\star}{2} - \frac{\ln(\frac{B(\tau)}{B(u)})}{\sigma\sqrt{\tau-u}}\right)\right] \end{aligned} \tag{3.6}$$

$$\begin{aligned} p_\tau(B(\tau), \tau) = \frac{\partial p(B(\tau), \tau)}{\partial \tau} =& -rKe^{-r\tau}N(-d_2') + qB(\tau)e^{-q\tau}N(-d_1') \\ & + \frac{B(\tau)\sigma e^{-q\tau}n(d_1')}{2\sqrt{\tau}} \end{aligned} \tag{3.7}$$

$$p_{B(\tau)}(B(\tau), \tau) = \frac{\partial p}{\partial B(\tau)}(B(\tau), \tau) = -e^{-q\tau}N(-d_1') \tag{3.8}$$

Proof. Total derivative equation (2.6) by Proposition 1, we get

$$-B_\tau(\tau) = \frac{\partial p}{\partial B(\tau)}B_\tau(\tau) + \frac{\partial p}{\partial \tau} + \lim_{u \to \tau^-} f(B(\tau), \tau; B(u), u) + \int_0^\tau \left[\frac{\partial f(\cdot)}{\partial B(\tau)}B_\tau(\tau) + \frac{\partial f(\cdot)}{\partial \tau}\right] du$$

doi:10.6342/NTU202302497

$$B_\tau(\tau) = -\frac{p_\tau\Big(B(\tau),\tau\Big) + \lim\limits_{u\to\tau^-} f\Big(B(\tau),\tau;B(u),u\Big) + \int_0^\tau \frac{\partial f}{\partial \tau}\Big(B(\tau),\tau;B(u),u\Big)du}{1 + p_{B(\tau)}\Big(B(\tau),\tau\Big) + \int_0^\tau \frac{\partial f}{\partial B(\tau)}\Big(B(\tau),\tau;B(u),u\Big)du}$$

For $p_\tau\Big(B(\tau),\tau\Big)$, $p_{B(\tau)}\Big(B(\tau),\tau\Big)$, $\frac{\partial f}{\partial B(\tau)}\Big(B(\tau),\tau;B(u),u\Big)$, and $\frac{\partial f}{\partial \tau}\Big(B(\tau),\tau;B(u),u\Big)$ is the simple calculus works. And by the implication of Lemma 1, we have

$$\lim_{u\to\tau^-} f(B(\tau),\tau;B(u),u) = rKe^{-r\cdot 0}N(0) - qB(\tau)e^{-q\cdot 0}N(0) = \frac{1}{2}(rK - qB(\tau))$$

□

Lemma 1. *Define* $d_1^\star = d_1\Big(B(\tau),\tau-u,B(u)\Big), \qquad d_2^\star = d_1^\star - \sigma\sqrt{\tau-u}$

$d_1^\star,\ d_2^\star \to 0 \quad as \quad u \to \tau^-$

Proof.

$$\begin{aligned}
\lim_{u\to\tau^-} d_1\Big(B(\tau),\tau-u,B(u)\Big) &= \lim_{u\to\tau^-} \frac{\ln\Big(B(\tau)\Big) - \ln\Big(B(u)\Big) + (r-q+\frac{\sigma^2}{2})(\tau-u)}{\sigma\sqrt{\tau-u}} \\
&= \lim_{u\to\tau^-} \frac{\ln\Big(B(\tau)\Big) - \ln\Big(B(u)\Big)}{\sigma\sqrt{\tau-u}} \\
&= \frac{\partial \ln\Big(B(\tau^-)\Big)}{\partial \tau} \cdot \lim_{u\to\tau^-} \frac{\sqrt{\tau-u}}{\sigma} \\
&= \frac{B_\tau(\tau^-)}{B(\tau^-)} \cdot 0 = 0
\end{aligned}$$

doi:10.6342/NTU202302497

$$\lim_{u\to\tau^-} d_2\Big(B(\tau), \tau - u, B(u)\Big) = \lim_{u\to\tau^-} d_1\Big(B(\tau), \tau - u, B(u)\Big) - \sigma\sqrt{\tau - u} = 0$$

By the proof in Chen et al. (2013), $B(\tau)$ is differentiable for $\tau > 0$. That illustrates the third equality of limit d_1. □

Note that $g\Big(B(\tau), \tau; B(u), u\Big)$, $h\Big(B(\tau), \tau; B(u), u\Big)$ does not have limit for $u \to \tau^-$. There might be a problem if the integral does not converge. In the Appendix, we will ensure that the integrals of $g\Big(B(\tau), \tau; B(u), u\Big)$, and $h\Big(B(\tau), \tau; B(u), u\Big)$ converge.

3.2 Numerical Results

For the formula in (3.1), we require $\{B(u) \mid 0 < u \leq \tau\}$ as an input to calculate the value of $B_\tau(\tau)$ and we already obtained this in the previous section. The numerical result of $B(\tau)_{Kim}$ and $B(\tau)_{FDM}$ are $B_\tau(\tau)_{Kim}$ and $B_\tau(\tau)_{FDM}$, respectively. However, the result is not good. In Figure 3.1, the $\widehat{B}_\tau(\tau)_{FDM}$ and $\widehat{B}_\tau(\tau)_{Kim}$ is the secant line slope of $B(\tau)_{Kim}$ and $B(\tau)_{FDM}$ in Figure 2.2a. Ideally, we hope that $B_\tau(\tau)_{Kim}$ and $B_\tau(\tau)_{FDM}$ are close to $\widehat{B}_\tau(\tau)_{FDM}$ and $\widehat{B}_\tau(\tau)_{Kim}$, but they don't. Below we will discuss the reason for non-accurate numerical results.

doi:10.6342/NTU202302497

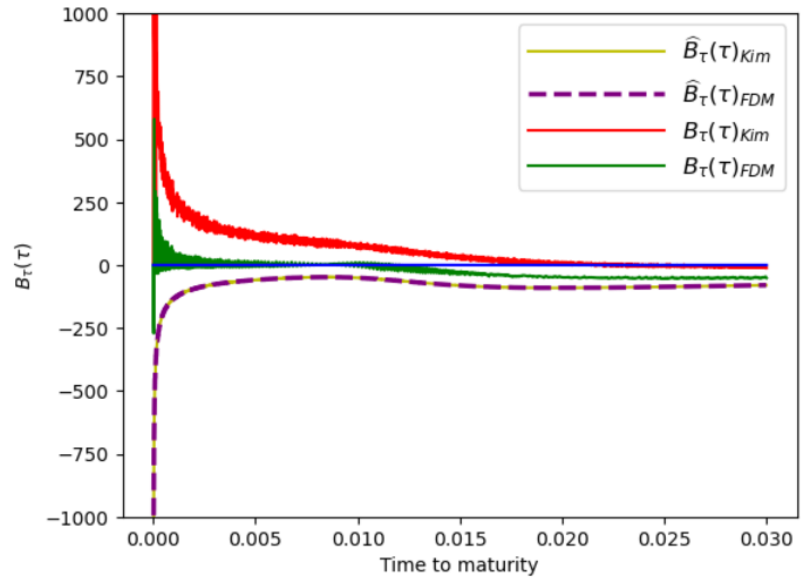


Figure 3.1: $B_\tau(\tau)$ comparison
With parameter $K = 50,\ r = 0.04,\ q = 0.045,\ \sigma = 0.3$

3.3 Discussion of Inaccurate Results

One possible reason is equation (4) in Kim (1990) for the American put version,

$$
\begin{aligned}
P(S_n, n\Delta t) &= p(S_n, n\Delta t) + O(\Delta t) \\
&\quad + \sum_{k=1}^{n-1} e^{-(n-k)r\Delta t} \int_0^{B_k} \left[(1 - e^{-r\Delta t})K - (1 - e^{-q\Delta t})S_k\right] \\
&\quad \times \psi(S_k, (n-k)\Delta t; S_n)\, dS_k
\end{aligned} \tag{3.9}
$$

where $\psi(S_k, (n-k)\Delta t; S_n)$ is the transition density function of the asset price at time t_k given the asset price S_n at time t_n. Equation (2.6) derives from it for

doi:10.6342/NTU202302497

$\Delta t \rightarrow 0$. $O(\Delta t)$ may not be an important term when calculating the numerical result of $B(\tau)$. However, the formula of $B_\tau(\tau)$ may lose an important term when we ignore $O(\Delta t)$ in deriving the numerical results. Calculating the approximate term of $O(\Delta t)$ as a function of τ may solve this problem.

Another possible reason is the accuracy of the trapezoid method of numerical integration. Because $h\Big(B(\tau), \tau; B(u), u\Big) \rightarrow \infty$ as $u \rightarrow \tau^-$. The numerical result may lose a large value for $\tau - u$ close to 0. See an example in Table 3.1.

Because $h\Big(B(\tau), \tau; B(u), u\Big)$ is $O\Big((\tau - u)^{-0.5}\Big)$ as $u \rightarrow \tau^-$ (see Appendix for the

$\tau - u$	10^{-8}	2×10^{-8}	3×10^{-8}	4×10^{-8}	5×10^{-8}	6×10^{-8}	7×10^{-8}	8×10^{-8}	9×10^{-8}	10^{-7}
$h(u)$	955.96	675.97	551.93	477.99	427.53	390.28	361.34	338.00	318.67	302.32

Table 3.1: $h\Big(B(\tau), \tau; B(u), u\Big)$ value at $\tau = 10^{-4}$

proof), we take a similar example to explain how trapezoid integration causes an error. Note that $\int_0^{0.0001} \frac{1}{\sqrt{x}}\, dx = 0.02$, but for the trapezoid method with 10000 points even distance numerical integration (without counting the point $x = 0$) the numerical result is 0.019803..., which is about 1% error with respect to the true value. For the case $\tau = 10^{-4}$, the numerical integration $\int_0^\tau h\Big(B(\tau), \tau; B(u), u\Big)\, du = 1.8987 \times 10^{-3}$ and the numerator of $B_\tau(\tau)$ is -1.304×10^{-5}. The 1% error may change the sign of $B_\tau(\tau)$. However, unless we know the approximate form of $h(\cdot)$ as $\tau - u$ is close to 0, we can't approximate the integration well in any numerical

doi:10.6342/NTU202302497

integration method.

For future work, we can try to use the secant line slope of the boundary as the accurate $B_\tau(\tau)$ compared with the formula result. We set the difference of the slope and the formula result as the cost function. Using deep learning or the gradient decent method, we can implicitly find the numerical formula of integration of $h\Big(B(\tau), \tau; B(u), u\Big)$. It may solve both of the possible problems that we propose.

doi:10.6342/NTU202302497

Chapter 4

Derivative of Early Exercise Boundary with respect to dividend yield

To explain the nonconvexity of the early exercise boundary, we focus on the behavior of the early exercise boundary change when the dividend yield changes and other variables are fixed.

doi:10.6342/NTU202302497

4.1 Derivation of the formula

Theorem 2. *For $\tau > 0$, the derivative of the early exercise boundary with respect to the dividend yield has the following form.*

$$B_q(\tau) = -\frac{p_q\Big(B(\tau),\tau\Big) + \int_0^\tau \Big[l\Big(B(\tau),\tau;B(u),u\Big)\cdot B_q(u) + m\Big(B(\tau),\tau;B(u),u\Big)\Big]\,du}{1 + p_{B(\tau)}\Big(B(\tau),\tau\Big) + \int_0^\tau g\Big(B(\tau),\tau;B(u),u\Big)du} \tag{4.1}$$

where

$$p_q\Big(B(\tau),\tau\Big) = B(\tau)e^{-q\tau}N(-d_1')\tau \tag{4.2}$$

$$l\Big(B(\tau),\tau;B(u),u\Big) = \frac{\partial f(\cdot)}{\partial B(u)} = \frac{rKe^{-r(\tau-u)}n(d_2^\star)}{B(u)\sigma\sqrt{\tau-u}} - \frac{qB(\tau)e^{-q(\tau-u)}n(d_1^\star)}{B(u)\sigma\sqrt{\tau-u}} \tag{4.3}$$

$$\begin{aligned} m\Big(B(\tau),\tau;B(u),u\Big) = \frac{\partial f(\cdot)}{\partial q} &= \frac{\sqrt{\tau-u}}{\sigma}\Big(-rKe^{-r(\tau-u)}n(d_2^\star) + qB(\tau)e^{-q(\tau-u)}n(d_1^\star)\Big) \\ &\quad + B(\tau)e^{-q(\tau-u)}N(-d_1^\star)\Big(q(\tau-u)-1\Big) \end{aligned} \tag{4.4}$$

$p_{B(\tau)}\Big(B(\tau),\tau\Big)$ *and* $g\Big(B(\tau),\tau;B(u),u\Big)$ *is the same as the formula in the previous section.*

Proof. Total derivative the equation (2.6) by q , we get

$$-B_q(\tau) = \frac{\partial p}{\partial B(\tau)}B_q(\tau) + \frac{\partial p}{\partial q} + \int_0^\tau \left[\frac{\partial f(\cdot)}{\partial B(\tau)}\cdot B_q(\tau) + \frac{\partial f(\cdot)}{\partial B(u)}\cdot B_q(u) + \frac{\partial f(\cdot)}{\partial q}\right]du$$

doi:10.6342/NTU202302497

Through rearranging equation and some simple calculus works, we prove it. □

Note that the integral of $l(B(\tau),\tau;B(u),u)$ may also have the problem of convergence. We will ensure that this is not a problem in the Appendix.

4.2 Methodology

Similar to the methodology in evaluating the numerical result of $B(\tau)$, we need the data of $\{B(u)|\ 0 \le u \le \tau\}$ and $\{B_q(u)|\ 0 \le u < \tau\}$ as an input to calculate the value of $B_q(\tau)$. For the first part of the data, we already know in Section 2. For the second part of the data, we first notice that the value of

$$B_q(0) = \frac{dB(0)}{dq} = \begin{cases} -\frac{rK}{q^2} & \text{if } \ q > r \\ 0 & \text{if } \ q < r \\ \text{undefined} & \text{if } \ q = r \end{cases} \tag{4.5}$$

Suppose we know the value of $\{B_q(k\Delta t_3)\ |\ 0 \le k \le N, k \in \mathbb{N}\}$, we want to find $E = B_q\Big((N+1)\Delta t_3\Big)$. For $0 < u \le N\Delta t_3$, we set $B_q(u)$ to be piece-wise linear function of $\{B_q(k\Delta t_3)\ |\ 0 \le k \le N\ , k \in \mathbb{N}\}$ Otherwise, for $u > N\Delta t_3$, we set $B_q(u)$ to be the value of piece-wise linear function of $B_q(N\Delta t_3)$ and E.

Then, we set $E_1 = B_q(N\Delta t_3)$ and $E_2 = B_q(N\Delta t_3) + \delta_2$ as upper bound and lower

doi:10.6342/NTU202302497

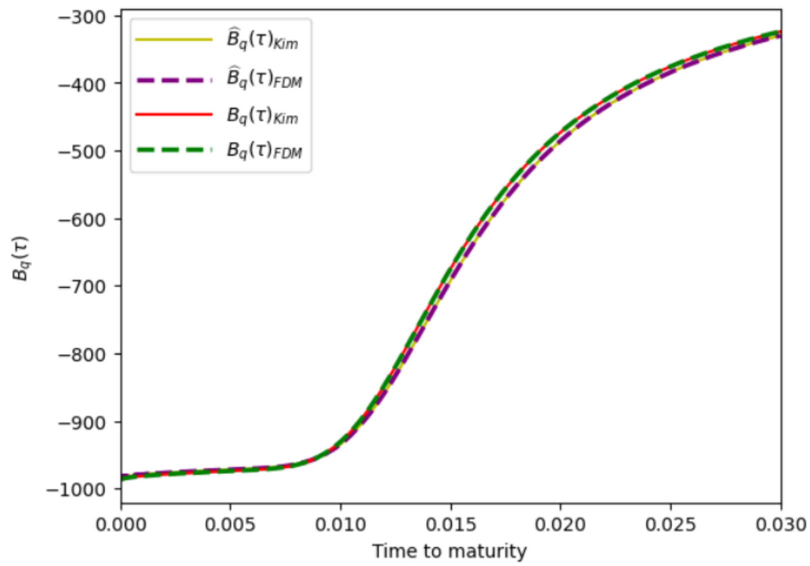
bound of $B_q\Big((N+1)\Delta t_3\Big)$. For $q > r$, δ_2 should be over 0 ; for $q < r$, δ_2 should be under 0. Using the bisection method, we can solve $E = B_q\Big((N+1)\Delta t_3\Big)$

4.3 Numerical Results

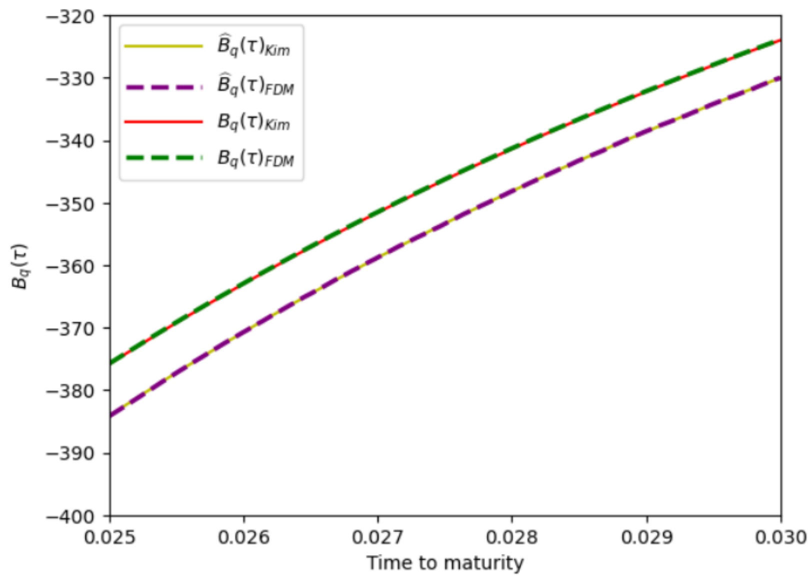
We establish the difference of $B(\tau)_{Kim}, B(\tau)_{FDM}$ with $q = 0.045$ and $q = 0.0451$, $\widehat{B}_q(\tau)_{Kim}$ and $\widehat{B}_q(\tau)_{FDM}$ as the benchmark.[3] By the formula (4.1), we obtain $B_q(\tau)_{Kim}$ and $B_q(\tau)_{FDM}$ using the input data $B(\tau)_{Kim}$ and $B(\tau)_{FDM}$. In Figure 4.1a, we find $B_q(\tau)_{Kim}$ and $B_q(\tau)_{FDM}$ are close to $\widehat{B}_q(\tau)_{Kim}$ and $\widehat{B}_q(\tau)_{FDM}$. It means the numerical result of the derivative with respect to the dividend yield is successful. RMSE of $B_q(\tau)_{Kim}$ and $B_q(\tau)_{FDM}$ is 6.20×10^{-2}. It means the numerical result is insensitive to the input data $B(\tau)$. RMSE of $B_q(\tau)_{Kim}$ and $\widehat{B}_q(\tau)_{Kim}$ is 9.67, which we can also see in 4.1b. We think it is because the difference $\Delta q = 0.0451 - 0.045 = 0.0001$ is not small enough to reveal the accurate data of the derivative. The lower Δq is, the lower RMSE of $B_q(\tau)_{Kim}$ and $\widehat{B}_q(\tau)_{Kim}$ will be.

[3] $\widehat{B}_q(\tau) = \Big(B(\tau, q = 0.045) - B(\tau, q = 0.0451)\Big)/0.0001$

doi:10.6342/NTU202302497



(a) $B_q(\tau)$ comparison with $\tau = 0 \sim 0.03$



(b) $B_q(\tau)$ comparison with $\tau = 0.025 \sim 0.03$

Figure 4.1: $B_q(\tau)$ comparison
With parameter $K = 50,\ r = 0.04,\ q = 0.045,\ \sigma = 0.3$

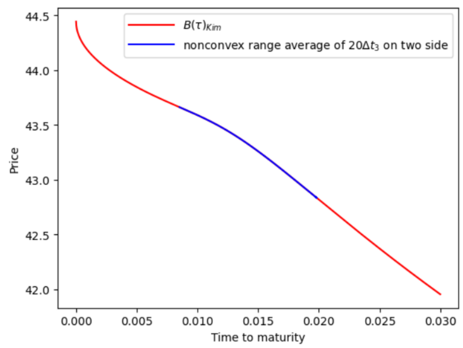
doi:10.6342/NTU202302497

Chapter 5

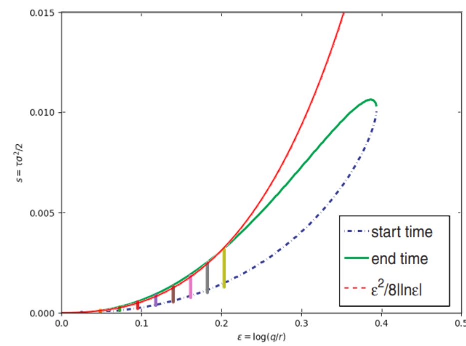
Nonconvexity Explanation

To discuss the nonconvex range of $B(\tau)$, we can check if $B(\tau) > \frac{B(\tau-\epsilon)+B(\tau+\epsilon)}{2}$ though we do not have the accurate result of $B_\tau(\tau)$. In Figure 5.1a, we can see the nonconvex range of $B(\tau)$. In Figure 5.1b, we compare the nonconvex result of different q with Figure 2.2 in Chen et al. (2013). The nonconvex range is almost the same as the numerical result of nonconvex range of Figure 2.2 in Chen et al. (2013). It says that $B(\tau)_{Kim}$ is accurate enough to check the convexity, though we can not derive $B_\tau(\tau)_{Kim}$. A closer look at the upper and lower bound of the nonconvex range, they get higher as q increases with other parameters remaining the same. For $q >> r$ ($q > 1.5r$), $B(\tau)$ regains convex. For the explanation of it, we should look at $B_q(\tau)$ for different q first.

doi:10.6342/NTU202302497



(a) $B(\tau)$ nonconvex range by checking if $B(\tau) > \frac{B(\tau-20\Delta t_3)+B(\tau+20\Delta t_3)}{2}$

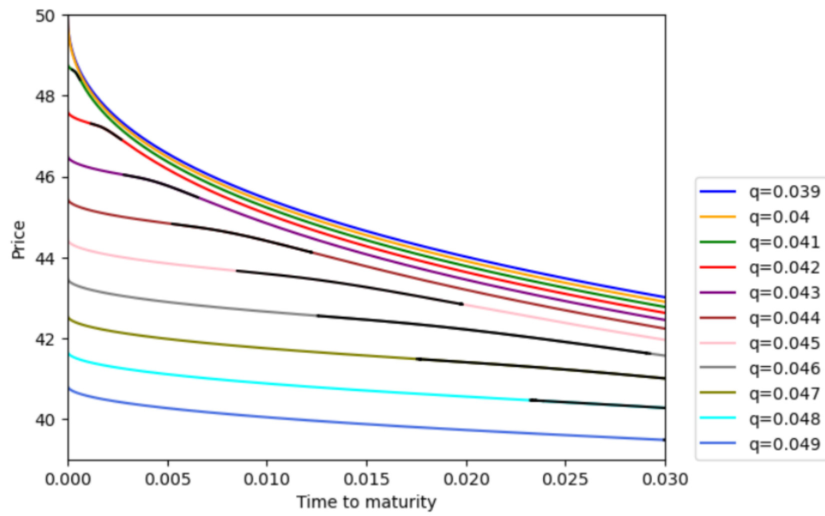


(b) The long bars are the nonconvexity result of $B(\tau)_{Kim}$ for different q.The teardrop region is the numerical result of nonconvex range in Chen et al. (2013)

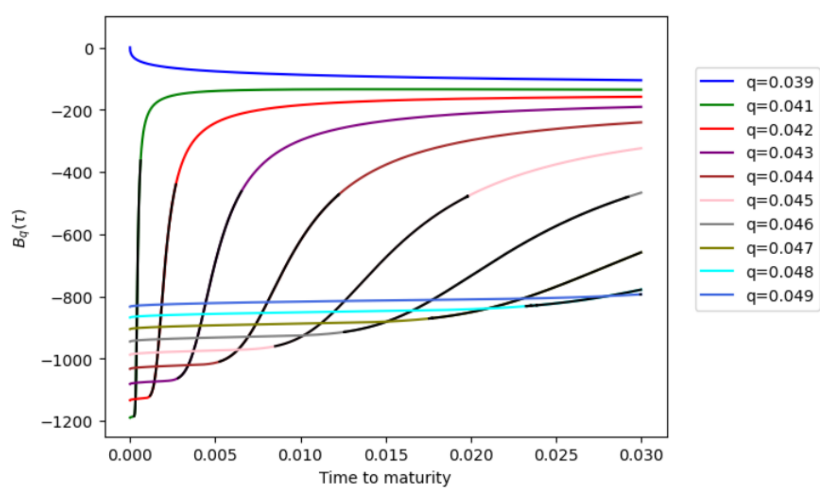
Figure 5.1: Nonconvex range discussion

In Figure 5.2b, we can see $B_q(\tau)$ is a decreasing function for $q < r$ at first. As q is slightly higher than r (i.e. $q = 0.041$), the strong power at $\tau = 0$ pulls $B(\tau)$ down like a rope, and $B_q(\tau)$ is $B(\tau)$'s speed of descend. As q continually increasing, the descending power of $B_q(0)$ slow down, but $B_q(\tau)$ start to strongly fall down from expiry to $\tau >> 0$ because of the momentum of $B_q(0)$ effect. We can also find that the nonconvex range is the steepest range of $B_q(\tau)$. It shows that the steepness of $B_q(\tau)$ is the reason of the nonconvexity of $B(\tau)$. That is when there is a significant difference in descending speeds of $B(\tau)$ between adjacent points with respect to q change, the convexity of $B(\tau)$ can not maintain.

doi:10.6342/NTU202302497



(a) $B(\tau)$ comparison and nonconvex range



(b) $B_q(\tau)$ comparison and the nonconvex range of $B(\tau)$

Figure 5.2: Nonconvexity explanation with different dividend yields

With parameter $K = 50,\ r = 0.04,\ \sigma = 0.3$. Nonconvex range of $B(\tau)$ with black color on it.

doi:10.6342/NTU202302497

Chapter 6

Conclusion

We give a novel method of evaluating the early exercise boundary. The $Kim's$ method can solve the disadvantage of the stepwise function in the finite difference method. We can also derive the derivative of the boundary by the same equation following the $Kim's$ method.

For the derivative of the early exercise boundary with respect to time, the numerical result is not good. We find the reason may be the error term $O(\Delta t)$ on the original statement in Kim (1990), causing the missing term in the numerical result of the derivative of the boundary. Or it may be the accuracy of the numerical integration is not good. But this does not mean the formula is useless, we believe that the formula can be used in deep learning to evaluate the derivative of the boundary

doi:10.6342/NTU202302497

in the future.

For the derivative of the early exercise boundary with respect to the dividend yield, the numerical result is really close to the difference of the early exercise boundary with respect to the dividend yield. We think the formula can even substitute the difference of the boundary with respect to the dividend yield for not having the problem of Δq error in the difference of the boundary.

We also give a view of the reason why the early exercise boundary loses convexity when $0 < q - r << 1$ by the numerical result of the early exercise boundary and its derivative with respect to the dividend yield. In Figure 5.2b, it says that $B_q(0)$ is the main reason for the steep part of $B_q(\tau)$ when $0 < q - r << 1$. The steep part of $B_q(\tau)$ makes $B(\tau)$ lose convexity. When the dividend yield continues to increase, the power of $B_q(0)$ slows down. The steep part of $B_q(\tau)$ disappears, causing $B(\tau)$ regains convexity.

doi:10.6342/NTU202302497

Appendix

A.1 Choice of finite difference method parameter ΔS and Δt

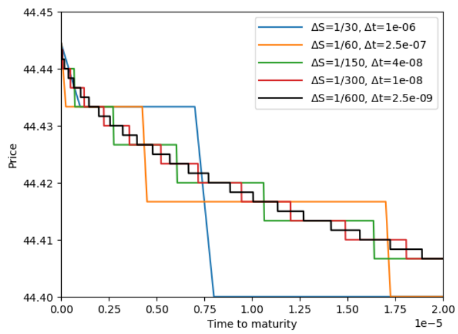
In Wilmott (1998), we know that we should fix the ratio $\frac{\Delta S^2}{\Delta t}$ to control the error of finite difference method to stably decrease as $\Delta t \to 0$.

Therefore, we try to narrow down ΔS and Δt at the same time at a reasonable time.

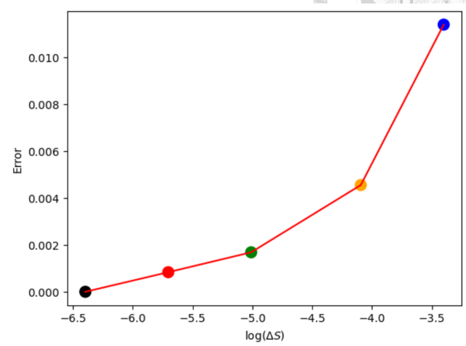
In Figure A.1, we can see the comparison of different ΔS, Δt with the same ratio $\frac{\Delta S^2}{\Delta t} = \frac{10000}{9}$. Finally we choose $\Delta S = \frac{1}{300}, \Delta t = 10^{-8}$ (the red one).

Note that if we consider the tree model case, we have $\frac{\Delta S}{S} = 2\sigma\sqrt{\Delta t}$ We take $\sigma = 0.3, S = 50$ for Figure 2.1 case. Then, we have $\frac{\Delta S^2}{\Delta t} = 900$. That is close to our choice.

doi:10.6342/NTU202302497



(a) $B(\tau)$ for finite difference method with different parameter comparison



(b) RMSE with respect to finite difference method with parameter $\Delta S = \frac{1}{600}$, $\Delta t = 2.5 \times 10^{-8}$

Figure A.1: Finite difference method $\Delta S, \Delta t$ choice

A.2 Convergence of the integral

For the convergence of the integration of (3.5), (3.6) and (4.3)

$$
\begin{aligned}
g\Big(B(\tau), \tau; B(u), u\Big) = \frac{\partial f(\cdot)}{\partial B(\tau)} &= -\frac{rKe^{-r(\tau-u)}n(d_2^\star)}{B(\tau)\sigma\sqrt{\tau-u}} + \frac{qe^{-q(\tau-u)}n(d_1^\star)}{\sigma\sqrt{\tau-u}} - qe^{-q(\tau-u)}N(-d_1^\star) \\
&= O\Big((\tau-u)^{-\frac{1}{2}}\Big) + O\Big((\tau-u)^{-\frac{1}{2}}\Big) + O(1) \\
&= O\Big((\tau-u)^{-\frac{1}{2}}\Big) \quad \text{as} \quad u \to \tau^-
\end{aligned}
$$

doi:10.6342/NTU202302497

$$
\begin{aligned}
h\Big(B(\tau),\tau;B(u),u\Big) = \frac{\partial f(\cdot)}{\partial \tau} &= -rKe^{-r(\tau-u)}\left[rN(-d_2^\star) + \frac{n(d_2^\star)}{\tau-u}\left(\frac{d_2^\star}{2} - \frac{\ln(\frac{B(\tau)}{B(u)})}{\sigma\sqrt{\tau-u}}\right)\right] \\
&\quad + qB(\tau)e^{-q(\tau-u)}\left[qN(-d_1^\star) + \frac{n(d_1^\star)}{\tau-u}\left(\frac{d_1^\star}{2} - \frac{\ln(\frac{B(\tau)}{B(u)})}{\sigma\sqrt{\tau-u}}\right)\right] \\
&= -rKe^{-r(\tau-u)}\left[rN(-d_2^\star) + k_2\Big(B(\tau),\tau;B(u),u\Big)\right] \\
&\quad + qB(\tau)e^{-q(\tau-u)}\left[qN(-d_1^\star) + k_1\Big(B(\tau),\tau;B(u),u\Big)\right] \\
&= O\Big((\tau-u)^{-\frac{1}{2}}\Big) \quad \text{as} \quad u \to \tau^-
\end{aligned}
$$

For the third equality, we know that

$$
\begin{aligned}
k_2\Big(B(\tau),\tau;B(u),u\Big) &= n(d_2^\star)\left(-\frac{\ln\Big(B(\tau)\Big) - \ln\Big(B(u)\Big)}{2\sigma\sqrt{\tau-u}^3} + \frac{r-q-\frac{\sigma^2}{2}}{2\sigma\sqrt{\tau-u}}\right) \\
&= n(d_2^\star)\left(O\Big((\tau-u)^{-\frac{1}{2}}\Big) + O\Big((\tau-u)^{-\frac{1}{2}}\Big)\right) \quad \text{(similar technique in Lemma1)} \\
&= O\Big((\tau-u)^{-\frac{1}{2}}\Big) \quad \text{as} \quad u \to \tau^-
\end{aligned}
$$

Similarly, $k_1\Big(B(\tau),\tau;B(u),u\Big) = O\Big((\tau-u)^{-\frac{1}{2}}\Big) \quad \text{as} \quad u \to \tau^-$

$$
\begin{aligned}
l\Big(B(\tau),\tau;B(u),u\Big) = \frac{\partial f(\cdot)}{\partial B(u)} &= O\Big((\tau-u)^{-\frac{1}{2}}\Big) - O\Big((\tau-u)^{-\frac{1}{2}}\Big) \\
&= O\Big((\tau-u)^{-\frac{1}{2}}\Big) \quad \text{as} \quad u \to \tau^-
\end{aligned}
$$

doi:10.6342/NTU202302497

And we know the fact that

$$\int_0^{\tau} (\tau - u)^{-\frac{1}{2}} du < \infty$$

Therefore, the integrals all converge.

A.3 $\alpha(\tau)$ regression

In Evans et.al. (2002) , it gives us an expression of $B(\tau)$

$$B(\tau) = Ke^{x_f(\tau)} = \frac{rK}{q} e^{-\alpha\left(\frac{\sigma^2}{2}\tau\right)\sigma\sqrt{2\tau}} \tag{A.1}$$

where $\alpha(\cdot)$ is an unknown function.

Furthermore, it gives an approximation

$$\alpha_0 := \lim_{\tau \to 0^+} \alpha(\tau) = 0.4517... \quad \text{for } q > r \tag{A.2}$$

We wonder the relationship between $\alpha(\tau)$ and σ, r, q, τ and find a regression

$$\alpha(\tau) = \beta_0 + \beta_1\left(\frac{\ln r}{\sqrt{\tau}}\right) + \beta_2\left(\frac{\ln q}{\sqrt{\tau}}\right) + \beta_3\left(\frac{1}{\sigma\sqrt{\tau}}\right) + \epsilon \tag{A.3}$$

doi:10.6342/NTU202302497

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.839
Model:                            OLS   Adj. R-squared:                  0.839
Method:                 Least Squares   F-statistic:                 2.169e+05
Date:                Fri, 05 May 2023   Prob (F-statistic):               0.00
Time:                        23:58:58   Log-Likelihood:             3.8344e+05
No. Observations:              125000   AIC:                        -7.669e+05
Df Residuals:                  124996   BIC:                        -7.668e+05
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.4516   4.54e-05   9936.288      0.000       0.451       0.452
x1          4.823e-05   8.51e-07     56.695      0.000    4.66e-05    4.99e-05
x2          5.821e-06   7.72e-07      7.540      0.000    4.31e-06    7.33e-06
x3          1.217e-05   1.69e-07     71.898      0.000    1.18e-05    1.25e-05
==============================================================================
Omnibus:                    25548.626   Durbin-Watson:                   0.082
Prob(Omnibus):                  0.000   Jarque-Bera (JB):           917485.949
Skew:                          -0.131   Prob(JB):                         0.00
Kurtosis:                      16.270   Cond. No.                     1.91e+03
==============================================================================
```

Figure A.2: $\alpha(\tau)$ regression result by the data of finite difference method

Each r, q, σ with 5 different value, and $0 < \tau \leq 10^{-4}$ with 1000 different value

The result is in Figure A.2 .

We can see the intercept $\hat{\beta}_0 = 0.4516$, which is really close to α_0 mentioned in Evans et.al (2002). Also, all t-values are over 5. It means that all variables are important and indispensable. Therefore, we can simply approximate the early exercise boundary near expiry using this regression formula.

doi:10.6342/NTU202302497

References

Chen, X., Chadam, J., Jiang, L., and Zheng, W. (2008). Convexity of the Exercise Boundary of the American Put Option on a Zero Dividend Asset. Mathematical Finance, 18, 185—197.

Chen, X., Cheng, H., and Chadam, J. (2013). Nonconvexity of the Optimal Exercise Boundary for American Put Option on a Dividend-Paying Asset. Mathematical Finance, 23, 169—185.

Evans, J. D., Kuske, R., and Keller, J. B. (2002). American Options on Assets with Dividends near Expiry. Mathematical Finance, 12, 219—237.

Hull, J. C. (2011). Options, Futures, and Other Derivatives (8th ed.). Pearson College Div.

Kim, I. J. (1990). The Analytic Valuation of the American Options. The Review of Financial Studies, 3, 547—572.

Wilmott, P. (1998). Derivatives: The Theory and Practice of Financial Engineer-

doi:10.6342/NTU202302497

ing (1st ed.). Wiley.

doi:10.6342/NTU202302497