Quantitative Finance

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rquf20

FX Open Forward

Julien Hok & Alex S.L. Tse

To cite this article: Julien Hok & Alex S.L. Tse (2024) FX Open Forward, Quantitative Finance, 24:8, 1037-1055, DOI: 10.1080/14697688.2024.2388802

To link to this article: https://doi.org/10.1080/14697688.2024.2388802

Published online: 27 Aug 2024.

Quantitative Finance, 2024
Vol. 24, No. 8, 1037–1055, https://doi.org/10.1080/14697688.2024.2388802

FX Open Forward

JULIEN HOK† and ALEX S.L. TSE *‡

†Investec Bank, 30 Gresham St, London EC2V 7QN, UK
‡Department of Mathematics, University College London, London WC1H 0AY, UK

(*Received 29 February 2024; accepted 27 July 2024*)

FX Open Forward is a derivative instrument where the contract holder has the obligation to purchase a specific amount of foreign currency under a fixed exchange rate by the contract expiry date. In contrast to a traditional forward contract, a distinctive feature of FX Open Forward is that the timing and notional size of the currency conversion can be freely chosen by the contract holder. Under a Black–Scholes model where interest rates can be negative, we provide a complete solution of the early exercise strategy of an FX Open Forward. When domestic rate and foreign rate are both positive (negative), the full contractual notional should be exercised when the spot FX level is sufficiently high (low). Unlike American options, the optimal waiting region of FX Open Forward is always connected even when interest rates are negative.

Keywords: FX Open Forward; American derivative; Optimal stopping; Free-boundary problem

JEL Classification: D81, F31, G12

1. Introduction

Forward contract is arguably one of the most basic derivative instruments. The long party of a forward has the obligation to purchase on the expiry date a specific amount of the underlying asset at a fixed delivery price agreed at the inception of the contract. Static replication of a forward contract is theoretically possible. This involves a long position of the underlying asset plus some investment in the money market account at time zero. Consequently, the pricing of a forward is relatively simple where its fair value is linear in the spot price of the asset and is insensitive to the underlying price dynamics.

Forward contract can serve as a convenient tool for investors to hedge against financial risk. For example, a multinational corporation receiving foreign income may wish to eliminate the exchange rate risk because of the requirement to report their profit in domestic currency. They can then enter an FX forward to convert the foreign income into the domestic currency at a locked exchange rate. One main limitation of FX forward, as a hedging tool, is that the settlement of the domestic currency against the foreign currency can only take place on the maturity date of the contract. But the hedger may not have perfect information about when exactly the foreign cash flows will be incurred when they first enter the contract. If the timing of the cash flows is drastically different from the contract expiry date, such mismatch can pose significant market risk and liquidity risk to the hedger.

FX Open Forward† is a popular alternative to the traditional FX 'Fixed' Forward. Instead of exchanging the notional on the contract expiry date only, FX Open Forward allows the long party of the contract to freely convert (i.e. exercise) any arbitrary fraction of the notional at the locked exchange rate at any time during the contract lifetime, subject to the obligation that the full contractual notational needs to be exercised by the expiry date. The additional flexibility with the size and the timing of the exercise allow the hedger to perform currency conversion 'on demand', eliminating the uncertainty of cash flows timing mismatch.

Despite the popularity of FX Open Forward, there appear to be very few studies investigating its pricing behaviours. To fill in the gap in the literature, in this paper we rigorously analyse the problem of pricing and optimal exercise of an FX Open Forward. Our first theoretical result is that the optimal exercise strategy to be adopted by the long party involves converting the entire notional at once under a very general arbitrage-free market model with a linear pricing rule. In other words, the flexibility to partially settle any fraction of the notional is indeed redundant. An FX Open Forward is therefore financially equivalent to an American forward contract. This result is perhaps not surprising in view of linear

*Corresponding author. Email: alex.tse@ucl.ac.uk

† The product is also known as flexible forward, flexi-forward, Open FX Forward, etc. Apparently there is no market consensus on the naming convention yet.

pricing of financial derivatives. If it is ever optimal to exercise a fraction of the notional at a certain time point, then the same optimality criterion should also hold for any remaining fraction of contract. The optimal exercise rule is thus an 'all-or-nothing' strategy: the long party of the contract either does nothing or settles the full notional at once. While this result appears to be well understood among practitioners, we confirm that such an optimal exercise rule is indeed robust to the modelling setup and does not rely on additional assumptions like Markovian property nor dynamic programming principle.

Our second contribution is to give a thorough analysis and description of the optimal exercise strategy of an FX Open Forward under a Black–Scholes model, covering both the cases of positive and negative interest rates. Since we have established that FX Open Forward can be regarded as an American forward, the problem can be analysed via standard theories of optimal stopping. We adopt the free-boundary approach to study the analytical behaviours of the optimal exercise boundary. The long party of the FX Open Forward should exercise the full notional when the spot FX is sufficiently high (low) when domestic and foreign interest rates are positive (negative). While the techniques behind are fundamentally similar to those for a standard American option and there indeed are cases that the pricing behaviours of FX Open Forward and American call option coincide, the optimal exercise strategy of FX Open Forward is drastically different from that of American call option when interest rates are negative. In general, the long party of an FX Open Forward has a strong incentive to early exercise when the contract is out-of-money to prevent the loss from further looming under negative interest rate. In contrast, it is never optimal to early exercise an out-of-money American call option because the payoff is floored at zero while there is a strictly positive probability of realising a gain if the spot FX rallies in the future.† To the best of our knowledge, the theoretical properties of the optimal exercise strategy of an FX Open Forward (or equivalently an American forward) have not been comprehensively documented in the literature to date.

All things considered, American forward is indeed simpler to be theoretically analysed relative to its option counterpart. Our contributions are therefore not about development of new techniques to solve a difficult optimal stopping problem. Instead, we are interested in unravelling when, how and why the early exercise strategies might differ fundamentally across these two types of American derivatives from economic as well as mathematical perspectives. By focusing on a 'simple' Black–Scholes model with constant interest rates, we can build intuitions towards the similarities and differences via more transparent theoretical properties and numerical analysis. Further practical modelling considerations and other interesting avenues of future research will be discussed in the concluding section.

We conclude the introduction by briefly discussing some related works. As far as we are aware of, the literature directly addressing FX Open Forward and American forward is very limited. Based on heuristic arguments, Kwok and Lau (2003) point out that the flexibility of exercising any arbitrary fraction of the notional during the lifetime of an American FX Forward is redundant, and they provide a numerical example showing how the optimal exercise boundary behaves under positive interest rates. Giribone and Ligato (2016) study numerical pricing of FX Open Forward by tree-based methods. Hölbl and Lovric (2019) propose a least-square Monte Carlo algorithm to evaluate FX Open Forward by relating the product to a swing option.‡ In parallel, there is a long strand of papers dedicated to careful theoretical analysis of American (put) option under Black–Scholes model. Earlier works on this topic include McKean (1965), Jacka (1991) and Myneni (1992), among others. A more complete and contemporary analysis of American option can be found in Battauz *et al.* (2015), covering the cases of negative interest rates as well which could result in disconnected waiting regions. Publications on Open Forward (American forward) contracts appear to be scarce, and we will show that the optimal exercise strategy can behave rather differently from that of an American call option. More generally, our findings also complement the growing literature of American derivatives under negative interest rates (see, e.g., Battauz *et al.* 2012, 2015, 2022a), where the consideration of forward payoff leads to new (albeit simpler) early exercising behaviours. FX Open Forward can also serve as an example of American derivative under which the near-maturity behaviours of the early exercise boundary can be analysed readily. For American option, such problem has been comprehensively studied in the literature (see Evans *et al.* 2002, Lamberton and Villeneuve 2003, Battauz *et al.* 2022b and the references therein). The technical subtlety with option is that one needs to make careful distinction whether the early exercise boundary approaches the strike price near maturity (at which the option payoff function is not differentiable), and the corresponding asymptotic expression of the boundary function might have different forms. But a forward payoff is smooth everywhere and hence the near-maturity behaviour is always in a parabolic form regardless of the model parameters used.

The rest of this paper is organised as follows. In Section 2, we consider a general complete and arbitrage-free market to demonstrate it is optimal for the FX Open Forward holder to exercise the entire notional at once, i.e. the contract is financially equivalent to an American forward. Section 3 presents the analysis of the FX Open Forward (American forward) in a classical Black–Scholes model. Both cases of infinite and finite maturity are considered, and we give a complete description of the optimal exercise strategy which property depends crucially on the signs of the domestic and foreign interest rates. Some numerical illustrations are presented in Section 4. Section 5 concludes. Further technical

† While we exclusively focus on FX derivatives in this paper, similar implications can be generalised to other asset classes, such as commodity derivatives, where a negative dividend yield can be interpreted as the warehouse storage cost of the physical underlying.

‡ Note that, however, there is an important fundamental difference between swing option and FX Open Forward. For swing option, there is typically a constraint on the maximal notional (known as the *daily contract quantity*) that can be exercised on each trading date, whereas for FX Open Forward one can opt to exercise the full contractual notional. Likewise, 'flexible forward' written on energy commodities (see, e.g. Dar *et al.* 2022) has similar restriction on the minimal and maximal amount over the power to be supplied over each time interval.

results and proofs not presented in the main body of the paper are collected in an appendix.

2. American derivative with partial early exercise

The key feature of an FX Open Forward is that the contract holder can early exercise any proportion of the underlying notional during the contract lifetime, provided that the sum of the early exercise proceeds must equal to the full notional by the expiry date. This product can therefore be broadly viewed as an American derivative with partial early exercise, whereas a standard American derivative requires the entire contract to be settled upon early exercise. In this section, we argue that an American derivative with partial early exercise is financially equivalent to a standard American derivative.

To develop intuitions, we will first focus on a discrete-time model. Consider a sequence of trading dates indexed by $n \in \{1, \ldots, M\}$. On some filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n=0,\ldots,M}, \mathbb{Q})$, let $S = (S_n)_{n=0,1,\ldots,M}$ be an $\{\mathcal{F}_n\}$-adapted process such that S_n is the price level of a risky asset on date n. We also introduce an $\{\mathcal{F}_n\}$-adapted domestic risk-free money market process $R = (R_n)_{n=0,1,\ldots,M}$ which is strictly positive with $R_0 \equiv 1$. One unit of the domestic currency deposited at time zero will grow to R_n at time n.

Consider a derivative product with payoff function $g(\cdot)$ and maturity date M written on N units of the risky asset (N will often be referred to as the 'notional', as we typically consider a unit of foreign currency as the underlying risky asset). On each date $n \in \{0, \ldots, M\}$, the derivative holder can partially exercise a fraction $\theta_n \in [0, N]$ of the contract where they will then receive an immediate payoff of $\theta_n g(S_n)$, which discounted value is $R_n^{-1}\theta_n g(S_n)$. An exercise strategy can be identified as an $\{\mathcal{F}_n\}$-adapted, non-negative process $\theta = (\theta_n)_{n=0,\ldots,M}$ such that $\sum_{n=0}^{M} \theta_n = N$. This equality constraint enforces the requirement that the full contractual notional must be entirely settled by the expiry date.

We assume the financial market is arbitrage-free, and the underlying probability measure $\mathbb{Q}$ is the associated risk-neutral measure.† The time-zero fair price of the contract is

$$\sup_{\theta \in \mathscr{A}} \mathbb{E}\left[\sum_{n=0}^{M} R_n^{-1}\theta_n g(S_n)\right]. \tag{1}$$

Throughout this paper, $\mathbb{E}(\cdot)$ denotes the expectation operator under the probability measure $\mathbb{Q}$. The supremum in (1) is taken over the class of admissible exercise strategies in the form of

$$\mathscr{A} := \Big\{\theta = (\theta_n)_{n=0,\ldots,M} : \theta \text{ is } \{\mathcal{F}_n\}\text{-adapted}, \theta_n \in [0, N] \text{ for all } n, \sum_{n=0}^{M} \theta_n = N\Big\}.$$

At the first sight, the possibility of partial exercise seems to create additional financial value to the contract holder relative to a standard American derivative written on N units of the asset. But it turns out that such flexibility is indeed redundant, as the following proposition shows.

Proposition 1 *Let $\mathscr{T}$ be the set of $\{\mathcal{F}_n\}$-stopping times valued on $\{0, 1, \ldots, M\}$. Then*

$$\sup_{\theta \in \mathscr{A}} \mathbb{E}\left[\sum_{n=0}^{M} R_n^{-1}\theta_n g(S_n)\right] = N \sup_{\tau \in \mathscr{T}} \mathbb{E}\left[R_\tau^{-1} g(S_\tau)\right].$$

Moreover, if τ^* is the optimiser to $\sup_{\tau \in \mathscr{T}} \mathbb{E}[R_\tau^{-1} g(S_\tau)]$, then

$$\sup_{\theta \in \mathscr{A}} \mathbb{E}\left[\sum_{n=0}^{M} R_n^{-1}\theta_n g(S_n)\right] = \mathbb{E}\left[\sum_{n=0}^{M} R_n^{-1}\theta_n^* g(S_n)\right]$$

under the choice of $\theta_n^* := N 1_{(n=\tau^*)}$.

Proof Define $J(\theta) := \mathbb{E}[\sum_{n=0}^{M} R_n^{-1}\theta_n g(S_n)]$ and $H(\tau) := \mathbb{E}[R_\tau^{-1} g(S_\tau)]$. For any given $\tau \in \mathscr{T}$, consider a process $\hat{\theta}$ defined via $\hat{\theta}_n = N 1_{(n=\tau)}$. Then by construction, $\hat{\theta} \in \mathscr{A}$ and hence

$$\begin{aligned}\sup_{\theta \in \mathscr{A}} J(\theta) \geq J(\hat{\theta}) &= N\mathbb{E}\left[\sum_{n=0}^{M} R_n^{-1} 1_{(n=\tau)} g(S_n)\right] \\ &= N\mathbb{E}\left[R_\tau^{-1} g(S_\tau)\right] = NH(\tau).\end{aligned}$$

As τ is arbitrary, taking supremum over $\tau \in \mathscr{T}$ leads to $\sup_{\theta \in \mathscr{A}} J(\theta) \geq N \sup_{\tau \in \mathscr{T}} H(\tau)$.

To show the reverse inequality, for any given $\theta \in \mathscr{A}$ and $x \in [0, N]$, define a random variable

$$\hat{\tau}(x) := \inf\left\{k \geq 0 : \sum_{i=0}^{k} \theta_i \geq x\right\}.$$

In words, $\hat{\tau}(x)$ represents the time when x out of N units of the underlying asset has been exercised. If we interpret $\hat{\tau}(x)$ as a stochastic process indexed by x, then $\hat{\tau}(x)$ is a left-continuous, increasing pure jump process with state space $\{0, 1, \ldots, M\}$ such that $\hat{\tau}(0) = 0$ and $\hat{\tau}(N) \leq M$. Importantly, $\hat{\tau}(x) \in \mathscr{T}$ for

† A discrete-time financial model is generally incomplete where the risk-neutral measure is not unique. One may interpret a discrete-time model as an approximation of a complete continuous-time model that arises within a specific numerical procedure (e.g. lattice tree) which determines $\mathbb{Q}$. Another common industrial practice to get around incompleteness is to fix a stochastic model of the underlying asset which is then calibrated using market prices of liquidly traded options. The pricing measure $\mathbb{Q}$ to be used is then jointly implied by the modelling choice and the market. In our exposition, we simply assume $\mathbb{Q}$ is exogenously given.

all x and $\theta_n = \int_0^N 1_{(\hat{\tau}(x)=n)}\,\mathrm{d}x$. Then

$$\begin{aligned}
J(\theta) &= \mathbb{E}\left[\sum_{n=0}^{M} R_n^{-1}\theta_n g(S_n)\right] \\
&= \mathbb{E}\left[\sum_{n=0}^{M} R_n^{-1}\left(\int_0^N 1_{(\hat{\tau}(x)=n)}\,\mathrm{d}x\right) g(S_n)\right] \\
&= \int_0^N \mathbb{E}\left[\sum_{n=0}^{M} R_n^{-1} 1_{(\hat{\tau}(x)=n)} g(S_n)\right]\mathrm{d}x \\
&= \int_0^N \mathbb{E}\left[R_{\hat{\tau}(x)}^{-1} g(S_{\hat{\tau}(x)})\right]\mathrm{d}x \\
&= \int_0^N H(\hat{\tau}(x))\,\mathrm{d}x \le \int_0^N \sup_{\tau\in\mathscr{T}} H(\tau)\,\mathrm{d}x = N \sup_{\tau\in\mathscr{T}} H(\tau).
\end{aligned}$$

Taking supremum over $\theta \in \mathscr{A}$ now yields $\sup_{\theta\in\mathscr{A}} J(\theta) \le N\sup_{\tau\in\mathscr{T}} H(\tau)$.

For τ^* being an optimiser to $\sup_{\tau\in\mathscr{T}} H(\tau)$, the optimality of θ^* defined via $\theta_n^* = N1_{(\tau^*=n)}$ now follows trivially from the fact that $J(\theta^*) = NH(\tau^*)$. ■

The financial intuition behind Proposition 1 is as follows. The fair price of an American derivative with the possibility to partially exercise is identical to the fair price of a standard American derivative with only one opportunity to early exercise the full notional. Moreover, the optimal partial exercise strategy involves settling the entire notional when it is optimal to exercise the corresponding standard American derivative. To solve problem (1), it is therefore sufficient to price a standard American derivative with payoff function $g(\cdot)$ and maturity M. If the corresponding optimal early exercise strategy is τ^*, then the optimal early exercise strategy for the contract with partial early exercise is to exercise the full notional N at time τ^*. This is a generic result independent of the payoff function as well as the underlying models of S and R. When specialising to the case of $g(s) = s - K$ where K is some fixed delivery price, the contract becomes an FX Open Forward which is the main focus of this paper. The extra flexibility of the partial exercise timing does not make an FX Open Forward more valuable than an American forward.†

The irrelevance of partial exercise is perhaps not too surprising. Intuitively, one should early exercise a unit of American derivative when its intrinsic value (the payoff received if the contract is exercised immediately) is larger than its continuation value (the fair value under the plan of not exercising today but follow the optimal exercise rule thereafter). Under linear pricing, American derivative on each unit of the underlying asset shares the same intrinsic and continuation value. Hence if it is ever optimal to early exercise one unit of the contract because its intrinsic value is higher than its continuation value, this must also be the case for all other units within the contract. The optimal exercise rule is therefore an 'all-or-nothing' rule.

We now show that the redundancy of partial exercise is also true in a continuous-time setup. Fix a maturity date of $T > 0$ and let the underlying probability space be $(\Omega, \mathcal{F}, \mathbb{Q})$ where $\mathbb{Q}$ again denotes the risk-neutral measure. Let $\{\mathcal{F}_t\}_{t\le T}$ be a right-continuous filtration of $\mathcal{F}$. Write $S = (S_t)_{t\le T}$ and $R = (R_t)_{t\le T}$ as the underlying processes for the price of the risky asset and the domestic money market account respectively, which are both strictly positive and $\{\mathcal{F}_t\}$-progressively measurable. Consider an American derivative written on N units of the risky asset with payoff function $g(\cdot)$ which expires at time T. The holder of the contract is allowed to partially exercise any arbitrary fraction of the underlying notional any time on $[0, T]$. Let Φ_t be the cumulative units of the underlying that has been exercised by time t. A feasible partial exercise strategy Φ is identified by $\Phi \in \mathcal{A}$ where

$$\begin{aligned}
\mathcal{A} := \{\Phi = (\Phi_t)_{0\le t\le T} : \Phi \text{ is adapted, right} \\
-\text{ continuous, non-negative, increasing with } \Phi_T = N\}.
\end{aligned}$$

The time-zero fair value of the contract is then given by

$$\sup_{\Phi\in\mathcal{A}} \mathbb{E}\left[\int_0^T R_u^{-1} g(S_u)\,\mathrm{d}\Phi_u\right]. \tag{2}$$

The following result is a continuous-time version of Proposition 1.

Proposition 2 *Let $\mathcal{T}$ be the set of $\{\mathcal{F}_t\}$-stopping times valued on* $[0, T]$. *Then*

$$\sup_{\Phi\in\mathcal{A}} \mathbb{E}\left[\int_0^T R_u^{-1} g(S_u)\,d\Phi_u\right] = N \sup_{\tau\in\mathcal{T}} \mathbb{E}\left[R_\tau^{-1} g(S_\tau)\right].$$

Moreover, if τ^* is the optimiser to $\sup_{\tau\in\mathcal{T}} \mathbb{E}[R_\tau^{-1} g(S_\tau)]$, then

$$\sup_{\Phi\in\mathcal{A}} \mathbb{E}\left[\int_0^T R_u^{-1} g(S_u)\,d\Phi_u\right] = \mathbb{E}\left[\int_0^T R_u^{-1} g(S_u)\,d\Phi_u^*\right]$$

under the choice of $\Phi_t^* := N1_{(\tau^*\ge t)}$.

Proof The proof is largely similar to that of Proposition 1. Define $J(\Phi) := \mathbb{E}[\int_0^T R_u^{-1} g(S_u)\,\mathrm{d}\Phi_u]$ and $H(\tau) := \mathbb{E}[R_\tau^{-1} g(S_\tau)]$. We could easily deduce $\sup_{\Phi\in\mathcal{A}} J(\Phi) \ge N\sup_{\tau\in\mathcal{T}} H(\tau)$ as $(N1_{(t\ge\tau)})_{0\le t\le T} \in \mathcal{A}$ for any $\tau \in \mathcal{T}$.

Now we show the reverse inequality. For any given $\Phi \in \mathcal{A}$, construct its extension $(\tilde{\Phi}_t)_{t\ge 0}$ via $\tilde{\Phi}_t := \Phi_{\min(t,T)}$. Define

$$\hat{\tau}(x) := \inf\left\{u \ge 0 : \tilde{\Phi}_u > x\right\}.$$

Economically, $\hat{\tau}(x)$ represents the first time when more than x out of N units of the underlying asset has been exercised. But mathematically, $x \to \hat{\tau}(x)$ is indeed the right-continuous inverse of $t \to \tilde{\Phi}_t$ with $\hat{\tau}(x) \in \mathcal{T}$ for all $x \in [0, N)$, and $\hat{\tau}(x) = +\infty$ for $x \ge N$. By the time-change formula for Lebesgue–Stieljes integral (see Proposition 1.4, Chapter V of Revuz and Yor 2013), we have

$$\int_{\hat{\tau}(0)}^{\hat{\tau}(N)} R_u^{-1} g(S_u)\,\mathrm{d}\tilde{\Phi}_u = \int_0^N R_{\hat{\tau}(x)}^{-1} g(S_{\hat{\tau}(x)})\,\mathrm{d}\tilde{\Phi}_{\hat{\tau}(x)}.$$

But since $\tilde{\Phi}_t = 0$ on $t < \hat{\tau}(0)$, $\tilde{\Phi}_t = N$ on $t \ge T$, and the processes $(x)_{x\in[0,N)}$ and $(\tilde{\Phi}_{\hat{\tau}(x)})_{x\in[0,N)}$ only differ from each other

† This conclusion might change under a pricing rule that does not satisfy linearity. An example is utility indifference pricing under an incomplete market. See, for example Henderson and Hobson (2009).

on a set with zero Lebesgue measure, we deduce

$$
\begin{aligned}
J(\Phi) &= \mathbb{E}\left[\int_0^T R_u^{-1} g(S_u)\,\mathrm{d}\Phi_u\right] \\
&= \mathbb{E}\left[\int_0^\infty R_u^{-1} g(S_u)\,\mathrm{d}\tilde{\Phi}_u\right] \\
&= \mathbb{E}\left[\int_{\hat{\tau}(0)}^{\hat{\tau}(N)} R_u^{-1} g(S_u)\,\mathrm{d}\tilde{\Phi}_u\right] \\
&= \mathbb{E}\left[\int_0^N R_{\hat{\tau}(x)}^{-1} g(S_{\hat{\tau}(x)})\,\mathrm{d}\tilde{\Phi}_{\hat{\tau}(x)}\right] \\
&= \mathbb{E}\left[\int_0^N R_{\hat{\tau}(x)}^{-1} g(S_{\hat{\tau}(x)})\,\mathrm{d}x\right] \\
&= \int_0^N H(\hat{\tau}(x))\,\mathrm{d}x \le \int_0^N \sup_{\tau\in\mathcal{T}} H(\tau)\,\mathrm{d}x = N \sup_{\tau\in\mathcal{T}} H(\tau).
\end{aligned}
$$

Taking supremum over $\Phi \in \mathcal{A}$ results in $\sup_{\Phi\in\mathcal{A}} J(\Phi) \le N \sup_{\tau\in\mathcal{T}} H(\tau)$. Finally, if τ^* solves $\sup_{\tau\in\mathcal{T}} H(\tau)$, then $J(\Phi^*) = NH(\tau^*)$ for $\Phi_t^* := N 1_{(\tau^*\ge t)}$ and hence Φ^* solves $\sup_{\Phi\in\mathcal{A}} \mathbb{E}[\int_0^T R_u^{-1} g(S_u)\,\mathrm{d}\Phi_u]$. ■

While the previous analysis focuses on a contract with finite maturity, it is not difficult to generalise the above conclusion to a perpetual contract with infinite maturity.

The statements in Propositions 1 and 2 are applicable to a generic modelling setup without relying on the dynamic programming principle. This allows one to confirm the optimality of the 'all-or-nothing' early exercise rule even under, for example, a highly non-Markovian model of stock price and interest rate. This style of optimal strategy is analogous to a 'bang-bang' control involving abrupt switch between the minimum and maximum action values in the control space. In quantitative finance applications, bang-bang strategy commonly arises in the pricing problems of swing options and storage contracts (see , e.g. Barrera-Esteve *et al.* 2006, Bardou *et al.* 2010, Daluiso *et al.* 2020 and references therein). The existence of a bang-bang control as an optimal strategy can facilitate simplification and effective design of numerical pricing procedures like quantisation, Longstaff–Schwarz least-square or neural network methods (Barrera-Esteve *et al.* 2006, Bardou *et al.* 2010). In our problem, this kind of bang-bang optimality allows us to reduce the pricing problem of FX Open Forward to that of a standard American derivative.

3. Valuation of FX Open Forward under Black–Scholes model

From the analysis in Section 2, the pricing of an FX Open Forward can be reduced to the pricing of an American forward. In this section, we will focus on a canonical continuous-time model to deduce carefully the form of the optimal early exercising strategy, whereas such analysis cannot be performed easily in a discrete-time model.

We consider a Black–Scholes model or more precisely a Garman–Kohlhagen model for a foreign exchange market. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge0}, \mathbb{Q})$ be a standard filtered probability space supporting a one-dimensional Brownian motion $B = (B_t)_{t\ge0}$, where $\mathbb{Q}$ is the risk-neutral measure with respect to the domestic risk-free instrument. The constant domestic and foreign risk-free rates are denoted by $r_d \in \mathbb{R}$ and $r_f \in \mathbb{R}$ respectively. Importantly, we do not make any assumption about the signs of r_d and r_f to capture the possibility that interest rates can be negative. Let $S = (S_t)_{t\ge0}$ be the price process of a foreign risky currency (quoted in terms of the domestic currency) which has dynamics of

$$\frac{\mathrm{d}S_t}{S_t} = (r_d - r_f)\,\mathrm{d}t + \sigma\,\mathrm{d}B_t, \tag{3}$$

where $\sigma > 0$ is the FX volatility. Sometimes we will use the notation $S^{t,s} = (S_u^{t,s})_{u\ge t}$ to denote the solution to (3) with initial data $S_t = s$, i.e.

$$
\begin{aligned}
S_u^{t,s} = s \exp\Bigg(\left(r_d - r_f - \frac{\sigma^2}{2}\right)(u-t) \\
+ \sigma(B_u - B_t)\Bigg), \quad u \ge t,\ s > 0.
\end{aligned}
$$

More generally, our framework can be extended to different asset classes beyond foreign exchange upon interpreting r_f as the dividend yield of some risky asset.

To price an FX Open Forward, it is sufficient to look at a standard American derivative with payoff function $g(s) := s - K$. Its time-t fair price is given by the solution to the optimal stopping problem

$$V(t,s) := \sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}\left[e^{-r_d(\tau-t)} g(S_\tau) \mid S_t = s\right], \tag{4}$$

where $\mathcal{T}_{t,T}$ is the set of all $\{\mathcal{F}_t\}$-stopping times valued in $[t,T]$. Our goal is to characterise the solution to (4) and describe the corresponding stopping time which attains the supremum. Our results can be extended to 'short-forward payoff' $g(s) := K - s$ easily using a trivial variation of the put-call symmetry. Throughout this paper, an American derivative with payoff function $s - K$ (resp. $K - s$) is referred to as a purchase (resp. sale) Open Forward. The main body of this paper focuses on purchase Open Forward, while sale Open Forward is considered in Appendix 2.

3.1. Perpetual FX Open Forward

We first provide some preliminary results for a perpetual version of the FX Open Forward when $T = +\infty$ in (4). Due to the time homogenous Markovian structure of the process S, the value of the perpetual contract does not depend on the current time and the problem becomes

$$V_\infty(s) := \sup_{\tau\in\mathcal{T}_{0,\infty}} \mathbb{E}\left[e^{-r_d\tau} g(S_\tau) \mid S_0 = s\right]. \tag{5}$$

In particular, the supremum is taken over all non-negative $\{\mathcal{F}_t\}$-stopping times.

THEOREM 1 *For problem* (5)*:*

(1) *If* $r_d \le 0 \le r_f$, $V_\infty(s) = s - K$ *and* $\tau^* = 0$ *is optimal to problem* (5).

(2) *If* $r_f < 0 \le r_d$, $V_\infty(s) = +\infty$ *and* $\tau^* = +\infty$ *is optimal to problem* (5).
(3) *If* $r_f = 0 < r_d$, $V_\infty(s) = s$ *and* $\tau^* = +\infty$ *is optimal to problem* (5).
(4) *If* $r_f = 0 = r_d$, $V_\infty(s) = s - K$ *which can be attained by any non-negative stopping time* τ.
(5) *If* $r_d > 0$ *and* $r_f > 0$,

$$V_\infty(s) = \begin{cases} \dfrac{b_2}{\gamma_2}\left(\dfrac{s}{b_2}\right)^{\gamma_2}, & s < b_2; \\ s - K, & s \ge b_2, \end{cases} \tag{6}$$

where $\gamma_2 > 1$ *is the larger root to the quadratic equation*

$$\frac{\sigma^2}{2}\gamma^2 + \left(r_d - r_f - \frac{\sigma^2}{2}\right)\gamma - r_d = 0 \tag{7}$$

and $b_2 := \frac{\gamma_2}{\gamma_2 - 1}K$. *The optimal early exercise strategy is* $\tau^* = \inf\{t \ge 0 : S_t \ge b_2\}$, *i.e. to early exercise the contract when the spot price level is sufficiently high.*
(6) *If* $r_d < 0$ *and* $r_f < 0$:
  (a) *If* $r_f \le r_d$, *or* $r_d < r_f$ *and* $\sigma > |\sqrt{-2r_d} - \sqrt{-2r_f}|$, *then* $V_\infty(s) = +\infty$. *The explicit constructions of* τ^* *which attain an infinite contract value can be found in the proof of this proposition in the appendix.*
  (b) *If* $r_d < r_f$ *and* $\sigma \le |\sqrt{-2r_d} - \sqrt{-2r_f}|$, *then*

$$V_\infty(s) = \begin{cases} s - K, & s \le b_1; \\ \dfrac{b_1}{\gamma_1}\left(\dfrac{s}{b_1}\right)^{\gamma_1}, & s > b_1, \end{cases} \tag{8}$$

where $\gamma_1 > 1$ *is the smaller root to the quadratic equation (7) and* $b_1 := \frac{\gamma_1}{\gamma_1 - 1}K$. *The optimal early exercise strategy is* $\tau^* = \inf\{t \ge 0 : S_t \le b_1\}$, *i.e. to early exercise the contract when the spot price level is sufficiently low.*

For a perpetual FX Open Forward, only Case (5) and Case(6)(b) in Theorem 1 lead to a non-trivial early exercise strategy where the necessary condition is $r_d r_f > 0$, i.e. domestic and foreign interest rates need to be non-zero and have the same sign. Under such a combination of parameters, there exists a meaningful trade-off between the extra foreign interests received (governed by r_f) versus the opportunity cost of paying the delivery price early (governed by r_d) when early exercise is considered.

In Case (5) where $r_d > 0$ and $r_f > 0$, the value of the contract indeed coincides with the value of a perpetual American call option which payoff function is $(S - K)^+$ (see, e.g., Merton 1973). When the maturity is infinity and $r_d > 0$, there is no incentive for the contract holder to early exercise the forward at a loss because any negative payout can be indefinitely deferred and discounted away. Hence the contract is financially equivalent to a perpetual American call option. When domestic interest rate is positive, it is well known that early exercise will only take place if foreign interest rate (dividend) is strictly positive. Indeed, on letting $r_f \downarrow 0$ in Case (5) one can verify that $\gamma_2 \downarrow 1$ and in turn $b_2 \uparrow \infty$. The optimal exercise region $[b_2, \infty)$ therefore vanishes and we have $\tau^* \uparrow +\infty$ because early exercise is now deferred indefinitely. In parallel, the corresponding contract value becomes

$$\begin{aligned} \lim_{\gamma_2 \downarrow 1} V_\infty(s) &= \lim_{\gamma_2 \downarrow 1} \frac{K}{\gamma_2 - 1}\left(\frac{\gamma_2 - 1}{\gamma_2 K}s\right)^{\gamma_2} \\ &= \lim_{\gamma_2 \downarrow 1} s^{\gamma_2} K^{1-\gamma_2}\gamma_2^{-\gamma_2}(\gamma_2 - 1)^{\gamma_2 - 1} = s. \end{aligned}$$

Thus we observe that Case (5) degenerates to Case (3) when the foreign interest rate r_f decreases to zero.

However, this equivalence does not hold when $r_d < 0$ as shown in Case (6). The compounding effect under negative domestic interest rate makes it more attractive to realise loss earlier as well as to realise gain later. Away from the mathematically degenerate Case (6)(a), the agent has an incentive to settle the contract before it becomes deeply out-of-money since otherwise any loss will be compounded and grow over time.

Interestingly, under negative rates the early exercise strategy of the FX Open Forward is drastically different from that of an American call option. The incentive to defer gains under negative interest rates also holds in the case of American options, as observed by Battauz *et al.* (2012, 2022b). Consequently, the holder of an American call (put) option will wait in general when the option is deeply in-the-money, i.e. when spot price is sufficiently high (low). However, unlike the Open Forward contract, there is no incentive for an option holder to take loss via early exercising an out-of-money contract because there is only financial upside in waiting until maturity. Such optionality creates another waiting region over sufficiently low (high) spot prices for call (put) option. This is precisely why negative interest rates result in a double-continuation region for American options in the works of Battauz *et al.* (2012, 2022b, 2022a), etc. For American call (put) option, the upper (lower) continuation region originates from the financing motive to postpone gains under negative interest rates while the lower (upper) continuation region originates from the optionality of the underlying derivative contract. In contrast, there is no optionality within an FX Open Forward because the contract payoff $S - K$ is linear. The early exercise strategy is thus simpler which is completely characterised by a single exercise threshold entirely driven by the financing incentive of gain-postponement/loss-anticipation.

3.2. *FX Open Forward with finite maturity*

Now we consider the case that the maturity of the contract is finite (i.e. $T < \infty$ in (4)). We first identify two cases where the early exercise strategy becomes degenerate.

LEMMA 1 *Consider problem* (4) *with finite maturity* $T < \infty$. *If* $r_d \ge 0 \ge r_f$, *it is optimal to defer the exercise of the contract until the maturity date* ($\tau^* = T$). *If* $r_f \ge 0 \ge r_d$, *it is optimal to early exercise the contract immediately* ($\tau^* = t$).

Proof Write $J(t,s;\tau) := \mathbb{E}[e^{-r_d(\tau-t)}g(S_\tau)\,|\,S_t = s]$. Then under $r_d \geq 0 \geq r_f$, for any $\tau \in \mathcal{T}_{t,T}$ we have

$$\begin{aligned}J(t,s;\tau) &= s\mathbb{E}[e^{-r_f(\tau-t)}e^{\sigma(B_\tau-B_t)-\frac{\sigma^2}{2}(\tau-t)}] - K\mathbb{E}[e^{-r_d(\tau-t)}]\\ &\leq se^{-r_f(T-t)}\mathbb{E}[e^{\sigma(B_\tau-B_t)-\frac{\sigma^2}{2}(\tau-t)}] - Ke^{-r_d(T-t)}\\ &= se^{-r_f(T-t)} - Ke^{-r_d(T-t)}.\end{aligned}$$

Taking supremum over $\tau \in \mathcal{T}_{t,T}$ leads to $V(t,s) = \sup_{\tau\in\mathcal{T}_{t,T}} J(t,s;\tau) \leq se^{-r_f(T-t)} - Ke^{-r_d(T-t)}$ and equality is attained under $\tau = T$. Hence we have $V(t,s) = J(t,s;\tau = T) = se^{-r_f(T-t)} - Ke^{-r_d(T-t)}$ and $\tau^* = T$ is optimal.

Similarly, if $r_f \geq 0 \geq r_d$, then

$$\begin{aligned}J(t,s;\tau) &= s\mathbb{E}[e^{-r_f(\tau-t)}e^{\sigma(B_\tau-B_t)-\frac{\sigma^2}{2}(\tau-t)}] - K\mathbb{E}[e^{-r_d(\tau-t)}]\\ &\leq s\mathbb{E}[e^{\sigma(B_\tau-B_t)-\frac{\sigma^2}{2}(\tau-t)}] - K = s - K\end{aligned}$$

for any $\tau \in \mathcal{T}_{t,T}$. Then we conclude $\sup_{\tau\in\mathcal{T}_{t,T}} J(t,s;\tau) \leq s - K$ and equality is attained under $\tau = t$. Hence $V(t,s) = J(t,s;\tau = t) = s - K$ and $\tau^* = t$ is optimal. ∎

The intuition is similar to the one in the perpetual cases: under positive (negative) domestic interest rate, the holder of the derivative prefers paying the deliver price K later (earlier). Furthermore, if the foreign interest rate is negative (positive), then holding the asset physically is disadvantageous (advantageous), and as such the holder should exercise the contract as late (early) as possible. Throughout the rest of this paper, we will impose the below standing assumption.

ASSUMPTION 1 *The domestic and foreign interest rates are both non-zero and have the same sign, i.e.* $r_d r_f > 0$.

3.2.1. Fundamental properties of the value function. Using the Markovian structure of the problem, one can see that the value function in (4) can be rewritten as

$$V(t,s) = \sup_{\tau\in\mathcal{T}_{0,T-t}} \mathbb{E}\left[e^{-r_d\tau}g(S_\tau)\,|\,S_0 = s\right]. \tag{9}$$

We first establish some theoretical properties of the value function $V(t,s)$ defined in (9).

LEMMA 2 *For the value function* $V : [0,T]\times(0,\infty) \to \mathbb{R}$ *defined in* (9)*:*

1. $V(t,s) \geq g(s) = s - K$ *for all* $(t,s) \in [0,T]\times(0,\infty)$.
2. *For any fixed* $s > 0$, $t \to V(t,s)$ *is decreasing with* $V(T,s) = g(s)$.
3. *For any fixed* $t \in [0,T]$, $s \to V(t,s)$ *is a continuous, convex and increasing function.*
4. *For any fixed* $t \in [0,T]$, *write* $f(s) := V(t,s) - s$ *for* $s \in (0,\infty)$. *Then* $f(s)$ *is bounded from below by* $-K$. *Moreover:*
   - (a) *If* $r_d > 0$ *and* $r_f > 0$, *then* $f(0) := \lim_{s\downarrow 0} f(s) = -Ke^{-r_d(T-t)}$, *and* $f(s) = -K$ *for all sufficiently large* s.
   - (b) *If* $r_d < 0$ *and* $r_f < 0$, *then* $f(0) := \lim_{s\downarrow 0} f(s) = -K$ *and* $\lim_{s\to\infty} f(s) = +\infty$.

Proof (1) It follows immediately from the fact that $\tau = 0$ is an admissible stopping time in (9).

(2) For $0 \leq t_1 \leq t_2 \leq T$, $V(t_1,s) = \sup_{\tau\in\mathcal{T}_{0,T-t_1}} \mathbb{E}[e^{-r_d\tau} g(_\tau)] \geq \sup_{\tau\in\mathcal{T}_{0,T-t_2}} \mathbb{E}[e^{-r_d\tau}g(S_\tau)] = V(t_2,s)$ since $\mathcal{T}_{0,T-t_2} \subseteq \mathcal{T}_{0,T-t_1}$. The claim of $V(T,s) = g(s)$ is trivial as the set $\mathcal{T}_{0,0}$ consists of the deterministic stopping time $\tau \equiv 0$ only.

(3) For a fixed $t \leq T$ and an arbitrary $\tau \in \mathcal{T}_{0,T-t}$, the map

$$\begin{aligned}s \to e^{-r_d\tau}g(S_\tau) = e^{-r_d\tau}\Bigg(s\exp\Bigg(\left(r_d - r_f - \frac{\sigma^2}{2}\right)\tau\\ + \sigma B_\tau\Bigg) - K\Bigg) =: w(\tau,s)\end{aligned}$$

is an increasing linear function. Therefore $s \to \mathbb{E}[w(s,\tau)]$ is also an increasing linear function. Hence $V(t,s) = \sup_{\tau\in\mathcal{T}_{0,T-t}} \mathbb{E}[w(\tau,s)]$ is the supremum of some increasing linear functions, which must be convex, increasing and continuous.

(4) The lower bound $f(s) \geq -K$ is simply due to $V(t,s) \geq g(s) = s - K$. Furthermore:

(a) Since $\tau = T - t \in \mathcal{T}_{0,T-t}$, we have

$$\begin{aligned}f(s) &= V(t,s) - s\\ &\geq \mathbb{E}\left[e^{-r_d(T-t)}(S_{T-t} - K)\right]\\ &\quad - s = s(e^{-r_f(T-t)} - 1) - Ke^{-r_d(T-t)}.\end{aligned}$$

To establish an upper bound, observe that

$$\begin{aligned}V(t,s) &= \sup_{\tau\in\mathcal{T}_{0,T-t}} \mathbb{E}\left[e^{-r_d\tau}(S_\tau - K)\right]\\ &= \sup_{\tau\in\mathcal{T}_{0,T-t}} \mathbb{E}\left[\left(se^{\left(-r_f-\frac{\sigma^2}{2}\right)\tau+\sigma B_\tau} - Ke^{-r_d\tau}\right)\right]\\ &\leq s\sup_{\tau\in\mathcal{T}_{0,T-t}} \mathbb{E}\left[e^{\left(-r_f-\frac{\sigma^2}{2}\right)\tau+\sigma B_\tau}\right]\\ &\quad - K\inf_{\tau\in\mathcal{T}_{0,T-t}} \mathbb{E}\left[e^{-r_d\tau}\right]\\ &\leq sC - Ke^{-r_d(T-t)}\end{aligned}$$

with $C := \sup_{\tau\in\mathcal{T}_{0,T-t}} \mathbb{E}[e^{(-r_f-\frac{\sigma^2}{2})\tau+\sigma B_\tau}] \leq 1$. Therefore $s(e^{-r_f(T-t)} - 1) - Ke^{-r_d(T-t)} \leq f(s) \leq (C-1)s - Ke^{-r_d(T-t)}$ and in turn $\lim_{s\downarrow 0} f(s) = -Ke^{-r_d(T-t)}$.

Meanwhile, $s - K \leq V(t,s) \leq V_\infty(s)$ where $V_\infty(s)$ is the fair price of the contract with infinite maturity. Hence

$$-K \leq f(s) \leq V_\infty(s) - s = -K$$

for $s \geq b_2$ where b_2 is defined in part (5) of Theorem 1 leading to $f(s) = -K$ for $s \geq b_2$.

(b) Note that

$$\begin{aligned} V(t,s) &= \sup_{\tau\in\mathcal{T}_{0,T-t}} \mathbb{E}\left[e^{-r_d\tau}(S_\tau - K)\right] \\ &= \sup_{\tau\in\mathcal{T}_{0,T-t}} \mathbb{E}\left[\left(se^{\left(-r_f-\frac{\sigma^2}{2}\right)\tau+\sigma B_\tau} - Ke^{-r_d\tau}\right)\right] \\ &\le s \sup_{\tau\in\mathcal{T}_{0,T-t}} \mathbb{E}\left[e^{\left(-r_f-\frac{\sigma^2}{2}\right)\tau+\sigma B_\tau}\right] \\ &\quad - K \inf_{\tau\in\mathcal{T}_{0,T-t}} \mathbb{E}\left[e^{-r_d\tau}\right] \\ &\le sC - K \end{aligned}$$

as $r_d < 0$, and we have defined $C := \sup_{\tau\in\mathcal{T}_{0,T-t}} \mathbb{E}[e^{(-r_f-\frac{\sigma^2}{2})\tau+\sigma B_\tau}] \le e^{-r_fT} < \infty$ under $r_f < 0$. Thus $-K \le f(s) \le (C-1)s - K$ and hence $\lim_{s\downarrow 0} f(s) = -K$.

For any deterministic $\tau = n$ where $n \in (0, T-t]$ is arbitrary, we have

$$\begin{aligned} V(t,s) &\ge e^{-r_dn}\left(s\mathbb{E}\left[\exp\left(\left(r_d - r_f - \frac{\sigma^2}{2}\right)n\right.\right.\right. \\ &\quad \left.\left.\left. +\ \sigma B_n\right)\right] - K\right) = se^{-r_fn} - Ke^{-r_dn} \end{aligned}$$

and hence $f(s) \ge (e^{-r_fn} - 1)s - Ke^{-r_dn} \to +\infty$ as $s \uparrow +\infty$ because $r_f < 0$. ∎

3.2.2. Theoretical characterisation of the optimal exercise strategy. From standard theories of optimal stopping of Markovian process (see, e.g., Krylov 2008 or Shiryaev 2007), the optimal stopping time associated with value function (4) is given by

$$\tau^* = \tau^*(t,s) = \inf\{u \ge t : V(u, S_u^{t,s}) = g(S_u^{t,s})\}.$$

Moreover, $(e^{-r_du}V(u, S_u^{t,s}))_{u\ge t}$ is a supermartingale and $(e^{-r_d(u\wedge\tau^*)}V(u\wedge\tau^*, S_{u\wedge\tau^*}^{t,s}))_{u\ge t}$ is a martingale.

Define the exercise (stopping) set as

$$\mathcal{D} := \{(t,s) \in [0,T] \times (0,\infty) : V(t,s) = g(s)\} \tag{10}$$

and the continuation set as

$$\mathcal{C} := \{(t,s) \in [0,T] \times (0,\infty) : V(t,s) > g(s)\}. \tag{11}$$

Then τ^* can be rewritten as $\tau^* = \inf\{u \ge t : (u, S_u^{t,s}) \in \mathcal{D}\}$. We now show that the continuation and exercise set can be separated by a single boundary function.

Proposition 3 (1) *If $r_d > 0$ and $r_f > 0$, there exists a strictly positive function $b : [0,T) \to (0,\infty)$ such that*

$$\begin{aligned} \mathcal{C} &:= \{(t,s) \in [0,T) \times (0,\infty) : s < b(t)\}, \\ \mathcal{D} &:= \{(t,s) \in [0,T) \times (0,\infty) : s \ge b(t)\} \\ &\quad \times \cup\{\{T\} \times (0,\infty)\}. \end{aligned}$$

(2) *If $r_d < 0$ and $r_f < 0$, there exists a non-negative function $b : [0,T) \to [0,\infty)$ such that*

$$\begin{aligned} \mathcal{C} &:= \{(t,s) \in [0,T) \times (0,\infty) : s > b(t)\}, \\ \mathcal{D} &:= \{(t,s) \in [0,T) \times (0,\infty) : s \le b(t)\} \\ &\quad \times \cup\{\{T\} \times (0,\infty)\}. \end{aligned}$$

Proof Using the definition of $f(s) := f(t,s) := V(t,s) - s$ in lemma 2 (where we suppress the argument t for brevity), the sets in (11) and (10) can be written as

$$\begin{aligned} \mathcal{C} &:= \{(t,s) \in [0,T] \times (0,\infty) : f(t,s) > -K\}, \\ \mathcal{D} &:= \{(t,s) \in [0,T] \times (0,\infty) : f(t,s) = -K\}. \end{aligned}$$

But for each fixed $t \in [0,T)$, $s \to f(s)$ is continuous convex function and $f(s) \ge -K$ for all $s \in (0,\infty)$. When $r_d > 0$ and $r_f > 0$, $f(0) := \lim_{s\downarrow 0} f(s) = -Ke^{-r_d(T-t)} > -K$ and $f(s) = -K$ for large s using part 4(a) of lemma 2. Hence there must exist a unique $b = b(t) \in (0,\infty)$ such that $f(s) > -K$ for $s < b(t)$ and $f(s) = -K$ for $s \ge b(t)$, i.e. $f(s) > -K \iff s < b(t)$ and $f(s) = -K \iff s \ge b(t)$. Similar arguments hold in the case of $r_d < 0$ and $r_f < 0$ due to part 4(b) of lemma 2 but we can only conclude that $b(t) \ge 0$ because we may have $f(s) > -K$ for all $s \in (0,\infty)$. ∎

Corollary 1 *The continuation set $\mathcal{C}$ is non-empty.*

Proof It follows immediately from the definition of $\mathcal{C}$ and the fact that $b(t) < \infty$. ∎

Under the case of positive rates, $b(t) \in (0,\infty)$ for all t. It means that at any time there always exists a strictly positive and finite critical price threshold at or above which early exercise is optimal. However, when rates are negative we can then only conclude $b(t)$ to be finite and non-negative. If $b(t) = 0$ for some fixed t, the set $\{s \in (0,\infty) : s \le b(t)\}$ becomes empty such that one should not exercise the Open Forward at time t for any price level.†

Generally speaking, when the rates are positive (negative), it is optimal to early exercise the contract when the spot price is sufficiently high (low). The financial intuitions are largely the same as in section 3.1: positive (negative) domestic interest rate encourages early realisation of gain (loss) which is achieved by early exercising the contract when the spot is sufficiently high (low). There exists a single time-dependent threshold $b(t)$ separating the waiting and continuation region. Unlike an American call option, disconnected waiting region (as in Battauz *et al.* 2015, 2022a) can never arise even when interest rates are negative because of the lack of optionality within an FX Open Forward as discussed in section 3.1.

† In the case of negative rates, if the model parameters are described by case (6)(b) of theorem 1, then $V_\infty(s) = s - K$ for $s \le b_1$ and in turn we can deduce $V(t,s) = s - K$ for all $s \le b_1$ and $t \in [0,T)$, for some given constant $b_1 > 0$. The second part of proposition 3 can then be sharpened to conclude that the boundary function $b(t)$ is indeed strictly positive. In such case, it is always optimal to early exercise the FX Open Forward when spot price is sufficiently low. Moreover, even if the model parameters do not satisfy the conditions in case (6)(b) of theorem 1, we can still conclude that $b(t)$ becomes strictly positive when t approaches T. See part 3 of proposition 4.

The proposition below lists the key behaviours of the early exercise boundary $b(t)$ with respect to the current time t.

PROPOSITION 4 *The early exercise boundary function $b : [0,T) \to [0,\infty)$ has the following properties:*

(1) *b is decreasing (resp. increasing) if $r_d > 0$ and $r_f > 0$ (resp. $r_d < 0$ and $r_f < 0$).*
(2) *$b(t) \geq \frac{r_d}{r_f}K$ (resp. $b(t) \leq \frac{r_d}{r_f}K$) if $r_d > 0$ and $r_f > 0$ (resp. $r_d < 0$ and $r_f < 0$) for all $t \in [0,T)$.*
(3) *b is continuous and $b(T) := \lim_{t\to T} b(t) = \frac{r_d}{r_f}K$.*

As the maturity is approaching, the premium brought by the possibility to early exercise decreases and hence the contract holder is more likely to early exercise. The continuation region shrinks and consequently the exercise threshold decreases when rates are positive or increases when rates are negative. When time to expiry is very small, the exercise boundary is approximately given by r_dK/r_f. The associated intuition is similar to the one given in Andersen and Lake (2021). Suppose the current spot is $S_{T-\Delta} = s$ where the contract will expire after Δ unit of time. The immediate payoff is $s - K$ if the contract is exercised now. Otherwise, the discounted risk-neutral expected payoff at the expiry date is $se^{-r_f\Delta} - Ke^{-r_d\Delta} \approx s(1 - r_f\Delta) - K(1 - r_d\Delta)$. Early exercise is therefore beneficial if and only if $s - K > s(1 - r_f\Delta) - K(1 - r_d\Delta)$, which is equivalent to $s > r_dK/r_f$ if $r_d > 0$ and $r_f > 0$, or $s < r_dK/r_f$ if $r_d < 0$ and $r_f < 0$.

For American call option under negative interest rates where double-continuation region arises, Battauz *et al.* (2015) characterise the optimal exercise behaviours by two different boundary functions, $\ell(t)$ and $u(t)$, such that the stopping set is in form of $\mathcal{D} = \{(t,s) \in [0,T) \times (0,\infty) : \ell(t) \leq s \leq u(t)\} \cup \{\{T\} \times (0,\infty)\}$. Theorem 3.3 of Battauz *et al.* (2015) shows that the upper and lower thresholds have different monotonicity and near-maturity properties. On comparing our proposition 4 against their results, one can see that the early exercise boundary $b(t)$ of FX Open Forward shares the same qualitative behaviours as that of the upper boundary $u(t)$ of an American call under negative interest rates. It is not entirely not surprising: as discussed in section 3.1, the upper continuation region (characterized by $u(t)$) of an American call option is driven by the incentive to defer gain under negative interest rates while the lower continuation region (characterized by $\ell(t)$) is driven by the optionality of the contract. Only the former consideration is relevant to FX Open Forward and hence we expect the optimal exercise region of the FX Open Forward behaves similarly as that of the upper continuation region of an American call.

We now state the asymptotic expression of b when $t \approx T$ in the following proposition.

PROPOSITION 5 *Let $y^* \approx -0.6387$ be a constant such that $-y^*$ is the unique solution to the equation $G(x) = 0$ on $x \in (0, 1/\sqrt{2}]$ with*

$$G(x) := \mathbb{E}\left[\int_0^1 (x - \sqrt{v}Z)1_{(Z \geq x(1-\sqrt{1-v})/\sqrt{v})}\, dv\right]$$

where Z is a standard normal random variable. Then

$$\lim_{t\uparrow T} \frac{b(T) - b(t)}{\sigma b(T)\sqrt{T-t}} = \begin{cases} y^*, & \text{if } r_d > 0,\ r_f > 0; \\ -y^*, & \text{if } r_d < 0,\ r_f < 0. \end{cases} \tag{12}$$

Proof The result follows from an application of theorem 1 in Lamberton and Villeneuve (2003) to the function

$$f(t,x) := e^{-rt}(e^{(r_d - r_f - \sigma^2/2)t + \sigma x} - K) \tag{13}$$

and a straightforward modification of theorem 2 in Lamberton and Villeneuve (2003) on noting that

$$\mathcal{L}f(t,x) := \frac{\partial f}{\partial t}(t,x) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,x) \tag{14}$$

is $C^{1\times 2}$ everywhere. ■

As a technical remark, the near-maturity analysis of the early exercise boundary with American forward is indeed much easier than that with American option. The linearity of forward payoff guarantees the smoothness of the operator defined in (14), and this in turn leads to a parabolic behaviour of the exercise boundary. When an American call option is involved, we have to replace the function f in (13) by

$$f(t,x) := e^{-r_dt}(e^{(r_d - r_f - \sigma^2/2)t + \sigma x} - K)^+. \tag{15}$$

Similar parabolic behaviour in form of (12) holds if the f in (15) is smooth near $(t = 0, x = \frac{1}{\sigma}\ln \tilde{b}(T))$ with $\tilde{b}(T)$ being the limiting value of the (an) early exercise boundary for an American call option, or equivalently $\tilde{b}(T) \neq K$. But there are many settings under which an American call option leads to an early exercise boundary with $\tilde{b}(T) = K$. For example, in the positive rates case, $\tilde{b}(T) = \max(K, \frac{r_d}{r_f}K)$ and as such $\tilde{b}(T) = K$ when $r_d \leq r_f$ (see Evans *et al.* 2002).† In the negative rates case, a double-continuation region arises and the lower exercise boundary tends to K (theorem 3.3 of Battauz *et al.* 2015). In those scenarios, the corresponding early exercise boundaries will have a different asymptotic expression in a parabolic–logarithmic form. Meanwhile, for American forward, the expression in (14) is robust to the model parameters and holds even when $r_d = r_f$ such that $b(T) = K$.

4. Numerical illustrations

This section is dedicated to the numerical experiments illustrating the technical results developed previously. Some comparisons between standard 'Fixed Forwards' versus Open Forwards, as well as American call/put options versus Open Forwards, are also studied. All numerical results in this section are obtained by finite difference methods.

† A contemporary analysis of the limiting value of the early exercise threshold can also be found in Battauz *et al.* (2022b).

4.1. *Forwards and Open Forwards*

A classical 'Fixed Forward' is a derivative contract between two parties to buy or sell an asset, at a specified future time T at a pre-agreed delivery price K. The party agreeing to buy the underlying asset assumes a long position (which we will refer to as a 'purchase forward'), and the party agreeing to sell the asset assumes a short position (which we will refer to as a 'sale forward'). The payoff at maturity is given by

$$f_T = \begin{cases} S_T - K, & \text{for a long position;} \\ K - S_T, & \text{for a short position,} \end{cases}$$

and the fair or present values (pv) of such contracts are

$$pv = \begin{cases} e^{-r_f T} S_0 - e^{-r_d T} K, & \text{for a long position;} \\ e^{-r_d T} K - e^{-r_f T} S_0, & \text{for a short position.} \end{cases}$$

The pv's of Fixed Forwards are model independent and in particular do not depend on the model volatility σ. Meanwhile, we have shown in propositions 1 and 2 that an Open Forward is equivalent to an American derivative with exercise payoff at any time t given by $S_t - K$ for a currency purchase or $K - S_t$ for a currency sale. Consequently, the pv of an Open Forward has an early exercise premium which depends on the volatility level σ. Figures 1(a,b) show the pv differences between forward and Open Forward for currency purchase and sale, respectively, as a function of volatility σ. Model and market parameters are $S_0 = 1, K = 1, T = 2, r_d = 5\%, r_f = 5\%$. We observe the early exercise premium associated with the Open Forward increases with the volatility level. At 20% volatility, the pv differences are about 56 bps. And at 35%, we get almost 1% premium, which is financially significant.

4.2. *American options and Open Forwards*

4.2.1. Positive interest rates. We now compare the early exercise boundaries between American call/put options and Open Forwards when interest rates are positive. Figure 2 is for American call and purchase Open Forward, while figure 3 is for American put and sale Open Forward. Model and market parameters are $S_0 = 1, K = 1, T = 2, r_d = 5\%, r_f = 5\%, \sigma = 30\%$. We observe that the boundaries are significantly different. The optimal exercise boundary for American call is higher than that of a purchase Open Forward. It means the holder of a purchase Open Forward will exercise earlier (i.e. early exercise over a strictly larger set of spot prices) in comparison to the holder of an American call option (figure 2 a). This can also be illustrated via figure 2(b), which shows the intrinsic and fair values of the American call and purchase Open Forward. Obviously, the payoff of American call dominates that of purchase Open Forward, and this ranking holds for their fair values as well. Since the early exercise boundary can be inferred from the critical price level at which the fair value function touches the payoff function, figure 2(b) shows that the early exercise boundary for purchase Open Forward is lower than that of American call option.

Similarly, the exercise boundary for American put is lower than that for a sale Open Forward (see figure 3 a). Thus the holder of a sale Open Forward will exercise earlier relative to the holder of American put option. Figure 3(b) provides a similar illustration of the early exercise boundaries via the fair versus the intrinsic values.

4.2.2. Negative interest rates. In a negative rate environment, it is known that in a Black–Scholes model there could exist double waiting region: exercise is optimally postponed when the option is deeply in-the-money or out-of-money (see, e.g. Battauz *et al.* 2015, Andersen and Lake 2021). For American put option, this phenomenon occurs when the following model parameter conditions are satisfied:

$$r_d < 0, \quad \mu - \frac{\sigma^2}{2} > 0. \tag{16}$$

and

$$\left(\mu - \frac{\sigma^2}{2}\right)^2 + 2r_d\sigma^2 > 0, \tag{17}$$

where $\mu = r_d - r_f$.

Figure 4(a) shows the present value of an American put option and its payoff where a double waiting region is clearly observed using the following model parameters from Battauz *et al.* (2015): $S_0 = 1, K = 1.2, T = 1, r_d = -4\%, r_f = -12\%, \sigma = 20\%$ (i.e. there are two disconnected sets of spot prices where the option fair value strictly dominates its payoff). In comparison, for Open Forward contract, disconnected waiting region can never arise even when interest rates are negative (see propositions 3 and A.5). An illustration is provided in figure 4(b) which shows the present value of a sale Open Forward option and its payoff function against spot price under the same model parameters. Here, with negative rates, exercise of the sale Open Forward is optimally postponed if the spot level is sufficiently low.

4.3. *Open Forwards exercise boundaries*

The signs of interest rates play an important role in the determination of the exercise boundary shapes. Figure 5 shows the exercise boundary as a function of time for purchase Open Forward. In figure 5(a), interest rates are positive given by $r_d = r_f = 5\%$. In figure 5(b), negative rates $r_d = r_f = -2\%$ are considered. Other model and market parameters are $S_0 = 1, K = 1, T = 2, \sigma = 30\%$. We observe a decreasing $b(t)$ exercise boundary under positive interest rates, $b(t) \geq 1 = \frac{r_d}{r_f}K$ and $b(t)$ is converging to 1 as t approaches the maturity of $T = 2$. When interest rates are negative, $b(t)$ is increasing, dominated by the $\frac{r_d}{r_f}K = 1$ and converging towards this bound as t tends to T. These results are in line with proposition 4 for the case of $r_d > 0, r_f > 0$. Similarly, figure 6 shows the exercise boundary as a function of time for sale Open Forward. In figure 6(a), $b(t)$ is increasing under positive interest rates $r_d = r_f = 5\%$, dominated by $\frac{r_d}{r_f}K = 1$ and converging to this bound as $t \to T = 2$. In figure 6(b), when interest rates are negatives, $b(t)$ is decreasing, dominating $\frac{r_d}{r_f}K = 1$ and approaching this bound when time to expiry shortens. These results are in line with proposition A.6 given in appendix 2.

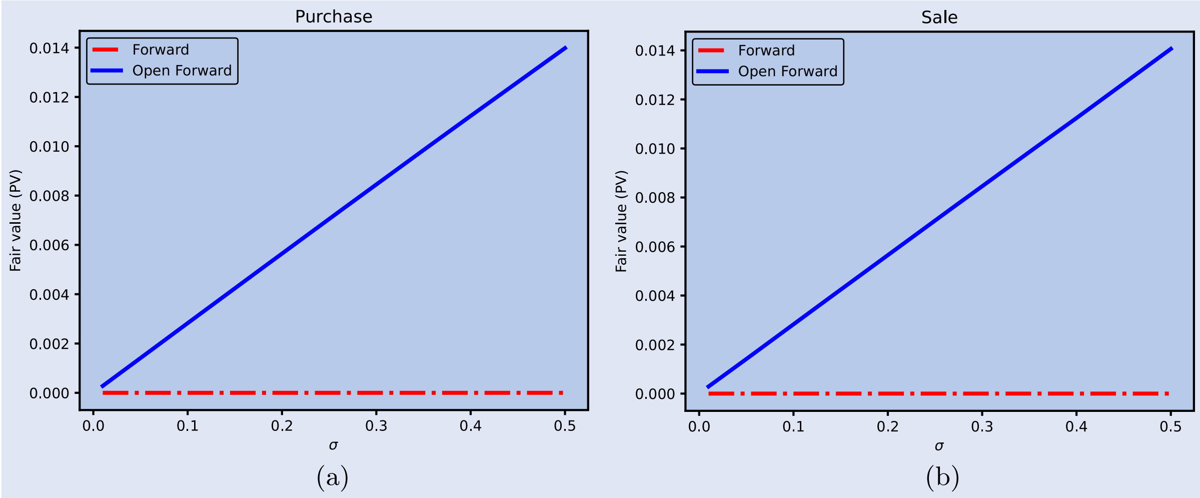

Figure 1. PV comparison between forwards and Open Forwards as functions of volatility σ. Model and market parameters: $S_0 = 1, K = 1, T = 2, r_d = 5\%, r_f = 5\%$. (a) Currency purchase forward and Open Forward and (b) Currency sale forward and Open Forward.

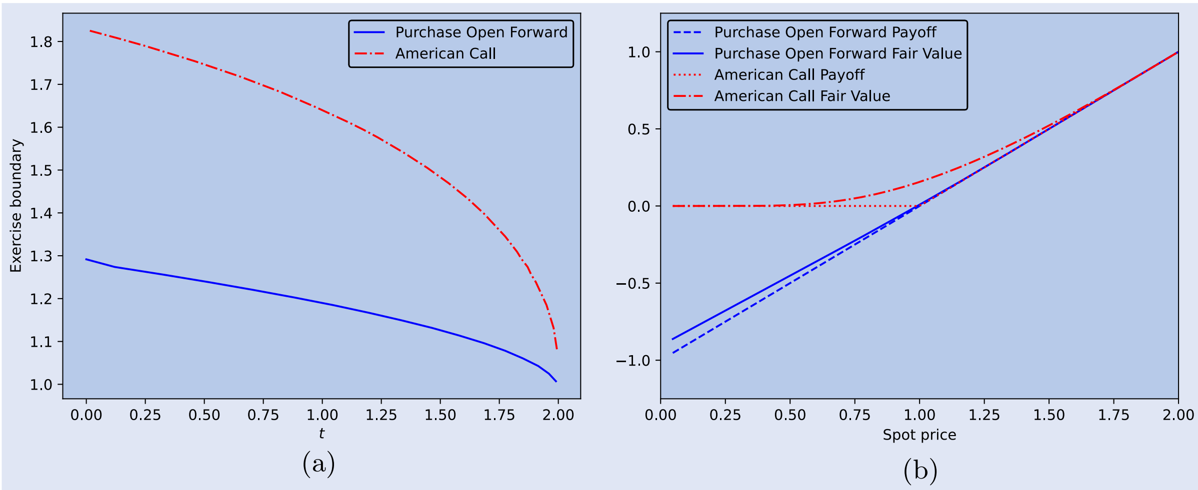

Figure 2. Comparison between American call and purchase Open Forward. Model and market parameters: $S_0 = 1, K = 1, T = 2, r_d = 5\%, r_f = 5\%, \sigma = 30\%$. (a) Exercise boundaries for American call and purchase Open Forward, above which early exercise is optimal and (b) Fair values of American call and purchase Open Forward as functions of spot price.

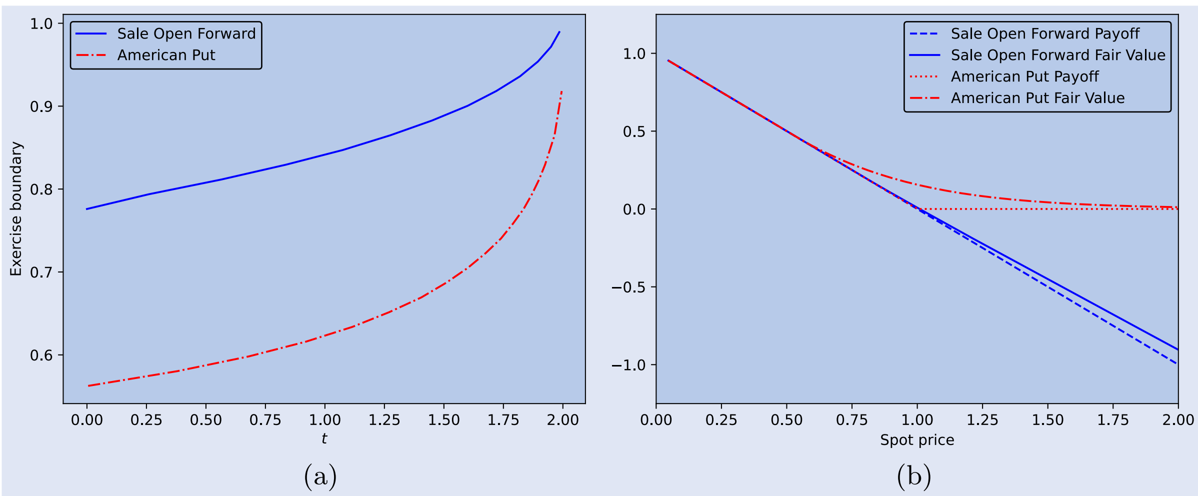

Figure 3. Comparison between American put and sale Open Forward. Model and market parameters: $S_0 = 1, K = 1, T = 2, r_d = 5\%, r_f = 5\%, \sigma = 30\%$. (a) Exercise boundaries for American put and sale Open Forward, below which early exercise is optimal and (b) Fair values of American put and sale Open Forward as functions of spot price.

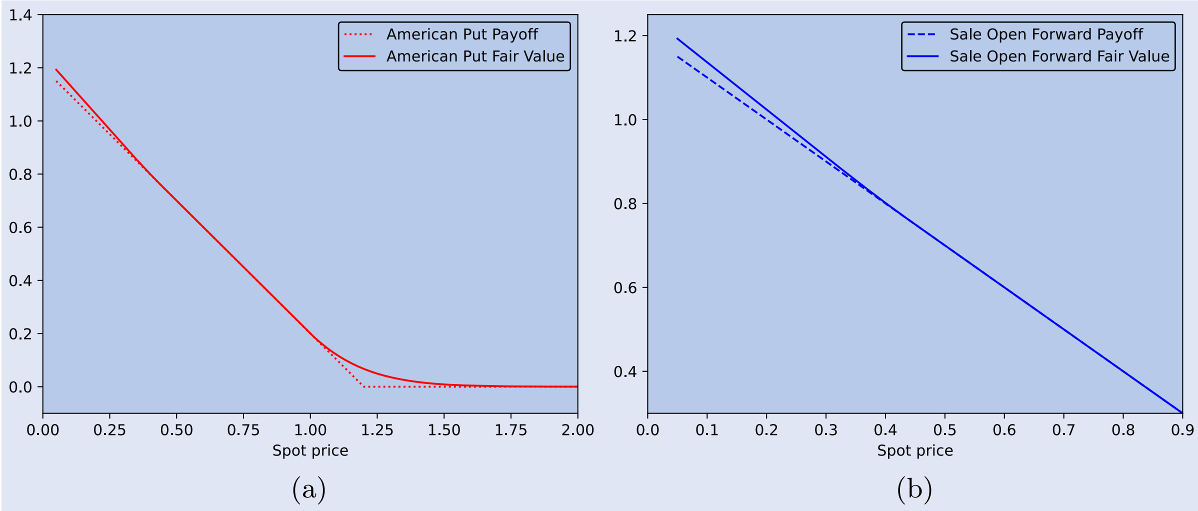

Figure 4. Comparison between American put and Sale Open Forward with negative rates. Model and market parameters: $S_0 = 1, K = 1.2, T = 1, r_d = -4\%, r_f = -12\%, \sigma = 20\%$. (a) American put payoff and fair value as functions of spot and (b) Sale Open Forward payoff and fair value as functions of spot.

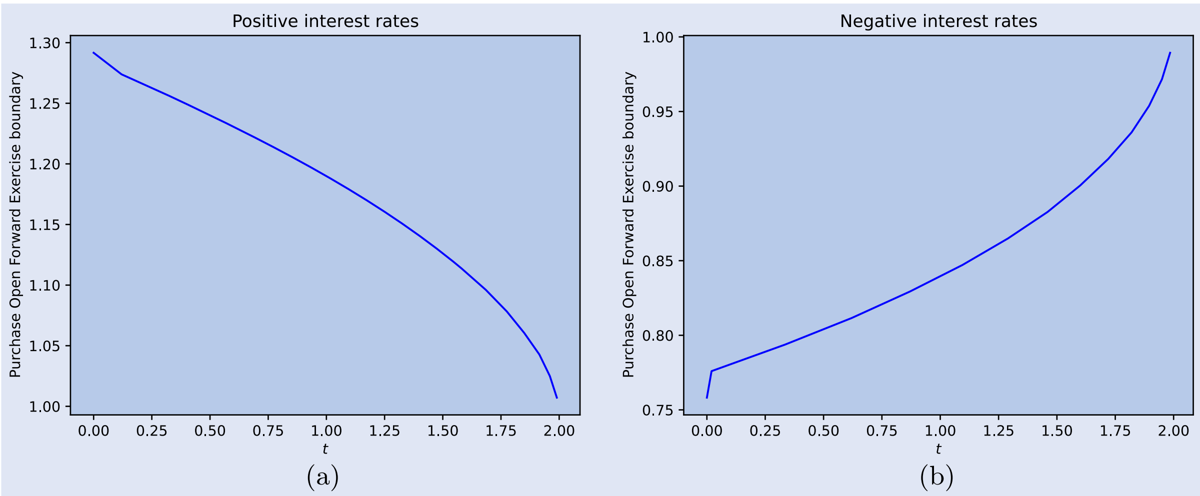

Figure 5. Exercise boundaries comparison for purchase Open Forward with positive rates ($r_d = r_f = 5\%$) and negative rates ($r_d = r_f = -2\%$). Others model and market parameters: $S_0 = 1, K = 1, T = 2, \sigma = 30\%$. (a) Exercise boundary of purchase Open Forward under positive rates and (b) Exercise boundary of negative Open Forward under positive rates.

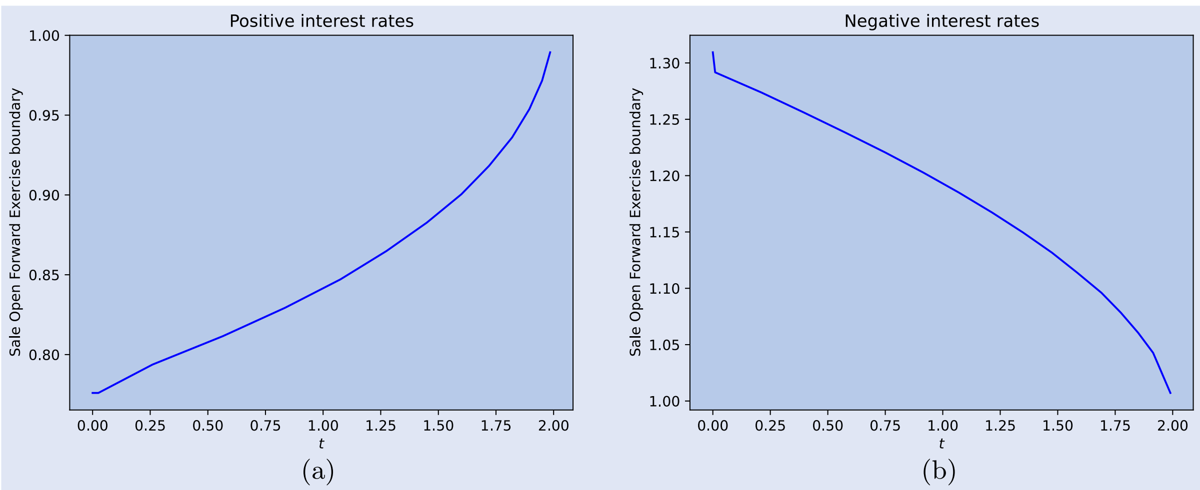

Figure 6. Exercise boundaries comparison for purchase Open Forward purchase with positive rates ($r_d = r_f = 5\%$) and negative rates ($r_d = r_f = -2\%$). Others model and market parameters: $S_0 = 1, K = 1, T = 2, \sigma = 30\%$. (a) Exercise boundary of sale Open Forward under positive rates and (b) Exercise boundary of sale Open Forward under negative rates.

5. Summary and concluding remarks

In this paper, we first show that the vested flexibility of converting any amount of domestic currency into foreign currency at any time within the contract life of an FX Open Forward is indeed redundant. The optimal strategy has a 'all-or-nothing'/'bang-bang' form where the contract holder always settles the full notional at once. This result simplifies significantly the problem and reduces the analysis to that of a standard American derivative. Then we focus on a Black–Scholes model to theoretically analyse the pricing and the optimal exercise strategy. An important theoretical insight is that the form of the exercise strategy crucially depends on the signs of the interest rates, and the optimal exercise strategy of FX Open Forward is significantly different from that of American call/put option when interest rates are negative because the lack of optionality precludes the continuation region over the out-of-money regime of spot price.

We outline a few possible interesting directions of future research. Inspired by the observation that the form of the early exercise strategy is heavily influenced by the interest rates' signs, one natural extension is to incorporate stochastic interest rates and study how the state of interest rates interacts with the early exercise decision. Such considerations are particularly important for models that allow for negative interest rates, which have gained popularity as negative interest rate regime has become an important stress testing scenario as recommended by the Basel Committee.

Another important modelling element to be considered is time-varying/stochastic volatility to better capture volatility skew/smile. Although we have observed that the pricing of FX Open Forward has rather muted vega under the Black–Scholes model, it will be interesting to further quantify how the early exercise strategy and the early exercise premium are affected by the volatility dynamics in some stochastic and/or local volatility models (such as Heston, Schobel and Zhu, Dupire and so on).

Finally, in an incomplete market featuring unhedgable risk factors or transaction costs, one may consider alternative pricing rules via say utility indifference pricing or minimisation of hedging error. When valuation is no longer performed using a single pricing measure, we expect that partial early exercise may be financially relevant again. Such setup is likely to offer an interesting and non-degenerate theoretical problem of stochastic control, while offering alternative predictions over the empirical settlement patterns of the contract holders.

Acknowledgments

We thank the two anonymous referees and the associate editor for their constructive comments which help improve the quality of this paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

No funding was received.

ORCID

Alex S.L. Tse http://orcid.org/0000-0001-6836-8096

References

Andersen, L. and Lake, M., Fast American option pricing: The double-boundary case. *Wilmott*, 2021, **2021**(116), 30–41.
Bardou, O., Bouthemy, S. and Pagès, G., When are swing options bang-band. *Int. J. Theor. Appl. Fin.*, 2010, **13**(6), 867–899.
Barrera-Esteve, C., Bergeret, F., Dossal, C., Gobet, E., Meziou, A., Munos, R. and Reboul-Salze, D., Numerical methods for the pricing of swing options: a stochastic control approach. *Methodol. Comput. Appl. Probab.*, 2006, **8**, 517–540.
Battauz, A., De Donno, M.D. and Sbuelz, A., Real options with a double continuation region. *Quant. Finance*, 2012, **12**(3), 465–475.
Battauz, A., De Donno, M. and Sbuelz, A., Real options and American derivatives: The double continuation region. *Manage. Sci.*, 2015, **61**(5), 1094–1107.
Battauz, A., De Donno, M. and Sbuelz, A., On the exercise of American quanto options. *North Am. J. Econ. Fin.*, 2022a, **62**, 101738.
Battauz, A., De Donno, M., Gajda, J. and Sbuelz, A., Optimal exercise of American put options near maturity: A new economic perspective. *Rev. Derivatives Res.*, 2022b, **25**, 23–46.
Daluiso, R., Nastasi, E., Pallavicini, A. and Sartorelli, G., Pricing commodity swing options. arXiv:2001.08906, 2020.
Dar, Z., Gupta, A. and Kar, K., Flexible forward contracts for renewable energy generators in deregulated electricity markets, 2022. Available at SSRN 4171016.
Detemple, J., American options: symmetry properties. In *Handbooks in Mathematical Finance: Topics in Option Pricing, Interest Rates and Risk Management*, edited by E. Jouini, J. Cvitanić and M. Musiela, pp. 67–104, 2001 (Cambridge University Press).
Evans, J.D., Kuske, R. and Keller, J.B., American options on assets with dividends near expiry. *Math. Fin.*, 2002, **12**(3), 219–237.
Giribone, P.G. and Ligato, S., Flexible-forward pricing through Leisen–Reimer trees: Implementation and performance comparison with traditional Markov chains. *Int. J. Fin. Eng.*, 2016, **3**(2), 1650010.
Henderson, V. and Hobson, D., Utility indifference pricing-an overview. In *Indifference Pricing: Theory and Applications*, edited by R. Carmona (Princeton University Press: Princeton, NJ).
Hölbl, M. and Lovric, G., Open fx forwards: Introducing a standard mark-to-model approach, 2019. Available at SSRN 3333174.
Jacka, S.D., Optimal stopping and the American put. *Math. Fin.*, 1991, **1**(2), 1–14.
Krylov, N.V., *Controlled Diffusion Processes*, 2008 (Springer Science & Business Media: Berlin).
Kwok, Y.K. and Lau, K.W., American currency forward, 2003. Available at https://repository.hkust.edu.hk/ir/bitstream/1783.1-299/2/American.pdf.
Lamberton, D. and Villeneuve, S., Critical price near maturity for an American option on a dividend-paying stock. *Ann. Appl. Probab.*, 2003, **13**(2), 800–815.
McKean, H.P., A free boundary problem for the heat equation arising from a problem in mathematical economics. *Indust. Manag. Rev.*, 1965, **6**, 32–39.
Merton, R.C., Theory of rational option pricing. *Bell J. Econ. Manage. Sci.*, 1973, **4**(1), 141–183.
Myneni, R., The pricing of the american option. *Ann. Appl. Probab.*, 1992, **2**(1), 1–23.

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

Pham, H., Optimal stopping, free boundary, and American option in a jump-diffusion model. *Appl. Math. Optim.*, 1997, **35**, 145–164.

Revuz, D. and Yor, M., *Continuous Martingales and Brownian Motion*, 2013 (Springer Science & Business Media: Berlin).

Shiryaev, A.N., *Optimal Stopping Rules*, 2007 (Springer Science & Business Media: New York).

Shreve, S.E., *Stochastic Calculus for Finance II: Continuous-time Models*, 2004 (Springer: Berlin).

Young, W., A note on monotone functions. *Q. J. Pure Appl. Math.*, 1910, **41**, 79–87.

Appendices

Appendix 1. Proofs

Before proving theorem 1, it is useful to first derive the moment generating function of T_α the first passage time of a drifting Brownian motion reaching level α. Standard textbooks typically derive the expression of $\mathbb{E}[e^{-r_d T_\alpha}]$ when $r_d > 0$ (e.g. Shreve 2004, theorem 8.3.2) but we need to extend the result to cover negative interest rate.

LEMMA A.1 *Let $X = (X_t)_{t\geq 0}$ be a drifting Brownian motion in form of $X_t := \nu t + \sigma B_t$ with $\nu, \sigma > 0$. Define the first passage time $T_\alpha := \inf\{t \geq 0 : X_t = \alpha\}$ for $\alpha > 0$. Then*

$$\mathbb{E}(e^{zT_\alpha}) = \begin{cases} \exp\left[\dfrac{\alpha\nu}{\sigma^2}\left(1-\sqrt{1-\dfrac{2\sigma^2}{\nu^2}z}\right)\right], & z \leq \dfrac{\nu^2}{2\sigma^2}; \\ +\infty, & z > \dfrac{\nu^2}{2\sigma^2}. \end{cases}$$

Proof It is known that $T_\alpha \sim \mathrm{IG}(\frac{\alpha}{\nu}, \frac{\alpha^2}{\sigma^2})$, where $\mathrm{IG}(\theta, \lambda)$ denotes an inverse Gaussian distribution with location parameter θ and shape parameter λ. The result follows from evaluating the moment generating function of such random variable. ■

Proof of theorem 1. We first argue that for any $\tau \in \mathcal{T}_{0,\infty}$, $\mathbb{E}[e^{\sigma B_\tau - \frac{\sigma^2}{2}\tau}] \leq 1$. Define $\tau_n := \min(\tau, n)$ for $n \in \mathbb{N}$ and $\tau \in \mathcal{T}_{0,\infty}$. Then using optimal sampling theorem, $\mathbb{E}[e^{\sigma B_{\tau_n} - \frac{\sigma^2}{2}\tau_n}] = 1$ for each n as τ_n is bounded. The claim follows from an application of Fatou's lemma to the sequence of random variables $(e^{\sigma B_{\tau_n} - \frac{\sigma^2}{2}\tau_n})_{n\in\mathbb{N}}$ and the fact that $\tau_n \uparrow \tau$ almost surely.

(1) With $r_d \leq 0 \leq r_f$,

$$e^{-r_d\tau}(S_\tau - K) = se^{-r_f\tau}e^{\sigma B_\tau - \frac{\sigma^2}{2}\tau} - Ke^{-r_d\tau} \leq se^{\sigma B_\tau - \frac{\sigma^2}{2}\tau} - K$$

and thus

$$V_\infty(s) \leq s\sup_{\tau\geq 0}\mathbb{E}\left[e^{\sigma B_\tau - \frac{\sigma^2}{2}\tau}\right] - K \leq s - K.$$

But we also have $V_\infty(s) \geq s - K$ as $\tau = 0$ is an admissible strategy. Hence $V_\infty(s) = s - K$ and $\tau^* = 0$ is an optimal early exercise rule.

(2) Since $\tau = T$ is an admissible strategy for any arbitrary $T > 0$,

$$\begin{aligned} V_\infty(s) &\geq \mathbb{E}[e^{-r_d T}(S_T - K)] \\ &= se^{-r_f T}\mathbb{E}\left(e^{\sigma B_T - \frac{\sigma^2}{2}T}\right) - Ke^{-r_d T} \\ &= se^{-r_f T} - Ke^{-r_d T} \to +\infty \end{aligned}$$

as $T \uparrow +\infty$ if $r_f < 0 \leq r_d$.

(3) When $r_f = 0 < r_d$, we have $e^{-r_d\tau}(S_\tau - K) \leq se^{\sigma B_\tau - \frac{\sigma^2}{2}\tau}$ and hence

$$V_\infty(s) \leq s\sup_{\tau\geq 0}\mathbb{E}\left[e^{\sigma B_\tau - \frac{\sigma^2}{2}\tau}\right] \leq s.$$

On the other hand, there exists a sequence of stopping times in form of $\tau_n = n$ under which $\mathbb{E}[e^{-r_d\tau_n}(S_{\tau_n} - K)] = s - Ke^{-r_d n} \to s$ as $n \to \infty$. Hence $V_\infty(s) = s$.

(4) If $r_f = 0 = r_d$, the value function simplifies to

$$V_\infty(s) = s\sup_{\tau\geq 0}\mathbb{E}\left[e^{\sigma B_\tau - \frac{\sigma^2}{2}\tau}\right] - K \leq s - K.$$

To show that $V_\infty(s) = s - K$, one just needs to show that there exists τ such that $\mathbb{E}[e^{\sigma B_\tau - \frac{\sigma^2}{2}\tau}] = 1$. This can for example be achieved by any constant stopping time.

(5) Denote by V_∞^C the function constructed in (6). It is straightforward to show that V_∞^C is a convex function satisfying

$$\begin{cases} \dfrac{\sigma^2 s^2}{2}\dfrac{\mathrm{d}^2 V_\infty^C}{\mathrm{d}s^2} + (r_d - r_f)s\dfrac{\mathrm{d}V_\infty^C}{\mathrm{d}s} - r_d V_\infty^C = 0 \text{ and} & \\ V_\infty^C(s) \geq s - K, & s < b_2; \\ \dfrac{\sigma^2 s^2}{2}\dfrac{\mathrm{d}^2 V_\infty^C}{\mathrm{d}s^2} + (r_d - r_f)s\dfrac{\mathrm{d}V_\infty^C}{\mathrm{d}s} - r_d V_\infty^C \leq 0 \text{ and} & \\ V_\infty^C(s) = s - K, & s \geq b_2. \end{cases}$$

Now, we can apply generalised Ito's lemma to $M_t := e^{-r_d t}V_\infty^C(S_t)$ which gives

$$\begin{aligned} \mathrm{d}M_t &= e^{-r_d t}\left\{\left[\frac{\sigma^2 S_t^2}{2}\partial_{ss}^2 V_\infty^C + (r_d - r_f)S_t\partial_s V_\infty^C - r_d V^C\right]\mathrm{d}t \right. \\ &\quad \left. + \sigma S_t \partial_s V_\infty^C \,\mathrm{d}B_t\right\} \\ &= e^{-r_d t}\left[(r_d K - r_f S_t)1_{\{S_t\geq b_2\}}\,\mathrm{d}t + \sigma S_t\partial_s V_\infty^C \,\mathrm{d}B_t\right]. \end{aligned}$$

Next, we show that $b_2 \geq r_d K / r_f$ which is equivalent to showing that $\gamma_2/(\gamma_2 - 1) > r_d/r_f$. Since $\gamma_2 > 1$, the claim is obviously true if $r_d \leq r_f$. It remains to consider the case of $r_d > r_f$. Let $\hat{\gamma}_2 := 1/\gamma_2$. As γ_2 is the larger, positive root to the quadratic equation (7), $\hat{\gamma}_2$ is the larger, positive root to the quadratic equation

$$f(\hat{\gamma}) := -r_d\hat{\gamma}^2 + \left(r_d - r_f - \frac{\sigma^2}{2}\right)\hat{\gamma} + \frac{\sigma^2}{2} = 0.$$

Then we have

$$f\left(\frac{r_d - r_f}{r_d}\right) = -\frac{\sigma^2}{2}\frac{r_f}{r_d} < 0$$

such that $1/\gamma_2 = \hat{\gamma}_2 < 1 - r_f/r_d$, and in turn $\gamma_2/(\gamma_2 - 1) > r_d/r_f$.

We therefore conclude that $(r_d K - r_f S_t)1_{\{S_t\geq b_2\}} \leq 0$ and as such M is a local supermartingale, and indeed a supermartingale since $M_t = e^{-r_d t}V_\infty^C(S_t)$ is bounded from below by $-K$. By optimal stopping theorem,

$$\mathbb{E}[e^{-r_d\tau}g(S_\tau)] \leq \mathbb{E}[M_\tau] \leq M_0 = V_\infty^C(s).$$

Taking supremum over τ leads to $V_\infty(s) \leq V_\infty^C(s)$.

To show the reverse inequality $V_\infty \geq V_\infty^C$, it is sufficient to show that the choice of

$$\begin{aligned} \tau^* &= \inf\{t \geq 0 : S_t^{0,s} \geq b_2\} \\ &= \inf\left\{t \geq 0 : \frac{1}{\sigma}\left(r_d - r_f - \frac{\sigma^2}{2}\right) + B_t \geq \frac{1}{\sigma}\ln\frac{b_2}{s}\right\} \end{aligned}$$

results in $\mathbb{E}[e^{-r_d\tau^*}(S_{\tau^*} - K)] = V_\infty^C(s)$. This can be easily verified with the help of lemma A.1.

(6) (a) If $r_f < r_d < 0$, then for any T we have

$$V_\infty(s) \geq \mathbb{E}[e^{-r_dT}(S_T - K)] = e^{-r_dT}(se^{(r_d-r_f)T} - K) \to +\infty$$

as $T \uparrow +\infty$.

In the corner case of $r_d = r_f < 0$, the above conclusion is still true for $s > K$ such that $V_\infty(s) = +\infty$ for all $s > K$. If $s \leq K$, then for some $b > K$ consider an early exercise strategy in form of

$$\tau := \begin{cases} 1 & \text{if } \tau_b > 1; \\ \tau_b + T & \text{if } \tau_b \leq 1, \end{cases}$$

where $T > 0$ is a constant and $\tau_b := \inf\{t \geq 0 : S_t = b\}$. In words, τ is a strategy where the contract holder either exercises at time 1 if S_t does not visit level b over the time interval $t \in [0, 1]$, or T unit of time after S reaches level b if it happens before time 1. Then

$$\begin{aligned} V_\infty(s) &\geq \mathbb{E}[e^{-r_d\tau}(S_\tau - K)] \\ &= \mathbb{E}\left[e^{-r_d(\tau_b+T)}1_{\{\tau_b\leq 1\}}(S_{\tau_b+T} - K)\right] \\ &\quad + \mathbb{E}\left[e^{-r_d}1_{\{\tau_b>1\}}(S_1 - K)\right] \\ &= \mathbb{E}\left[e^{-r_d\tau_b}1_{\{\tau_b\leq 1\}}\right][e^{-r_dT}(be^{(r_d-r_f)T} - K)] \\ &\quad + \mathbb{E}\left[e^{-r_d}1_{\{\tau_b>1\}}(S_1 - K)\right] \\ &\geq e^{-r_dT}(b-K)\mathbb{E}\left[1_{\{\tau_b\leq 1\}}\right] - Ke^{-r_d} \\ &= e^{-r_dT}(b-K)\mathbb{P}(\tau_b \leq 1) - Ke^{-r_d} \to +\infty \end{aligned}$$

as $T \uparrow +\infty$. Here we have used the fact that S visits any arbitrary level within the time interval $[0, 1]$ with strictly positive probability. Specifically, $\mathbb{P}(\tau_b \leq 1) \geq \mathbb{P}(S_1 \geq b) = \Phi(\frac{\ln(s/b)-\sigma^2/2}{\sigma}) > 0$. Hence we can conclude $V_\infty(s) = +\infty$ for all s.

Suppose $r_d < r_f$ and $\sigma > |\sqrt{-2r_d} - \sqrt{-2r_f}|$. As $r_d < r_f$, we have $|\sqrt{-2r_d} - \sqrt{-2r_f}|^2 < 2(r_f - r_d) < |\sqrt{-2r_d} + \sqrt{-2r_f}|^2$. There are two possibilities. If $|\sqrt{-2r_d} - \sqrt{-2r_f}| < \sigma < \sqrt{2(r_f - r_d)} < |\sqrt{-2r_d} + \sqrt{-2r_f}|$, then consider the quadratic function (in σ^2)

$$f(\sigma^2) := \frac{\sigma^4}{4} + (r_d + r_f)\sigma^2 + (r_d - r_f)^2.$$

The two roots of the equation $f(\sigma^2) = 0$ are given by

$$\sigma_\pm^2 = \frac{-(r_d + r_f) \pm \sqrt{4r_dr_f}}{1/2} = (\sqrt{-2r_d} \pm \sqrt{-2r_f})^2.$$

Hence $f(\sigma^2) < 0 \iff \sigma_-^2 < \sigma^2 < \sigma_+^2 \iff |\sqrt{-2r_d} - \sqrt{-2r_f}| < \sigma < |\sqrt{-2r_d} + \sqrt{-2r_f}|$. Thus $|\sqrt{-2r_d} - \sqrt{-2r_f}| < \sigma < |\sqrt{-2r_d} + \sqrt{-2r_f}|$ implies

$$\begin{aligned} &\left(r_d - r_f - \frac{\sigma^2}{2}\right)^2 + 2r_d\sigma^2 \\ &= \frac{\sigma^4}{4} + (r_d + r_f)\sigma^2 + (r_d - r_f)^2 = f(\sigma^2) < 0. \end{aligned}$$

For any $s > K$, pick some $b \in (K, s)$ and consider an early exercise strategy in form of

$$\begin{aligned} \tau_b &:= \inf\{t \geq 0 : S_t \leq b\} \\ &= \inf\left\{t \geq 0 : -\frac{1}{\sigma}\left(r_d - r_f - \frac{\sigma^2}{2}\right)t \right. \\ &\qquad \left. - B_t \geq \frac{1}{\sigma}\ln\frac{s}{b} > 0\right\}, \end{aligned}$$

i.e. the contract is early exercised when the positively drifting Brownian motion $-\frac{1}{\sigma}(r_d - r_f - \frac{\sigma^2}{2})t - B_t$ first reaches level $\frac{1}{\sigma}\ln\frac{s}{b}$. Note that τ_b is finite and in turn $S_{\tau_b} = b$ almost surely as a positively drifting Brownian motion reaches any positive level in finite time with probability one. Then we conclude

$$\begin{aligned} V_\infty(s) &\geq \mathbb{E}[e^{-r_d\tau_b}(S_{\tau_b} - K)] \\ &= (b-K)\mathbb{E}(e^{-r_d\tau_b}) \to +\infty \end{aligned}$$

because $\mathbb{E}(e^{-r_d\tau_b})$ diverges to infinity under $r_d + \frac{1}{2\sigma^2}(r_d - r_f - \frac{\sigma^2}{2})^2 < 0$ due to lemma A.1. Hence $V_\infty(s) = +\infty$ for all $s > K$, and the same conclusion extends to any $s > 0$ since S_t reaches any arbitrary level above K with positive probability as per the argument in the proof of part 6(a).

The second possibility is that we have $|\sqrt{-2r_d} - \sqrt{-2r_f}| < \sqrt{2(r_f - r_d)} \leq \sigma$. Then $r_d - r_f + \frac{\sigma^2}{2} \geq 0$, and

$$\begin{aligned} &\mathbb{E}[e^{-r_d\tau}(S_\tau - K)] \\ &= \mathbb{E}\left[e^{\sigma B_\tau - \frac{\sigma^2}{2}\tau}e^{-r_f\tau}\left(s - Ke^{\left(-r_d+r_f+\frac{\sigma^2}{2}\right)\tau - \sigma B_\tau}\right)\right] \\ &= \tilde{\mathbb{E}}\left[e^{-r_f\tau}\left(s - Ke^{\left(-r_d+r_f+\frac{\sigma^2}{2}\right)\tau - \sigma B_\tau}\right)\right] \\ &= \tilde{\mathbb{E}}\left[e^{-r_f\tau}\left(s - Ke^{-\left(r_d-r_f+\frac{\sigma^2}{2}\right)\tau - \sigma\tilde{B}_\tau}\right)\right], \end{aligned}$$

where we have defined a new measure $\tilde{\mathbb{Q}}$ via the Radon–Nikodym derivative $\frac{d\tilde{\mathbb{Q}}}{d\mathbb{Q}}|_{\mathcal{F}_t} = e^{\sigma B_t - \frac{\sigma^2}{2}t}$ and $\tilde{B}_t := B_t - \sigma t$ is a Brownian motion under $\tilde{\mathbb{Q}}$. For any $b > 0$, consider an early exercise strategy in form of $\tau_b := \inf\{t \geq 0 : -(r_d - r_f + \frac{\sigma^2}{2})\tau - \sigma\tilde{B}_\tau \leq -b\}$. Note that $r_d - r_f + \frac{\sigma^2}{2} \geq 0$ implies $\tau_b < \infty$ almost surely and in turn $\mathbb{E}[e^{-r_d\tau_b}(S_{\tau_b} - K)] = (s - Ke^{-b})\tilde{\mathbb{E}}[e^{-r_f\tau_b}]$. If $r_f + \frac{1}{2\sigma^2}(r_d - r_f + \frac{\sigma^2}{2}) < 0$, then $\tilde{\mathbb{E}}[e^{-r_f\tau_b}] = +\infty$ for any $b > 0$ and hence $V_\infty(s) = +\infty$ on $s > Ke^{-b}$, and in turn $V_\infty(s) = +\infty$ for all s. Else if $r_f + \frac{1}{2\sigma^2}(r_d - r_f + \frac{\sigma^2}{2}) \geq 0$,

$$\begin{aligned} \tilde{\mathbb{E}}\left[e^{-r_f\tau_b}\right] = \exp\Bigg(-\frac{b}{\sigma}\Bigg(&-\frac{1}{\sigma}\left(r_d - r_f + \frac{\sigma^2}{2}\right) \\ &+ \sqrt{\frac{1}{\sigma^2}\left(r_d - r_f + \frac{\sigma^2}{2}\right)^2 + 2r_f}\Bigg)\Bigg) \end{aligned}$$

by lemma A.1. But $r_f < 0$ and hence

$$\begin{aligned} &-\frac{1}{\sigma}\left(r_d - r_f + \frac{\sigma^2}{2}\right) \\ &+ \sqrt{\frac{1}{\sigma^2}\left(r_d - r_f + \frac{\sigma^2}{2}\right)^2 + 2r_f} < 0 \end{aligned}$$

such that $\tilde{\mathbb{E}}[e^{-r_f\tau_b}] \to +\infty$ as $b \uparrow +\infty$. Therefore, $V_\infty(s) \geq (s - Ke^{-b})\tilde{\mathbb{E}}[e^{-q\tau_b}] \to +\infty$ as $b \uparrow +\infty$.

(b) With some algebra, one can check that the conditions of $r_d < 0$, $r_f < 0$, $r_d < r_f$ and $\sigma \leq |\sqrt{-2r_d} - \sqrt{-2r_f}|$ ensure that equation (7) admits two distinct positive roots with the smaller root γ_1 being larger than one. The rest of the arguments is largely similar to part (5) of the proof. We can show that V_∞^C, defined as the function constructed in (8), satisfies

$$\begin{cases} \dfrac{\sigma^2 s^2}{2}\dfrac{\mathrm{d}^2 V_\infty^C}{\mathrm{d}s^2} + (r_d - r_f) \\ \quad s\dfrac{\mathrm{d}V_\infty^C}{\mathrm{d}s} - r_d V_\infty^C = 0 \text{ and} \\ \quad V_\infty^C(s) \geq s - K, & s > b_1; \\ \dfrac{\sigma^2 s^2}{2}\dfrac{\mathrm{d}^2 V_\infty^C}{\mathrm{d}s^2} + (r_d - r_f) \\ \quad s\dfrac{\mathrm{d}V_\infty^C}{\mathrm{d}s} - r_d V_\infty^C \leq 0 \text{ and} \\ \quad V_\infty^C(s) = s - K, & s \leq b_1. \end{cases}$$

Generalised Ito's lemma applied to $M_t := e^{-r_d t} V^C(S_t)$ now gives

$$\begin{aligned} \mathrm{d}M_t &= e^{-r_d t}\left\{\left[\frac{\sigma^2 S_t^2}{2}\partial_{ss}^2 V_\infty^C + (r_d - r_f)S_t\partial_s V_\infty^C \right.\right. \\ &\qquad \left.\left. - \; r_d V^C\right]\mathrm{d}t + \sigma S_t \partial_s V_\infty^C \,\mathrm{d}B_t\right\} \\ &= e^{-r_d t}\left[(r_d K - r_f S_t)1_{\{S_t \leq b_1\}}\,\mathrm{d}t + \sigma S_t\partial_s V_\infty^C\,\mathrm{d}B_t\right]. \end{aligned}$$

However, the argument in part (5) of the proof does not apply because with $r_d < 0$ we no longer have $(e^{-r_d t} V_\infty^C(S_t))_{t\geq 0}$ being uniformly bounded by $-K$ from below.† Nonetheless, the quadratic variation of the stochastic integral $\int_0^t e^{-r_d u} S_u \partial_s V_\infty^C(S_u)\,\mathrm{d}B_u$ is

$$\begin{aligned} &\int_0^t e^{-2r_d u} S_u^2 (\partial_s V_\infty^C)^2\,\mathrm{d}u \\ &\quad = \int_0^t e^{-2r_d u} S_u^2 (\partial_s V_\infty^C)^2 1_{\{S_u \leq b_1\}}\,\mathrm{d}u \\ &\qquad + \int_0^t e^{-2r_d u} S_u^2 (\partial_s V_\infty^C)^2 1_{\{S_u > b_1\}}\,\mathrm{d}u \\ &\quad = \int_0^t e^{-2r_d u} S_u^2 1_{\{S_u \leq b_1\}}\,\mathrm{d}u \\ &\qquad + \int_0^t e^{-2r_d u} S_u^2 \left(\frac{S_u}{b_1}\right)^{2\gamma_1 - 2} 1_{\{S_u > b_1\}}\,\mathrm{d}u \\ &\quad \leq b_1^2 \frac{1 - e^{-2r_d t}}{2r_d} + b_1^{2-2\gamma_1}\int_0^t e^{-2r_d u} S_u^{2\gamma_1}\,\mathrm{d}u. \end{aligned}$$

Then since

$$\begin{aligned} &\mathbb{E}\left[\int_0^t e^{-2r_d u} S_u^{2\gamma_1}\,\mathrm{d}u\right] \\ &\quad = \int_0^t s^{2\gamma_1} e^{2[-r_d + (r_d - r_f - \sigma^2/2)\gamma_1 + \sigma^2\gamma_1^2]u}\,\mathrm{d}u \\ &\quad = s^{2\gamma_1}\int_0^t e^{\sigma^2\gamma_1 u}\,\mathrm{d}u = s^{2\gamma_1}\frac{e^{\sigma^2\gamma_1 t} - 1}{\sigma^2\gamma_1} \end{aligned}$$

where we have used the fact that γ_1 is a solution to (7), we conclude $\mathbb{E}[\int_0^t e^{-2r_d u} S_u^2 (\partial_s V_\infty^C)^2\,\mathrm{d}u] < \infty$ for all t. Hence the stochastic integral $\int_0^t e^{-r_d u} S_u \partial_s V_\infty^C(S_u)\,\mathrm{d}B_u$ is a true martingale. M is then a supermartingale. The proof is concluded by following the same arguments in part (5).

■

† This kind of technical difficulty does not arise for American option which payoff is always non-negative.

In what follows, we collect some regularity results of the value function. We only briefly sketch the proofs as they are standard and can be easily adapted from similar proofs for American options.

PROPOSITION A.1 *The map $(t, s) \to V(t, s)$ is continuous.*

Proof Since separate continuity implies joint continuity if the function is monotonous in one of the arguments (Young 1910), in view of part 2 and 3 of lemma 2 it is sufficient to show that $t \to V(t, s)$ is continuous. This can be done by following the same line of arguments in chapter 25.2.3 of Peskir and Shiryaev (2006) in the context of American put option. ■

PROPOSITION A.2 *On the continuation set $\mathcal{C}$, $V(t, s) \in C^{1\times 2}$ and it satisfies*

$$\mathcal{L}V := \frac{\partial V}{\partial t} + \frac{\sigma^2 s^2}{2}\frac{\partial^2 V}{\partial s^2} + (r_d - r_f)s\frac{\partial V}{\partial s} - r_d V = 0. \qquad \text{(A1)}$$

Proof This can be established from the facts that $(e^{-r_d(u\wedge\tau^*)}V(u \wedge \tau^*, S_{u\wedge\tau^*}^{t,s}))_{u\geq t}$ is a martingale for $\tau^* = \inf\{u \geq t : (u, S_u^{t,s}) \notin \mathcal{C}\}$ together with standard existence and uniqueness results of a parabolic PDE with operator $\mathcal{L}$. See, for example, proposition 2.6 of Jacka (1991). ■

PROPOSITION A.3 *For any $t \in [0, T]$, the map $s \to V(t, s)$ is C^1. In particular, if $b(t) > 0$ then smooth pasting holds at $s = b(t)$ such that $\partial_s V(t, b(t)) = 1$.*

Proof Since $V(t, s) = s - K$ on $\mathcal{D}$ and it is known that $V(t, s)$ is $C^{1\times 2}$ on $\mathcal{C}$, it is sufficient to show that $s \to V(t, s)$ is C^1 at $s = b(t)$ in the case of $b(t) > 0$. Hence one just needs to show $\partial_s^+ V(t, b(t)) = \partial_s^- V(t, b(t)) = 1$. This can be verified by following the arguments in chapter 25.2.4 of Peskir and Shiryaev (2006). ■

PROPOSITION A.4 *$V(t, s)$ is strictly convex in s on $\mathcal{C}$.*

Proof Similar proofs can be found in corollary 3.1 of Pham (1997), for example. The idea is that if there existed $(\bar{t}, \bar{s}) \in \mathcal{C}$ such that $\partial_{ss}V(\bar{t}, \bar{s}) = 0$, then the consideration of a suitable Dirichlet problem and Feynman–Kac formula would lead to $\partial_{ss}V(t, s) = 0$ for all $(t, s) \in \mathcal{C}$. This contradicts the definition of the continuation set $\mathcal{C}$ which by Corollary 1 is non-empty. ■

We are now ready to prove proposition 4 in the main text.

Proof of proposition 4. We prove the result for the case of $r_d < 0$ and $r_f < 0$, where the case of $r_d > 0$ and $r_f > 0$ can be handled similarly. The proof follows closely to Jacka (1991) but some extra care is needed to handle the negative interest rates.

(1) Note that $(t, b(t)) \in \mathcal{D}$ when $b(t) > 0$. Now fix t_1, t_2 such that $0 \leq t_1 \leq t_2 < T$. If $b(t_1) = 0$, then obviously $b(t_2) \geq 0 = b(t_1)$. Else if $b(t_1) > 0$, the fact that $V(t, s)$ is decreasing in t leads to $b(t_1) - K = V(t_1, b(t_1)) \geq V(t_2, b(t_1)) \geq b(t_1) - K$ and hence $V(t_2, b(t_1)) = b(t_1) - K$, i.e. $(t_2, b(t_1)) \in \mathcal{D}$. Thus we must have $b(t_2) \geq b(t_1)$.

(2) For $(t, s) \in \mathcal{D}$ where $\mathcal{D}$ denotes the interior of $\mathcal{D}$, the value function is given by $V(t, s) = s - K$ which is $C^{1\times 2}$ on $\mathcal{D}$. As $(e^{-r_d t}V(t, S_t))_{t\geq 0}$ is a supermartingale, we can apply Ito's lemma to deduce

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 s^2}{2}\frac{\partial^2 V}{\partial s^2} + (r_d - r_f)s\frac{\partial V}{\partial s} - r_d V \leq 0.$$

Hence we have $(r_d - r_f)s - r_d(s - K) \leq 0 \implies -r_f s + r_d K \leq 0$. This inequality is true for all $(t, s) \in \mathcal{D} \iff s < b(t)$ and therefore $-r_f b(t) + r_d K \leq 0$ and finally $b(t) \leq \frac{r_d}{r_f}K$ as $r_f < 0$.

(3) If $b(t) = 0$ for all $t \in [0, T)$ then we are done. Otherwise, define

$$t_0 := \inf\{t \in [0, T) : b(t) > 0\}$$

such that $b(t) > 0$ on $t \in (t_0, T)$ due to the monotonicity of b. Obviously $b(t) = 0$ for all $t \in [0, t_0)$ on which b is continuous. It remains to show that b is continuous on (t_0, T) and $b(t_0) = 0$.

Fix some small $\delta > 0$. The monotonicity of b implies that b has right limit at any $t \in [t_0 + \delta, T)$ and in particular $b(t) \le b(t+)$. Now, consider a sequence $(t_n)_{n\ge 1}$ where $t_n \downarrow t$ such that $b(t_n) \ge b(t_0 + \delta) > 0$. Then $(t_n, b(t_n)) \in \mathcal{D}_\delta$ for each n, where

$$\begin{aligned}\mathcal{D}_\delta &:= \mathcal{D} \cap \left\{[t_0 + \delta, T] \times \left[b(t_0 + \delta), \frac{r_d}{r_f}K\right]\right\} \\ &= \{(t, s) \in [t_0 + \delta, T] \\ &\quad \times \left[b(t_0 + \delta), \frac{r_d}{r_f}K\right] : V(t, s) = s - K\right\}.\end{aligned}$$

Since V is continuous, $\mathcal{D}_\delta$ is a closed set which implies $(t, b(t+)) \in \mathcal{D}_\delta \subset \mathcal{D}$ and in turn $b(t+) \le b(t)$. Hence we conclude $b(t) = b(t+)$, i.e. b is right continuous on $[t_0 + \delta, T)$. As $\delta > 0$ is arbitrary, b is right continuous on (t_0, T).

We now show that $b(t)$ is left continuous on $t \in (t_0, T)$. The monotonicity of b implies the existence of a left-limit everywhere. Hence using the closeness of $\mathcal{D}_\delta$ again, $(t, b(t-)) \in \mathcal{D}_\delta \subset \mathcal{D}$ for any $t \in (t_0 + \delta, T)$. Assume on contrary that there exists $\bar{t} \in (t_0 + \delta, T)$ such that $b(\bar{t}-) < b(\bar{t})$. Define $\xi := \frac{1}{2}(b(\bar{t}) + b(\bar{t}-))$ such that $b(\bar{t}-) < \xi < b(\bar{t}) \le r_d K/r_f$. Choose an arbitrary $u \in (t_0 + \delta, \bar{t})$ such that $0 < b(u) \le b(\bar{t}-) < \xi < b(\bar{t}) \le r_d K/r_f$. As $(u, b(u)) \in \mathcal{D}$, value matching and smooth pasting properties of the value function lead to $V(u, b(u)) = b(u) - K$ and $V_s(u, b(u)) = 1$. Now,

$$\begin{aligned}V(u, \xi) - (\xi - K) &= [V(u, \xi) - \xi] - [V(u, b(u)) - b(u)] \\ &= \int_{b(u)}^{\xi} (V_s(u, y) - 1)\,\mathrm{d}y \\ &= \int_{b(u)}^{\xi}\int_{b(u)}^{y} V_{ss}(u, z)\,\mathrm{d}z\,\mathrm{d}y. \end{aligned} \tag{A2}$$

For any $(t, s) \in \mathcal{C}$, $V(t, s)$ satisfies $-\frac{\partial V}{\partial t} - \frac{\sigma^2 s^2}{2}\frac{\partial^2 V}{\partial s^2} - (r_d - r_f)s\frac{\partial V}{\partial s} + r_d V = 0$. Hence for any $u \in (t_0 + \delta, \bar{t})$

$$\begin{aligned}&\lim_{s\downarrow b(u)} \frac{\sigma^2 s^2}{2}\frac{\partial^2 V}{\partial s^2}(u, s) \\ &\quad = \lim_{s\downarrow b(u)}\left[r_d V(u, s) - (r_d - r_f)s\frac{\partial V}{\partial s}(u, s) - \frac{\partial V}{\partial t}(u, s)\right] \\ &\quad \ge \lim_{s\downarrow b(u)}\left[r_d V(u, s) - (r_d - r_f)s\frac{\partial V}{\partial s}(u, s)\right] \\ &\quad = r_d(b(u) - K) - (r_d - r_f)b(u) \\ &\quad = r_f b(u) - r_d K > r_f\xi - r_d K > 0\end{aligned}$$

since $\frac{\partial V}{\partial t} \le 0$, V and $\frac{\partial V}{\partial S}$ are continuous in s with $V(t, b(t)) = b(t) - K$, $\frac{\partial V}{\partial s}(t, b(t)) = 1$ for any $t \in (t_0, T)$, and $b(u) < \xi < r_d K/r_f$. As $\frac{\partial^2 V}{\partial s^2} > 0$ on $\mathcal{C}$ due to Proposition A.4, the map $(t, s) \to \frac{\sigma^2 s^2}{2}\frac{\partial^2 V}{\partial s^2}$ attains a strictly positive minimum, say $\epsilon > 0$, over the compact set $\mathrm{cl}([0, T] \times (0, \xi) \cap \mathcal{C})$. Then (A2) leads to

$$\begin{aligned}V(u, \xi) - (\xi - K) &\ge \frac{2r_f^2\epsilon}{r_d^2 K^2\sigma^2}\int_{b(u)}^{\xi}\int_{b(u)}^{y} \mathrm{d}z\,\mathrm{d}y \\ &= \frac{r_f^2\epsilon}{r_d^2 K^2\sigma^2}(\xi - b(u))^2.\end{aligned}$$

On sending $u \uparrow \bar{t}$, using the continuity of V we have

$$V(\bar{t}, \xi) - (\xi - K) \ge \frac{r_f^2\epsilon}{r_d^2 K^2\sigma^2}(\xi - b(\bar{t}-))^2 > 0,$$

but this contradicts the fact that $0 < \xi < b(\bar{t})$ which would have implied $(\bar{t}, \xi) \in \mathcal{D}$ and in turn $V(\bar{t}, \xi) = \xi - K$. We therefore conclude $b(t)$ is left-continuous on (t_0, T) as $\delta > 0$ is arbitrary.

The conclusion that $b(t_0) = 0$ follows from a similar argument. Suppose on contrary that $b(t_0) > 0$ and set $\xi := \frac{b(t_0)}{2} \in (0, b(t_0))$. Then for any arbitrary $u \in [0, t_0)$ we can use the same previous arguments to establish

$$V(u, \xi) - (\xi - K) \ge \frac{r_f^2\epsilon}{r_d^2 K^2\sigma^2}(\xi - b(u))^2$$

for some $\epsilon > 0$. Using the fact that $b(t_0-) = 0$, taking limit $u \uparrow t_0$ leads to $V(t_0, \xi) - (\xi - K) > 0$. But this contradicts the fact that $0 < \xi < b(t_0)$ which would have implied $V(t_0, \xi) = \xi - K$.

Finally, we show that $b(T) := \lim_{t\to T} b(t) = \frac{r_d}{r_f}K$. Since $b(t) \le r_d K/r_f$ for all $t \in [0, T)$, it is sufficient to argue that $b(T) \ge r_d K/r_f$. Suppose on contrary that $b(T) < r_d K/r_f$. Write $\xi := \frac{1}{2}(\frac{r_d K}{r_f} + b(T)) \in (b(T), \frac{r_d K}{r_f})$. Using the same arguments for the left-continuity proof, we can deduce $V(u, \xi) - (\xi - K) > \frac{r_f^2\epsilon}{r_d^2 K^2\sigma^2}(\xi - b(u))^2$ for any $u < T$ and some $\epsilon > 0$. But $\lim_{t\uparrow T} V(t, s) = V(T, s) = s - K$, contradiction will again be obtained when we let $u \uparrow T$.

■

Appendix 2. Put-call symmetry and extension to sale FX Open Forward

Put-call symmetry is a convenient tool to establish a linkage between the pricing results of call and put options (both European and American). See Detemple (2001) and the references therein. This symmetry result also holds for forward payoffs.

Write $J(t, s, K, r_d, r_f; \tau) := \mathbb{E}^{(t,s)}[e^{-r_d(\tau - t)}(S_\tau - K)]$ which is the time-t risk-neutral expected value of a purchase FX Open Forward with maturity T under an exercise strategy of $\tau \in \mathcal{T}_{t,T}$, initial spot s, delivery price K, domestic interest rate r_d and foreign interest rate r_f. Define $\hat{J}(t, s, K, r_d, r_f; \tau) = \mathbb{E}^{(t,s)}[e^{-r_d(\tau - t)}(K - S_\tau)]$ similarly for a sale FX Open Forward. Now,

$$\begin{aligned}\hat{J}(t, s, K, r_d, r_f; \tau) &= \mathbb{E}^{(t,s)}\left[e^{-r_d(\tau - t)}(K - S_\tau)\right] \\ &= \mathbb{E}^{(t,s)}\left[e^{-r_d(\tau - t)}(K - se^{\left(r_d - r_f - \frac{\sigma^2}{2}\right)(\tau - t) + \sigma(B_\tau - B_t)})\right] \\ &= \mathbb{E}^{(t,s)}\left[e^{-r_f(\tau - t)}e^{\sigma(B_\tau - B_t) - \frac{\sigma^2}{2}(\tau - t)}\right. \\ &\quad \left.\times\left(Ke^{-\left(r_d - r_f - \frac{\sigma^2}{2}\right)(\tau - t) - \sigma(B_\tau - B_t)} - s\right)\right] \\ &= \tilde{\mathbb{E}}^{(t,s)}\left[e^{-r_f(\tau - t)}\left(Ke^{\left(r_f - r_d - \frac{\sigma^2}{2}\right)(\tau - t) - \sigma(\tilde{B}_\tau - \tilde{B}_t)} - s\right)\right] \\ &= J(t, K, s, r_f, r_d; \tau),\end{aligned} \tag{A3}$$

where we have defined a new measure $\tilde{\mathbb{Q}}$ via the Radon–Nikodym derivative

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}}{\mathrm{d}\mathbb{Q}}\right|_{\mathcal{F}_t} = e^{\sigma B_t - \frac{\sigma^2}{2}t}$$

under which $\tilde{B}_t := B_t - \sigma t$ is a Brownian motion under $\tilde{\mathbb{Q}}$. Upon taking supremum over stopping times τ, we have

$$\sup_{\tau\in\mathcal{T}_{t,T}} \hat{J}(t,s,K,r_d,r_f;\tau) = \sup_{\tau\in\mathcal{T}_{t,T}} J(t,K,s,r_f,r_d;\tau). \tag{A4}$$

In other words, the pricing properties of a sale FX Open Forward can simply be inferred from that of a purchase FX Open Forward upon swapping spot price with delivery level, and domestic interest rate with foreign interest rate.

A.1. Perpetual sale FX Open Forward

Using put-call symmetry, we now state a mirror version of theorem 1. Let

$$\hat{V}_\infty(s) := \sup_{\tau\in\mathcal{T}_{0,\infty}} \mathbb{E}\left[e^{-r_d\tau}(K - S_\tau) \mid S_0 = s\right] \tag{A5}$$

be the fair price of a perpetual sale FX Open Forward.

Theorem A.1 *For problem* (A5)*:*

1. *If* $r_f \le 0 \le r_d$, $\hat{V}_\infty(s) = K - s$ *and* $\tau^* = 0$ *is optimal to problem* (A5).
2. *If* $r_d < 0 \le r_f$, $\hat{V}_\infty(s) = +\infty$ *and* $\tau^* = +\infty$ *is optimal to problem* (A5).
3. *If* $r_d = 0 < r_f$, $\hat{V}_\infty(s) = K$ *and* $\tau^* = +\infty$ *is optimal to problem* (A5).
4. *If* $r_d = 0 = r_f$, $\hat{V}_\infty(s) = K - s$ *which can be attained by any non-negative stopping time* τ.
5. *If* $r_d > 0$ *and* $r_f > 0$,
$$\hat{V}_\infty(s) = \begin{cases} -\dfrac{b_1}{\gamma_1}\left(\dfrac{s}{b_1}\right)^{\gamma_1}, & s > b_1; \\ K - s, & s \le b_1, \end{cases} \tag{A6}$$
where $\gamma_1 < 0$ *is the smaller root to the quadratic equation* (7) *and* $b_1 := \frac{\gamma_1}{\gamma_1 - 1}K$*. The optimal early exercise strategy is* $\tau^* = \inf\{t \ge 0 : S_t \le b_1\}$*, i.e. to early exercise the contract when the spot price level is sufficiently low.*
6. *If* $r_d < 0$ *and* $r_f < 0$*:*
   (a) *If* $r_d \le r_f$*, or* $r_f < r_d$ *and* $\sigma > |\sqrt{-2r_d} - \sqrt{-2r_f}|$*, then* $\hat{V}_\infty(s) = +\infty$.
   (b) *If* $r_f < r_d$ *and* $\sigma \le |\sqrt{-2r_d} - \sqrt{-2r_f}|$*, then*
$$\hat{V}_\infty(s) = \begin{cases} K - s, & s \ge b_2; \\ -\dfrac{b_2}{\gamma_2}\left(\dfrac{s}{b_2}\right)^{\gamma_2}, & s < b_2, \end{cases} \tag{A7}$$
where $\gamma_2 < 0$ *is the larger root to the quadratic equation (7) and* $b_2 := \frac{\gamma_2}{\gamma_2 - 1}K$*. The optimal early exercise strategy is* $\tau^* = \inf\{t \ge 0 : S_t \ge b_2\}$*, i.e. to early exercise the contract when the spot price level is sufficiently high.*

Proof Results follow upon direct replacement of r_d by r_f and s by K in Theorem 1. A useful observation is that if $\hat{\gamma}_1 < \hat{\gamma}_2$ are the roots to the equation

$$\frac{\sigma^2}{2}\gamma^2 + \left(r_f - r_d - \frac{\sigma^2}{2}\right)\gamma - r_f = 0,$$

then we have $\gamma_1 = 1 - \hat{\gamma}_2$ and $\gamma_2 = 1 - \hat{\gamma}_1$ where $\gamma_1 < \gamma_2$ are the roots to the quadratic equation in (7). ■

A.2. Finite-maturity sale FX Open Forward

For a finite-maturity purchase FX Open Forward, we know that there exists a function $b = b(t)$ that divides the exercise region and continuation region. Similar results hold for a sale FX Open Forward. Moreover, the optimal exercise boundaries of these two contracts are explicitly linked.

Before we state the results, we first present a useful scaling property of the optimal exercise boundary. Define by $b(t) = b(t;K,r_d,r_f)$ the optimal exercise boundary function of a purchase FX Open Forward under delivery level K, domestic interest rate r_d and foreign interest rate r_f.

Lemma A.2 *The optimal exercise boundary for a purchase FX Open Forward is homogeneous in K with degree 1, i.e.* $b(t;K,r_d,r_f) = Kb(t;1,r_d,r_f)$ *for any* $K > 0$.

Proof Starting from the definition, we have

$$\begin{aligned} V(t,s) = V(t,s;K) &:= \sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}[e^{-r_d(\tau-t)}(S_\tau^{t,s} - K)] = K \\ &\times \sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}\left[e^{-r_d(\tau-t)}\left(\frac{S_\tau^{t,s}}{K} - 1\right)\right] \\ &= K \sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}\left[e^{-r_d(\tau-t)}\left(S^{t,s/K} - 1\right)\right]. \end{aligned} \tag{A8}$$

Suppose $r_d > 0$ and $r_f > 0$. The optimal exercise time of problem (A8) is given by

$$\begin{aligned} \tau^* &= \inf\left\{u \in [t,T] : S_u^{t,s/K} \ge b(u;1,r_d,r_f)\right\} \\ &= \inf\left\{u \in [t,T] : \frac{S_u^{t,s}}{K} \ge b(u;1,r_d,r_f)\right\} \\ &= \inf\left\{u \in [t,T] : S_u^{t,s} \ge Kb(u;1,r_d,r_f)\right\}. \end{aligned}$$

But on the other hand, it is also known that the optimiser of $\sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}[e^{-r_d(\tau-t)}(S_\tau^{t,s} - K)]$ is

$$\inf\left\{u \in [t,T] : S_u^{t,s} \ge b(u;K,r_d,r_f)\right\}.$$

We therefore must have $b(t;K,r_d,r_f) = Kb(t;1,r_d,r_f)$. The argument is the same for the case of $r_d < 0$ and $r_f < 0$. ■

Proposition A.5 *Denote by*

$$\hat{V}(t,s) := \sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}\left[e^{-r_d(\tau-t)}(K - S_\tau) \mid S_t = s\right] \tag{A9}$$

the time-t fair value of a sale FX Open Forward. Let

$$\hat{\mathcal{D}} := \inf\{(t,s) \in [0,T] \times (0,\infty) : \hat{V}(t,s) = K - s\} \text{ and}$$
$$\hat{\mathcal{C}} := \inf\{(t,s) \in [0,T] \times (0,\infty) : \hat{V}(t,s) > K - s\}$$

be the associated stopping set and continuation set.

1. *If* $r_d > 0$ *and* $r_f > 0$*, then there exists a strictly positive function* $\hat{b}(t) = \hat{b}(t;K,r_d,r_f) \in (0,\infty)$ *such that*
$$\hat{\mathcal{C}} := \{(t,s) \in [0,T) \times (0,\infty) : s > \hat{b}(t)\},$$
$$\hat{\mathcal{D}} := \{(t,s) \in [0,T) \times (0,\infty) : s \le \hat{b}(t)\} \cup \{\{T\} \times (0,\infty)\}.$$
2. *If* $r_d < 0$ *and* $r_f < 0$*, then there exists a strictly positive function (which may take value of infinity)* $\hat{b}(t) = \hat{b}(t;K,r_d,r_f) \in (0,\infty]$ *such that*
$$\hat{\mathcal{C}} := \{(t,s) \in [0,T) \times (0,\infty) : s < \hat{b}(t)\},$$
$$\hat{\mathcal{D}} := \{(t,s) \in [0,T) \times (0,\infty) : s \ge \hat{b}(t)\} \cup \{\{T\} \times (0,\infty)\}.$$

Moreover, write $\hat{b}(t) = b(t;K,r_d,r_f)$ *as the optimal exercise boundary function of a sale FX Open Forward under delivery level K, domestic interest rate* r_d *and foreign interest rate* r_f*, then*

$$\hat{b}(t;K,r_d,r_f) = \begin{cases} \dfrac{K}{b(t;1,r_f,r_d)}, & b(t;1,r_f,r_d) > 0; \\ +\infty, & b(t;1,r_f,r_d) = 0. \end{cases} \tag{A10}$$

Proof Starting from (A4) and (A3), we have

$$\hat{V}(t,s) := \sup_{\tau\in\mathcal{T}_{t,T}} \hat{J}(t,s,K,r_d,r_f;\tau) = \sup_{\tau\in\mathcal{T}_{t,T}} J(t,K,s,r_f,r_d;\tau)$$
$$= \sup_{\tau\in\mathcal{T}_{t,T}} \tilde{\mathbb{E}}^{(t,s)}\left[e^{-r_f(\tau-t)}(Ke^{\left(r_f-r_d-\frac{\sigma^2}{2}\right)(\tau-t)-\sigma(\tilde{B}_\tau-\tilde{B}_t)}-s)\right]. \tag{A11}$$

The optimiser associated with (A11) can be inferred from the baseline result of a purchase FX Open Forward. Suppose $r_d < 0$ and $r_f < 0$. Then by proposition 3, the optimiser is then

$$\tau^* = \inf\left\{u\in[t,T] : Ke^{\left(r_f-r_d-\frac{\sigma^2}{2}\right)(u-t)-\sigma(\tilde{B}_u-\tilde{B}_t)} \le b(u;s,r_f,r_d)\right\}$$
$$= \inf\left\{u\in[t,T] : Ke^{\left(r_f-r_d+\frac{\sigma^2}{2}\right)(u-t)-\sigma(B_u-B_t)} \le b(u;s,r_f,r_d)\right\}$$
$$= \inf\left\{u\in[t,T] : e^{\left(r_d-r_f-\frac{\sigma^2}{2}\right)(u-t)+\sigma(B_u-B_t)} b(u;s,r_f,r_d) \ge K\right\}$$
$$= \inf\left\{u\in[t,T] : S_u^{t,s} b(u;1,r_f,r_d) \ge K\right\}$$
$$= \inf\left\{u\in[t,T] : S_u^{t,s} \ge \hat{b}(t;K,r_d,r_f)\right\},$$

where we used lemma A.2 in the second last equality. The optimal strategy is thus to exercise the sale FX Open Forward when the spot price S_t is at or above $\hat{b}(t) = \hat{b}(t;K,r_d,r_f)$. The case of $r_d > 0$ and $r_f > 0$ can be handled similarly. As a remark, in the case of negative rates $b(t)$ may take value of zero under which it is suboptimal to early exercise the purchase FX Open Forward at all spot levels. This translates into $\hat{b}(t) = +\infty$ where financially it means that at time t one should defer exercising the sale FX Open Forward at all spot levels. ∎

PROPOSITION A.6 *View $t \to \hat{b}(t) = \hat{b}(t;K,r_d,r_f)$ as a function of t. Then:*

1. *$\hat{b}$ is increasing (resp. decreasing) if $r_d > 0$ and $r_f > 0$ (resp. $r_d < 0$ and $r_f < 0$).*
2. *$\hat{b}(t) \le \frac{r_d}{r_f}K$ (resp. $b(t) \ge \frac{r_d}{r_f}K$) if $r_d > 0$ and $r_f > 0$ (resp. $r_d < 0$ and $r_f < 0$) for all $t \in [0,T)$.*
3. *$\hat{b}$ is continuous and $\hat{b}(T) := \lim_{t\to T}\hat{b}(t) = \frac{r_d}{r_f}K$.*

In the case of $r_d > 0$ and $r_f > 0$, the domain of $\hat{b}$ is $[0,T)$. In the case of $r_d < 0$ and $r_f < 0$, the domain of $\hat{b}$ is restricted to (t_0,T), where

$$t_0 := \inf\{t\in[0,T] : b(t;1,r_f,r_d) > 0\}.$$

Proof It follows immediately from proposition 4 and (A10). ∎