NON-TRANSFERABLE NON-HEDGEABLE EXECUTIVE STOCK OPTION PRICING

David B. Colwell,[a] David Feldman,[b] Wei Hu[c,d]

Latest Revision July 30, 2014

Abstract. We introduce a method of valuating non-transferable non-hedgeable (NTNH) contingent claims and use it to price executive stock options (ESOs). We use NTNH constraints to break local co-linearity, caused by derivative assets, that renders common portfolio optimization methods ineffective. We are, thus, able to translate portfolios that include NTNH derivatives into portfolios of primary assets (only). We achieve this by replicating derivatives using primary assets, and then integrating NTNH constraints into a single rectangular constraint. Solving the constrained portfolio optimization facilitates the identification of stochastic discount factors that price any contingent claim in such portfolios. Implementing our method, we find subjective prices of NTNH European and American ESOs, both for block exercise and for a continuum of exercise ratios. We identify executives' optimal exercise policies. We use these to find the objective price/cost of ESOs to firms. We run a simulation study of price sensitivities and changes to optimal exercise policies with respect to model parameters and obtain policy implications regarding ESOs incentivizing efficiency.

Keywords: Executive Stock Options, Constrained Portfolio Optimization, Stochastic Discount Factor, Non-Hedgeable, Non-Transferable

JEL Codes: G11, G13, C02, C61

[a] School of Banking and Finance, Australian School of Business, The University of New South Wales, UNSW Sydney, NSW 2052, Australia; Phone: +61-(0)2-9385-5851; email: d.colwell@unsw.edu.au.

[b] School of Banking and Finance, Australian School of Business, The University of New South Wales, UNSW Sydney, NSW 2052, Australia; Phone: +61-(0)2-9385-5847; email: d.feldman@unsw.edu.au.

[c] Corresponding Author, Department of Finance and Banking, School of Economics and Finance, Curtin University of Technology, Perth, WA 6101, Australia; phone: +61-(0)8-9266-9012; email: wei.hu@curtin.edu.au.

[d] We thank Menachem Brenner, Wolfgang Buehler, Ben Goldys, Linda Pesante, Marek Rutkowski, James Sørlie, and Zvi Wiener for helpful discussions and thank participants of the Australasian Finance and Banking Conference, Portuguese Finance Network International Conference, International Risk Management Conference, and The Econometric Society Australasian Meeting. The SSRN link for this paper is http://ssrn.com/abstract=2446385.

1 Introduction

Executive stock options (ESOs) are call options granted by firms to employees as a form of compensation in addition to base salary, bonuses, and retirement savings. ESOs typically become "vested" (exercisable) over time [Murphy (1999)]. ESOs are generally non-transferrable and non-hedgeable (NTNH) for at least two reasons: NTNH features incentivize executives, and they prevent negative signalling to investors. Holders of NTNH ESOs cannot transfer the ESOs to third parties at any price, nor are they allowed to hedge their value by positions in underlying assets or other derivatives. ESOs are typically forfeited if the executive leaves the firm before vesting.

There are two approaches to pricing contingent claims: arbitrage and equilibrium. The difficulty in pricing ESOs lies in their NTNH feature. On one hand, NTNH constraints nullify arbitrage pricing, disabling arbitrage strategies. On the other hand, NTNH adds constraints to a portfolio's optimization.

The state of the art in this area of research includes the stochastic discount factor (SDF) approach and the utility-based approach. Ingersoll (2006) was one of the first to use the SDF approach. He used the constrained optimization method of Cvitanić and Karatzas (1992) and Karatzas and Kou (1996) to solve the optimal constrained portfolio of the undiversified executive, identified an appropriate SDF, and derived a closed-form subjective ESO fair price. However, his approach includes two restrictive features. First, the manager is infinitely lived; second, the value of the executives' firm shares and options holdings as a fraction of their total personal wealth must remain constant over a period. In the real world, however, the value of this fraction fluctuates stochastically[1] and the corresponding constraints are stochastic intervals.

A utility-based approach relaxes the restrictive assumptions of Ingersoll (2006) at the cost of not deriving a SDF, hence losing a compact, concise, general pricing framework. Two comprehensive studies in this field are Carpenter, Stanton and Wallace (2009) and Leung and Sircar (2009). In these articles, an executive's goal is to maximize the expected utility of terminal wealth by choosing the optimal exercise time and an optimal trading strategy before and after exercise. Traded assets include all the primary assets and some non-transferable options. Solving the free-boundary problem, the authors obtain the executives' continuation region and the critical stock price boundary, above which the option holders exercise and below which they await. Carpenter, Stanton and Wallace (2009) allow only for a single block exercise of the option, while Leung and Sircar (2009) allow a discrete partial exercise.

Here, we aim to take the merits of both approaches. We use a partial equilibrium approach—constrained portfolio optimization. We show that NTNH constraints break the local

[1] There is no chance, of course, that managers' firm's shares and options holdings value are perfectly correlated with the total value of their personal portfolio.

co-linearity caused by derivative assets in solving portfolio optimization problems. Thus, we are able to translate portfolios that include NTNH derivatives into portfolios that consist of primary assets only by replicating (vanilla) derivatives using primary assets and then integrating the NTNH constraints into a single rectangular constraint. Solving the constrained portfolio optimization problem identifies a SDF that gives the subjective price of any contingent claim written on the primary assets in this portfolio. Furthermore, by using executives' optimal exercise policies, we derive the objective price of ESOs from the firms' perspectives.

We obtain the subjective and objective prices of NTNH American ESOs with block exercise policy and under continuous partial exercise policy. The optimal exercise rate is the one that equates the marginal utility increase, from relaxation of the HTNH constraints induced by reduction in the number of outstanding NTNH ESOs, equal to the marginal utility decrease from loss of option time value. The optimal exercise rate is the one maximizing the expected utility of terminal wealth. However, because of the NTNH constraints, this rate is not the one that maximizes the value of the NTNH ESOs.

We simulate NTNH ESOs and execute a comprehensive comparative statics analysis of prices and optimal exercise policies of NTNH American ESOs with continuous partial exercise. We identify policy implications that enhance the ESOs' incentivizing efficiency.

Accurate ESO pricing is important for several reasons. First, it is a requirement of the Financial Accounting Standard Board (FASB). Though once optional, the reporting of their cost became mandatory in 2004 because of the extensive use of ESOs.[2]

Second, accurate and subjective ESO pricing is essential for understanding the incentives they induce. Understanding this and the sensitivity of ESO values to stock prices allows us to design ESOs by choosing features, such as vesting period, maturity, and strike price to optimally induce incentives.

Third, pricing of NTNH ESOs facilitates solving other pricing problems. This is because many asset pricing problems can be transformed into a NTNH or other non-transferable contingent claims pricing with different constraints. Examples of the problems with non-transferable constraints include asset pricing under transaction costs and valuation of pensions, human capital, or real estate investment [Detemple and Sundaresan (1999)].

Other methods have been used in attempts to price ESOs. The FASB proposes a simple method of ESO pricing, substituting the ESOs' expiration dates with exogenously specified expected time to exercise. Huddart and Lang (1996) conclude that this approach overstates the cost of the ESOs to the firm. The drawback of the FASB approach is that it does not properly reflect executives' optimal ESO exercise times conditional on possible stock price paths. It

[2] Please see Statement of Financial Accounting Standards (SFAS) No. 123 (revised 2004), paragraph B79 and B80. See also International Financial Reporting Standard 3 (IFRS 2), Share-based Payment 3, and Securities Exchange Act of 1934 (as amended through P.L. 111-257, approved October 5, 2010), Sec. 16 (C).

ignores optimal exercise behaviour, for example. Hitting time models extend the FASB model by, first, replacing the exogenously specified expected exercise time with an endogenous first time the stock price hits an (exogenously given deterministic) optimal exercise boundary and, then, aggregating over possible stock price paths (Hull and White (2004) and Cvitanić et al. (2008)). While the exercise boundaries in the above models are independent of model parameters, such as the stocks' drift rates and volatilities, Leung and Sircar (2009) show that optimal exercise boundaries are highly sensitive to these parameters.

Extensions of hitting time models are default models, so named because they share features of credit risk valuation models. These default models define early exercise as the first arrival of an exogenous counting process. Jennergren and Näslund (1993) and Carpenter (1998) use the first jump time of an exogenous Poisson process, which serves as a proxy for executives exercising the options early due to liquidity shocks, diversification needs, voluntary or involuntary job termination, or other events relevant to the executives but not to unrestricted option holders. Carr and Linetsky (2000) extend the default model by making the intensity of the Poisson process depend on the firm's stock price and time.

The above three models—FASB, hitting time, and default—all focus on finding objective prices. They price the ESOs to firms, assuming exogenously given executives' exercise policies, including estimated times to exercise, the exercise boundary, or the exercise process. The price is the expected payoff at the exercise time discounted by the SDF derived from the unconstrained portfolio optimization problem. In contrast, we develop here the executives' subjective fair price of ESOs and identify endogenous optimal exercise policies. We then use these subjective prices under endogenous optimal exercise policies to price ESOs to firms.

The rest of the paper is organized as follows, Section 2 develops a theoretical ESO pricing model and prices European options; Section 3 numerically prices American ESOs under a continuous partial exercise policy; and Section 4 concludes.

2 ESOs pricing model

2.1 Constrained portfolio optimization

Cvitanić and Karatzas (1992) develop a duality technique to solve the constrained consumption/investment problems by transforming the original constrained problems into *auxiliary unconstrained problems*. This approach assumes the support function of the constraint to be bounded below in the duality and existence proofs. However, in ESO pricing problem settings, the aforementioned condition does not hold. To implement the Cvitanić and Karatzas (1992) approach in pricing ESOs, we relax the bounded below condition into a less strict one.

In the market $\mathcal{M}$ there is a traded bond whose price, $S_0(t)$, appreciates at a deterministic risk-free rate of interest, $r(t)$, thus evolving according the differential equation,

$$dS_0(t) = S_0(t)r(t)dt, \;\; S_0(0) = 1. \tag{1}$$

Uncertainty in the market is driven by a d dimensional standard Brownian Motion $W(t) = (W_1(t), \dots, W_d(t))^T$ in $\Re^d$, defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and we denote by $\{\mathcal{F}_t\}$ the $\mathbb{P}$-augmentation of the natural filtration $\mathcal{F}_t^W = \sigma(W(s); 0 \le s \le t)$, with time span $[0, T]$ for some finite $T > 0$.

Primary asset prices $S_i(t), i = 1, \dots, d$ follow the dynamics of

$$dS_i(t) = S_i(t)\left[b_i(t)dt + \sum_{j=1}^{d} \sigma_{ij}(t)dW_j(t)\right], \;\; S_i(0) = s_i, i, j = 1,2, \dots, d. \tag{2}$$

Here $\sigma(t) \triangleq \left(\sigma_{ij}(t)\right)$ is a $d \times d$ volatility matrix, $b(t) \triangleq \left(b_i(t)\right)$, a $d \times 1$ drift rate vector. Let $\mathcal{K} \triangleq \{K(t, \omega); (t, \omega) \in [0, T] \times \Omega\}$ be a family of closed, convex, nonempty subsets of $\mathcal{R}^d$; and let $\delta(v(t)) \equiv \delta\left(v(t, \omega)\middle|K(t, \omega)\right) \triangleq \{\sup_{\varrho \in K}(-\varrho^T v(t)) : \Re^d \to \Re \cup \{+\infty\}\,; (t, \omega) \in [0, T] \times \Omega\}$ be the corresponding family of support functions, where $\widetilde{K}(t, \omega) \triangleq \{v(t) \in \Re^d;\; \delta(v(t)|K) < \infty\}$ is the effective domain of the support function and $v(t) \equiv v(t, \omega) = \left(v_1(t, \omega), \dots, v_d(t, \omega)\right)^T$. Let $\mathcal{H}$ denote the Hilbert space of $\{\mathcal{F}_t\}$-progressively measurable processes v with values in $\Re^d$, and with the inner product $< v_1, v_2 > \triangleq \mathbb{E}\int_0^T (v_1(t))^T v_2(t)dt$. Without loss of generality, we assume that the ESO is written on $S_1(t)$.

Let $U(\cdot)$ be a strictly increasing, strictly concave, of class C^1 utility function, satisfying,

$$U'(0+) \triangleq \lim_{x \downarrow 0} U'(x) = \infty, \;\; U'(\infty) \triangleq \lim_{x \to \infty} U'(x) = 0\,. \tag{3}$$

Let $x_t \in (0, \infty)$ be the initial, at $t = 0$, or realized, at $t > 0$, executive's wealth. Let the vector $\pi(t, \omega)$ represent the proportion of wealth invested, at time t, in each of the primary assets (henceforth, the portfolio process). $\mathcal{A}_0(x_t, t, T)$ is the set of admissible portfolio processes, ensuring that the portfolio wealth process, denoted by $X^{x_t, \pi}(s)$, is finite for $s \in [t, T]$ and satisfies $\mathbb{E}\left[U^-\left(X^{x_t, \pi}(T)\right)|\mathcal{F}_t\right] < \infty$.

We now define an auxiliary market $\mathcal{M}_v$. There, the risk-free rate is $r_v(t) = r(t) + \delta(v(t)), \forall\, t,\; t \in [0, T]$; the vector of drift rates of the primary assets is $b_v(t) = b(t) + v(t) + \delta(v(t))\mathbf{1}$, where $\mathbf{1} \triangleq (1\,1 \dots 1)'$, and $\mathcal{A}_v(x_t, t, T)$ is the admissible set of portfolio processes $\pi_v(t, \omega)$, ensuring that the portfolio wealth in this market, denoted by $X_v^{x_t, \pi_v}(s)$, is finite for $s \in [t, T]$ and satisfies $\mathbb{E}\left[U^-\left(X_v^{x_t, \pi_v}(T)\right)|\mathcal{F}_t\right] < \infty$.

It turns out that in the auxiliary market the constraints are not binding and, thus, can be ignored. In particular, the role of the vector $v(t)$ is to modify the original market, changing the expected returns of some assets in order to make them more (or less) attractive. So, for example, if one asset has a short-sale constraint, then $v(t)$ would increase the expected return on that asset until it is not optimal to short the asset. Proposition 1 below (see also Cvitanić and

Karatzas (1992), Proposition 8.3) then gives conditions under which the optimal unconstrained portfolio in the auxiliary market is also the optimal constrained portfolio in the original market. Note that the total adjustment to the optimal portfolio expected return in the auxiliary market is zero.

The *constrained optimization problem* in the original market $\mathcal{M}$ is to maximize, for all $0 < x < \infty$, the derived utility

$$J(x_t; \pi) \triangleq \mathbb{E}\left[U\left(X^{x_t,\pi}(T)\right) | \mathcal{F}_t\right], \qquad \forall\, t \in [0, T] \tag{4}$$

by choosing π over the class (here, ℓ is Lebesgue measure).

$$\mathcal{A}(x_t,, K(t,\omega), t, T) \triangleq \{\pi \in \mathcal{A}_0(x_t, t\, T);\ \pi(t,\omega) \in K(t,\omega)\ for\ \ell \otimes \mathbb{P} - a.e.(t,\omega)\}. \tag{5}$$

The unconstrained problem in the auxiliary market $\mathcal{M}_v$ (the *auxiliary unconstrained problem*) is to maximize, for all $0 < x < \infty$,

$$J_v(x; \pi) \triangleq \mathbb{E}\left[U\left(X_v^{x_t,\pi_v}(T)\right) | \mathcal{F}_t\right] \tag{6}$$

by choosing π over the class $\mathcal{A}_v(x_t, t, T)$.

The *dual problem of the auxiliary unconstrained problem* is to minimize

$$\tilde{J}(y_t; v) \triangleq \mathbb{E}\left[\tilde{U}\left(y_t H_v(T)\right) | \mathcal{F}_t\right] \tag{7}$$

by choosing v over the class

$$\mathfrak{D}_t = \ \mathfrak{D}_t(K) \triangleq \left\{v \in \mathcal{H};\ \mathbb{E}\int_t^T \delta\left(v(s)\right)ds < \infty\right\}, \tag{8}$$

where $\tilde{U}(y) \triangleq max_{x>0}[U(x) - xy] = U(I(y)) - yI(y)$, $I(\cdot)$ is the inverse function of $U'(\cdot)$, $H_v(t) \triangleq exp\left\{-\int_0^t r_v(s)\right\} exp\left\{-\int_0^t \theta_v(s)dW(s) - \frac{1}{2}\int_0^t \|\theta_v(s)\|^2 ds\right\}$, is the stochastic discount factor (SDF) in the auxiliary market $\mathcal{M}_v$, and $\theta_v(s) \triangleq \sigma^{-1}(t)[b_v(t) - r_v(t)]$. If we define $\gamma_v(t) \triangleq exp\left\{-\int_0^t r_v(s)ds\right\}$ (sometimes known as the reciprocal of the money market account) and $Z_v(t) \triangleq exp\left\{-\int_0^t \theta_v(s)dW(s) - \frac{1}{2}\int_0^t \|\theta_v(s)\|^2 ds\right\}$, then $H_v(t) = \gamma_v(t) \times Z_v(t)$. If $v = 0$, then $H_0(T) \triangleq H_v(T)|_{v=0}$ is the SDF in the original market $\mathcal{M}$. $y_t \equiv \mathcal{Y}_v(x_t)$ is the inverse function of $x_t = \mathcal{X}_v(y_t) \triangleq \mathbb{E}\left[H_v(T) I\left(y_t H_v(T)\right) | \mathcal{F}_t\right]$. We call the optimal solution $\hat{v}$; thus,

$$\hat{v} \triangleq \underset{v \in \mathfrak{D}_t}{\operatorname{argmin}}\, \tilde{J}(y_t; v) = \underset{v \in \mathfrak{D}_t}{\operatorname{argmin}}\, \mathbb{E}\left[\tilde{U}\left(y_t H_v(T)\right) | \mathcal{F}_t\right]. \tag{9}$$

As Cvitanić and Karatzas (1992) point out (p. 768), the dual approach "is of great importance here…because, as it turns out, it is far easier to prove existence of optimal policies in the dual, rather than in the primal, problem." We assume that the following conditions from Cvitanić and Karatzas (1992) are satisfied. The effective domain of the support function is the same for all (t, ω); that is,

$$\widetilde{K}(t,\omega) \equiv \widetilde{K}. \tag{10}$$

$$\delta\big(v(t)\big|K(t,\omega)\big), 0 \le t \le T, \text{ is } \{\mathcal{F}_t\} \text{-progressively measurable.} \tag{11}$$

$$\delta(v(t)|K(t,\omega)) \text{ is continuous on } \widetilde{K} \text{ for all } (t,\omega). \tag{12}$$

$$\text{For some } \alpha \in (0,1), \gamma \in (1,\infty), \text{ we have } \alpha U'(X) \ge U'(\gamma X), \forall X \in (0,\infty). \tag{13}$$

$$X \to XU'(X) \text{ is non-decreasing on } (0,\infty). \tag{14}$$

$$U(\infty) = \infty. \tag{15}$$

$$U(0_+) > -\infty. \tag{16}$$

$$\forall y \in (0,\infty), \exists v \in \mathfrak{D} \text{ such that } \tilde{J}(y;v) < \infty. \tag{17}$$

We are now ready to state Proposition 1.

Proposition 1. For every $\{\mathcal{F}_t\}-$progressively measurable process $v(t) \equiv v(t,\omega)$, $0 \le t \le T$, such that $v(t) \in \widetilde{K}(t,\omega)$ for $\ell \otimes \mathbb{P} - a.e. (t,\omega)$, assume that the following conditions hold:

$$\forall y \in (0,\infty), \lim_{\|v\|\to\infty} \tilde{J}(y;v) = \infty \tag{18}$$

and

$$\forall v \in (0,\infty), \lim_{y\downarrow 0} \tilde{J}(y;v) = \infty. \tag{19}$$

Then, when $\pi_{\hat{v}}(t,\omega) \in K(t,\omega)$ and $\delta\big(\hat{v}(t,\omega)\big) + \big(\pi_{\hat{v}}(t,\omega)\big)^T \hat{v}(t,\omega) = 0$ are satisfied, $\pi_{\hat{v}}(t,\omega)$ maximizes the *constrained optimization problem* in the original market $\mathcal{M}$ and the *auxiliary unconstrained problem* in auxiliary market $\mathcal{M}_{\hat{v}}$. Also, $\hat{v}(t,\omega)$ minimizes the *dual problem of the auxiliary unconstrained problem* in the auxiliary market $\mathcal{M}_{\hat{v}}$.

Proof. The proof directly follows from Theorem 7.4, Proposition 8.3, Proposition 12.1, Proposition 12.2, Theorem 12.4, Theorem 13.1 and Proposition 13.2 in Cvitanić and Karatzas (1992).

Remark. If $K(t,\omega)$ contains the origin, then the support function of the constraint is bounded below and, clearly, Conditions (18) and (19) are satisfied. However, the NTNH constraints imply that $K(t,\omega)$ does not contain the origin. The existence of a minimum to Equation (7) can still be guaranteed, however, if these two conditions are first imposed then verified.

2.2 NTNH European ESO pricing

There are three types of prices associated with ESOs: the *market price*, the *subjective price,* and the *objective price*.

Market price ($BS(t), 0 \le t \le T$) – the Black-Scholes price without taking the NTNH features into account. We call the corresponding SDF the *market SDF*.

Subjective price ($\hat{p}(t), 0 \le t \le T$) – the value of the ESO to its grantee, under optimal exercise while fully considering the NTNH features. Because of the NTNH constraints, the corresponding subjective SDF is different from the market SDF, and the subjective ESO price is lower than the market one.

Objective price ($\tilde{p}(t), 0 \leq t \leq T$) – the ESO cost for firms to fulfill their obligation upon exercise. As NTNH constraints do not apply to firms, the objective SDF and the market one are same; on the other hand, the objective and subjective prices represent the same payoff. For European ESOs, the market price and objective price are the same. For American ESOs, the objective price is lower than the market price due to executives' suboptimal (early) exercise. The objective price is higher than the subjective price because firms can replicate the option without the NTNH constraint. Thus, the objective price falls between the market price and subjective price. The difference between the objective and subjective prices is the amount paid by firms but not appreciated by executives. We call it the *dead weight cost*. It measures the efficiency loss to firms in exchange for incentivising their executives.

We now introduce terminology describing the constraints. We call an ESO n*on-transferable* if its owner is forbidden from transferring it to a third party, at any time, under any price. We call an ESO *non-hedgeable* if its owner is forbidden from selling shares of the ESO issuer stock that they do not own.

We assume that executives' portfolios consist of n shares of NTNH ESOs with exercise price k and maturity T,[3] and a sub-portfolio composed of primary assets only. We name the sub-portfolio *outside wealth* and denote its initial value by x_s and its value process by $X(t), t \geq s$. We denote the initial value of *total wealth* as $x_t^n \equiv x_t + nBS(t)$ and the value process of total wealth by $\mathbb{X}(t)$. We denote $x \equiv x_0$ and $x^n \equiv x_0^n$.

In this section, we focus on pricing European NTNH ESOs, ignoring early exercise. The idea behind our pricing of NTNH contingent claims is replacing the n shares of non-transferable contingent claim with a replicating portfolio composed of primary assets only. Because markets are complete, we can see firms as having two equivalent choices. One is giving executives n option payoffs at time T, which we denote $nB(T)$; the other is giving executives, at time 0, a cash amount of $nBS(0)$ and making them use this sum to replicate (using a self-financing strategy) the option payoffs $nB(T)$ at time T. Under the latter interpretation, because the ESO cannot be hedged, their position in S_1 can never be less than $n\phi(d_1)$ shares[4] of S_1. This puts the problem of pricing NTNH options in the framework of a constrained portfolio optimization problem. Holding n shares of a non-transferable European option is equivalent to holding $n\phi(d_1)$ shares of firm stock and $-n\phi(d_2)\mathrm{e}^{-r(T-t)}k$ dollars of savings account, where k is the strike price of the

[3] For simplicity and brevity, we often depress dependency on parameters such as T.

[4] $\phi(\cdot)$ is the cumulative density function of the standard normal distribution $d_1 = \frac{1}{\sigma\sqrt{T-t}}\left[ln\left(\frac{S}{k}\right) + \left(r + \frac{\sigma^2}{2}\right)(T - t)\right]$, $d_2 = d_1 - \sigma\sqrt{T-t}$.

ESOs. Recalling that under the non-hedgeable constraint, the number of shares of the firm's stock in the total portfolio should be greater than or equal to $n\phi(d_1)$ shares, the combined constraint for the portfolio process can be written as follows:

$$\forall t, \pi(t,\omega) \in K(t,\omega) = \left[\frac{n\phi(d_1)S_1(t)}{\mathbb{X}(t)}, \infty\right) \times (-\infty, \infty)^{d-1}, \tag{20}$$

where $S_1(t)$ is the firm's stock price underlying the ESO. Then, we identify $\widetilde{K}$ and $\delta(v(t))$ as

$$\widetilde{K} = [0,\infty) \times \{0\}^{d-1}, \tag{21}$$

$$\delta(v(t)) \equiv -\frac{n\phi(d_1)S_1(t)}{\mathbb{X}(t)} \times v_1(t) \text{ on } \widetilde{K}, \tag{22}$$

which also confirms the fulfillment of Conditions (10)-(12). We assume $v(t) \in \mathfrak{D}_t$, then $\mathbb{E}\int_0^T \delta\big(v(t)\big|K(t,\omega)\big)dt = \mathbb{E}\int_0^T [-\frac{n\phi(d_1)S_1(t)}{\mathbb{X}(t)} \times v_1(t)]dt < \infty, \forall v(t) \in \widetilde{K}$.

We note that our use of the Black-Scholes options pricing formula implies that the volatility of the underlying asset is deterministic.

Assumption 1. $\forall t \in [0,T]$, and $\forall\, x_t \in (0,\infty)$, the indirect utility function, $M(x_t,t)$, is

$$M(x_t,t) \triangleq \underset{\pi \in \mathcal{A}(x_t, K(t,\omega),t,T)}{esssup} \mathbb{E}\Big[\Big(U\big(X^{x_t,\pi}(T)\big)\Big|\mathcal{F}_t\Big)\Big]. \tag{23}$$

We denote the solution of Equation (23) under the constraint $K(t,\omega)$ of Equation (20) as π^*. We assume $M'(\cdot,t) > 0$ on $(0,\infty)$.

We have now developed a sufficient mathematical structure to identify, in the following theorem, the subjective price of a NTNH European ESO, characterizing the SDF as the real-world marginal rate of substitution.

Theorem 1. Under Assumption 1, the time t subjective price of NTNH *European ESOs* with terminal payoff $B(T)$ is uniquely determined by

$$\hat{p}(t) \equiv \hat{p}(x_t,n,t) = \frac{\mathbb{E}\left[B(T)U'\left(X^{x_t,\pi^*}(T) + nB(T)\right)|\mathcal{F}_t\right]}{M'(x_t+nBS(t),t)}, \forall x_t > 0, \tag{24}$$

where $X^{x_t,\pi^*}(T)$ represents the optimal outside wealth for the constrained problem.

Proof. See Appendix A.

Theorem 2. i) Under Assumption 1, the time t subjective price of NTNH *European ESOs* with terminal payoff $B(T)$ is uniquely determined by

$$\hat{p}(t) \equiv \hat{p}(x_t,n,t) = \frac{\mathbb{E}[H_{\hat{v}}(T)B(T)|\mathcal{F}_t]}{H_{\hat{v}}(t)} = \frac{\mathbb{E}^{\hat{v}}[\gamma_{\hat{v}}(T)B(T)|\mathcal{F}_t]}{\gamma_{\hat{v}}(t)}, \tag{25}$$

where $\mathbb{E}^{\hat{v}}(\cdot)$ is an expectation under the martingale measure in the auxiliary market $\mathcal{M}_{\hat{v}}$, for $\hat{v}$ as defined in Equation (9).

ii) The time t objective price of this NTNH European ESO is equal to the market price:

$$\tilde{p}(t) \equiv \tilde{p}(t) = BS(t) = \frac{\mathbb{E}[H_0(T)B(T)|\mathcal{F}_t]}{H_0(t)} = \frac{\mathbb{E}^0[\gamma_0(T)B(T)|\mathcal{F}_t]}{\gamma_0(t)}, \tag{26}$$

where $\mathbb{E}^0(\cdot)$ is expectation under the martingale measure in the auxiliary market $\mathcal{M}_0$ (where $\hat{v} = 0$).

Proof. See Appendix A.

Theorem 2 identifies the auxiliary markets $\mathcal{M}_{\hat{v}}$ and $\mathcal{M}_0$ as the corresponding "risk-neutral markets" to constrained and unconstrained pricing in the original market, respectively.

The number of ESO shares granted affects the subjective price in two ways. First, by affecting the initial total wealth. [5] Second, as n increases the constraint $K(t,\omega) = [\frac{n\phi(d_1)S_1(t)}{\mathbb{X}(t)}, \infty) \times (-\infty,\infty)^{d-1}$ becomes more binding, inducing a "more aggressive" SDF. In the following subsection, we show the effect of this dependence of the subjective price on the number of ESO shares and on optimal exercise policies for American ESOs (block exercise and a continuum of exercise ratios), which, in turn, affect the outstanding number of ESO shares, affecting ESO prices, and so on.

2.3 NTNH American ESO pricing with block exercise policy

The ESOs' exercise time τ, τ: $(\Omega, \mathcal{F}, \mathbb{P}) \to [0,T]$, is a stopping time with respect to the filtration $\{\mathcal{F}_t\}$. Let Γ be the class of $\{\mathcal{F}_t\}$ stopping times. Before τ, the NTNH constraint is $K(t,\omega) = [n\phi(d_1)S_1(t)/\mathbb{X}(t), \infty) \times (-\infty,\infty)^{d-1}$, and we denote the corresponding SDF $H_f(t)$. Because of the executives' general restriction from short selling their firm's stock, even after τ, although there are no NTNH ESOs in the portfolio anymore, the portfolio cannot have negative weights in the firm's stock. Hence, the constraint becomes $K(t,\omega) = [0,\infty) \times (-\infty,\infty)^{d-1}$, and we denote the corresponding SDF $H_\lambda(t)$. The post-exercise total wealth at τ is the sum of outside wealth $X^{x_t,\pi}(\tau)$ and the intrinsic value of ESOs $nB(\tau)$. Hence,

$$K(t,\omega) = \left[\frac{n\phi(d_1)S_1(t)}{\mathbb{X}(t)}, \infty\right) \times (-\infty,\infty)^{d-1} 1_{\{t<\tau\}} + [0,\infty) \times (-\infty,\infty)^{d-1} 1_{\{t\geq\tau\}}. \tag{27}$$

We have now developed sufficient mathematical structure to identify, in the following theorem, the subjective price of a NTNH American ESO under block exercise policy, characterizing the SDF as the real-world marginal rate of substitution.

Theorem 3. Under a block exercise policy, the subjective price of a NTNH American ESO is uniquely determined by

$$\hat{p}(t) \equiv \hat{p}(x_t,n,t) = \frac{\mathbb{E}\left[M'\left(X^{x_t,\pi^*}(\tau^*) + nB(\tau^*), \tau^*\right)B(\tau^*)|\mathcal{F}_t\right]}{L'(x_t,n,t)}, \quad \forall x_t > 0, \tag{28}$$

where $\pi^* = \pi^*(s,\omega)$ is the solution to Problem (30) below for $s \in (t,\tau^*)$ but is the solution to

[5] If the constraint is a closed, convex cone or the utility is logarithmic, then according to Cvitanić and Karatzas (1992), the initial wealth does not appear in the pricing formula. The non-hedgeability constraint is a closed, convex cone; however, the NTNH constraint is not.

Problem (29) below for $s \in (\tau, T)$, and $K(t, \omega)$ is as in Equation (27):

$$M(x_\tau, \tau) \triangleq \underset{\pi \in \mathcal{A}(x_\tau, K(\tau,\omega), \tau, T)}{esssup} \mathbb{E}\left[\left(U\left(X^{x_\tau,\pi}(T)\right)\middle|\mathcal{F}_\tau\right)\right], \tag{29}$$

and

$$L(x_t, n, t) \triangleq \underset{\tau \in \Gamma}{esssup} \underset{\pi \in \mathcal{A}(x_t, K(t,\omega), t, \tau)}{esssup} \mathbb{E}[M(X^{x_t,\pi}(\tau) + nB(\tau), \tau)|\mathcal{F}_t]\,. \tag{30}$$

Also, τ^* is the optimal exercise time; that is, the optimal stopping time in Problem (30).

Proof. See Appendix B.

We are now ready to develop a martingale pricing expression of a NTNH American ESO under block exercise.

Theorem 4. The subjective and objective prices of a NTNH American ESO under block exercise are uniquely determined as follows.

i) The subjective price is

$$\hat{p}(t) \equiv \hat{p}(x_t, n, t) = \frac{\mathbb{E}[H_f(\tau^*)B(\tau^*)|\mathcal{F}_t]}{H_f(t)} = \frac{\mathbb{E}^f[\gamma_f(\tau^*)B(\tau^*)|\mathcal{F}_t]}{\gamma_f(t)}, \tag{31}$$

where τ^* is the optimal exercise time identified by the optimal stopping time, Problem (30), and f minimizes the dual problem $\tilde{J}(y_t; v) \triangleq \mathbb{E}\left[\tilde{U}\left(y_t H_v(\tau^*)\right)\right]$ over $\mathfrak{D}_t$ with $K(t, \omega)$ as in Equation (27) for $t < \tau^*$.

ii) The objective price of this NTNH American ESO is

$$\tilde{p}(t) \equiv \tilde{p}(x_t, n, t) = \frac{\mathbb{E}[H_0(\tau^*)B(\tau^*)|\mathcal{F}_t]}{H_0(t)} = \frac{\mathbb{E}^0[\gamma_0(\tau^*)B(\tau^*)|\mathcal{F}_t]}{\gamma_0(t)}. \tag{32}$$

Proof. See Appendix B.

Remark. Equation (31) implies that the subjective price of a NTNH American ESO under block exercise is determined by $H_f(\tau^*)$, the SDF before the optimal exercise time τ^* with constraint $K(t, \omega) = [n\phi(d_1)S_1(t)/\mathbb{X}(t), \infty) \times (-\infty, \infty)^{d-1}$. In contrast, because optimal stopping times are determined backward, τ^* is determined by $H_\lambda(t)$, the SDF after τ^*. Hence, the subjective price $\hat{p}(x_t, n, t)$ is determined by both the auxiliary market $\mathcal{M}_f$ before τ^* and the auxiliary market $\mathcal{M}_\lambda$ after τ^*. Thus, these two determinants are time separable.

Remark. Note that the subjective price $\hat{p}(x_t, n, t)$ is not equal to $\mathbb{E}[H_l(T)B(T)|\mathcal{F}_t]/H_l(t)$, where the auxiliary market $\mathcal{M}_l$ is defined as the one that encompasses both auxiliary markets $\mathcal{M}_f$ (before optimal exercise) and $\mathcal{M}_\lambda$ (after optimal exercise). The equality does not hold because it uses the terminal payoff rather than the one at exercise.

Remark. Further note that even after improving the previous pricing method, the subjective price $\hat{p}(x_t, n, t)$ is not equal to $esssup_\tau\left\{\mathbb{E}\left[H_f(\tau)B(\tau)|\mathcal{F}_t\right]\right\}/H_f(t)$. This is because executives with NTNH American ESOs do not simply maximize the subjective prices of their options. Under NTNH constraints, the optimal ESO exercise policy can be obtained only by maximizing the expected utility of the terminal wealth. As optimal exercise under NTNH constraints is a function of the executive's wealth relative to outside wealth, separation (between maximizing the option value and maximizing the expected utility) is lost. With no constraints, however, separation exists; thus, maximizing the subjective price of the options and maximizing the expected utility of total terminal wealth are equivalent.

2.4 NTNH American ESO pricing with continuous partial exercise policy

Leung and Sircar (2009) extended the block exercise policy into a discrete partial exercise one. We now determine the optimal exercise policy under the most general conditions: executives may exercise continuously in time and chose exercise rates from a continuum. We call this exercise policy *continuous partial exercise.*

Under continuous partial exercise, the dynamics of the total wealth process is

$$\begin{gathered} d\mathbb{X}(t) = \big(\mathbb{X}(t)r(t)\big)dt + \mathbb{X}(t)\big(\pi(t)\big)^T\sigma(t)dW_0(t) - \dot{n}(t)\big(BS(t) - B(t)\big)dt, \\ \mathbb{X}(0) = x + nBS(0), \qquad \dot{n}(t) \in \mathcal{N} \end{gathered} \tag{33}$$

where $W_0(t) \equiv W(t) + \int_0^t \theta(s)\,ds$, $\theta(t) \equiv \sigma^{-1}(t) \times [b(t) - r(t)\mathbf{1}]$, and $\dot{n}(t) \equiv \dot{n}(t,\omega)$ is the time t optimal exercise rate that maximizes the expected utility of terminal wealth. The constant n is the number of ESO shares granted at time zero. The control variable, the random exercise rate process, is denoted by $\dot{n}(t)$. It belongs to the set $\mathcal{N}(t) \triangleq \{\dot{n}(t)\colon \int_0^t \dot{n}(s)ds \leq n\}$, the collection of all feasible exercise rates. We define $\widehat{N}(t) \triangleq \int_0^t \dot{n}(s)ds$ to be the accumulated number of ESOs that have been exercised at time t. The constraint of the portfolio process is

$$\pi(t) \in K(t,\omega) = \big[\big(n - \widehat{N}(t)\big)\phi(d_1)S_1(t)/\mathbb{X}(t), \infty\big) \times (-\infty,\infty)^{d-1}. \tag{34}$$

The derived utility of an executive maximizing expected utility of terminal wealth by choosing optimal portfolio and exercise rate processes is

$$\mathbb{J}\big(X^{*,\dot{n}^*}(t), \widehat{N}(t), S(t), t\big) = esssup_{(\pi,\dot{n})\in(\mathcal{A}(x_t,K(t,\omega),t,T),\mathcal{N})}\, \mathbb{E}\big[U\big(\mathbb{X}^{x_t,\pi}(T)\big)|\mathcal{F}_t\big], \tag{35}$$

where $K(t,\omega)$ is described in Equation (34).

For a given $\dot{n}(t)$, we know how to solve the portfolio optimization problem by employing the duality technique. As the indirect utility function is a monotone increasing function of initial wealth, the time t optimal portfolio process generates the ex-ante optimal initial wealth for the next period $t + \Delta t$; otherwise, the existence of another portfolio process, which leads to higher wealth, would contradict optimality. Hence, we can simplify the optimization problem by separating it into the following two stages:

$$\mathbb{J}\big(\mathbb{X}^{*,\dot{n}^*}(t), \widehat{N}(t), S(t), t\big) = \underset{\dot{n}\in\mathcal{N}(t)}{esssup}\ \underset{\pi\in\mathcal{A}(x_t,K(t,\omega),t,T)}{esssup}\ \mathbb{E}\big[U\big(\mathbb{X}^{x,\pi,\dot{n}}(T)\big)|\mathcal{F}_t\big], \tag{36}$$

where $\mathbb{X}^{*,\dot{n}}(t) \equiv \mathbb{X}^{x,\pi^*,\dot{n}}(t)$ is the total wealth process generated by the optimal portfolio process π^* for a given initial wealth x and under an exercise rate process $\dot{n}(t)$.

The following proposition identifies the first-order condition (FOC) necessary to determine the optimal exercise process.

Proposition 2. Assume the firm's stock pays no dividend. The FOC of the (viscosity) solution to the optimization problem in Equation (36) is

$$-\big(BS(t)-B(t)\big)\frac{\partial \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t),\widehat{N}(t),S(t),t)}{\partial \mathbb{X}^{*,\dot{n}}(t)}+\frac{\partial \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t),\widehat{N}(t),S(t),t)}{\partial \widehat{N}(t)}=0. \tag{37}$$

Proof. See Appendix C.

It is well known that under a non-negative interest rate, an American call option written on a non-dividend-paying stock, under constant strike price, should never be exercised before its expiry date (see, e.g., Rutkowski 2009). In other words, under usual circumstances, an American call option price is always greater than or equal to its intrinsic value. We will now explain why early exercise might become optimal. As executives exercise the NTNH ESOs, the newly received cash is subject *only* to short sale constraints of the firm's stock, and the other portfolio constraints are relaxed. Updating the corresponding SDF and auxiliary market definitions induces appropriate pricing. The optimal exercise rate sets the marginal increase of subjective value due to the constraint relaxation to exactly offset the marginal loss of opportunity value due to early exercise.

We are now ready to develop a martingale pricing expression of an NTNH American ESO under optimal continuous exercise using a continuum of exercise ratios.

Theorem 5. i) The subjective price of the NTNH *American ESO* with continuous partial exercise policy is

$$\hat{p}(t)\equiv\hat{p}(x_t,t,T)=\mathbb{E}\left[\int_t^T H_{f_{\widehat{N}^*}}(s)\dot{n}^*(s)B(s)ds\,|\mathcal{F}_t\right], \tag{38}$$

where $\dot{n}^*(t)$ solves Equation (37), and $v=f_{\widehat{N}^*}(s)$ minimizes dual problem (7), under $K(t,\omega)$ described in Equation (34).

ii) The objective price of this NTNH American ESO is

$$\tilde{p}(t)\equiv\tilde{p}(x_t,t,T)=\mathbb{E}\left[\int_t^T H_0(s)\dot{n}^*(s)B(s)ds\,|\mathcal{F}_t\right]. \tag{39}$$

Proof. This follows immediately from the definition of $H_{f_{\widehat{N}^*}}(s)$ and Theorem 2.

Remark. Our results, obviously, depend on π^*, and $\dot{n}^*$. While solving for π^*, we are able to use methods developed earlier (see Proposition 1). We now have to develop methods to solve for $\dot{n}^*$. Note that while both loss of option time value due to early exercise and an underlying stock paying dividend are wealth leakages to option holders, they are essentially different. To compensate for the latter, it is sufficient to deduct the present value of all future dividends from the stock price without affecting the SDF. However, exercising NTNH ESOs affects the portfolio process constraint at the time and affects the corresponding SDF. Thus, we cannot model negative jumps due to the loss of option time value the way we model negative jumps in stock prices due to dividends.

At any time t, the indirect utility is the conditional expectation of the utility generated by the terminal total wealth under the optimal exercise rate process and optimal portfolio process. The common approach is to solve for the optimal controls backward. However, here, total wealth is determined by d geometric Brownian motions requiring many branches of the multinomial tree. In addition, the accumulated number of the firm's stock shares $\widehat{N}(t)$ is another state variable. This makes the complexity of the problem computationally infeasible.

Moreover, working backward, at any time t, the time t total wealth is determined by the optimal portfolio process, $\pi^*(s)$, $0 \leq s \leq t$. Hence, using a binomial tree cannot solve the problem. Moreover, the FOC (37) is determined not only by the information given in the set $\{\mathcal{F}_s, s \in [0, t]\}$ but also by the future optimal portfolio process, which is determined by $\{\mathcal{F}_s, s \in [t, T]\}$. Therefore, the forward-looking Monte Carlo simulation cannot solve the problem either.

Under logarithmic preferences, however, the substitution and income effects offset each other, mathematically implying that the cross partial derivatives of the derived utility function are zero. Now solving the FOC (37)does not require future optimal knowledge of the optimal portfolio process π^*, in another manifestation of the so called myopia of logarithmic preferences. In this case, we can, thus, use Monte Carlo simulation to solve for $\dot{n}^*$, satisfying FOC (37).

3 Simulation experiment

In this section, we show that under logarithmic preferences, Conditions (18) and (19) are satisfied, and the FOC (37) is fully determined by the information available up to time t, not depending on future optimization outcomes. Thus, the forward-looking Monte Carlo simulation becomes useful. Below we describe the simulation method, present the numerical results of NTNH ESO pricing, and perform a sensitivity analysis.

3.1 Solving for the optimal exercise rate

Executives are granted n_{option} shares of unvested NTNH ESOs and n_{stock} shares of unvested (restricted) stocks. The ESOs and the stocks vest on future dates t_{vo} and t_{vs}, respectively. For any $t \geq t_{vo}$, we call the total wealth before exercise $pre\_\mathbb{X}(t)$. After solving FOC (37) for the optimal exercise rate $\dot{n}^*(t)$, we can write the total wealth after exercise as $post\_\mathbb{X}(t) = pre\_\mathbb{X}(t) - \dot{n}(t)dt(BS(t) - B(t))$, and the accumulated number of shares from exercised ESOs as $\widehat{N}(t)$. Upon exercise, the constraint $K(t, \omega)$ changes and, in turn, induces changes in $f_{\widehat{N}}(t)$, the optimal control of the dual problem, Equation (36). The updated $f_{\widehat{N}}(t)$ determines the corresponding SDF in the original market and the drift rate adjustments that convert the original market price processes to auxiliary market ones. We can now use $f_{\widehat{N}}(t)$ to solve for the new, after exercise, value of the optimal portfolio process, $\pi^*(t)$, making the appropriate primary asset weight adjustments. The changes in the primary asset prices, from $S(t)$ to $S(t+1)$), affect next period's wealth $pre\_\mathbb{X}(t+1)$, which then is adjusted to $post\_\mathbb{X}(t+1)$,after time $t+1$ exercise, and then further adjusted to $\pi^*(t+1)$, and so on. Figure 1 below describes this process. We calculate the following values for each time point: pre-exercise total wealth, $pre\_\mathbb{X}(t)$, optimal exercise rate, $\dot{n}^*(t)$, post-exercise total wealth, $post\_\mathbb{X}(t)$, updated constraints, $\widehat{N}(t-1), f_{\widehat{N}}(t-1)$, SDF, and the optimal portfolio process, $\pi^*(t)$.

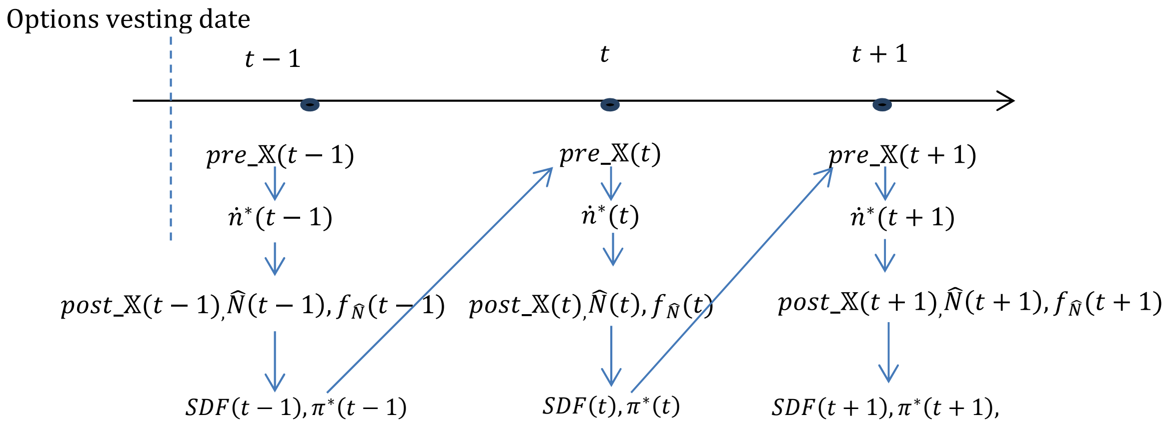


Figure 1. Sample path generating process for Monte Carlo simulation

The following Lemma states that under logarithmic preferences Conditions (18) and (19) are satisfied.

Lemma 1. Given $K(t,\omega) = [a,\infty) \times (-\infty,\infty)^{d-1}$, where $a \equiv a(t,\omega) > 0$, if $U(\cdot) = ln(\cdot)$, then for any $y \in (0,\infty)$, we have, $\lim_{\|v\|\to\infty} \tilde{J}(y;v) = \infty$, and for any $v \in (0,\infty)$, $\lim_{y\downarrow 0} \tilde{J}(y;v) = \infty$.

Proof. See Appendix D.

Finally, the proposition below identifies the logarithmic preferences case of FOC (37) that we use to solve for the optimal exercise rate, $\dot{n}^*(t)$.

Proposition 3. Under logarithmic preferences $U(\cdot) = ln(\cdot)$, which have an Arrow-Pratt relative risk-aversion coefficient of 1, the FOC (37) with respect to the optimal continuous partial exercise rate $\dot{n}(t)$ becomes

$$-\frac{BS(t)-B(t)}{pre\_\mathbb{X}(t)} + ln\big(post\_\mathbb{X}(t)\big) \times \frac{\partial\left(-\left(\frac{\left(n_{option}-\widehat{N}(t)\right)N\left(d_1(t)\right)S_1(t)+n_{stock}1_{t<t_{vs}}S_1(t)}{post\_\mathbb{X}(t)}\right)f_{\widehat{N}}(t)+\frac{1}{2}\left\|\theta_{f_{\widehat{N}}}(t)\right\|^2\right)}{\partial\widehat{N}(t)} = 0, \tag{40}$$

where $f_{\widehat{N}}(t) = \text{argmin}_{v\in\tilde{K}(t,\omega)}\left[-2\frac{(n-\widehat{N}(t))N(d_1)S_1(t)}{pre\_\mathbb{X}(t)}v + \|\sigma^{-1}(t)\times[b(t)-r(t)\mathbf{1}+v]\|^2\right]$,

$\theta_v(t) = \sigma^{-1}(t) \times [b(t) - r(t)\mathbf{1} + v(t)]$, and $\|\cdot\|$ is the Euclidean norm.

Proof. See Appendix E**.**

3.2 Simulation result and sensitivity analysis

In this section, we explicitly solve a family of examples of NTNH American call option prices under continuous partial exercise policies. We assume that executives' outside wealth in addition to the firm's restricted stocks is the market index. We compare our results with vanilla Black-Scholes prices and study price sensitivity to spot prices, interest rate, volatility, maturity, instantaneous correlation between primary assets' price processes, drift rates, vesting periods, and the initial endowments of unvested stocks, NTNH American ESOs, and cash. We define "incentives" as the delta of subjective price $\hat{p}(t)$, i.e., the change in subjective price per unit of share price change, $\partial\hat{p}(t)/\partial S_1(t)$ (see Table 1). We define the deadweight cost as the discrepancy between objective price and subjective price $\tilde{p}(t) - \hat{p}(t)$ (see Table 2). And we define "efficiency" as the ratio of incentives to deadweight cost, $\frac{\partial\hat{p}(t)/\partial S_1(t)}{\tilde{p}(t)-\hat{p}(t)}$ (see Table 2).

The intuition behind the term incentives is straightforward, as it is the sensitivity of the value to the executive with respect to the share price. We call the difference between the objective price and the subjective one "deadweight cost," as it is a price that the firm pays but the executive does not benefit from. The term "efficiency" is intuitively appealing as it increases in incentives, decreases in deadweight cost, and gives the ratio of incentives per unit of deadweight cost.

Firms will be interested in the following comparative statics.

1. Figures 2(a,c) demonstrate that NTNH ESOs induce strong incentives to improve performance. Not only do NTNH ESOs' subjective prices increase as a function of stock prices, they also increase in the stock drift rates, forming a positive feedback cycle and further intensifying incentives.
2. The objective price is always greater than the subjective price, and both prices move in the same direction when parameters change. See Table 2 and Figure 2.
3. Longer maturity adds value to both subjective and objective prices. See Figure 2(f).
4. Subjective and objective NTNH ESO prices decrease as the options become a larger proportion of their endowments of total initial wealth (of options, stocks, and cash). See Figure 2(m) to Figure 2(o).
5. Subjective and objective NTNH ESO prices are also affected by the drift rate and volatility of the market index. As the index's drift rate increases, the subjective and objective prices decrease. Similarly, as the index's volatility decreases, these prices decrease. See Table 2. Intuitively, as the index becomes more attractive, i.e., as its drift rate increases or its volatility decreases (all else being equal), the position in the ESO becomes less desirable. In

fact, for our parameter values, the subjective price is actually more sensitive to the volatility of the index than it is to the volatility of the stock. See Figure 2(d,h).

6. As the correlation between returns on the underlying stock and returns on the market index increases, the subjective and objective prices increase. For example, as the correlation approaches one, the call option on the stock effectively becomes a call option on the market index, which would be quite desirable. On the other hand, when the correlation is low, the ESO gives the executive a concentration of wealth in an asset that is "far from" the optimal market index. See Figure 2(b).
7. We argue that if the subjective price of the ESO increases (all else being equal), then the firm can grant fewer ESOs to the executives. So, for example, following the insights of the previous two points, if there is a decrease in the drift rate of the market index, then the subjective price increases, and the firms can grant fewer ESOs. Similar conclusions hold for the other variables.
8. Efficiency decreases in vesting period, maturity, stock volatility, and the proportion of the executive's endowment of total wealth. See Table 2.
9. The ESO's strike price has little effect on the option's efficiency.

Executives as well as firms will be interested in the following comparative statics.

10. Executives, on average, exercise more shares of deeper in-the-money options. See Figures 3(a).
11. Higher stock drift rates induce an executive to hold NTNH ESOs longer before exercise. See Figure 3(c).
12. A higher index drift rate induces an executive to exercise the NTNH ESOs more quickly. Intuitively, if the index is more "attractive" (all else being equal), the executive exercises the options in order to invest more money in the index. See Figure 3(d).
13. Low correlation between returns on the index and returns on the stock leads to the ESOs being exercised more quickly. See Figure 3(b). The rough intuition here is that a low correlation makes the ESO less "attractive" (all else being equal), causing executives to exercise them more quickly in order to invest in the index.
14. Shorter maturity options are exercised at a higher rate than longer maturity options, however, cumulatively, till maturity, executives exercise higher percentage of longer maturity options. See Figure 3(f).
15. Higher stock volatility induces, on average, earlier exercise. See Figure 3(g).
16. Higher index volatility induces, on average, later exercise. See Figure 3(h).
17. Lower proportions of option endowments out of initial total wealth, induce, on average, the exercise of fewer options. See Figure 3(m) to Figure 3(o).

The above findings imply the following efficiency-enhancing policy implications for firms: 1) rather than granting long term ESOs, roll over short term ESOs by granting long-term reload[6] options, 2) reduce the firm's stock price volatility, 3) control the proportion of ESO endowments out of executives' total wealth, and 4) shorten vesting periods.

4 Conclusion

We identify a method for pricing non-transferable, non-hedgeable (NTNH) European and American call options under both block and continuous partial exercise. We implement this methodology to price executive stock options (ESOs) and run a sensitivity analysis that suggests efficiency-enhancing ESO policy implications. Our pricing methodology is simple and universal. It can be used to solve a category of pricing problems, including pensions, human capital, real estate, and intellectual property.

Future work might investigate executives hedging ESOs using primary assets that are correlated with the ESOs' underlying assets. While the benefit of such hedging is clear, there are costs, both in terms of lower performance relative to that of unrestricted portfolios and in terms of additional transaction costs. These costs clearly depend on the level of the executive's outside wealth relative to the ESOs' value, and an interesting problem is identifying the value of outside wealth that is required to mitigate the non-hedgeability consequences.

We could also price NTNH ESOs accounting for executives' job termination. This would be comparable to pricing default NTNH-contingent claims with the credit name being the executives' job termination. We could also apply our NTNH ESO valuation method to more complex ESO pricing; for example, NTNH reload ESOs. Upon reload, executives receive additional shares at market value equal to the current intrinsic value of (all) their ESOs and additional re-grant of new at the money NTNH ESOs of the same maturity as the original ones.

[6] For re-load options, additional, at the money, stock options are granted upon the exercise of the previously granted options.

Table 1

Incentive of ESO

Spot price (stock)	Black-Scholes price	Subjective price	Objective price	Incentive
8	5.82	4.75	5.07	0.87
10	7.59	6.49	6.95	0.79
12	9.41	7.92	8.51	0.71

This table displays the benchmark Black-Scholes price, subjective price, and objective price of NTNH ESO. The total portfolio includes ESO, restricted firm stock, and market index. The benchmark parameter values are as follows: Maturity=15 years. Spot price of stock=$10; Spot price of index=$6; Strike price=$10; Volatility of stock=50%; Volatility of index=30%. Correlation between stock and index=60%; Drift rate of stock=15%; Drift rate of index=8%; Risk-free rate=4%; Initial option endowment=200 shares; Initial stock endowment=200 shares; Initial cash endowment=$1000; Vesting period of firm stock=1 year; Vesting period of option=2 year; Number of steps=30; Montecarlo repeating times=10000. The Benchmark Black-Scholes price, subjective price, and objective price of NTNH ESO are 7.59, 6.49 and 6.95, respectively (see the middle row with Spot stock price being $10). By changing the spot price of stock to a lower (higher) level $8 ($10), we investigate the incentive of option $\partial\hat{p}(t)/\partial S(t)$, which is defined as the delta of the subjective price.

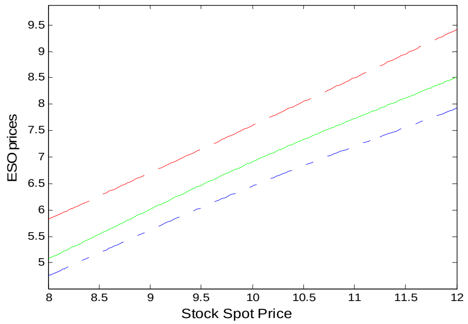
Table 2

Price sensitivity analysis and determinants of ESO's efficiency

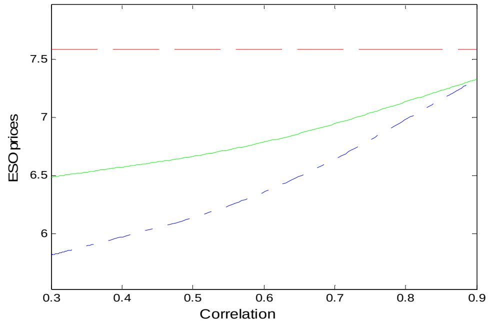
	Black-Scholes price	Subjective price	Objective price	Incentive	Dead weight cost	Efficiency
Correlation						
30%	7.59	5.82	6.49	0.79	0.67	1.18
60%	7.59	6.29	6.75	0.79	0.46	1.73
90%	7.59	7.33	7.33	0.79	0.00	254.76
Drift rate (stock)						
8%	7.59	2.96	4.06	0.79	1.10	0.72
15%	7.59	6.29	6.75	0.79	0.46	1.73
20%	7.59	7.28	7.39	0.79	0.11	7.24
Drift rate (index)						
6%	7.59	6.70	7.03	0.79	0.33	2.40
8%	7.59	6.29	6.75	0.79	0.46	1.73
20%	7.59	2.80	3.56	0.79	0.76	1.05
Risk-free rate						
2%	7.15	6.03	6.46	0.79	0.43	1.83
4%	7.59	6.29	6.75	0.79	0.46	1.73
6%	7.99	6.46	6.95	0.79	0.49	1.62
Maturity						
5	4.82	4.03	4.53	0.79	0.50	1.60
15	7.59	6.29	6.75	0.79	0.46	1.73
25	8.78	7.14	7.79	0.79	0.65	1.21
Volatility (stock)						
0.2	5.20	4.39	4.58	0.79	0.19	4.13
0.5	7.59	6.29	6.75	0.79	0.46	1.73
0.9	9.40	6.30	8.33	0.79	2.03	0.39
Volatility (index)						
0.1	7.59	2.75	3.90	0.79	1.15	0.69
0.3	7.59	6.29	6.75	0.79	0.46	1.73
0.6	7.59	6.69	7.03	0.79	0.33	2.39
Strike price						
6	8.24	7.33	7.79	0.79	0.46	1.73
10	7.59	6.29	6.75	0.79	0.46	1.73
14	7.10	5.41	5.86	0.79	0.45	1.77
Spot price (index)						
2	7.59	6.29	6.75	0.79	0.46	1.73
6	7.59	6.29	6.75	0.79	0.46	1.73
10	7.59	6.29	6.75	0.79	0.46	1.73
Vesting period (stock)						
0.5	7.59	6.48	6.75	0.79	0.27	2.95
1	7.59	6.29	6.75	0.79	0.46	1.73
1.5	7.59	6.10	6.75	0.79	0.65	1.22
Vesting period (option)						
1	7.59	6.18	6.61	0.79	0.43	1.84
2	7.59	6.29	6.75	0.79	0.46	1.73
3	7.59	6.38	6.88	0.79	0.50	1.60
Initial endowment (option)						
20	7.59	7.35	7.59	0.79	0.24	3.28
200	7.59	6.29	6.75	0.79	0.46	1.73
2000	7.59	2.85	4.35	0.79	1.51	0.53
Initial endowment (stock)						
20	7.59	4.51	5.23	0.79	0.72	1.10
200	7.59	6.29	6.75	0.79	0.46	1.73
2000	7.59	7.25	7.59	0.79	0.35	2.29
Initial endowment (cash)						
500	7.59	5.84	6.42	0.79	0.58	1.37
1000	7.59	6.29	6.75	0.79	0.46	1.73
10000	7.59	7.59	7.59	0.79	0.00	5320.38

This table lists the results of ESO price sensitivity analysis; namely, how does parameter change affect the change of ESO prices? The table also displays all the determinants affecting the ESO's efficiency, which is defined as incentive per unit of dead weight cost. The dead weight cost is the objective price net of the subjective price. Each panel lists three levels (low, benchmark, high) of results from above to below. We use 0.79, which is the benchmark incentive in Table 1, throughout the sensitivity analysis.

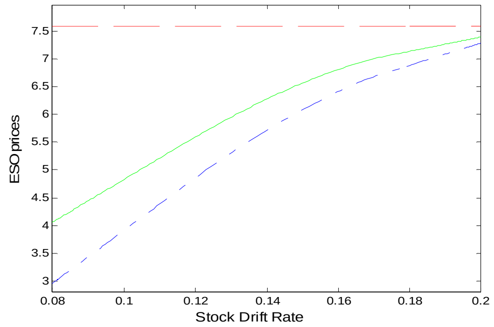
(a) Price sensitivity w.r.t spot price of stock



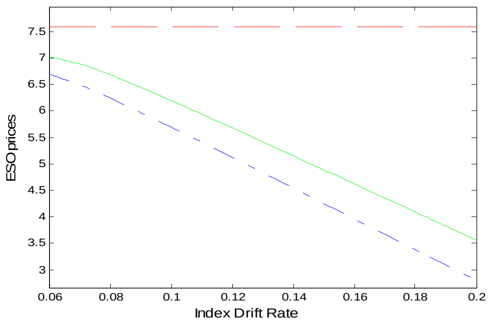
(b) Price sensitivity w.r.t correlation



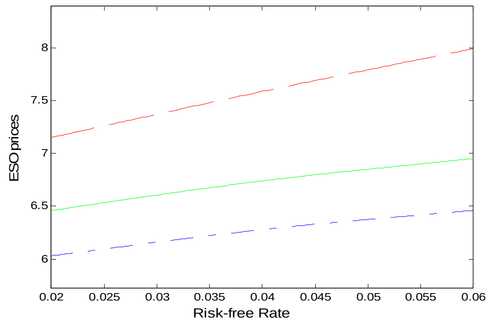
(c) Price sensitivity w.r.t drift rate of stock



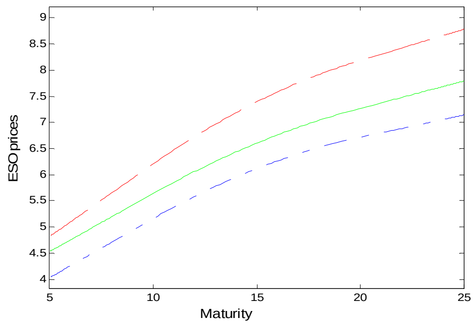
(d) Price sensitivity w.r.t drift rate of index



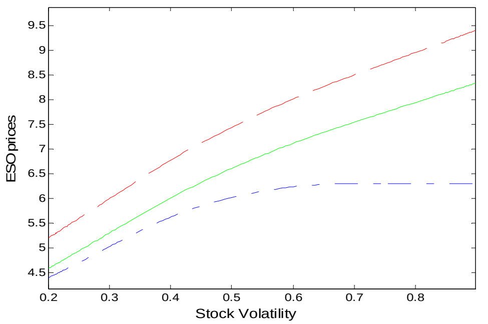
(e) Price sensitivity w.r.t risk-free rate



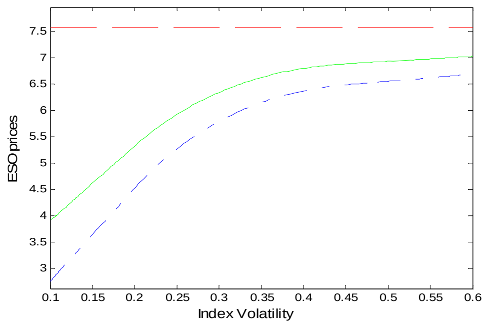
(f) Price sensitivity w.r.t maturity



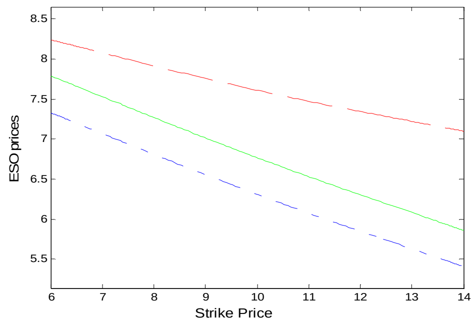
(g) Price sensitivity w.r.t stock volatility



(h) Price sensitivity w.r.t index volatility



(i) Price sensitivity w.r.t strike price



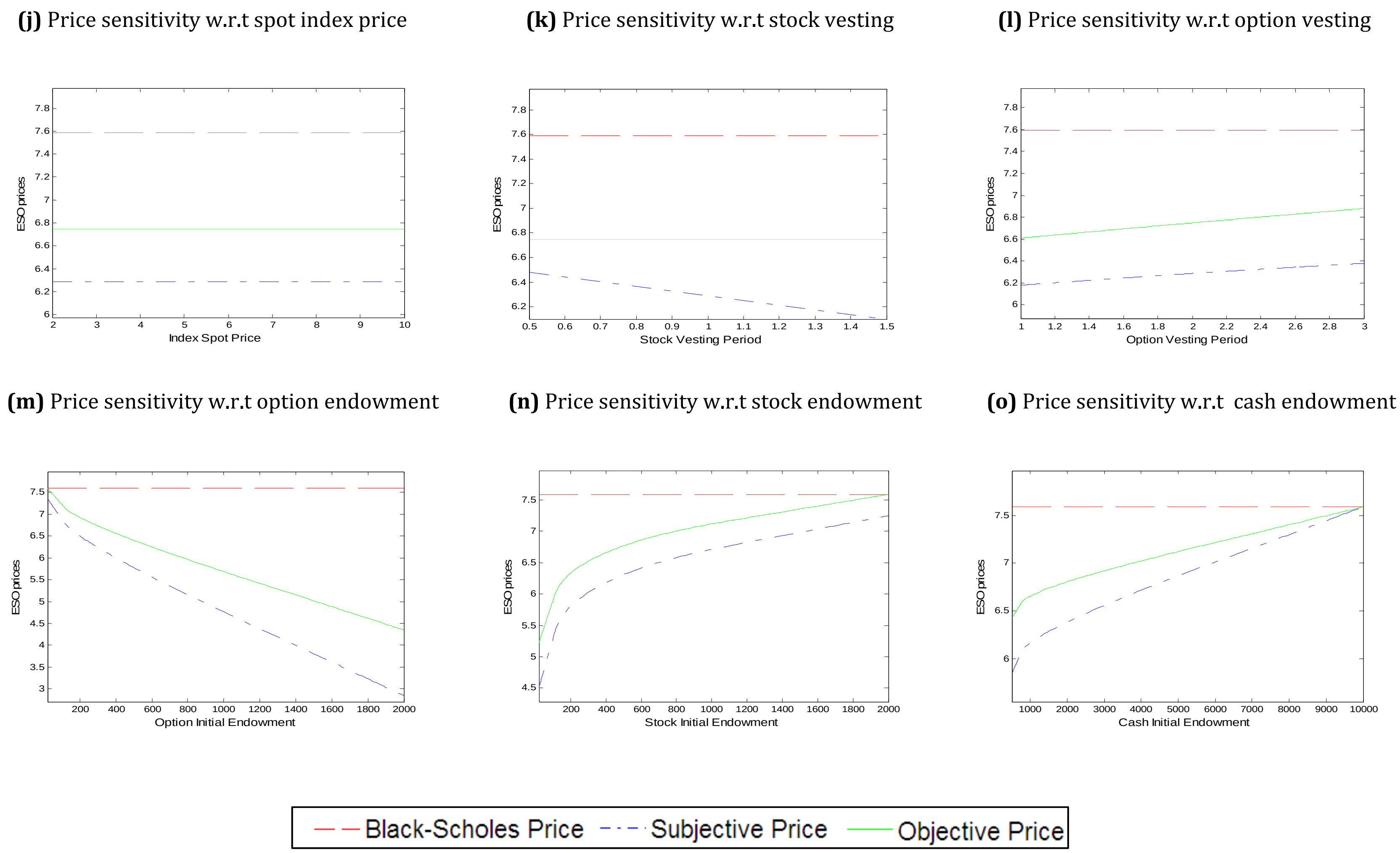
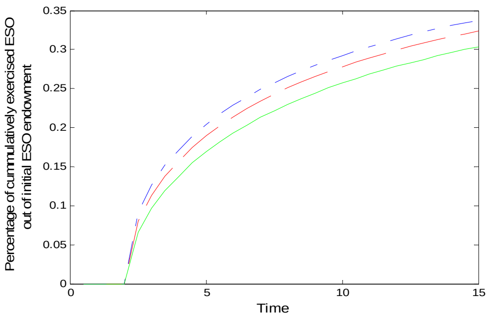
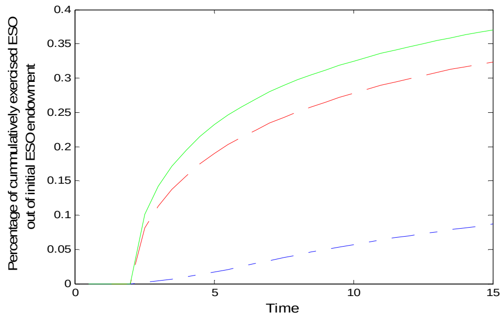


Figure 2. Price sensitivity. This figure displays the patterns of how parameters' changes affect ESO price changes (i.e., Black-Scholes, subjective, and objective prices). Each curve interpolates three parameter values specified in Tables 1 and 2, low level, benchmark level, and high level.

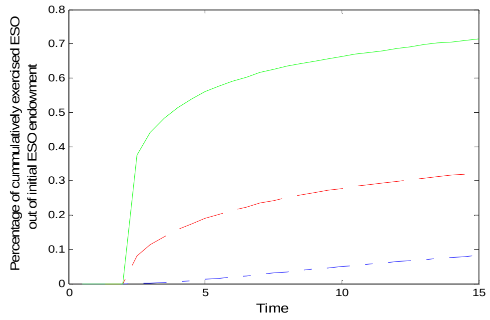
(a) Stock spot price



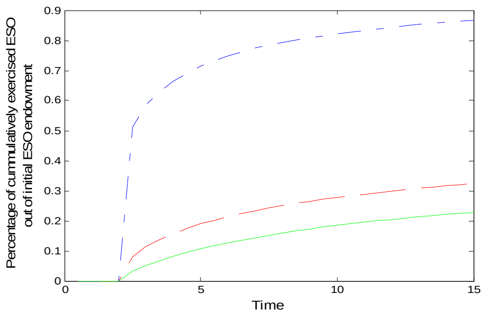
(b) Correlation



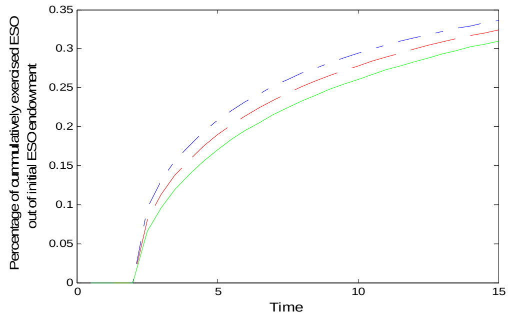
(c) Stock drift rate



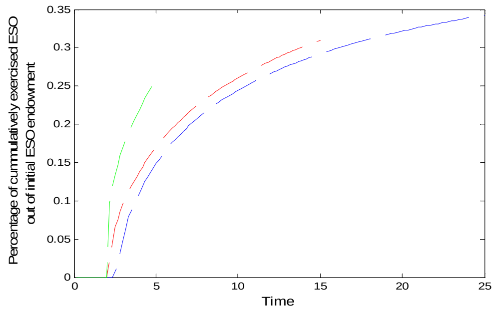
(d) Index drift rate



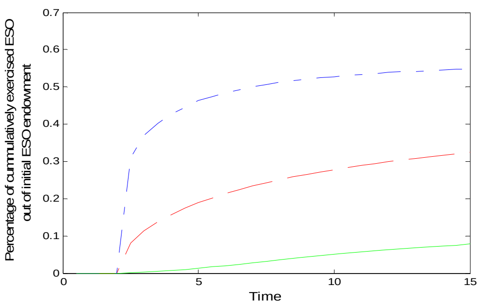
(e) Risk-free rate



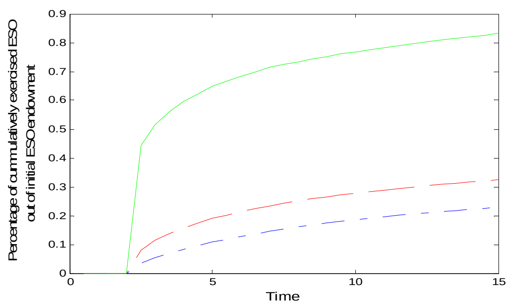
(f) Maturity



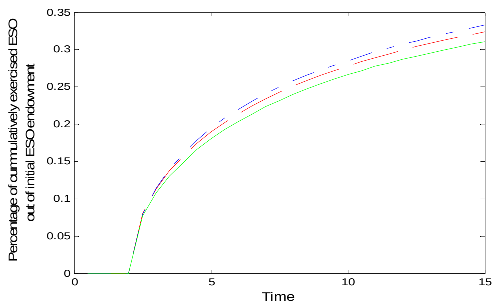
(g) Stock volatility



(h) Index volatility



(i) Strike price



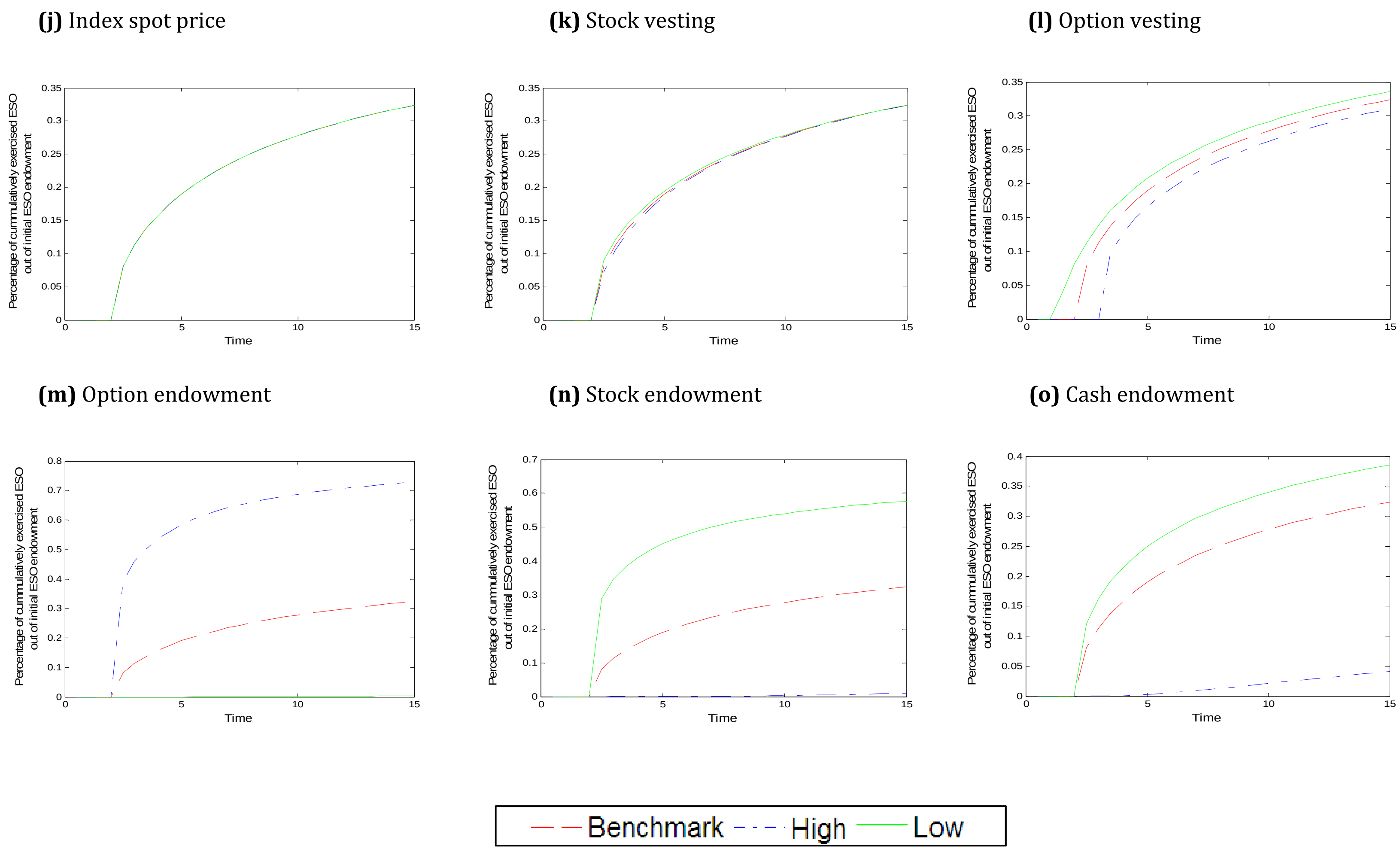


Figure 3. Exercise policies. This figure displays how different levels (Benchmark, High, Low) of the parameters affect the average exercise policies throughout the option's life. The Y axis represents the percentage of the averaged cumulative numbers of exercised options over 10,000 Monte Carlo sample paths out of the initial ESO endowment. The X axis represents running time in years with maturity of 15, except for in Figure 3(f), where maturity is the changing parameter.

Appendix A. Proofs of Theorem 1 and Theorem 2

A.1. Proof of Theorem 1

The utility function U: $(0,\infty) \to \Re$ is strictly increasing, and strictly concave, of class C^1 and satisfies, $U'(0^+) \triangleq lim_{\varkappa \downarrow 0} U'(\varkappa) = \infty$, $U'(\infty) \triangleq lim_{\varkappa \to \infty} U'(\varkappa) = 0$.

The following inequality exists due to the concavity of the utility function:

$$U(\varkappa) + (\wp - \varkappa)U'(\varkappa) \geq U(\wp) \geq U(\varkappa) + (\wp - \varkappa)U'(\wp), \qquad \forall\, 0 < \varkappa < \wp < \infty. \tag{A.1}$$

For any initial outside wealth ς, we introduce the models,

$$M(\varsigma, 0) \triangleq \sup_{\pi \in \mathcal{A}(\varsigma, K(0,\omega), 0, T)} \mathbb{E}[U(X^{\varsigma,\pi}(T)], \qquad 0 < \varsigma < \infty \tag{A.2}$$

$$\breve{M}(\varsigma, n, 0) \triangleq \sup_{\pi \in \mathcal{A}(\varsigma, K(0,\omega), 0, T)} \mathbb{E}[U(X^{\varsigma,\pi}(T) + nB(T)], \qquad 0 < \varsigma < \infty \tag{A.3}$$

$$\Psi(\alpha, \hat{p}, \varsigma) \triangleq \sup_{\pi \in \mathcal{A}(\varsigma - \alpha, K(0,\omega), 0, T)} \mathbb{E}\left[U\left(X^{\varsigma-\alpha,\pi}(T) + \frac{\alpha}{\hat{p}}B(T)\right)\right], \qquad 0 < \varsigma < \infty, \tag{A.4}$$

where $\mathcal{A}(x, K(t,\omega), t, T)$ is defined in Equation (5). For any $\epsilon > 0$, write $\alpha = n\hat{p} - \epsilon$, then

$$\Psi(\alpha, \hat{p}, \varsigma) \geq \mathbb{E}\left[U\left(X_*^{\varsigma-\alpha}(T) + nB(T) - \tfrac{\epsilon}{\hat{p}}B(T)\right)\right],$$

where $X_*^{\varsigma-\alpha}(T) \equiv X^{\varsigma-\alpha,\pi^*}(T)$ and π^* solves $\breve{M}(\varsigma - \alpha, n, 0)$. Here we abandon $\frac{\epsilon}{\hat{p}}$ shares of the option in exchange for ϵ dollars in cash.

Let the initial outside wealth be $\varsigma = n\hat{p} + x$. Then, thanks to the first inequality in (A.3), we have

$$\begin{aligned} &\mathbb{E}\left[U\left(X_*^{\varsigma-\alpha}(T) + nB(T) - \frac{\epsilon}{\hat{p}}B(T)\right)\right] \\ &\qquad \geq \mathbb{E}\left[U\big(X_*^{\varsigma-\alpha}(T) + nB(T)\big) - \frac{\epsilon}{\hat{p}}B(T)U'\left(X_*^{\varsigma-\alpha}(T) + \frac{n\hat{p} - \epsilon}{\hat{p}}B(T)\right)\right]. \end{aligned} \tag{A.5}$$

Since $\varsigma - \alpha = x + \epsilon \geq x$ and $\varsigma \mapsto X_*^{\varsigma}(T)$ is non-decreasing, we get

$$\Psi(\alpha, \hat{p}, \varsigma) \geq \mathbb{E}\big[U\big(X_*^{\varsigma-\alpha}(T) + nB(T)\big)\big] - \frac{\epsilon}{\hat{p}}\mathbb{E}\left[U'\left(X_*^{x}(T) + \frac{n\hat{p} - \epsilon}{\hat{p}}B(T)\right)B(T)\right].$$

By the definition of $\breve{M}(\varsigma - \alpha, n, 0)$,

$$\Psi(\alpha, \hat{p}, \varsigma) \geq \breve{M}(\varsigma - \alpha, n, 0) - \frac{\epsilon}{\hat{p}}\mathbb{E}\left[U'\left(X_*^{x}(T) + \frac{n\hat{p} - \epsilon}{\hat{p}}B(T)\right)B(T)\right].$$

Also, $\Psi(n\hat{p}, \hat{p}, x) = \breve{M}(\varsigma - n\hat{p}, n, 0)$. So,

$$-\Psi(\alpha, \hat{p}, \varsigma) \leq -\breve{M}(\varsigma - \alpha, n, 0) + \tfrac{\epsilon}{\hat{p}}\mathbb{E}\left[U'\left(X_*^{x}(T) + \tfrac{n\hat{p}-\epsilon}{\hat{p}}B(T)\right)B(T)\right], \text{ and}$$

$$\Psi(n\hat{p}, \hat{p}, x) - \Psi(\alpha, \hat{p}, \varsigma) \leq \breve{M}(\varsigma - n\hat{p}, n, 0) - \breve{M}(\varsigma - \alpha, n, 0) + \tfrac{\epsilon}{\hat{p}}\mathbb{E}\left[U'\left(X_*^{x}(T) + \tfrac{n\hat{p}-\epsilon}{\hat{p}}B(T)\right)B(T)\right].$$

So, for $n\hat{p} > \alpha$,

$$\frac{\Psi(n\hat{p},\hat{p},x)-\Psi(\alpha,\hat{p},\varsigma)}{n\hat{p}-\alpha} \le \frac{\breve{M}(\varsigma-n\hat{p},n,0)-\breve{M}(\varsigma-\alpha,n,0)}{n\hat{p}-\alpha}+\frac{\epsilon}{\hat{p}(n\hat{p}-\alpha)}\mathbb{E}\left[U'\left(X_*^x(T)+\frac{n\hat{p}-\epsilon}{\hat{p}}B(T)\right)B(T)\right].$$

In general, $\underset{n\to\infty}{limsup}(a_n+b_n)\le \underset{n\to\infty}{limsup}(a_n)+\underset{n\to\infty}{limsup}(b_n)$.

Taking the *limsup* of both sides, it follows that

$$\underset{\alpha\uparrow n\hat{p}}{limsup}\frac{\Psi(n\hat{p},\hat{p},x)-\Psi(\alpha,\hat{p},\varsigma)}{n\hat{p}-\alpha}\le \underset{\alpha\uparrow n\hat{p}}{limsup}\frac{\breve{M}(\varsigma-n\hat{p},n,0)-\breve{M}(\varsigma-\alpha,n,0)}{n\hat{p}-\alpha} +\underset{\alpha\uparrow n\hat{p}}{limsup}\frac{\epsilon}{\hat{p}(n\hat{p}-\alpha)}\mathbb{E}\left[U'\left(X_*^x(T)+\frac{n\hat{p}-\epsilon}{\hat{p}}B(T)\right)B(T)\right].$$

Now, using the fact that $\alpha=n\hat{p}-\epsilon$, we write this as

$$\underset{\epsilon\downarrow 0}{limsup}\frac{\Psi(n\hat{p},\hat{p},x)-\Psi(n\hat{p}-\epsilon,\hat{p},\varsigma)}{\epsilon} \le \underset{\epsilon\downarrow 0}{limsup}\frac{\breve{M}(\varsigma-n\hat{p},n,0)-\breve{M}(\varsigma-\alpha,n,0)}{\epsilon}+\underset{\epsilon\downarrow 0}{limsup}\frac{\epsilon}{\hat{p}\epsilon}\mathbb{E}\left[U'\left(X_*^x(T)+\frac{n\hat{p}-\epsilon}{\hat{p}}B(T)\right)B(T)\right].$$

As $\epsilon\downarrow 0$, $X_*^x(T)+\frac{n\hat{p}-\epsilon}{\hat{p}}B(T)$ increases, and $U^{'}\left(X_*^x(T)+\frac{n\hat{p}-\epsilon}{\hat{p}}B(T)\right)$ decreases. Thus, by the Monotone Convergence Theorem

$$\underset{\epsilon\downarrow 0}{limsup}\frac{\Psi(n\hat{p},\hat{p},x)-\Psi(n\hat{p}-\epsilon,\hat{p},\varsigma)}{\epsilon}\le -\breve{M}'(\varsigma-n\hat{p},n,0)+\frac{1}{\hat{p}}\mathbb{E}\left[B(T)U'\left(X_*^x(T)+nB(T)\right)\right]$$
$$=-\breve{M}'(x,n,0)+\tfrac{1}{\hat{p}}\mathbb{E}\left[B(T)U'\left(X_*^x(T)+nB(T)\right)\right].$$

Now we consider the case, $\epsilon<0$, or equivalently, $n\hat{p}<\alpha$. From the second inequality in (A.3), we have

$$U\left(X_*^{\varsigma-\alpha}(T)+\frac{n\hat{p}-\epsilon}{\hat{p}}B(T)\right)\ge U\left(X_*^{\varsigma-\alpha}(T)+nB(T)\right)-\frac{\epsilon}{\hat{p}}B(T)U'\left(X_*^{\varsigma-\alpha}(T)+\frac{n\hat{p}-\epsilon}{\hat{p}}B(T)\right).$$

In this case, $\varsigma-\alpha=x+\epsilon\le x$ and since $\varsigma\mapsto X_*^{\varsigma}(T)$ is non-decreasing, we get

$$U\left(X_*^{\varsigma-\alpha}(T)+\frac{n\hat{p}-\epsilon}{\hat{p}}B(T)\right)\ge U\left(X_*^{\varsigma-\alpha}(T)+nB(T)\right)-\frac{\epsilon}{\hat{p}}B(T)U'\left(X_*^{\varsigma}(T)+\frac{n\hat{p}-\epsilon}{\hat{p}}B(T)\right).$$

Taking expectations gives us the inequality in (A.5). Using the definitions of $\Psi(\alpha,\hat{p},\varsigma)$ and $\breve{M}(\varsigma-\alpha,n,0)$, and dividing by $n\hat{p}-\alpha<0$, it is not hard to see that

$$\frac{\Psi(n\hat{p},\hat{p},x)-\Psi(\alpha,\hat{p},\varsigma)}{n\hat{p}-\alpha} \ge \frac{\breve{M}(\varsigma-n\hat{p},n,0)-\breve{M}(\varsigma-\alpha,n,0)}{n\hat{p}-\alpha}+\frac{\epsilon}{\hat{p}(n\hat{p}-\alpha)}\mathbb{E}\left[U'\left(X_*^x(T)+\frac{n\hat{p}-\epsilon}{\hat{p}}B(T)\right)B(T)\right].$$

Again, in general, $\underset{n\to\infty}{liminf}(a_n+b_n)\ge \underset{n\to\infty}{liminf}(a_n)+\underset{n\to\infty}{liminf}(b_n)$.

Taking the *liminf* of both sides, it follows that

$$\liminf_{\alpha\downarrow n\hat{p}} \frac{\Psi(n\hat{p},\hat{p},x)-\Psi(\alpha,\hat{p},\varsigma)}{n\hat{p}-\alpha}$$
$$\geq \liminf_{\alpha\downarrow n\hat{p}} \left\{\frac{\breve{M}(\varsigma-n\hat{p},n,0)-\breve{M}(\varsigma-\alpha,n,0)}{n\hat{p}-\alpha}+\frac{\epsilon}{\hat{p}(n\hat{p}-\alpha)}\mathbb{E}\left[U'\left(X_*^x(T)+\frac{n\hat{p}-\epsilon}{\hat{p}}B(T)\right)B(T)\right]\right\}$$
$$\geq \liminf_{\alpha\downarrow n\hat{p}} \frac{\breve{M}(\varsigma-n\hat{p},n,0)-\breve{M}(\varsigma-\alpha,n,0)}{n\hat{p}-\alpha}$$
$$+\liminf_{\alpha\downarrow n\hat{p}} \frac{\epsilon}{\hat{p}(n\hat{p}-\alpha)}\mathbb{E}\left[U'\left(X_*^x(T)+\frac{n\hat{p}-\epsilon}{\hat{p}}B(T)\right)B(T)\right].$$

Now, using the fact that $\alpha = n\hat{p}-\epsilon$, we write this as

$$\liminf_{\epsilon\uparrow 0} \frac{\Psi(n\hat{p},\hat{p},x)-\Psi(n\hat{p}-\epsilon,\hat{p},\varsigma)}{\epsilon}$$
$$\geq \limsup_{\epsilon\uparrow 0} \frac{\breve{M}(\varsigma-n\hat{p},n,0)-\breve{M}(\varsigma-\alpha,n,0)}{\epsilon}+\limsup_{\epsilon\uparrow 0}\frac{\epsilon}{\hat{p}\epsilon}\mathbb{E}\left[U'\left(X_*^x(T)+\frac{n\hat{p}-\epsilon}{\hat{p}}B(T)\right)B(T)\right]$$
$$= -\breve{M}'(x,n,0)+\tfrac{1}{\hat{p}}\mathbb{E}\big[B(T)U'\big(X_*^x(T)+nB(T)\big)\big].$$

Hence,

$$\limsup_{\alpha\downarrow n\hat{p}} \frac{\Psi(n\hat{p},\hat{p},x)-\Psi(\alpha,\hat{p},x)}{n\hat{p}-\alpha}$$
$$\leq -\breve{M}'(x,n,0)+\frac{1}{\hat{p}}\mathbb{E}\big[B(T)U'\big(X_*^x(T)+nB(T)\big)\big] \leq \liminf_{\alpha\uparrow n\hat{p}} \frac{\Psi(n\hat{p},\hat{p},x)-\Psi(\alpha,\hat{p},x)}{n\hat{p}-\alpha}.$$

If the derivative $\frac{\partial\Psi(\alpha,\hat{p},x)}{\partial\alpha}$ exists, then the fair subjective price of the NTNH ESO is set to be the solution of the equation, $\frac{\partial\Psi(\alpha,\hat{p},x)}{\partial\alpha}\Big|_{\alpha=n\hat{p}} = 0$.

Intuitively, this means that the executive is indifferent between holding the current portfolio or removing a small amount of money, α, out of the portfolio and buying more options. If Ψ is not differentiable, then the fair subjective price is set to be the *weak solution* (see definition 7.2 in Karatzas and Kou 1996) of the above equation. In either case, the inequalities above show that $\hat{p}$ is determined by

$$-\breve{M}'(x,n,0)+\tfrac{1}{\hat{p}}\mathbb{E}\big[B(T)U'\big(X_*^x(T)+nB(T)\big)\big]=0.$$

The subjective price of the ESO at time zero is

$$\hat{p}(x,n,0)\equiv\hat{p}=\frac{\mathbb{E}\big[B(T)U'\big(X_*^x(T)+nB(T)\big)\big]}{\breve{M}'(x,n,0)}=\frac{\mathbb{E}\big[B(T)U'\big(X_*^x(T)+nB(T)\big)\big]}{M'(x+nBS(0),0)}.$$

Thanks to the Markov property, the subjective price of the ESO at time t is

$$\hat{p}(x_t,n,t)=\frac{\mathbb{E}\left[B(T)U'\left(X^{x_t,\pi^*}(T)+nB(T)\right)|\mathcal{F}_t\right]}{M'(x_t+nBS(t),\ t)},\qquad \forall x_t>0.$$

□

A.2. Proof of Theorem 2

Theorem 7.4 in Karatzas and Kou (1996) requires their equation (7.24), which is based on the assumption that δ is bounded from below. Although we cannot assume that δ is bounded from below in this paper, it is a sufficient condition rather than a necessary one. This assumption is used to prove that the conditions $\forall y \in (0,\infty), \lim_{\|v\|\to\infty} \tilde{J}(y;v) = \infty$ and $\forall v \in (0,\infty), \lim_{y\downarrow 0} \tilde{J}(y;v) = \infty$ hold. Hence, we directly make these less strict assumptions, given in equations (18) and (19) respectively, so that the results in Theorem 7.4 still hold. We can then prove Theorem 2 directly:

$$\hat{p}(x_t, n, t) = \frac{\mathbb{E}[H_{\hat{v}}(T)B(T)|\mathcal{F}_t]}{H_{\hat{v}}(t)} = \frac{\mathbb{E}^{\hat{v}}[\gamma_{\hat{v}}(T)B(T)|\mathcal{F}_t]}{\gamma_{\hat{v}}(t)},$$

where the second equality is obtained by abstract Bayes' theorem, and $\mathbb{E}^{\hat{v}}(\cdot)$ is under the corresponding martingale measure with respect to the auxiliary market $\mathcal{M}_{\hat{v}}$. According to Theorem 9.1 of Cvitanić and Karatzas (1992), page 780, there exists a portfolio process satisfying the constraints, such that the initial $\hat{p}(x_t, n, t)$ will evolve to $B(T)$ at the terminal date. Since the firm will repay the executive when the European call option is exercised at its maturity, and the NTNH constraints are not for the firm, the SDF for the objective price of the NTNH ESO is simply the market SDF $H_0(T)$. Again, thanks to abstract Bayes' theorem, the objective price of the NTNH ESO with terminal payoff $B(T)$ at time t is equal to the market price:

$$\tilde{p}(t) = p(t) = \frac{\mathbb{E}[H_0(T)B(T)|\mathcal{F}_t]}{H_0(t)} = \frac{\mathbb{E}^0[\gamma_0(T)B(T)|\mathcal{F}_t]}{\gamma_0(t)}.$$

$\square$

Remark: The total initial wealth for the executive is x amount of cash and n shares of contingent claims remaining fixed in the portfolio. Those n shares can be replicated by a sub-portfolio of underlying assets. The outside initial wealth x can be invested freely as long as there is no negative position in the firm's stock. The combined non-transferable and non-hedgeable constraint is that the number of shares of the firm's stock in the portfolio has to be no less than n shares. Thus, we have translated the problem into a traditional constrained portfolio optimization problem. We can get the optimal portfolio strategy within the constraints and, correspondingly, the optimal terminal wealth under constraints. It is natural to define the price for the optimal terminal wealth $X_*^x(T) + nB(T)$ as the total initial wealth, which is

$$x + nBS(0) = \mathbb{E}^{\hat{v}}\left[\gamma_{\hat{v}}(T)\left(X^{x,\pi^*}(T) + nB(T)\right)\right].$$

From Theorem 2,

$$\hat{p}(x, n, t) = \frac{\mathbb{E}^{\hat{v}}[\gamma_{\hat{v}}(T)B(T)|\mathcal{F}_t]}{\gamma_{\hat{v}}(t)}.$$

Then the subjective price of outside wealth $\mathbb{E}^{\hat{v}}\left[\gamma_{\hat{v}}(T)X^{x,\pi^*}(T)\right]$ will be $x + nBS(0) - n\hat{p}(x, n, t)$.

Appendix B. Proofs of Theorem 3 and Theorem 4

B.1. Proof of Theorem 3

First, we prove that the indirect utility $M(\cdot, t)$ satisfies

$$M(\varkappa,t)+(\wp-\varkappa)M'(\varkappa,t)\geq M(\wp,t)\geq M(\varkappa,t)+(\wp-\varkappa)M'(\wp,t),\ \forall\ 0<\varkappa<\wp<\infty. \quad \text{(B.1)}$$

According to Theorem 7.3 in Karatzas and Kou (1996), with their Assumption 6.2 being relaxed by our Assumptions (18) and (19) from Proposition 1, and assuming that the constraint set K is closed, convex and satisfies the regularity conditions, then for any $x>0$, there exists a process $\hat{v}\equiv\hat{v}_x\in\mathfrak{D}$ and a portfolio process $\pi^*\in\mathcal{A}(x,K(0,\omega),0,T)$ for the constrained portfolio optimization problem of $M(x,0)=sup_{(\pi,C)\in\mathcal{A}(x,K(0,\omega),0,T)}\mathbb{E}\left[U\left(X^{x,\pi}(T)\right)\right]$, $0<x<\infty$, with corresponding terminal wealth $X^{x,\pi^*}(T)=I\left(\mathcal{Y}_{\hat{v}}(x)\gamma_{\hat{v}}(T)Z_{\hat{v}}(T)\right)$ $a.s.$, where $y_t\equiv\mathcal{Y}_{\hat{v}}(x_t)$ is the inverse function of $\mathcal{X}_{\hat{v}}(y_t)\triangleq\mathbb{E}[H_{\hat{v}}(T)I(y_tH_{\hat{v}}(T))|\mathcal{F}_t]$, and $I(\cdot)$ is the inverse function of $U'(\cdot)$. π^* is optimal for the problem. The value function $M(\cdot,t)$ is continuously differentiable, and its derivative can be represented as $M'(x,t)\equiv M_x(x,t)=\mathcal{Y}_{\hat{v}}(x)>0,\ \forall x>0$.

Note that, according to Assumption 7.2 in Karatzas and Kou (1996), we have $\forall\pi\in\mathcal{A}(\varkappa,K(t,\omega),t,T)$, and $\forall\pi\in\mathcal{A}(\wp,K(t,\omega),t,T)$, $\frac{X^{\wp,\pi}(T)}{X^{\varkappa,\pi}(T)}=\frac{\wp}{\varkappa}$, which implies that

$$\frac{\partial X^{\varkappa,\pi^*}(T)}{\partial\varkappa}=\lim_{\varepsilon\to 0}\frac{X^{\varkappa+\varepsilon,\pi^*}(T)-X^{\varkappa,\pi^*}(T)}{\varepsilon}=\lim_{\varepsilon\to 0}\frac{\left(\frac{\varkappa+\varepsilon}{\varkappa}-1\right)X^{\varkappa,\pi^*}(T)}{\varepsilon}=\frac{X^{\varkappa,\pi^*}(T)}{\varkappa}.$$

Also, the concavity of the utility function, $U(\cdot)$, gives us the inequality,

$$U(\varkappa)+(\wp-\varkappa)U'(\varkappa)\geq U(\wp)\geq U(\varkappa)+(\wp-\varkappa)U'(\wp),\ \ \forall\ 0<\varkappa<\wp<\infty.$$

Therefore,

$$U(X^{\varkappa,\pi}(T)+\left(X^{\wp,\pi}(T)-X^{\varkappa,\pi}(T)\right)U'\left(X^{\varkappa,\pi}(T)\right)\geq U(X^{\varkappa,\pi}(T)+\left(X^{\wp,\pi}(T)-X^{\varkappa,\pi}(T)\right)U'\left(X^{\wp,\pi}(T)\right),$$

from which we can see that

$$\mathbb{E}\left(U(X^{\varkappa,\pi^*}(T)\right)+(\wp-\varkappa)\mathbb{E}\left(\frac{X^{\varkappa,\pi^*}(T)}{\varkappa}U'\left(X^{\varkappa,\pi^*}(T)\right)\right)\geq\mathbb{E}\left(U\left(X^{\wp,\pi^*}(T)\right)\right)$$
$$\geq\mathbb{E}\left(U(X^{\wp,\pi^*}(T)\right)+(\wp-\varkappa)\mathbb{E}\left(\frac{X^{\wp,\pi^*}(T)}{\wp}U'\left(X^{\wp,\pi^*}(T)\right)\right).$$

Note that

$$\frac{X^{\varkappa,\pi^*}(T)}{\varkappa}U'\left(X^{\varkappa,\pi^*}(T)\right)=\frac{\partial X^{\varkappa,\pi^*}(T)}{\partial\varkappa}U'\left(X^{\varkappa,\pi^*}(T)\right)=\frac{dU\left(X^{\varkappa,\pi^*}(T)\right)}{d\varkappa}=lim_{\varepsilon\to 0}\frac{U\left(X^{\varkappa+\varepsilon,\pi^*}(T)\right)-U\left(X^{\varkappa,\pi^*}(T)\right)}{\varepsilon},$$

since $U(\cdot)$ is a utility function, being strictly increasing and strictly concave of class C^1. Also, it is not hard to see that if $\varepsilon_2>\varepsilon_1>0$, then $0\leq\frac{U\left(X^{\varkappa+\varepsilon_1,\pi^*}(T)\right)-U\left(X^{\varkappa,\pi^*}(T)\right)}{\varepsilon_1}\leq\frac{U\left(X^{\varkappa+\varepsilon_2,\pi^*}(T)\right)-U\left(X^{\varkappa,\pi^*}(T)\right)}{\varepsilon_2}$.

Then, by the monotone convergence theorem, we have,

$$\mathbb{E}\left(\lim_{\varepsilon\downarrow 0}\frac{U\left(X^{\varkappa+\varepsilon,\pi^*}(T)\right)-U\left(X^{\varkappa,\pi^*}(T)\right)}{\varepsilon}\right)=\lim_{\varepsilon\downarrow 0}\mathbb{E}\left(\frac{U\left(X^{\varkappa+\varepsilon,\pi^*}(T)\right)-U\left(X^{\varkappa,\pi^*}(T)\right)}{\varepsilon}\right).$$

If $\varepsilon_2<\varepsilon_1<0$, then $0\leq\frac{U\left(X^{\varkappa+\varepsilon_1,\pi^*}(T)\right)-U\left(X^{\varkappa,\pi^*}(T)\right)}{\varepsilon_1}\leq\frac{U\left(X^{\varkappa+\varepsilon_2,\pi^*}(T)\right)-U\left(X^{\varkappa,\pi^*}(T)\right)}{\varepsilon_2}$. Again by the monotone convergence theorem, we have

$$\mathbb{E}\left(\lim_{\varepsilon\uparrow 0}\frac{U\left(X^{\varkappa+\varepsilon,\pi^*}(T)\right)-U\left(X^{\varkappa,\pi^*}(T)\right)}{\varepsilon}\right)=\lim_{\varepsilon\uparrow 0}\mathbb{E}\left(\frac{U\left(X^{\varkappa+\varepsilon,\pi^*}(T)\right)-U\left(X^{\varkappa,\pi^*}(T)\right)}{\varepsilon}\right).$$

Then,

$$\mathbb{E}\left(\lim_{\varepsilon\to 0}\frac{U\left(X^{\varkappa+\varepsilon,\pi^*}(T)\right)-U\left(X^{\varkappa,\pi^*}(T)\right)}{\varepsilon}\right)=\lim_{\varepsilon\to 0}\mathbb{E}\left(\frac{U\left(X^{\varkappa+\varepsilon,\pi^*}(T)\right)-U\left(X^{\varkappa,\pi^*}(T)\right)}{\varepsilon}\right).$$

Hence,

$$\mathbb{E}\left(U(X^{\varkappa,\pi^*}(T)\right)+(\wp-\varkappa)\frac{\partial\mathbb{E}\left(U\left(X^{\varkappa,\pi^*}(T)\right)\right)}{\partial\varkappa}\geq\mathbb{E}\left(U\left(X^{\wp,\pi^*}(T)\right)\right)$$

$$\geq\mathbb{E}\left(U(X^{\varkappa,\pi^*}(T)\right)+(\wp-\varkappa)\frac{\partial\mathbb{E}\left(U\left(X^{\wp,\pi^*}(T)\right)\right)}{\partial\varkappa};$$

that is,

$M(\varkappa,t)+(\wp-\varkappa)M'(\varkappa,t)\geq M(\wp,t)\geq M(\varkappa,t)+(\wp-\varkappa)M'(\wp,t),\ \forall\ 0<\varkappa<\wp<\infty.$

So far, we have proved that the indirect utility function of the portfolio is concave. The indirect utility function of the portfolio that we define for pricing the NTNH ESOs is exactly the same except that the total initial wealth includes two parts, the net initial wealth x and n shares of non-transferable ESOs worth $nBS(0)$.

In our model, the n shares of the NTNH American ESOs are included in the total wealth portfolio, and the marginal $\frac{\epsilon}{p}$ share is outside of the executive's total wealth; they are assumed shares for finding the indifference price. Since we are interested in the price of the ESO within the total wealth, we let the exercise time of the extra $\frac{\epsilon}{p}$ shares be the same as the former; otherwise it is a multi-option and multi-optimal exercise time problem, which we do not discuss in this paper.

We define

$$\Psi(\epsilon,\hat{p},x)\triangleq\sup_{\tau\in\Gamma}\ \sup_{\pi\in\mathcal{A}\left(x+\epsilon+nBS(0)-\frac{\epsilon}{\hat{p}}BS(0),K(0,\omega),0,\tau\right)}\mathbb{E}\left[M\left(X^{x+\epsilon,\pi}(\tau)+nB(\tau)-\frac{\epsilon}{\hat{p}}B(\tau),\tau\right)\right]$$

$$=\sup_{\tau\in\Gamma}\ \sup_{\pi\in\mathcal{A}\left(x+\epsilon+nBS(0)-\frac{\epsilon}{\hat{p}}BS(0),K(0,\omega),0,\tau\right)}\mathbb{E}\left[M\left(\mathbb{X}^{x+\epsilon+nBS(0)-\frac{\epsilon}{\hat{p}}BS(0),\pi}(\tau),\tau\right)\right].$$

Here, $X^{x+\epsilon,\pi}(\tau)$ represents the outside wealth, the wealth apart from the ESO, and $\mathbb{X}^{x+\epsilon+nBS(0)-\frac{\epsilon}{\hat{p}}BS(0),\pi}$ represents the total wealth that evolves from the initial total wealth $x+\epsilon+nBS(0)-\frac{\epsilon}{\hat{p}}BS(0)$. The stochastic constraint $K(t,\omega)$ in this model is $\left[\frac{\left(n-\frac{\epsilon}{\hat{p}}\right)\phi(d_1(t,\omega))S_1(t)}{\mathbb{X}(t)},\infty\right)\times(-\infty,\infty)^{d-1}$, which is a function of ϵ. In theory; there is no problem with having a stochastic constraint. This only makes the drift rate and the risk-free rate, $r_v(t)$, stochastic in the auxiliary market. Moreover, please be aware that the constraint $K(t,\omega)$ in Karatzas and Kou (1996) is not a function of ϵ, while it is a function of ϵ in this case.

Let τ be any stopping time. As noted in the proof of Theorem 3, the following inequality exists, due to the concavity of the indirect utility function M:

$$M(\varkappa,\tau)+(\wp-\varkappa)M'(\varkappa,\tau)\geq M(\wp,\tau)\geq M(\varkappa,\tau)+(\wp-\varkappa)M'(\wp,\tau),\forall\ 0<\varkappa<\wp<\infty. \quad \text{(B.2)}$$

Define

$$\Psi(\epsilon,\hat{p},x)\triangleq \sup_{\tau\in\Gamma}\ \sup_{\pi\in\mathcal{A}\left(x+\epsilon+nBS(0)-\frac{\epsilon}{\hat{p}}BS(0),K(0,\omega),0,\tau\right)}\mathbb{E}\left[M\left(X^{x+\epsilon+nBS(0)-\frac{\epsilon}{\hat{p}}BS(0),\pi}(\tau),\tau\right)\right].$$

We next define

$$L(x,n,0)\triangleq \sup_{\tau\in\Gamma}\ \sup_{\pi\in\mathcal{A}\left(x+\epsilon+nBS(0)-\frac{\epsilon}{\hat{p}}BS(0),K(0,\omega),0,\tau\right)}\mathbb{E}[M(X^{x,\pi}(\tau)+nB(\tau),\tau)]. \quad \text{(B.3)}$$

Let τ^* and π^* solve equation (B.3) and fix $\epsilon>0$. From the first inequality in (B.2), we have

$$M\left(X^{x+\epsilon,\pi^*}(\tau^*)+nB(\tau^*)-\frac{\epsilon}{\hat{p}}B(\tau^*),\tau\right)+\frac{\epsilon}{\hat{p}}B(\tau^*)M'\left(X^{x+\epsilon,\pi^*}(\tau^*)+nB(\tau^*)-\frac{\epsilon}{\hat{p}}B(\tau^*),\tau^*\right)$$
$$\geq M\left(X^{x+\epsilon,\pi^*}(\tau^*)+nB(\tau^*),\tau^*\right).$$

Because $\epsilon>0$ and X is non-decreasing, we get

$$M\left(X^{x+\epsilon,\pi^*}(\tau^*)+nB(\tau^*)-\frac{\epsilon}{\hat{p}}B(\tau^*),\tau^*\right)$$
$$\geq M\left(X^{x+\epsilon,\pi^*}(\tau^*)+nB(\tau^*),\tau^*\right)-\frac{\epsilon}{\hat{p}}B(\tau^*)M'\left(X^{x,\pi^*}(\tau^*)+nB(\tau^*)-\frac{\epsilon}{\hat{p}}B(\tau^*),\tau^*\right).$$

The rest of the proof is similar to the proof of Theorem 1.

The reason we employ τ^* is that we made an assumption at the beginning, that the marginal $\left(\frac{\epsilon}{\hat{p}}\text{ shares}\right)$ ESOs will have the same optimal exercise time as the n shares of non-transferable constrained contingent claims by solving $\sup_{\tau\in\Gamma}\mathbb{E}\left[M\left(X^{x,\pi^*}(\tau)+nB(\tau),\tau\right)\right]$.

Similar to the proof of Theorem 1, then, we have

$$\frac{\partial\Psi(0,\hat{p},x)}{\partial\epsilon}=L'(x,n,0)-\frac{1}{\hat{p}}\sup_{\tau\in\Gamma}\left\{\mathbb{E}\left[B(\tau)M'\left(X^{x,\pi^*}(\tau)+nB(\tau),\tau\right)\right]\right\}=0.$$

The fair price is $\hat{p}(x,n,0)=\frac{\mathbb{E}\left[M'\left(X^{x,\pi^*}(\tau^*)+nB(\tau^*),\tau^*\right)B(\tau^*)\right]}{L'(x,n,0)}$. By Markov property,

$$\hat{p}(x_t,n,t)=\frac{\mathbb{E}\left[M'\left(X^{x_t,\pi^*}(\tau^*)+nB(\tau^*),\tau^*\right)B(\tau^*)|\mathcal{F}_t\right]}{L'(x_t,n,t)},\quad \forall x_t>0.$$

□

B.2. Proof of Theorem 4

Step1. Assume that for any stopping time $\tau\leq T$ and for any $v,\hat{v}$, we have that $\mathcal{X}_{v,\tau,\hat{v}}(y)\triangleq\mathbb{E}\left[H_{\hat{v}}(\tau,T)H_v(0,\tau)I\left(yH_{\hat{v}}(\tau,T)H_v(0,\tau)\right)+n(BS(\tau)-B(\tau))H_v(0,\tau)\right]$ is finite for every $y\in(0,\infty)$. Under this assumption, the function $\mathcal{X}_{v,\tau,\hat{v}}$: $(0,\infty)\to(0,\infty)$ is continuous and strictly decreasing, with $\mathcal{X}_{v,\tau,\hat{v}}(0+)=\infty$. We let $\mathcal{Y}_{v,\tau,\hat{v}}(\cdot)$ denote its inverse and introduce the following random variables.

For brevity, we denote $x^n = x + nBS(0)$, and we define, $\xi^{x^n}_{v,\tau,\grave{v}} \triangleq I\left(\mathcal{Y}_{v,\tau,\grave{v}}(x^n)H_{\grave{v}}(\tau,T)H_v(0,\tau)\right)$.

It is obvious from the above definition that

$$\mathbb{E}\left[\xi^{x^n}_{v,\tau,\grave{v}}H_{\grave{v}}(\tau,T)H_v(0,\tau) + n(BS(\tau) - B(\tau))H_v(0,\tau)\right] = x^n,$$

which guarantees that the initial outside wealth is x and the total initial wealth is x plus the Black-Scholes price of n shares of NTNH ESOs.

Please note that at the exercise time, the constraints are relaxed and there is a negative jump with respect to the wealth process. Holding n shares of non-transferable contingent claims is equivalent to holding the replicating portfolio. After being exercised, the shares become cash with the amount equal to the intrinsic value of the contingent claims and can be reinvested into any primary assets under the constraint of no shorting selling of the firm's stock. Then the NTNH constraint is relaxed. On the other hand, the amount of cash is equal to the intrinsic value of the contingent claim, which is lower than the market price of the replicating portfolio or the Black-Scholes price of the contingent claim. That means exercising the contingent claim will make the wealth portfolio lose value immediately. Hence, there is a trade-off between those two. Finding the optimal exercise time involves balancing the benefit and loss, and maximizing the current expected indirect utility at the exercise time. (To be clear, there is an increase in what we call outside wealth, but a decrease in total wealth.)

If the intrinsic value is equal to zero, early exercise will make the option holder worse off by losing time value in exchange for zero dollars of intrinsic value. Note that a non-zero intrinsic value could be reinvested without the non-transferability constraint.

In the auxiliary market, $\mathcal{M}_{v,\tau,\grave{v}}$, the dynamic of the total wealth process is

$$\begin{aligned} d\mathbb{X}_{v,\tau,\grave{v}}(t) &= \left(r(t)\mathbb{X}_{v,\tau,\grave{v}}(t)\right)dt + \mathbb{X}_{v,\tau,\grave{v}}(t)\big[(\delta(\grave{v}(t)) + \pi^*_{\grave{v}}(t)\grave{v}(t))1_{\{t>\tau\}} + (\delta(v(t)) + \\ &(\pi_v(t))^T v(t))1_{\{t<\tau\}}\big]dt + \mathbb{X}_{v,\tau,\grave{v}}(t)\pi_{v,\tau,\grave{v}}(t)\sigma(t)dW_0(t) - \left(BS(t) - B(t)\right)1_{\{t=\tau\}}, \end{aligned} \tag{B.4}$$

where $\grave{v}(t,\omega) \in \widetilde{K}$, for $\mathcal{E}\otimes P - a.e.(t,\omega) \in [0,\tau] \times \Omega$, and $v(t,\omega) \in \widetilde{K}$, for $\mathcal{E}\otimes P - a.e.(t,\omega) \in [\tau,T] \times \Omega$. Also, $\mathcal{E}$ represents the Effros Measure of t on the random closed set, e.g. $[0,\tau]$ and $[\tau,T]$ instead of the Lebesgue measure (See Molchanov (2005) page 25, Definition 2.1). For any y, the adjustment of the drift rate $f(t,\omega)$ and $\lambda(t,\omega)$ are the optimal $v(t,\omega)$ and $\grave{v}(t,\omega)$, respectively, to obtain the following infimum for a given stopping time τ:

$$\inf_{v,\grave{v}\in\mathfrak{D}} \tilde{J}(y;v,\grave{v}) \triangleq \mathbb{E}\left[\widetilde{U}\left(yH_{\grave{v}}(\tau,T)H_v(0,\tau)\right)\right] - y\mathbb{E}\left[H_v(0,\tau)n\left(BS(\tau) - B(\tau)\right)\right],$$

where $\mathfrak{D}$ is determined by $K(t,\omega)$ of (27).

Step2. We prove the following lemma. (Cf. Proposition 8.3 of Cvitanić and Karatzas (1992).)

Lemma B.1. For $\pi_\lambda(t,\omega) \in K(t,\omega)$ and $\pi_f(t,\omega) \in K(t,\omega)$, suppose

$$\delta(\lambda(t)) + (\pi_\lambda(t))^T\lambda(t) = 0, \tag{B.5}$$

$$\delta(f(t)) + \left(\pi_f(t)\right)^T f(t) = 0. \tag{B.6}$$

Then, where $\mathbb{X}_f(\cdot)$ and $\mathbb{X}_\lambda(\cdot)$ are the wealth dynamics in the auxiliary markets $\mathcal{M}_f$ and $\mathcal{M}_\lambda$, respectively, $\pi_f \in \mathcal{A}(x^n, K(0,\omega), 0, \tau\text{-})$ and $\pi_\lambda \in \mathcal{A}(\mathbb{X}_\lambda(\tau\text{+}), K(\tau\text{+},\omega), \tau\text{+}, T)$, the optimal portfolio processes for the constrained optimization problems (29) and (30) in the original market, satisfy for any given $v, \grave{v} \in \mathfrak{D}$, $\mathbb{E}[U(\xi_{f,\tau,\lambda})] \le V_{v,\tau,\grave{v}}(x) \triangleq ess\ sup_{\tau\in\Gamma}\ ess\ sup_{\pi_v\in\mathcal{A}_v(x,0,\tau)}\ \mathbb{E}\left[M\left(X_v^{x,\pi_v}(\tau) + nB(\tau), \tau\right)\right]$, where $M_{\grave{v}}(x_\tau, \tau) \triangleq esssup_{\pi_{\grave{v}}\in\mathcal{A}_{\grave{v}}(x_\tau,\tau,T)}\ \mathbb{E}\left[U(X_{\grave{v}}^{x_\tau,\pi_{\grave{v}}}(T))|\mathcal{F}_\tau\right]$.

Proof. For simplicity, we define $\pi_{v,\tau,\grave{v}} \triangleq \pi_v 1_{\{t<\tau\}} + \pi_{\grave{v}} 1_{\{t>\tau\}}$ and $\mathcal{M}_{v,\tau,\grave{v}} \triangleq \mathcal{M}_v 1_{\{t<\tau\}} + \mathcal{M}_{\grave{v}} 1_{\{t>\tau\}}$. By replacing v, $\grave{v}$ with f and λ, respectively, in (B.4), we have

$$\begin{aligned} d\mathbb{X}_{f,\tau,\lambda} &= \left(r(t)\mathbb{X}_{f,\tau,\lambda}(t)\right)dt \\ &\quad +\mathbb{X}_{f,\tau,\lambda}(t)\left[(\delta(\lambda(t)) + (\pi_\lambda(t))^T\lambda(t))1_{\{t>\tau\}} + \left(\delta(f(t)) + \left(\pi_f(t)\right)^T f(t)\right)1_{\{t<\tau\}}\right]dt \\ &\quad +\mathbb{X}_{f,\tau,\lambda}(t)\left(\pi_{f,\tau,\lambda}\right)^T(t)\sigma(t)dW_0(t) - \left(BS(t) - B(t)\right)1_{\{t=\tau\}}, \end{aligned}$$

$$X_{f,\tau,\lambda}(0) = x^n,\ \mathbb{X}_{f,\tau,\lambda}(T) = \xi_{f,\tau,\lambda}^{x_n}(T),$$

where $\xi_{f,\tau,\lambda}^{x_n} \triangleq I\left(\mathcal{Y}_{f,\tau,\lambda}(x)H_\lambda(\tau,T)H_f(0,\tau)\right)$, and $\mathcal{Y}_{f,\tau,\lambda}(x)$ is the inverse function of

$$\mathcal{X}_{f,\tau,\lambda}(y) = \mathbb{E}\left[H_\lambda(\tau,T)H_f(0,\tau)I\left(yH_\lambda(\tau,T)H_f(0,\tau)\right) + n(BS(\tau) - B(\tau))H_f(0,\tau)\right].$$

In different markets, original or auxiliary, the drift rate is different. The risk-free rate is different as well. Although the Black-Scholes price $BS(t)$ at any time t is independent of the drift rate at time t, it is a function of the risk-free rate at time t and the spot price at time t. Standing at any time before t, the spot price at time t is a function of the drift rate of the underlying asset. Hence, it is necessary for us to specify in which market the Black-Scholes price will be taken in the above pricing model. And $BS(t)$ in this research is the Black-Scholes price in the original market. When (B.5) and (B.6) are satisfied, the wealth dynamic in the auxiliary market $\mathcal{M}_{f,\tau,\lambda}$ is as follows:

$$d\mathbb{X}_{f,\tau,\lambda}(t) = \left(r(t)\mathbb{X}_{f,\tau,\lambda}(t)\right)dt + \mathbb{X}_{f,\tau,\lambda}(t)\left(\pi_{f,\tau,\lambda}(t)\right)^T\sigma(t)dW_0(t) - n\left(BS(t) - B(t)\right)1_{\{t=\tau\}},$$

$$\mathbb{X}(0) = x^n.$$

Compare it with the wealth dynamic in the original market,

$$d\,\mathbb{X}(t) = \left(\,\mathbb{X}(t)r(t)\right)dt + \,\mathbb{X}(t)\left(\pi(t)\right)^T\sigma(t)dW_0(t) - n\left(BS(t) - B(t)\right)1_{\{t=\tau\}},\quad \mathbb{X}(0) = x^n.$$

It is obvious that $\mathbb{X}_{f,\tau,\lambda}(t)$ is also a wealth process corresponding to $\pi_{f,\tau,\lambda}(t)$ in the original market $\mathcal{M}$. Furthermore, from this and $\pi_\lambda(t,\omega)$ and $\pi_f(t,\omega) \in K(t,\omega)$of (27), we conclude that $\pi_f \in \mathcal{A}(x^n, K(0,\omega), 0, \tau\text{-})$, $\pi_\lambda \in \mathcal{A}(\mathbb{X}_{f,\tau,\lambda}(\tau\text{+}), K(\tau\text{+},\omega), \tau\text{+}, T)$, and $V_{f,\tau,\lambda}(x) = \mathbb{E}\left[U\left(\xi_{f,\tau,\lambda}\right)\right] \le$

$L(x,n,0)$, where $L(x,n,0)$ is defined in equation (B.3). By the fact that $\left[\left(\delta(\lambda(t)) + (\pi_\lambda(t))^T\lambda(t)\right)1_{\{t\geq\tau\}} + \left(\delta(f(t)) + \left(\pi_f(t)\right)^T f(t)\right)1_{\{t<\tau\}}\right] \geq 0$, for any given stopping time τ and for any given $v, \grave{v}, \mathbb{X}_{v,\tau,\grave{v}}(t) > \mathbb{X}(t) \geq 0, \forall 0 < t < T$. Hence, we have

$$\mathbb{E}\left(M(\mathbb{X}_{v,\tau,\grave{v}}(\tau),\tau)\right) > \mathbb{E}(M(\mathbb{X}(\tau),\tau) \tag{B.7}$$

$$\mathbb{E}\left(U\left(\mathbb{X}_{v,\tau,\grave{v}}(T)\right)\middle|\mathcal{F}_\tau\right) > \mathbb{E}(U(\mathbb{X}(T))|\mathcal{F}_\tau) \tag{B.8}$$

Then, $\forall v, \grave{v} \in \mathfrak{D}$, from (B.7), we have $\mathcal{A}(x^n, K(0,\omega), \ 0, \tau\text{-}) \subset \mathcal{A}_v(x^n, 0, \tau\text{-})$, and from (B.8), we have $\mathcal{A}(x_{\tau+}, K(\tau+,\omega), \ \tau+, T) \subset \mathcal{A}_{\grave{v}}((\mathbb{X}_{v,\tau,\grave{v}}(\tau+), \tau+, T)$ and, further, $L(x,n,0) \leq V_{v,\tau,\grave{v}}(x)$. Because $V_{f,\tau,\lambda}(x)$ is a special case of $V_{v,\tau,\grave{v}}(x)$, it is trivial that $L(x,n,0) \leq V_{f,\tau,\lambda}(x)$. Hence, we get $L(x,n,0) = V_{f,\tau,\lambda}(x)$, which indicates that the optimal unconstrained portfolio process $\pi_{f,\tau,\lambda}(t)$ in the auxiliary market $\mathcal{M}_{f,\tau,\lambda}$, *is* the optimal constrained portfolio *process* in the original market, $\mathcal{M}$.

Let $X_{v,\tau,\grave{v}}^{x^n,\pi_{f,\tau,\lambda}}$ be the wealth process with initial wealth x^n, portfolio process $\pi_{f,\tau,\lambda}$ in market $\mathcal{M}_{v,\tau,\grave{v}}$. We have $\mathbb{X}_{v,\tau,\grave{v}}^{x^n,\pi_{f,\tau,\lambda}}(t) \geq \mathbb{X}_{f,\tau,\lambda}(t) > 0, \ \forall\, 0 < t < T$. Thus, $\pi_f \in \mathcal{A}_v(x_n, 0, \tau\text{-})$ before τ, and $\pi_\lambda \in \mathcal{A}_{\grave{v}}((\mathbb{X}_{v,\tau,\grave{v}}(\tau+), \tau+, T)$ after τ, which indicates, $V_{f,\tau,\lambda}(x) \leq V_{v,\tau,\grave{v}}(x)$.

Step 3. The proof of the theorem is based on the following logic. Assume we are able to solve the corresponding optimal portfolio process $\pi_{v,\tau,\grave{v}}$ for any auxiliary market $\mathcal{M}_{v,\tau,\grave{v}}$. If the auxiliary market is not arbitrarily chosen, but satisfies the condition

$\left[\left(\delta(f(t)) + \left(\pi_f(t)\right)^T f(t)\right)1_{\{t<\tau\}} + (\delta(\lambda(t)) + (\pi_\lambda(t))^T\lambda(t))1_{\{t\geq\tau\}}\right] = 0$,

then the optimal portfolio process for the unconstrained auxiliary market is the optimal constrained original market. But, do we know how to solve the optimal portfolio process in an unconstrained auxiliary market $\mathcal{M}_{f,\tau,\lambda}$? Because in this case, there is a negative jump at the optimal exercise time, the problem is not exactly the same as the one in Cvitanić and Karatzas (1992).

To answer that question, we need to make sure $\mathbb{E}\left(U(\xi_{f,\tau,\lambda}^{x^n})\right)$ is well defined first, namely $\mathbb{E}[U^-(\xi_{f,\tau,\lambda}^{x^n})] < \infty$. According to the useful inequality $U(I(y)) \geq U(x) + y[I(y) - x]$, we have,

$$U(\xi_{f,\tau,\lambda}^{x^n}) \geq U(1) + \mathcal{Y}_{f,\tau,\lambda}(x^n)H_\lambda(\tau,T)H_f(0,\tau)[\xi_{f,\tau,\lambda}^{x^n} - 1],$$

$$-U(\xi_{f,\tau,\lambda}^{x^n}) \leq -U(1) - \mathcal{Y}_{f,\tau,\lambda}(x^n)H_\lambda(\tau,T)H_f(0,\tau)[\xi_{f,\tau,\lambda}^{x^n} - 1],$$

$$max(-U(\xi_{f,\tau,\lambda}^{x^n}), 0) \leq max(-U(1) - \mathcal{Y}_{f,\tau,\lambda}(x^n)H_\lambda(\tau,T)H_f(0,\tau)[\xi_{f,\tau,\lambda}^{x^n} - 1], 0).$$

If $U(\xi_{f,\tau,\lambda}^{x^n}) \geq 0$, then, $U^-(\xi_{f,\tau,\lambda}^{x^n}) = 0 \leq |U(1)| + \mathcal{Y}_{f,\tau,\lambda}(x^n)H_\lambda(\tau,T)H_f(0,\tau)$.

If $U(\xi_{f,\tau,\lambda}^{x^n}) \leq 0$, then,

$U^-(\xi_{f,\tau,\lambda}^{x^n}) = -U(\xi_{f,\tau,\lambda}^{x^n}) \leq -U(1) + \mathcal{Y}_{f,\tau,\lambda}(x^n)H_\lambda(\tau,T)H_f(0,\tau) - \mathcal{Y}_{f,\tau,\lambda}(x^n)H_\lambda(\tau,T)H_f(0,\tau)\xi_{f,\tau,\lambda}^{x^n}$

$$\leq |U(1)| + \mathcal{Y}_{f,\tau,\lambda}(x^n) H_\lambda(\tau,T) H_f(0,\tau).$$

Hence we have $\mathbb{E}\left(U(\xi_{f,\tau,\lambda}^{x^n})\right) \leq |U(1)| + \mathcal{Y}_{f,\tau,\lambda}(x^n)\mathbb{E}\left(H_\lambda(\tau,T)H_f(0,\tau)\right) < \infty$, where the last inequality is proved as follows. First, note that $\mathcal{Y}_{f,\tau,\lambda}(x^n)$ is a monotonically continuous function and, hence, $\forall x^n > 0$ belongs to a closed interval, $\mathcal{Y}_{f,\tau,\lambda}(x^n)$ and is bounded. Also,

$$H_\lambda(\tau,T)H_f(0,\tau)$$

$$= exp\left\{-\int_\tau^T r_\lambda(s) + \theta_\lambda(s)dW(s) + \frac{1}{2}\|\theta_\lambda(s)\|^2 ds\right\} exp\left\{-\int_0^\tau r_f(s) + \theta_f(s)dW(s) + \frac{1}{2}\left\|\theta_f(s)\right\|^2 ds\right\}$$

$$= exp\left\{-\int_\tau^T r(s) + \delta\big(\lambda(s)\big) + \frac{1}{2}\|\theta(s) + \sigma^{-1}(s)\lambda(s)\|^2 ds\right\} exp\left\{-\int_0^\tau r(s) + \delta\big(f(s)\big) + \frac{1}{2}\|\theta(s) + \sigma^{-1}(s)f(s)\|^2 ds\right\}.$$

In this case, for $t < s \leq \tau$, $K(s,\omega) = [\frac{n\phi(d_1)S_1(s)}{\mathbb{X}(s)}, \infty) \times (-\infty,\infty)^{d-1}$, and $\delta(v(s)) = -\frac{n\phi(d_1)S_1(t)}{\mathbb{X}(t)} f_1(s)$ on $\widetilde{K} = [0,\infty) \times \{0\}^{d-1}$; for $\tau < s < T$, $K(s,\omega) = [0,\infty) \times (-\infty,\infty)^{d-1}$, and $\delta(v(s)) = 0$, on $\widetilde{K} = [0,\infty) \times \{0\}^{d-1}$. Then,

$H_\lambda(\tau,T)H_f(0,\tau) =$
$exp\left\{-\int_\tau^T r(s) + \frac{1}{2}\|\theta(s) + \sigma^{-1}(s)\lambda(s)\|^2 ds\right\} exp\left\{-\int_0^\tau r(s) - \frac{n\phi(d_1)S_1(s)}{\mathbb{X}(s)} f_1(s) + \frac{1}{2}\|\theta(s) + \sigma^{-1}(s)f(s)\|^2 ds\right\}$,

where f_1 is the first element of the vector f. Now,

$$r(s) - \frac{n\phi(d_1)S_1(s)}{\mathbb{X}(s)} f_1(s) + \frac{1}{2}\|\theta(s) + \sigma^{-1}(s)f(s)\|^2$$

$$= r(s) - \frac{n\phi(d_1)S_1(s)}{\mathbb{X}(s)} f_1(s) + \frac{1}{2}\left\{\left(\theta_1(s) + \sigma_{11}^{-1}(s) * f_1(s)\right)^2 + \left(\theta_2(s) + \sigma_{21}^{-1}(s) * f_1(s)\right)^2 + \cdots + \left(\theta_d(s) + \sigma_{d1}^{-1}(s) * f_1(s)\right)^2\right\}$$

$$= r(s) - \frac{n\phi(d_1)S_1(s)}{\mathbb{X}(s)} f_1(s) + \frac{1}{2}\left\{\left(\theta_1(s)\right)^2 + \left(\theta_2(s)\right)^2 + \cdots \left(\theta_d(s)\right)^2 + 2 * \theta_1(s) * \sigma_{11}^{-1}(s) * f_1(s) + 2 * \theta_2(s) * \sigma_{21}^{-1}(s) * f_1(s) + \cdots + 2 * \theta_d(s) * \sigma_{d1}^{-1}(s) * f_1(s) + \left(\sigma_{11}^{-1}(s) * f_1(s)\right)^2 + \left(\sigma_{21}^{-1}(s) * f_1(s)\right)^2 + \cdots + \left(\sigma_{d1}^{-1}(s) * f_1(s)\right)^2\right\},$$

$$= Af_1(s)^2 + Bf_1(s) + C,$$

where

$$A \equiv \left(\sigma_{11}^{-1}(s)\right)^2 + \left(\sigma_{21}^{-1}(s)\right)^2 + \cdots + \left(\sigma_{d1}^{-1}(s)\right)^2 > 0,$$

$$B \equiv \theta_1(s)\sigma_{11}^{-1}(s) + \theta_2(s)\sigma_{21}^{-1}(s) + \cdots + \theta_d(s)\sigma_{d1}^{-1}(s) - \frac{n\phi(d_1)S_1(s)}{\mathbb{X}(s)},$$

$$C = \frac{1}{2}\left\{\left(\theta_1(s)\right)^2 + \left(\theta_2(s)\right)^2 + \cdots \left(\theta_d(s)\right)^2\right\} + r(s).$$

Hence, $r(s) - af_1(s) + \frac{1}{2}\|\theta(s) + \sigma^{-1}(s)f(s)\|^2$ is bounded from below by $\mathbb{B} \equiv -(B^2 - 4AC)/4A$, which is a random variable. Let us write $\grave{B} \equiv \theta_1(s)\sigma_{11}^{-1}(s) + \theta_2(s)\sigma_{21}^{-1}(s) + \cdots + \theta_d(s)\sigma_{d1}^{-1}(s)$, and note that $\grave{B}$ is deterministic.

Since $0 < \frac{n\phi(d_1)S_1(s)}{\mathbb{X}(s)} < n$, we have

$$B^2 = \left(\grave{B} - \frac{\phi(d_1)S_1(s)}{\mathbb{X}(s)}\right)^2 = \grave{B}^2 - 2\grave{B}\frac{n\phi(d_1)S_1(s)}{\mathbb{X}(s)} + \left(\frac{n\phi(d_1)S_1(s)}{\mathbb{X}(s)}\right)^2$$
$$\leq max\left(\grave{B}^2 - 2\grave{B}n + n^2, \grave{B}^2 + n^2\right).$$

Denote $max\left(\grave{B}^2 - 2\grave{B}n + n^2,\ \grave{B}^2 + n^2\right)$ as $\grave{\mathbb{B}}$, hence, $r(s) - af_1(s) + \frac{1}{2}\|\theta(s) + \sigma^{-1}(s)f(s)\|^2$ is bounded from below by $\grave{\mathbb{B}}$, a deterministic variable. Similarly, $r(s) + \frac{1}{2}\|\theta(s) + \sigma^{-1}(s)\lambda(s)\|^2$ is a bounded from below by a deterministic variable, $\widehat{\mathbb{B}} \equiv -\left(\hat{B}^2 - 4AC\right)/4A$. Then,

$\mathbb{E}\left(H_\lambda(\tau,T)H_f(0,\tau)\right) = \mathbb{E}\left(exp\left\{-\int_\tau^T r(s) + \frac{1}{2}\|\theta(s) + \sigma^{-1}(s)\lambda(s)\|^2\, ds\right\} exp\left\{-\int_0^\tau r(s) - \frac{n\phi(d_1)S_1(s)}{\mathbb{X}(s)} f_1(s) + \frac{1}{2}\|\theta(s) + \sigma^{-1}(s)f(s)\|^2\, ds\right\}\right) < exp\{\grave{\mathbb{B}} * (T-\tau)\} exp\{\widehat{\mathbb{B}} * (T-\tau)\} < \infty.$

Now, consider any arbitrary portfolio process, $\pi(t) \in \mathcal{A}_v(x, 0, \tau\text{-})$ for $t \in [0,\tau)$and $\pi(t) \in \mathcal{A}_{\grave{v}}(x_{\tau+}, \tau+, T)$ for $t \in (\tau, T]$. Then,

$$U(\xi_{v,\tau,\grave{v}}^{x^n}) \geq U\left(\mathbb{X}^{x^n,\pi}(T)\right) + \mathcal{Y}_{v,\tau,\grave{v}}(x^n)H_{\grave{v}}(\tau,T)H_v(0,\tau)[\xi_{v,\tau,\grave{v}}^{x^n} - \mathbb{X}^{x^n,\pi}(T)],$$

almost surely; therefore,

$$\mathbb{E}[U(\xi_{v,\tau,\grave{v}}^{x^n})] \geq \mathbb{E}\left[U\left(\mathbb{X}^{x^n,\pi}(T)\right)\right] + \mathcal{Y}_{v,\tau,\grave{v}}(x^n)\mathbb{E}\left[H_{\grave{v}}(\tau,T)H_v(0,\tau)\left(\xi_{\grave{v}}^{x^n} - \mathbb{X}^{x^n,\pi}(T)\right)\right]$$
$$= \mathbb{E}\left[U\left(\mathbb{X}^{x^n,\pi}(T)\right)\right] + \mathcal{Y}_{v,\tau,\grave{v}}(x^n)$$
$$\times \mathbb{E}\left[H_{\grave{v}}(\tau,T)H_v(0,\tau)\left(\xi_{\grave{v}}^{x^n} - \mathbb{X}^{x^n,\pi}(T)\right) + H_v(0,\tau)n(BS(\tau) - B(\tau))\right.$$
$$\left.- H_v(0,\tau)n(BS(\tau) - B(\tau))\right]$$
$$= \mathbb{E}\left[U\left(\mathbb{X}^{x^n,\pi}(T)\right)\right] + \mathcal{Y}_{v,\tau,\grave{v}}(x^n)\mathbb{E}\left[x^n - \left(H_{\grave{v}}(\tau,T)H_v(0,\tau)\mathbb{X}^{x^n,\pi}(T) + H_v(0,\tau)n(BS(\tau) - B(\tau))\right)\right]$$
$$\geq \mathbb{E}\left[U\left(\mathbb{X}^{x^n,\pi}(T)\right)\right].$$

We can see that the above proof depends on the equality $\mathbb{E}[\xi_{\grave{v}}^{x^n}H_{\grave{v}}(\tau,T)H_v(0,\tau)+H_v(0,\tau)n(BS(\tau) - B(\tau))] = x^{n,}$ which follows from the definition of $\xi_{\grave{v}}^{x^n}$, and the inequality $\mathbb{E}[H_{\grave{v}}(\tau,T)H_v(0,\tau)\mathbb{X}^{x^n,\pi}(T) + H_v(0,\tau)n(BS(\tau) - B(\tau))] \leq x^n$, in addition to the convexity of the direct and indirect utilities. Because in the auxiliary market $\mathcal{M}_{v,\tau,\dot{v}}$,, the dynamic of the wealth process is

$$d\mathbb{X}_{v,\tau,\dot{v}}(t) = \left(r(t)\mathbb{X}_{v,\tau,\dot{v}}(t)\right)dt$$
$$+ \mathbb{X}_{v,\tau,\dot{v}}(t)\left[(\delta(\dot{v}(t)) + \pi_{\dot{v}}^*(t)\dot{v}(t))1_{\{t>\tau\}} + (\delta(v(t)) + \pi_v^*(t)v(t))1_{\{t<\tau\}}\right]dt$$
$$+ \mathbb{X}_{v,\tau,\dot{v}}(t)\pi_{v,\tau,\dot{v}}{}^*(t)\sigma(t)dW_0(t) - \left(BS(t) - B(t)\right)1_{\{t=\tau\}}.$$

Without negative jumps, $-\left(BS(t) - B(t)\right)$, it is well-known that $H_{\dot{v}}(\tau,T)H_v(0,\tau)\mathbb{X}^{x^n,\pi}(T)$ is a super-martingale, hence by adding the negative jump, again it becomes a super-martingale.

Step 4. So far, we have studied for an arbitrary exercise time τ how to find the optimal portfolio process. The next question is how to determine the optimal exercise time τ^*.

Thanks to optimal stopping time theory, we obtain

$\tau^* = inf\{t \le u \le T\colon J(u, X_u) = M(X_u + nB(u), u)\}.$

Step 5. In the final step, we prove the equivalent martingale-based expression of the subjective a NTNH ESO price. We know that

$$\mathcal{X}_{f,\tau,\lambda}(y) \triangleq \mathbb{E}\left[H_\lambda(\tau,T)H_f(0,\tau)I\left(yH_\lambda(\tau,T)H_f(0,\tau)\right) + H_f(0,\tau)n(BS(\tau) - B(\tau))\right]$$

is finite, for every $y \in (0,\infty)$. Under this assumption, the function $\mathcal{X}_{f,\tau,\lambda}\colon (0,\infty) \to (0,\infty)$ is continuous and strictly decreasing, with $\mathcal{X}_{f,\tau,\lambda}(0+) = \infty$ and $\mathcal{X}_{f,\tau,\lambda}(\infty) = 0$. We let $\mathcal{Y}_{f,\tau,\lambda}(\cdot)$ denote its inverse and introduce the random variables

$\xi_{f,\tau,\lambda}^{x^n} \triangleq I\left(\mathcal{Y}_{f,\tau,\lambda}(x^n)H_\lambda(\tau,T)H_f(0,\tau)\right),$

$U'\left(\xi_{f,\tau,\lambda}^{x^n}\right) \triangleq U'\left(I\left(\mathcal{Y}_{f,\tau,\lambda}(x^n)H_\lambda(\tau,T)H_f(0,\tau)\right)\right) = \mathcal{Y}_{f,\tau,\lambda}(x^n)H_\lambda(\tau,T)H_f(0,\tau).$

We also know that $\xi_{f,\tau,\lambda}^{x^n} \triangleq \xi_\lambda^{X_*^{x^n}(\tau)} = I\left(\mathcal{Y}_\lambda\left(X_*^{x^n}(\tau) - n\left(BS(\tau) - B(\tau)\right)\right)H_\lambda(\tau,T)\right)$, where $X_*^{x^n}(\tau)$ denotes $X_{f,\tau,\lambda}^{x^n,\pi^*_{f,\tau,\lambda}}$, and $U'\left(\xi_{f,\tau,\lambda}^{x^n}\right) = U'\left(\xi_\lambda^{X_*^{x^n}(\tau)}\right) = \mathcal{Y}_\lambda\left(X_*^{x^n}(\tau) - n(BS(\tau) - B(\tau))\right)H_\lambda(\tau,T).$

We have $\mathcal{Y}_{f,\tau,\lambda}(x_n)H_\lambda(\tau,T)H_f(0,\tau) = \mathcal{Y}_\lambda\left(X_*^{x^n}(\tau) - n(BS(\tau) - B(\tau))\right)H_\lambda(\tau,T),$

$$H_f(0,\tau) = \frac{\mathcal{Y}_\lambda\left(X_*^{x^n}(\tau) - n(BS(\tau) - B(\tau))\right)}{\mathcal{Y}_{f,\tau,\lambda}(x^n)} = \frac{M'\left(X^{x,\pi^*}(\tau) + nB(\tau),\tau\right)}{L'(x,n,0)},$$

since the payoff is the sum of the NTNH American ESOs and outside wealth. The price of the NTNH American Contingent Claim is defined as its expectation of the payoff at the optimal exercise time, discounted by the SDF. Hence, $\hat{p}(x,n,0)\ \mathbb{E}\left[H_f(\tau^*)B(\tau^*)\right] = \mathbb{E}^f\left[\gamma_f(\tau^*)B(\tau^*)\right].$

Due to the Markov property of stock price, we have

$$\hat{p}(x_t,n,t) = \frac{\mathbb{E}\left[H_f(\tau^*)B(\tau^*)|\mathcal{F}_t\right]}{H_f(t)} = \frac{\mathbb{E}^f\left[\gamma_f(\tau^*)B(\tau^*)|\mathcal{F}_t\right]}{\gamma_f(t)}.$$

□

Appendix C. Proof of Proposition 2

The primitive objects of a stochastic control problem are the set $Z \subset \Re^3$ of states $Z(t) \triangleq (\mathbb{X}(t), N(t), S_1(t))$, where $\mathbb{X}(t)$ is the total wealth process, $N(t)$ is the accumulated number of ESO shares that have been exercised, and $S_1(t)$ is the stock's price.

In the Merton Model, the total wealth is the only state variable. The reason the primary assets' price vector $S(t)$ is not a state variable is that the terminal wealth is sufficient to determine the utility, and terminal wealth is linear in $S(t)$. Also, the control variable portfolio process does not affect the stock price of the next period.

In our partial exercise model, as the stock price affects whether the option holder will exercise the option at any particular time in the post-vesting period before the expiration date, the stock price is necessarily be included in the state variable set. In the case that the NTNH ESO is European, the stock price is still a necessary state variable because it determines whether the investor will exercise the option at the maturity. Additionally, the exercise payoff will determine the terminal wealth at the same time.

The state variable set also includes the cumulative number of shares of exercised options $\widehat{N}(t)$. On one hand, $n - \widehat{N}(t)$ is the maximum of number of shares that can be exercised, which determines the opportunity set of the control variable $\dot{n}(t)$. On the other hand, the control variable $\dot{n}(t)$ determines the cumulative number of shares of exercised options.

The dynamics of the system are as follows:

$$\begin{aligned} d\mathbb{X}(t) = \sum_{i=1}^{d} \pi_{f_{\widehat{N}}}^{i}(t)\,\mathbb{X}(t)\left\{b_i(t)dt + \sum_{j=1}^{d}\sigma_{ij}(t)dW_j(t)\right\} + \left\{1 - \sum_{i=1}^{d}\pi_{f_{\widehat{N}}}^{i}(t)\right\}\mathbb{X}(t)r(t)dt \\ - \dot{n}(t)\big(BS(t) - B(t)\big)dt, \end{aligned} \tag{C.1}$$

$$\mathbb{X}(0) = x^n \triangleq x + nBS(0),$$

$$dN(t) = \dot{n}(t)dt, \qquad N(0) = \int_0^{t=0} \dot{n}(s)ds = 0, \tag{C.2}$$

$$\frac{dS_1(t)}{S_1(t)} = b_1 dt + \sigma_{11} dW_1(t), S_1(0) = s_1. \tag{C.3}$$

The unique solution is taken as $\mathbb{X}^{*,\dot{n}}(t)$, where $\pi_{f_{\widehat{N}}}^{*}(t)$ is the optimal portfolio process given the constraint, and the drift adjustment $v(t)$ with respect to the auxiliary market $\mathcal{M}_v$ associated with the constraint $K(t,\omega)$ in (34) is $f_{\widehat{N}}(t)$.

Let us summarize some terminology and notation.

- The constraint on the state variable $Z(t) = (\mathbb{X}(t), N(t), S_1(t))$:
  $Z(t_+) \in \mathcal{Z}(t_+) = \big([0,\infty), [\widehat{N}(t), n], [0,\infty)\big), \forall t \in [0,T].$
- The Control variable $\dot{n}(t)$: the exercise rate at any time depends on the current stock price; hence, it is a function of the sample path, namely a random variable. Sometimes we denote it

as $\dot{n}(t,\omega)$.

- The set of feasible controls: $\mathcal{N} \triangleq \{\dot{n}(t,\omega) | \dot{n}(t,\omega): [0,T] \times \Omega \to \mathcal{R}^+, \dot{n}(t,\omega)\ is\ \{\mathcal{F}_t\}\text{-}$ $progessively\ measurable, \forall\ 0 \le t \le T, almost\ surely;\ \int_0^T \dot{n}(t,\omega)dt \le n\Big\}$.
- A running reward function: $f \equiv 0$.
- A terminal reward function $R: \mathcal{Z} \longrightarrow \Re$. (In our case, $R(Z) = U(\mathbb{X})$).
- The cost functional measuring the performance of the control: $\mathbb{E}\left[\int_0^T f(t, Z(t), \dot{n}(t))dt + F(Z(T))\right] = \mathbb{E}\left[U\left(\mathbb{X}^{*,\dot{n}}(T)\right)\right]$, where $\mathbb{X}^{*,\dot{n}}(T) \equiv \mathbb{X}^{x^n,\pi^*,\dot{n}}(T)$.
- The exercise rate $\dot{n}(\cdot)$ is called an *admissible control*, and $\left(Z(\cdot), \dot{n}(\cdot)\right)$ is called an admissible pair, if $\dot{n}(\cdot) \in \mathcal{N}$, $Z(\cdot)$ are the unique solutions of equations (C.1) to (C.3) under $\dot{n}(\cdot)$, and the state constraint is satisfied.
- The set of admissible controls is denoted as $\mathcal{V}_{ad}[0,T]$.
- The function $t \to f\left(t, Z(t), \dot{n}(t)\right)$ is in $\mathcal{L}^1(0,T)$. (In our case, $f \equiv 0$.)
- A controlled drift function $g: \mathcal{N} \times \mathcal{Z} \longrightarrow \Re^3$. In our case, $g(\dot{n}(t), Z(t)) = \Big(\Big(\mathbb{X}^{*,\dot{n}}(t)r(t) + \sum_{i=1}^d \pi_{f_N}^i(t)\, \mathbb{X}^{*,\dot{n}}(t)\left(b_i(t) - r(t)\right) - \dot{n}(t)\left(BS(t) - B(t)\right)\Big), \dot{n}(t),\ b_1(t)S_1(t)\Big)^T$.
- A controlled diffusion function h: $\mathcal{N} \times \mathcal{Z} \longrightarrow \Re^{3\times d}$: $h(\dot{n}(t), Z(t)) = \left(X^{*,\dot{n}}(t)\left(\pi(t)\right)^T \sigma(t), 0, \sigma_1 S_1(t)\right)^T$.

We assume that the primitives $(\mathcal{N}, \mathcal{Z}, \mathcal{V}_{ad}[0,T], g, h, f, R\ )$ are as above. Then, given any initial state $z \in \mathcal{Z}$, the utility of any admissible control $\dot{n}(t)$ is well defined as $V^q(z) = \mathbb{E}\left[R(Z_T^{\dot{n}})\right] = \mathbb{E}\left[U(\mathbb{X}_T^{*,\dot{n}})\right]$. The indirect utility at time $t = 0$ at initial state z is then $\mathbb{J}(z,0) = sup_{\dot{n}(t) \in \mathcal{V}_{ad}[0,T]}\ \mathbb{E}\left[U(\mathbb{X}_T^{*,\dot{n}})\right]$. We postulate that $\mathbb{J}(z,0)$ is in $C^{2,1}(\mathcal{Z},[0,T])$. According to dynamic programming principle,

$$\mathbb{J}(z,0) = \mathbb{J}\left(\left(\mathbb{X}^{*,\dot{n}}(0), \widehat{N}(0), S_1(0)\right), 0\right) = sup_{\dot{n} \in \mathcal{V}_{ad}[0,t]}\ \mathbb{E}\left[\mathbb{J}(Z_t, t)\right].$$

By Taylor's theorem, we have

$$\mathbb{J}(z,dt) = \mathbb{J}(z,0) + \sup_{\dot{n} \in \mathcal{V}_{ad}[0,T]} \mathbb{E}\{\mathbb{J}_Z(Z_t,t)g(\dot{n},Z_t)dt + \mathbb{J}_t(Z_t,t)dt$$
$$+\frac{1}{2}tr[h(\dot{n},Z_t)h(\dot{n},Z_t)'\mathbb{J}_{ZZ}(Z_t,t)]dZ \cdot dZ + O((dt)^2)\Big\}.$$

Hence we obtain the Bellman equation

$$0 = \frac{\partial \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t), \widehat{N}(t), S_1(t), t)}{\partial t}dt + \frac{\partial \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t), \widehat{N}(t), S_1(t), t)}{\partial \mathbb{X}^{*,\dot{n}}(t)}d\mathbb{X}^{*,\dot{n}}(t) + \frac{\partial \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t), \widehat{N}(t), S_1(t), t)}{\partial \widehat{N}(t)}d\widehat{N}(t) +$$
$$\frac{\partial \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t), \widehat{N}(t), S_1(t), t)}{\partial S_1(t)}dS_1(t) + \frac{1}{2}\frac{\partial^2 \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t), \widehat{N}(t), S_1(t), t)}{\partial^2 \mathbb{X}^{*,\dot{n}}(t)}d\mathbb{X}^{*,\dot{n}}(t)d\mathbb{X}^{*,\dot{n}}(t) +$$
$$\frac{\partial^2 \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t), \widehat{N}(t), S_1(t), t)}{\partial \mathbb{X}^{*,\dot{n}}(t)\partial \widehat{N}(t)}d\mathbb{X}^{*,\dot{n}}(t)d\widehat{N}(t) + \frac{1}{2}\frac{\partial^2 \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t), \widehat{N}(t), S_1(t), t)}{\partial^2 \widehat{N}(t)}d\widehat{N}(t)d\widehat{N}(t) +$$

$\frac{1}{2}\frac{\partial^2 \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t),\widehat{N}(t),S_1(t),t)}{\partial^2 S_1(t)} dS_1(t)dS_1(t) + \frac{\partial^2 \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t),\widehat{N}(t),S_1(t),t)}{\partial \mathbb{X}^{*,\dot{n}}(t)\partial S_1(t)} d\mathbb{X}^{*,\dot{n}}(t)dS_1(t) +$

$\frac{\partial^2 \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t),\widehat{N}(t),S_1(t),t)}{\partial \widehat{N}(t)\partial S_1(t)} d\widehat{N}(t)dS_1(t).$

We know that

$d\widehat{N}(t) = \dot{n}(t)dt,\ \ d\widehat{N}(t)d\widehat{N}(t) = (\dot{n}(t))^2(dt)^2 \in o(dt),\ \ d\mathbb{X}^{*,\dot{n}}(t)d\widehat{N}(t) \in o(dt)$

$dS_1(t)d\widehat{N}(t) \in o(dt), dS_1(t)dS_1(t) = (b_1S_1dt + \sigma_{11}S_1dW_1)(b_1S_1dt + \sigma_{11}S_1dW_1) = \sigma_{11}^2 S_1^2 dt$

$d\mathbb{X}^{*,\dot{n}}(t)d\mathbb{X}^{*,\dot{n}}(t) =$

$\left(\sum_{i=1}^d \pi^i_{f_{\widehat{N}}}(t)\, \mathbb{X}^{*,\dot{n}}(t)\{\sum_{j=1}^d \sigma_{ij}(t)dW_j(t)\}\right)\left(\sum_{i=1}^d \pi^i_{f_{\widehat{N}}}(t)\, \mathbb{X}^{*,\dot{n}}(t)\{\sum_{j=1}^d \sigma_{ij}(t)dW_j(t)\}\right) =$

$\sum_{i=1}^d\sum_{j=1}^d\sum_{m=1}^d\sum_{n=1}^d\left\{\left(\mathbb{X}^{*,\dot{n}}(t)\right)^2 \pi^i_{f_{\widehat{N}}}(t)\pi^m_{f_{\widehat{N}}}(t)\sigma_{ij}(t)\sigma_{mn}(t)dt\right\} = \left(\mathbb{X}^{*,\dot{n}}(t)\right)^2 \pi'(t)\sigma'(t)\sigma(t)\pi(t)d$

$d\mathbb{X}^{*,\dot{n}}(t)dS_1(t) = \left(\sum_{i=1}^d \pi^i_{f_{\widehat{N}}}(t)\, \mathbb{X}^{*,\dot{n}}(t)\{\sum_{j=1}^d \sigma_{ij}(t)dW_j(t)\}\right)(b_1S_1dt + \sigma_{11}S_1dW_1) =$

$\mathbb{X}^{*,\dot{n}}(t)\left\{\sum_{i=1}^d \pi^i_{f_{\widehat{N}}}(t)\sigma_{i1}(t)\right\}(\sigma_{11}S_1dt).$

Taking expectations on both sides and dividing both sides by dt, we get

$0 = sup_{\dot{n}\in\mathcal{V}_{ad}[0,T]}\left\{\frac{\partial \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t),\widehat{N}(t),S_1(t),t)}{\partial t} + \frac{\partial \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t),\widehat{N}(t),S_1(t),t)}{\partial X^{*,\dot{n}}(t)}\left(\mathbb{X}^{*,\dot{n}}(t)r(t) +\right.\right.$

$\left.\sum_{i=1}^d \pi^i_{f_N}(t)\, \mathbb{X}^{*,\dot{n}}(t)\big(b_i(t) - r(t)\big) - \dot{n}(t)\big(BS(t) - B(t)\big)\right) +$

$\frac{\partial \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t),\widehat{N}(t),S_1(t),t)}{\partial \widehat{N}(t)}\dot{n}(t) + \frac{\partial J(\mathbb{X}^{*,\dot{n}}(t),\widehat{N}(t),S_1(t),t)}{\partial S_1(t)} b_1S_1(t) +$

$\frac{1}{2}\frac{\partial^2 \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t),\widehat{N}(t),S_1(t),t)}{\partial (\mathbb{X}^{*,\dot{n}}(t))^2}\left(\mathbb{X}^{*,\dot{n}}(t)\right)^2 \big(\pi(t)\big)^T\big(\sigma(t)\big)^T\sigma(t)\pi(t) +$

$\left.\frac{1}{2}\frac{\partial^2 \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t),\widehat{N}(t),S_1(t),t)}{\partial (S_1(t))^2}\sigma_{11}^2S_1^2 + \frac{\partial^2 \mathbb{J}(\mathbb{X}^{*,\dot{n}}(t),\widehat{N}(t),S_1(t),t)}{\partial \mathbb{X}^{*,\dot{n}}(t)\partial S_1(t)}\mathbb{X}^{*,\dot{n}}(t)\{\sum_{i=1}^d \pi^i_{f_N}(t)\sigma_{i1}(t)\}\sigma_{11}S_1\right\}$

This is the Hamilton-Jacobi-Bellman equation. We take the derivative with respect to the control variable $\dot{n}(t)$ and set it equal to zero. Then we obtain the FOC, which implies the optimal $\dot{n}(t)$:

$$-\big(BS(t) - B(t)\big)\frac{\partial \mathbb{J}\big(\mathbb{X}^{*,\dot{n}}(t),\widehat{N}(t),S_1(t),t\big)}{\partial \mathbb{X}^{*,\dot{n}}(t)} + \frac{\partial \mathbb{J}\big(\mathbb{X}^{*,\dot{n}}(t),\widehat{N}(t),S_1(t),t\big)}{\partial \widehat{N}(t)} = 0$$

□

Appendix D. Proof of Lemma 1.

Suppose $U(\cdot) = ln(\cdot)$ and $K(t,\omega) = [a,\infty)\times(-\infty,\infty)^{d-1}$, where $a = a(t,\omega) > 0$, so that $\delta(v(s)) = -av_1(s)$ on $\widetilde{K} = [0,\infty)\times\{0\}^{d-1}$, and $\mathcal{Y}_v(x) \triangleq y = \frac{T+1}{x}$, where $v(s) \triangleq \big(v_1(s),\dots,v_d(s)\big)$, and x is the initial wealth at time $t = 0$. Then $\forall y = \mathcal{Y}_v(x) = \frac{T+1}{x} \in (0,\infty)$; note that

$\tilde{J}(y;v) = \tilde{J}(\mathcal{Y}_v(x);v) = \mathbb{E}\big[\widetilde{U}\big(\mathcal{Y}_v(x)H_v(T)\big)\big] = -\mathbb{E}\big[1 + ln\big(\mathcal{Y}_v(x)H_v(T)\big)\big]$

$= \mathbb{E}\left[-1 - ln\left(\frac{T+1}{x}H_v(T)\right)\right] = -\left[1 + ln\frac{T+1}{x}\right] + \mathbb{E}\left[ln\frac{1}{H_v(T)}\right]$

$= -[1 + lny] + \mathbb{E}\int_0^T\left[r(s) + \delta\big(v(s)\big) + \frac{1}{2}\|\theta(s) + \sigma^{-1}(s)v(s)\|^2\right]ds.$

To prove that $\lim_{\|v\|\to\infty} \tilde{J}(y;v) = \infty$, we need only to prove that $\lim_{\|v\|\to\infty} \Big[r(s) + \delta\big(v(s)\big) + \frac{1}{2}\|\theta(s) + \sigma^{-1}(s)v(s)\|^2\Big] = \infty$.

We have $\forall x \in (0,\infty)$,

$\lim_{\|v\|\to\infty} \Big[r(s) + \delta\big(v(s)\big) + \frac{1}{2}\|\theta(s) + \sigma^{-1}(s)v(s)\|^2\Big] = \lim_{\|v\|\to\infty} \Big[r(s) - av(s) + \frac{1}{2}\|\theta(s) + \sigma^{-1}(s)v(s)\|^2\Big] = \infty.$

On the other hand, for any fixed $v \in \widetilde{K}$, $\mathbb{E}\int_0^T \Big[r(s) + \delta\big(v(s)\big) + \frac{1}{2}\|\theta(s) + \sigma^{-1}(s)v(s)\|^2\Big]ds$ is a constant (i.e., independent of y) denoted as C; and so

$\lim_{y\downarrow 0} \tilde{J}(y;v) = \lim_{y\downarrow 0}(-[1 + lny] + C) = \infty.$

□

Appendix E. Proof of Proposition 3

Without loss of generality, we assume that the current time is $t = 0$. The total initial wealth is $\mathbb{X}(0)$. If we define the initial marginal utility with respect to total initial wealth as y, then we have $\mathcal{X}_{f_{\widehat{N}}}(y) = \mathbb{E}\left(H_{f_{\widehat{N}}}(T)I\left(yH_{f_{\widehat{N}}}(T)\right)\right)$.

Under the log utility assumption, we have $y = \mathcal{Y}_v(x) = \frac{1}{x^n}$. Now we have simplified the left-hand side of the FOC, and next we work on the right-hand side. The relationship between total initial wealth and total initial marginal utility implies

$\mathbb{J}\big(\mathbb{X}^{*,\dot{n}}(t), \widehat{N}(t), S_1(t), t\big) = \mathbb{E}\left(U\left(I\left(\mathcal{Y}_{f_{\widehat{N}}}(x^n)H_{f_{\widehat{N}}}(T)\right)\right)\right)$, that is,

$\mathbb{J}\big(\mathbb{X}^{*,\dot{n}}(t), \widehat{N}(t), S_1(t), t\big) = \mathbb{E}\left(ln\left(\frac{x^n}{H_{f_{\widehat{N}}}(T)}\right)\right) = ln(x^n) + \left(\int_0^t \left\{r(s) + \delta(f_{\widehat{N}}(s)) + \frac{1}{2}\left\|\theta_{f_{\widehat{N}}}(s)\right\|^2\right\}ds\right).$

Substituting the above results into the FOC, we obtain the following simplified FOC:

$$-\frac{BS(t) - B(t)}{\mathbb{X}(t)} + l\,n\big(\mathbb{X}(t)\big)\frac{\partial\left(-\frac{\left(n - \widehat{N}(t)\right)\phi\big(d_1(t)\big)S_1(t)}{\mathbb{X}(t)}f_{\widehat{N}}(t) + \frac{1}{2}\left\|\theta_{f_{\widehat{N}}}(t)\right\|^2\right)}{\partial\widehat{N}(t)} = 0.$$

We rewrite it more explicitly by differentiating between the pre-exercise total wealth at time t and the after-exercise total wealth at time t. By considering the vesting period of the stock and option, respectively, then we have

$$-\frac{BS(t) - B(t)}{pre_{\mathbb{X}(t)}} + l\,n\big(post_{\mathbb{X}(t)}\big)$$
$$\times\frac{\partial\left(-\left(\frac{\left(n_{option} - \widehat{N}(t)\right)\phi\big(d_1(t)\big)S_1(t) + n_{stock}1_{t<t_v}S_1(t)}{post_{\mathbb{X}(t)}}\right)f_{\widehat{N}}(t) + \frac{1}{2}\left\|\theta_{f_{\widehat{N}}}(t)\right\|^2\right)}{\partial\widehat{N}(t)} = 0,$$

where $post\_\mathbb{X}(t) = pre\_\mathbb{X}(t) - dN(t)(BS(t) - B(t))$, and t is after the end of the vesting period for the ESOs . Also, $\widehat{N}(t) \equiv 0$, before the end of the vesting period for the ESOs.

$\square$

REFERENCES

Carpenter, J., 1998. The Exercise and Valuation of Executive Stock Options. *Journal of Financial Economics* 48, 127–158.

Carpenter, J., Stanton, R., Wallace, N., 2010. Optional Exercise of Executive Stock Options and Implications for Firm Cost. *Journal of Financial Economics* 98, 315–337.

Carr, P., Linetsky, V., 2000. The Valuation of Executive Stock Options in an Intensity Based Framework. *European Finance Review* 4, 211–230.

Cvitanić, J., Karatzas, I., 1992. Convex Duality in Constrained Portfolio Optimization. *Annals of Applied Probability* 2, 767–818.

Cvitanić, J., Wiener, Z., Zapatero, F., 2008. Analytic Pricing of Executive Options. *Review of Financial Studies* 21, 683–724.

Detemple, J., Sundaresan, S., 1999. Nontraded Asset Valuation with Portfolio Constraints: A Binomial Approach. *The Review of Financial Studies* 12, 835–872.

Heston, S., 1993. A Closed-Form Solution for Option with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies* 6, 327-343.

Huddart, S., Lang, M., 1996. Employee Stock Option Exercises: An Empirical Analysis. *Journal of Accounting and Economics* 21, 5–43.

Hull, J., White, A., 2004. How to Value Employee Stock Options. *Financial Analysts Journal* 60, 114–119.

Ingersoll, J., 2006. The Subjective and Objective Evaluation of Incentive Stock Options. *Journal of Business* 79, 453–487.

Jennergren, L., Näslund, B., 1993. A Comment on "Valuation of Executive Stock Options and The FASB Proposal''. *Accounting Review* 68, 179–183.

Karatzas, I, and Kou, S. G., 1996. On The Pricing of Contingent Claims under Constraints. *The Annals of Applied Probability* 6, .2, 321–369.

Leung, T., Sircar, R., 2009. Accounting for Risk Aversion, Vesting, Job Termination Risk, and Multiple Exercises in Valuation of Employee Stock Options. *Mathematical Finance* 19, 99–128.

Molchanov, I., 2005. *Theory of Random Set*s *(Probability and its Applications)*. Springer-Verlag London Ltd., London.

Murphy, K. J., 1999. *Executive Compensation*. Handbook of Labor Economics 3, 2485–2563.