SIAM J. CONTROL OPTIM.
Vol. 43, No. 1, pp. 58–85

PERPETUAL CONVERTIBLE BONDS*

MIHAI SÎRBU†, IGOR PIKOVSKY‡, AND STEVEN E. SHREVE†

Abstract. A firm issues a convertible bond. At each subsequent time, the bondholder must decide whether to continue to hold the bond, thereby collecting coupons, or to convert it to stock. The firm may at any time call the bond. Because calls and conversions often occur far from maturity, it is not unreasonable to model this situation with a perpetual convertible bond, i.e., a convertible coupon-paying bond without maturity. This model admits a relatively simple solution, under which the value of the perpetual convertible bond, as a function of the value of the underlying firm, is determined by a nonlinear ordinary differential equation.

Key words. convertible bonds, stochastic calculus, viscosity solutions

AMS subject classifications. 90A09, 60H30, 60G44

DOI. 10.1137/S0363012902412458

1. Introduction.

Firms raise capital by issuing debt (bonds) and equity (shares of stock). The convertible bond is intermediate between these two instruments. A convertible is a bond in the sense that it entitles its owner to receive coupons plus the return of the principle at maturity. However, prior to maturity, the holder may "convert" the bond, surrendering it for a preset number of shares of stock. The price of the bond is thus dependent on the price of the firm's stock. Finally, prior to maturity, the firm may "call" the bond, forcing the bondholder to either surrender it to the firm for a previously agreed price or else convert it for stock as above.

After issuing a convertible bond, the firm's objective is to exercise its call option in order to maximize the value of shareholder equity. The bondholder's objective is to exercise his conversion option in order to maximize the value of the bond. If stock and convertible bonds are the only assets issued by a firm, then the value of the firm is the sum of the value of these two types of assets. In idealized markets where the Miller–Modigliani [17], [18] assumptions hold, changes in corporate capital structure do not affect firm value. In particular, the value of the firm does not change at the time of conversion, and the only change in the value of the firm at the time of call is a reduction by the call price paid to the bondholder if the bondholder surrenders rather than converts the bond. By acting to maximize the value of equity, the firm is in fact minimizing the value of the convertible bond. By acting to maximize the value of the bond, the bondholder is in fact minimizing the value of equity. This creates a two-person, zero-sum game.

Brennan and Schwartz [5] and Ingersoll [11] address the convertible bond pricing problem via the arbitrage pricing theory developed by Merton [16] and underlying the option pricing formula of Black and Scholes [4]. This leads to the conclusion that the firm should call as soon as the conversion value of the bond (the value the bondholder would receive if he converts the bond to stock) rises to the call price. There has been

*Received by the editors July 30, 2002; accepted for publication (in revised form) October 16, 2003; published electronically May 25, 2004. This material is based upon work supported by the National Science Foundation under grants DMS-0103814 and DMS-0139911.

http://www.siam.org/journals/sicon/43-1/41245.html

†Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA, 15213 (msirbu@andrew.cmu.edu, shreve@cmu.edu).

‡Global Modeling and Analytics, Credit Suisse First Boston, One Cabot Square, London, E14 4QJ, UK (igor.pikovsky@csfb.com).

considerable discussion whether firms call bonds at this time; see, e.g., [1], [2], [8], [12].

In the Brennan and Schwartz [5] model, dividends and coupons are paid at discrete dates. Between these dates the value of the firm is a geometric Brownian motion and the price of the convertible bond is governed by the linear second-order partial differential equation developed by Black and Scholes [4]. Brennan and Schwartz [6] generalize that model to allow random interest rates and debt senior to the convertible bond. In Ingersoll [11], coupons are paid out continuously, and for most of the results obtained, dividends are zero. Again, the bond price is governed by a linear second-order partial differential equation. In [5] the bond should not be converted except possibly immediately prior to a dividend payment; in [11] the bond should not be converted except possibly at maturity. Therefore, neither of these papers needs to address the free boundary problem which would arise if early conversion were optimal.

The present paper assumes that a firm's value comprises equity and convertible bonds. To simplify the discussion, we assume the equity is in the form of a single share of stock, and there is a single convertible bond. We assume the value of the issuing firm has constant volatility, the bond continuously pays a coupon at a fixed rate, and the firm equity pays a dividend at a rate which is a fixed fraction of the equity value. In particular, payments are always up to date and there is no issue of accrued interest at the time of a call, default, or conversion. Default occurs if the coupon payments cause the firm value to fall to zero, in which case the bond has zero recovery. In this model, both the bond price and the stock price are functions of the underlying firm value. As pointed out by [3], this means that the stock price does not have constant volatility. Furthermore, because the stock price is the difference between firm value and bond price and because dividends are paid proportionally to the stock price, the differential equation characterizing the bond price as a function of the firm value is *nonlinear*. The development of a mathematical methodology to treat this nonlinearity is the rationale for this paper.

To simplify the analysis, we assume the bond is *perpetual*, i.e., it never matures. This removes the time parameter from the problem, and the free boundary problems associated with optimal call and optimal conversion become "free point" problems. Perpetual bonds are the asymptotic case of finite-maturity bonds; work along the lines of this paper on these bonds is forthcoming. Also, as noted by Ingersoll [11], perpetual convertible bonds are unknown in the market, but they are close relatives of preferred stock, which does trade. Preferred stock does not mature, it can often be called by the issuing firm, and it can be converted to common stock by its owner.

In the time-independent setting of this paper, it is possible to place the convertible bond pricing problem on a firm theoretical foundation. Indeed, the price we obtain is shown to be the only arbitrage-free price in a perfectly liquid market in which the bond, the stock, and a constant-interest-rate money market can be traded. To establish this we first make the assumption that the respective parties adopt not necessarily optimal call and conversion strategies and derive the corresponding no-arbitrage bond price (Theorem 2.1). We then pose the determination of optimal call and conversion strategies as a two-person, zero-sum game and show that the game has a value (Theorem 2.4). We give a full description of the bond price as a function of the firm value in Theorem 2.5. One of the conclusions of that theorem is that it can be optimal to call the bond before the conversion price has reached the call price.

2. The model. We consider a firm whose value at time $t \geq 0$ is denoted by $X(t)$. We assume that prior to call or conversion of the convertible bond, the evolution of

$X(t)$ is governed by the stochastic differential equation

$$dX(t) = h(X(t))\,dt - c\,dt + \sigma X(t)\,dW(t), \tag{2.1}$$

where W is a one-dimensional Brownian motion on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, h is a Lipschitz continuous function satisfying $h(0) = 0$, and c and σ are positive constants. We denote by $\{\mathcal{F}(t); t \geq 0\}$ the filtration generated by the Brownian motion W, augmented by the null sets in $\mathcal{F}$.

At time t, the firm has a debt $D(t)$, and so the equity value is

$$S(t) = X(t) - D(t). \tag{2.2}$$

The debt is in the nature of a convertible bond, which pays coupons at the constant rate c. The bond never matures. The firm's dividend policy is to pay continuously to shareholders at a rate δ times the equity, where $\delta > 0$.

At any time, the owner of the convertible bond may *convert* it for stock. Upon conversion the bondholder will be issued new stock so that his share of the total equity of the company is the *conversion factor* γ, where $0 < \gamma < 1$. To simplify the discussion, let us assume that before conversion the firm has one share of stock outstanding. We are denoting by $X(t)$ the value of the firm and by $D(t)$ the size of the debt before conversion. Therefore, (2.2) gives the price of firm's single share of stock before conversion. Upon conversion, the firm issues new stock and the former bondholder becomes a stockholder. The total value of the firm's outstanding stock is $X(t)$, and the value of the stock owned by the former bondholder is $\gamma X(t)$. Therefore, the price of the share of the stock outstanding before conversion is now $(1-\gamma)X(t)$.

At any time, the firm may *call* the bond, which requires the bondholder to either immediately surrender it for the fixed *conversion price* $K > 0$ or else immediately convert it as described above. If the bond is surrendered, no new stock is issued and the price of the firm's single outstanding share becomes $X(t) - K$. In this model the firm may not call the bond if $X(t) < K$; i.e., there is no provision for reissuing debt.

Equation (2.1) describes the evolution of $X(t)$ only before call or conversion. Prior to conversion or call, the firm value $X(t)$ may drop to zero, in which case the firm declares bankruptcy and coupons and dividends cease.

There is a constant interest rate r, and we assume $\delta < r$. Prior to call or conversion of the bond, there are three tradable assets: the firm's stock, the convertible bond, and a money market paying interest r. We assume that all these are infinitely divisible and there are no transaction costs. Thus, the value $V(t)$ of a portfolio which holds $\Delta_1(t)$ shares of stock and $\Delta_2(t)$ convertible bonds at time t and finances this by investing or borrowing at interest rate r evolves according to the stochastic differential equation

$$\begin{aligned} dV(t) = {} & \Delta_1(t)\big(dS(t) + \delta S(t)\big) + \Delta_2(t)\big(dD(t) + cdt\big) \\ & + r\big(V(t) - \Delta_1(t)S(t) - \Delta_2(t)D(t)\big)\,dt. \end{aligned} \tag{2.3}$$

An arbitrage arises if one can begin with $V(0) = 0$ and use $\{\mathcal{F}(t)\}$-adapted processes Δ_1 and Δ_2 so that at some bounded stopping time τ at or before the minimum of the time of call, the time of conversion and the time of bankruptcy, $V(\tau) \geq 0$ almost surely and $V(\tau) > 0$ with positive probability. We restrict ourselves to trading strategies $\Delta_1(t)$, $\Delta_2(t)$ which cause $V(t)$ to be uniformly bounded from below for $0 \leq t \leq \tau$. Our goal is to price the convertible bond, under the assumption that the firm issuing the bond and the bondholder behave optimally, in a way which precludes arbitrage.

If the bond is called, the bondholder surrenders it for the call price K if K exceeds the conversion value $\gamma X(t)$ and converts it if $\gamma X(t) > K$. If $\gamma X(t) = K$, the bondholder is indifferent between surrender and conversion. Thus, if the bond is called when the firm value is $X(t)$, then the value of the bond is $\max\{K, \gamma X(t)\}$. If the bond has not been called, we assume the bondholder adopts a rule of the form: "convert as soon as the value of the firm equals or exceeds C_o." For the firm, we consider call strategies of the form "call the first time the value of the firm equals or exceeds C_a." The firm must choose $C_a \geq K$; if $C_a < K$, the firm would call when the firm value was insufficient to pay the call price. The firm and bondholder each choose a strategy, characterized by positive constants $C_a \geq K$ and $C_o > 0$. Once C_a, C_o are chosen, we want to find the price of the bond as a smooth function of the value of the firm such that no arbitrage can occur.

To set the notation, for an arbitrary number $a > 0$ we define the *nonlinear* differential operator acting on functions $f \in C[0,a] \cap C^2(0,a)$ by

$$\mathcal{N}f(x) \triangleq rf(x) - (rx - c)f'(x) + \delta(x - f(x))f'(x) - \frac{1}{2}\sigma^2 x^2 f''(x). \tag{2.4}$$

We shall see that this differential operator corresponds to the stochastic differential equation for the firm value

$$dX(t) = \big(rX(t) - c\big)\,dt - \delta\big(X(t) - f(X(t))\big)\,dt + \sigma X(t)\,dW(t), \tag{2.5}$$

rather than (2.1) posited above. This turns out to be the so-called *risk-neutral* evolution of the value of the firm. Under the risk-neutral evolution, the firm value has mean rate of change r reduced by the coupon and dividend payments. The volatility σ is the same as in (2.1). An interesting feature of this model is that the function f appearing in (2.5), which determines the evolution of the "state" under the risk neutral measure for this problem, must be determined by optimality considerations. It is not known a priori.

THEOREM 2.1. *Suppose $C_a \geq K$ and $C_o > 0$ are chosen (not necessarily optimally) by the firm and bondholder, respectively, and set*

$$a_* \triangleq \min\{C_a, C_o\}, \quad \tau_* \triangleq \inf\{t \geq 0; X(t) \notin (0, a_*)\}. \tag{2.6}$$

Assume $X(0) \in (0, a_*)$ and

$$D(t) = f(X(t)), \qquad 0 \leq t \leq \tau_*, \tag{2.7}$$

for a function $f \in C[0,a_*] \cap C^2(0,a_*)$ satisfying the boundary conditions

$$f(0) = 0, \quad f(a_*) = \begin{cases} \gamma a_* & \text{if } 0 < C_o < C_a, \\ \max\{K, \gamma a_*\} & \text{if } K \leq C_a \leq C_o. \end{cases} \tag{2.8}$$

If there is no arbitrage, then

$$\mathcal{N}f(x) = c \text{ for } 0 < x < a_*. \tag{2.9}$$

Conversely, if the function f satisfies (2.8) *and* (2.9) *and the derivative f' is bounded on $(0, a_*)$, then there is no arbitrage.*

Proof. Assume that the price of the bond is $D(t) = f(X(t))$ for a function $f \in C[0,a_*] \cap C^2(0,a_*)$ satisfying (2.8). In particular, the value of the equity is $S(t) = X(t) - f(X(t))$ for $0 \leq t \leq \tau_*$. Taking (2.3) into account, we see that the value

$V(t)$ of a self-financing portfolio starting with initial capital $V(0)=0$ and containing $\Delta_1(t)$ shares of stock and $\Delta_2(t)$ units of convertible bond evolves according to

$$\begin{aligned} dV(t) = \Delta_1(t)\big[d\big(X(t)-f(X(t))\big)+\delta\big(X(t)-f(X(t))\big)dt\big] \\ + \Delta_2(t)\big[df(X(t))+c\,dt\big] \\ + r\big[V(t)-\Delta_1(t)\big(X(t)-f(X(t))\big)-\Delta_2(t)f(X(t))\big]dt. \end{aligned}$$

Therefore,

$$\begin{aligned} &d(e^{-rt}V(t)) \\ &\quad = e^{-rt}\big[\Delta_1(t)\big(1-f'(X(t))\big)+\Delta_2(t)f'(X(t))\big]dX(t) \\ &\qquad + e^{-rt}\Delta_1(t)\left[-\frac{1}{2}\sigma^2X^2(t)f''(X(t))-(r-\delta)X(t)+(r-\delta)f(X(t))\right]dt \\ &\qquad + e^{-rt}\Delta_2(t)\left[\frac{1}{2}\sigma^2X^2(t)f''(X(t))+c-rf(X(t))\right]dt. \end{aligned} \tag{2.10}$$

We choose $\Delta_1(t) = f'(X(t))\mathrm{sgn}(\mathcal{N}f(X(t))-c)$ and $\Delta_2(t) = -(1-f'(X(t)))$ sgn $(\mathcal{N}f(X(t))-c)$, so that $\Delta_1(t)(1-f'(X(t)))+\Delta_2(t)f'(X(t))=0$. With these choices (2.10) becomes $d(e^{-rt}V(t)) = |\mathcal{N}f(X(t))-c|\,dt$. This equation shows that the portfolio value $V(t)$ is bounded from below by $V(0)=0$ and provides an arbitrage unless $\mathcal{N}f(x)=c$ for $0<x<a_*$.

We now prove the converse. Assume $D(t)=f(X(t))$ for $0\le t\le\tau_*$, and f satisfies (2.8) and (2.9). Let $\tau\le\tau_*$ be a bounded stopping time. Since $\frac{h(X(t))}{X(t)}$ and $\frac{f(X(t))}{X(t)}$ are bounded for $0\le t\le\tau_*$, we can use Girsanov's theorem to construct an equivalent probability measure $\widetilde{\mathbb{P}}$ such that

$$\int_0^t \frac{h(X(s))}{X(s)}ds+\sigma W(t) = rt-\delta\int_0^t\left(1-\frac{f(X(s))}{X(s)}\right)ds+\sigma\widetilde{W}(t) \tag{2.11}$$

for $0\le t\le\tau$, where $\widetilde{W}$ is a Brownian motion under $\widetilde{\mathbb{P}}$. The differential of the value of the firm may be rewritten as

$$dX(t) = rX(t)dt-\delta(X(t)-f(X(t)))dt-cdt+\sigma X(t)d\widetilde{W}(t), \qquad 0\le t\le\tau. \tag{2.12}$$

Let us consider the value $V(t)$ starting with initial capital $V(0)=0$ corresponding to a self-financing trading strategy $\Delta_1(t)$, $\Delta_2(t)$ for $0\le t\le\tau$. We can write the evolution of $V(t)$ as

$$\begin{aligned} d(e^{-rt}V(t)) = \Delta_1(t)\big(d(e^{-rt}S(t))+\delta e^{-rt}S(t)dt\big) \\ + \Delta_2(t)\big(d(e^{-rt}D(t))+ce^{-rt}dt\big). \end{aligned} \tag{2.13}$$

Since $D(t)=f(X(t))$, $S(t)=X(t)-f(X(t))$, and the function f is smooth, we can apply Itô's formula to obtain

$$\begin{aligned} &d(e^{-rt}S(t))+\delta e^{-rt}S(t)dt \\ &\qquad = e^{-rt}\left(\mathcal{N}f(X(t))-c\right)dt+e^{-rt}(1-f'(X(t)))\sigma X(t)d\widetilde{W}(t), \end{aligned} \tag{2.14}$$

$$\begin{aligned} &d(e^{-rt}D(t))+ce^{-rt}dt \\ &\qquad = -e^{-rt}\left(\mathcal{N}f(X(t))-c\right)+e^{-rt}f'(X(t))\sigma X(t)d\widetilde{W}(t). \end{aligned} \tag{2.15}$$

We assume $\mathcal{N}f(x) - c = 0$ for $0 < x < a_*$, and taking into account (2.14), (2.15), and (2.13), we conclude that $e^{-r(t\wedge\tau)}V(t\wedge\tau)$ is a local martingale under $\widetilde{\mathbb{P}}$. But V is uniformly bounded below, and Fatou's lemma implies $\widetilde{\mathbb{E}}[e^{-r\tau}V(\tau)] \le V(0) = 0$. This means it is impossible to have $\widetilde{\mathbb{P}}\{V(\tau) \ge 0\} = 1$ and $\widetilde{\mathbb{P}}\{V(\tau) > 0\} > 0$. Since $\widetilde{\mathbb{P}}$ is equivalent to the probability measure $\mathbb{P}$, no arbitrage exists. □

In the remainder of this section we state the principal results of the paper. Their proofs are provided in section 7.

To compute the "no arbitrage" price of the convertible bond for some (not necessarily optimal) call and conversion levels, we need an existence and uniqueness result for boundary value problems associated with (2.9).

THEOREM 2.2. *Let y_1 be a positive number and $0 < y_1 \le x_1$. Then there exists a unique solution $f \in C[0, x_1] \cap C^2(0, x_1)$ of the boundary value problem*

$$\begin{cases} \mathcal{N}f(x) = c \text{ for } x \in (0, x_1), \\ f(0) = 0,\ f(x_1) = y_1. \end{cases} \tag{2.16}$$

Furthermore, the derivative f' is bounded on $(0, x_1)$. If $y_1 < x_1$, then $f'(x) < 1$ for all $x \in (0, x_1)$.

Taking into account Theorem 2.1, Theorem 2.2, and the discussion regarding the price of the bond at call or conversion time, we see that once the call and conversion levels have been set, the "no-arbitrage" price of the convertible bond is

$$D(t) = f(X(t), C_a, C_o), \tag{2.17}$$

where the function $f(x, C_a, C_o)$ is given in the next definition.

DEFINITION 2.3.

(i) *If $0 < C_o < C_a$, define $f(x, C_a, C_o)$ for $0 \le x \le C_o$ to be the unique solution of the equation $\mathcal{N}f = c$ on $(0, C_o)$ satisfying the boundary conditions $f(0) = 0$, $f(C_o) = \gamma C_o$. For $x \ge C_o$, define*

$$f(x, C_a, C_o) = \begin{cases} \gamma x, & C_o \le x < C_a, \\ \max\{K, \gamma x\}, & x \ge C_a. \end{cases}$$

(ii) *If $K \le C_a \le C_o$, define $f(x, C_a, C_o)$ for $0 \le x \le C_a$ to be the unique solution of the equation $\mathcal{N}f = c$ on $(0, C_a)$ satisfying the boundary conditions $f(0) = 0$, $f(C_a) = \max\{K, \gamma C_a\}$. For $x \ge C_a$, define $f(x, C_a, C_o) = \max\{K, \gamma x\}$.*

Equation (2.17) provides a bond price once the call and conversion levels C_a and C_o have been chosen. The firm wishes to minimize the value of the bond (in order to maximize the value of equity), and the bondholder wishes to maximize the value of the bond. This creates a two-person game, and according to the next theorem, this game has a value.

THEOREM 2.4. *There exist $C_a^* \ge K$ and $C_o^* > 0$ such that for each $x \ge 0$, we have*

$$f(x, C_a^*, C_o^*) = \inf_{C_a \ge K} f(x, C_a, C_o^*) = \sup_{C_o > 0} f(x, C_a^*, C_o). \tag{2.18}$$

Equation (2.18) implies the following equalities, so we can define

$$f_*(x) \triangleq f(x, C_a^*, C_o^*) = \sup_{C_o > 0} \inf_{C_a \ge K} f(x, C_a, C_o) = \inf_{C_a \ge K} \sup_{C_o > 0} f(x, C_a, C_o). \tag{2.19}$$

This is the price of the bond as a function of the underlying firm value x, and C_a^* and C_o^* are the optimal call and optimal conversion levels, respectively.

THEOREM 2.5. *The function* f_* *is in* $C[0,\infty)$ *and is described by one of three cases. There are two constants* $0 \le K_1 < K_2$ *depending on* $r, \delta, \sigma, c,$ *and* γ.

(i) *If* $K > K_2$, *then* $f_* \in C^1(0,\infty)$ *and satisfies*

$$0 < f_*'(x) < 1 \text{ for } x > 0. \tag{2.20}$$

In this case,

$$C_o^* = \min\{x > 0; f_*(x) = \gamma x\} = \frac{K_2}{\gamma},$$

f_* *restricted to* $(0, C_o^*)$ *is the unique classical solution of* $\mathcal{N} f_* = c$ *on* $(0, C_o^*)$ *with boundary conditions* $f_*(0) = 0$ *and* $f_*(C_o^*) = \gamma C_o^*$,

$$f_*(x) = \gamma x \text{ for } x \ge C_o^*, \tag{2.21}$$

and $C_a^* = \frac{K}{\gamma} > C_o^* = \frac{K_2}{\gamma}$.

(ii) *If* $K_1 \le K \le K_2$, *then* f_* *restricted to* $(0, K/\gamma)$ *is the unique classical solution of* $\mathcal{N} f_* = c$ *on* $(0, K/\gamma)$ *with the boundary conditions* $f_*(0) = 0$ *and* $f_*(K/\gamma) = K$. *We have*

$$0 < f_*'(x) < 1 \text{ for } 0 < x < \frac{K}{\gamma}, \tag{2.22}$$

$$f_*(x) = \gamma x \text{ for } x \ge \frac{K}{\gamma}. \tag{2.23}$$

In this case, $C_o^* = C_a^* = \frac{K}{\gamma}$.

(iii) *If* $K_1 > 0$, *there is a third case. A sufficient condition for* $K_1 > 0$ *is* $0 < \gamma < \frac{1}{2}$. *In the third case,* $0 < K < K_1$, f_* *restricted to* $(0, K/\gamma)$ *is continuously differentiable,* $C_a^* \in (K, K/\gamma)$, *and* f_* *restricted to* $(0, C_a^*)$ *is the unique solution of* $\mathcal{N} f_* = c$ *on* $(0, C_a^*)$ *with the boundary conditions* $f_*(0) = 0$, $f_*(C_a) = K$. *We have*

$$0 < f_*'(x) < 1 \text{ for } 0 < x < C_a^*, \tag{2.24}$$

$$f_*(x) = \begin{cases} K, & C_a^* \le x \le \frac{K}{\gamma}, \\ \gamma x, & x \ge \frac{K}{\gamma}. \end{cases} \tag{2.25}$$

In particular, $f_*'(C_a^*-) = 0$ *and* $K < C_a^* < C_o^* = \frac{K}{\gamma}$.

From Theorem 2.5 we see that the firm debt at time t is $D(t) = f_*(X(t))$, and (2.2) becomes

$$S(t) = X(t) - f_*(X(t)). \tag{2.26}$$

So long as $x \in (0, C_a^* \wedge C_o^*)$, the function $F(x) \triangleq x - f_*(x)$ is strictly increasing because of (2.20), (2.22), and (2.24) and hence has an inverse F^{-1}. We may invert (2.26) to obtain $X(t) = F^{-1}(S(t))$, and thereby obtain a formula for the market price of the convertible bond in terms of the equity of the firm: $D(t) = f_*(F^{-1}(S(t)))$. In all three cases of Theorem 2.5, the firm should call as soon as $D(t)$ rises to the call price K. In cases (i) and (ii), this is the first time the conversion value of the bond rises to the call price. In case (iii), the call should occur before the conversion value rises to the call price. The owner of the bond should convert as soon as $D(t) - \gamma F^{-1}(S(t))$ falls to zero, i.e., as soon as the difference between the bond price and the bond's conversion value falls to zero.

3. Generation of candidate functions. Theorem 2.5 asserts that for small values of x, the value $f_*(x)$ of the convertible bond satisfies the second-order ordinary differential equation $\mathcal{N}f(x) = c$. Not only is this equation nonlinear, it is also singular at $x = 0$. Rather than solving the differential equation $\mathcal{N}f(x) = c$ directly, we generate a one-parameter family of solutions to the variational inequality

$$\min\{\mathcal{N}f(x) - c, f(x) - \gamma x\} = 0. \tag{3.1}$$

To do this, we first construct for a fixed function $g \in C[0,a]$, a solution to the variational inequality

$$\min\{\mathcal{L}_g f(x) - c, f(x) - \gamma x\} = 0, \tag{3.2}$$

subject to boundary conditions $f(0) = 0$, $f(a) = \gamma a$. Here, the *linear* differential operator $\mathcal{L}_g$ is defined by

$$\mathcal{L}_g f(x) \triangleq rf(x) - (rx - c)f'(x) + \delta(x - g(x))f'(x) - \frac{1}{2}\sigma^2 x^2 f''(x). \tag{3.3}$$

In section 6 we prove existence of a function g for which the solution to this equation is g itself.

Definition 3.1. *Let $a \in (0,\infty)$ be given. Denote $\mathcal{D}_a = [0,a]$ and let $\mathcal{G}_a$ be the set of continuous functions $g\colon \mathcal{D}_a \to \mathbb{R}$ which are continuously differentiable on $(0,a)$ and satisfy*

$$\begin{aligned} &g(0) = 0, \quad g(a) = \gamma a,\\ &g(x) \geq \gamma x, \quad -M_a \leq g'(x) < 1 \quad \forall x \in (0,a), \end{aligned}$$

where M_a will be defined in Proposition 5.6. We denote by $\overline{\mathcal{G}}_a$ the closure of $\mathcal{G}_a$ with respect to the supremum norm in $C[0,a]$.

Denote $\mathcal{D}_\infty = [0,\infty)$ and let $\mathcal{G}_\infty$ be the set of continuous functions $g\colon \mathcal{D}_\infty \to \mathbb{R}$ which are continuously differentiable on $(0,\infty)$ and satisfy

$$\begin{aligned} &g(0) = 0, \quad g(x) = \gamma x \quad \forall x \in [b_g,\infty),\\ &g(x) \geq \gamma x, \quad 0 \leq g'(x) < 1 \quad \forall x \in (0,\infty), \end{aligned}$$

where b_g is a finite number depending on the function g. Let (C_γ, d) be the complete metric space of continuous functions on $\mathcal{D}_\infty$ which satisfy $\lim_{x\to\infty}[g(x) - \gamma x] = 0$, and d is the supremum metric. We denote by $\overline{\mathcal{G}}_\infty$ the closure of $\mathcal{G}_\infty$ in (C_γ, d).

For $a \in (0,\infty]$, $g \in \overline{\mathcal{G}}_a$, and $x \in \mathcal{D}_a$, we define $X^x(t)$ by $X^x(0) = x$ and

$$dX^x(t) = rX^x(t)\,dt - \delta\left(X^x(t) - g(X^x(t))\right)dt - c\,dt + \sigma X^x(t)\,dW(t) \tag{3.4}$$

for $0 \leq t \leq \tau_0^x \wedge \tau_a^x$, where $\tau_y^x \triangleq \inf\{t \geq 0; X^x(t) = y\}$. We then set

$$T_a g(x) \triangleq \sup_{0\leq\tau\leq\tau_0^x\wedge\tau_a^x} \mathbb{E}\left[\int_0^\tau e^{-ru}c\,du + \mathbb{I}_{\{\tau<\infty\}}e^{-r\tau}\gamma X^x(\tau)\right], \tag{3.5}$$

where the supremum is over stopping times τ which satisfy $0 \leq \tau \leq \tau_0^x \wedge \tau_a^x$.

We interpret the objects in Definition 3.1 as follows. Suppose we have a function g which maps the value of the firm into the value of convertible bond. Then $S(t)$ in (2.2) is given by $S(t) = X(t) - g(X(t))$. As we have already seen in the proof of

Theorem 2.1 (see (2.12)), under a “risk-neutral” measure, we expect the value of the firm to have mean rate of growth equal to the interest rate r, reduced by the dividend and coupon payments. In other words, if $g(x)$ is the value of the bond when x is the value of the firm, then the evolution of the value of the firm should be given by (3.4).

The fortunes of the firm, which depend on the function g and the initial condition x, may result in bankruptcy at time τ_0^x. If bankruptcy never occurs, then $\tau_0^x = \infty$. The bondholder collects dividends at rate c until bankruptcy occurs or until he converts the bond to stock. He may convert at any stopping time $\tau \le \tau_a^x$; if he has not converted by the time τ_a, he must do so at this time. The parameter a in this restriction on the stopping time τ will allow us to construct a one-parameter family of solutions to (2.4), and we shall later see that the correct choice of the parameter a depends on the call price K. However, in this interpretation of the function $T_a g$, we do not permit the firm to call. Since the conversion option is worthless after bankruptcy, we assume without loss of generality that $0 \le \tau \le \tau_0^x$. Upon conversion, the bondholder receives stock valued at $\gamma X^x(\tau)$. It follows that the risk-neutral value of a conversion strategy τ is $\mathbb{E}\left[\int_0^\tau e^{-ru} c\, du + \mathbb{I}_{\{\tau<\infty\}} e^{-r\tau}\gamma X^x(\tau)\right]$, and $T_a g(x)$ is the value of the optimal conversion strategy, if it exists.

We began this discussion with the supposition that $g(x)$ is the value of the convertible bond when x is the value of the firm. But the value of the convertible bond should be the risk-neutral discounted value of coupons collected plus the risk-neutral discounted value of the stock received upon conversion. In other words, we seek a function $f \in \overline{\mathcal{G}}_a$ such that $T_a f = f$. Such a function will satisfy (2.9), at least for small values of x.

In section 4 we prove continuity of the function $T_a g$. In section 5 we show that, like g, the function $T_a g$ is in $\overline{\mathcal{G}}_a$, and we state the *Hamilton–Jacobi–Bellman equation* (3.2) satisfied by $T_a g$. In section 6, we show that the mapping $T_a\colon \overline{\mathcal{G}}_a \to \overline{\mathcal{G}}_a$ has a unique fixed point, which we call f_a. Section 7 shows that for each call price K, there is a value of a so that f_a is a part of the function described in Theorem 2.5. This enables us to prove Theorems 2.4 and 2.5. Finally, at the end of section 7 we also prove Theorem 2.2.

4. Continuity of candidate functions. Let $a \in (0,\infty]$ and $g \in \overline{\mathcal{G}}_a$ be given, and define $T_a g$ by (3.5). If a is finite, we extend g to be constant on $(-\infty, 0]$ and on $[a,\infty)$. Since the extended g is Lipschitz, we may use (3.4) to define $X^x(t)$ for all $t \ge 0$. The assumptions on g ensure that for some $\eta > 0$, $\delta(x - g(x)) + c \ge \eta x$ for all $x \ge 0$. We now set $Z(t) = \exp\left\{-\sigma W(t) - \frac{1}{2}\sigma^2 t\right\}$, so that $d(Z(t)X^x(t)) \le (r - \sigma^2 - \eta) Z(t) X^x(t)\, dt$ for all $0 \le t \le \tau_0^x$. Integration yields

$$Z(t)X^x(t) \le x + \left(r - \sigma^2 - \eta\right)\int_0^t Z(u)X^x(u)\, du, \quad 0 \le t \le \tau_0^x,$$

and an application of Gronwall’s inequality gives the bound

$$X^x(t) \le \frac{x}{Z(t)} e^{\left(r-\sigma^2-\eta\right)t} = x \exp\left\{\sigma W(t) + \left(r - \frac{1}{2}\sigma^2 - \eta\right) t\right\},\ 0 \le t \le \tau_0^x. \tag{4.1}$$

LEMMA 4.1. *The function $T_a g$ satisfies the bounds*

$$\gamma x \le T_a g(x) \le \frac{c}{r} + \gamma x \quad \forall x \in \mathcal{D}_a. \tag{4.2}$$

Proof. The lower bound in (4.2) follows from taking $\tau \equiv 0$ in (3.5). For the upper bound, we apply the optional sampling theorem and Fatou’s lemma to the martingale

$\exp\left\{\sigma W(t) - \frac{1}{2}\sigma^2 t\right\}$ and use (4.1) to obtain for any stopping time τ satisfying $0 \leq \tau \leq \tau_0^x$

$$\begin{aligned}\text{(4.3)} \qquad \mathbb{E}e^{-r\tau}X^x(\tau) &\leq x\mathbb{E}\exp\left\{\sigma W(\tau) - \frac{1}{2}\sigma^2\tau\right\} \\ &\leq x \liminf_{t\to\infty} \mathbb{E}\exp\left\{\sigma W(t\wedge\tau) - \frac{1}{2}\sigma^2(t\wedge\tau)\right\} \;=\; x.\end{aligned}$$

Therefore,

$$T_a g(x) \leq \int_0^\infty e^{-ru} c\,du + \gamma \sup_{0\leq\tau\leq\tau_0^x} \mathbb{E}e^{-r\tau}X^x(\tau) \leq \frac{c}{r} + \gamma x.$$

LEMMA 4.2. *For all $y \geq 0$, τ_y^x is almost surely continuous in x at all $x \geq 0$.*

Proof. It is possible to choose for each initial condition a version of the process $X^x(t)$, $t \geq 0$, such that $X^x(t)$ is jointly continuous in (t,x), almost surely (see [15, Theorem 4.2.5]). Because of the uniqueness of the solution to (3.4), we have for $0 \leq \xi < x \leq y$ that $X^\xi(t) \leq X^x(t)$, $0 \leq t < \infty$, almost surely; if these processes ever coalesce, they would henceforth coincide. This implies that $\lim_{\xi\uparrow x} \tau_y^\xi \geq \tau_y^x$. On the other hand, $\tau_y^x = \inf\{t \geq 0; X^x(t) > y\}$, which implies that $\lim_{\xi\uparrow x} \tau_y^\xi \leq \tau_y^x$. Therefore,

$$\text{(4.4)} \qquad \lim_{\xi\uparrow x} \tau_y^\xi = \tau_y^x.$$

By a similar argument, we conclude

$$\text{(4.5)} \qquad \lim_{\xi\downarrow x} \tau_y^\xi = \tau_y^x.$$

Combining (4.4) and (4.5), we see that, almost surely, $\lim_{\xi\to x} \tau_y^\xi = \tau_y^x$, $0 \leq x < y$, and (4.4) holds for $x = y$. A similar argument shows that $\lim_{\xi\to x} \tau_y^\xi = \tau_y^x$, $0 \leq y < x$, and (4.5) holds for $x = y$. □

Using (4.1) to bound $e^{-rt}X^x(t)$, $\lim_{t\to\infty}\exp\{\sigma W(t) - \frac{1}{2}\sigma^2 t\} = 0$, joint continuity of $X^x(t)$ in (t,x), and Lemma 4.2, we conclude that the process

$$\text{(4.6)} \qquad Y^x(t) \triangleq \int_0^{t\wedge\tau_0^x\wedge\tau_a^x} e^{-ru}c\,du + \mathbb{I}_{\{t\wedge\tau_0^x\wedge\tau_a^x<\infty\}} e^{-r(t\wedge\tau_0^x\wedge\tau_a^x)}\gamma X^x(t\wedge\tau_0^x\wedge\tau_a^x)$$

is jointly continuous in $(t,x) \in [0,\infty] \times \mathcal{D}_a$, almost surely. In particular, we have continuity at time $t = \infty$, where

$$Y^x(\infty) \triangleq \int_0^{\tau_0^x\wedge\tau_a^x} e^{-ru}c\,du + \mathbb{I}_{\{\tau_0^x\wedge\tau_a^x<\infty\}} e^{-r(\tau_0^x\wedge\tau_a^x)}\gamma X^x(\tau_0^x\wedge\tau_a^x).$$

LEMMA 4.3. *The function $T_a g$ is lower semicontinuous on $\mathcal{D}_a$.*

Let τ be any nonnegative stopping time. Lemma 4.2 implies that $\tau\wedge\tau_0^x\wedge\tau_a^x$ is almost surely continuous in x. The function

$$h_{\tau,a} = \mathbb{E}\left[\int_0^{\tau\wedge\tau_0^x\wedge\tau_a^x} e^{-ru}c\,du + \mathbb{I}_{\{\tau\wedge\tau_0^x\wedge_a^x<\infty\}} e^{-r(\tau\wedge\tau_0^x\wedge\tau_a^x)}\gamma X^x(\tau\wedge\tau_0^x\wedge\tau_a^x)\right]$$

is thus lower semicontinuous (Fatou's lemma), and $T_a g(x) = \sup_\tau h_{\tau,a}(x)$, the supremum of lower semicontinuous functions, is lower semicontinuous. □

We know from inequality (4.1) that

$$\sup_{0\le t\le\infty} Y^x(t) \le \frac{c}{r} + \gamma x \sup_{t\ge 0} \exp\left\{\sigma W(t) - \left(\eta + \frac{1}{2}\sigma^2\right) t\right\} = \frac{c}{r} + \gamma x e^{\sigma W^*}, \tag{4.7}$$

where $W^* = \sup_{t\ge 0}\left[W(t) - \left(\frac{\sigma}{2} + \frac{\eta}{\sigma}\right) t\right]$. According to [13, Exercise 5.9, Chapter 3], W^* has density

$$\mathbb{P}\{W^* \in db\} = 2\left(\frac{\sigma}{2} + \frac{\eta}{\sigma}\right) \exp\left\{-2\left(\frac{\sigma}{2} + \frac{\eta}{\sigma}\right) b\right\} db, \qquad b > 0. \tag{4.8}$$

This means that $\mathbb{E}e^{\sigma W^*} < \infty$, so we obtain

$$\mathbb{E} \sup_{0\le t\le\infty} Y^x(t) < \infty. \tag{4.9}$$

In light of Lemmas 4.1 and 4.3, the set

$$\mathcal{S}_g \triangleq \{x \in \mathcal{D}_a; T_a g(x) = \gamma x\} = \{x \in \mathcal{D}_a : T_a g(x) \le \gamma x\}$$

is closed, contains the origin, and contains a if a is finite. We define

$$\tau_*^x \triangleq \inf\{t \ge 0; X^x(t) \in \mathcal{S}_g\}, \tag{4.10}$$

a stopping time satisfying $\tau_*^x \le \tau_0^x \wedge \tau_a^x$. Since inequality (4.9) holds, it is known from the general theory of optimal stopping that the process

$$Z^x(t) \triangleq \int_0^{t\wedge\tau_0^x\wedge\tau_a^x} e^{-ru} c\, du + \mathbb{I}_{\{t\wedge\tau_0^x\wedge\tau_a^x<\infty\}} e^{-r(t\wedge\tau_0^x\wedge\tau_a^x)} T_a g(X^x(t\wedge\tau_0^x\wedge\tau_a^x)) \tag{4.11}$$

is a supermartingale for $0 \le t \le \infty$; the stopped process $Z^x(t\wedge\tau_*^x)$, $0 \le t \le \infty$, is a martingale; and τ_*^x is an optimal stopping time, i.e.,

$$T_a g(x) = \mathbb{E}\left[\int_0^{\tau_*^x} e^{-ru} c du + \mathbb{I}_{\{\tau_*^x<\infty\}} e^{-r\tau_*^x} \gamma X^x(\tau_*^x)\right] = \mathbb{E}Y^x(\tau_*^x). \tag{4.12}$$

To prove this, one can first show, using the Markov property, that the process $\{Z^x(t)\}_{0\le t\le\infty}$ is the *Snell envelope* of $\{Y^x(t)\}_{0\le t\le\infty}$, i.e.,

$$Z^x(t) = \operatorname{ess\,sup}_{\tau\ge t} \mathbb{E}\left[Y^x(\tau)|\mathcal{F}_t\right], \tag{4.13}$$

and then appeal to [14, Appendix D] . Another way to prove it is to combine Theorem 1, page 124, and Theorem 3, page 127, from [20].

LEMMA 4.4. *Assume $a = \infty$. We have*

$$\gamma x \le T_\infty g(x) \le x \quad \forall x \in \mathcal{D}_\infty, \tag{4.14}$$

and there is a number $b > 0$ such that

$$T_\infty g(x) = \gamma x \quad \forall x \in [b, \infty). \tag{4.15}$$

If $a \in (0, \infty)$, we have

$$\gamma x \le T_a g(x) \le x \quad \forall x \in \mathcal{D}_a. \tag{4.16}$$

Proof. We shall construct a number $b > 0$ and a function $\varphi\colon [0,\infty) \mapsto \mathbb{R}$ such that

$$\gamma x \le \varphi(x) \le x \ \forall x \in [0,b], \quad \varphi(x) = \gamma x \ \forall x \in [b,\infty); \tag{4.17}$$

φ'' is defined and continuous on $[0,\infty)$, except at $\sqrt{b}$ and b, but has one-sided derivatives at these points; φ' is defined, bounded, and continuous on $[0,\infty)$ except at $\sqrt{b}$, but has one-sided derivatives at this point which satisfy

$$D^-\varphi(\sqrt{b}) - D^+\varphi(\sqrt{b}) > 0, \tag{4.18}$$

$$\mathcal{L}_g\varphi(x) \ge c \quad \forall x \in [0,\infty) \setminus \{\sqrt{b}, b\}. \tag{4.19}$$

Once b and φ are constructed, we choose an arbitrary $x \ge 0$. With $X(t) = X^x(t)$, the extension of Itô's rule to continuous, piecewise C^2 functions [13, Chapter 3, Theorem 7.1 and Corollary 7.2] implies that

$$d\left(e^{-rt}\varphi(X(t))\right) = -e^{-rt}\mathcal{L}_g\varphi(X(t))\,dt - e^{-rt}\left(D^-\varphi(\sqrt{b}) - D^+\varphi(\sqrt{b})\right)d\Lambda(t) \\ + e^{-rt}\sigma X(t)\varphi'(X(t))\,dW(t),$$

where $\Lambda(t)$ is the (nondecreasing) local time of X at $\sqrt{b}$. From (4.18) and (4.19), we see that

$$d\left(e^{-rt}\varphi(X(t))\right) \le -e^{-rt}c\,dt + e^{-rt}\sigma X(t)\varphi'(X(t))\,dW(t).$$

Hence, for any stopping time $\tau \le \tau_0^x$ and any deterministic time T, we have

$$\mathbb{E}e^{-r(\tau\wedge T)}\varphi(X(\tau\wedge T)) \le \varphi(x) - \mathbb{E}\int_0^{\tau\wedge T} e^{-rt}c\,dt,$$

where we have used the boundedness of φ' and (4.1) to ensure that the expectation of the Itô integral is zero. This last inequality implies

$$\varphi(x) \ge \mathbb{E}\left[\int_0^{\tau\wedge T} e^{-rt}c\,dt + \mathbb{I}_{\{\tau<\infty\}}e^{-r(\tau\wedge T)}\gamma X(\tau\wedge T)\right].$$

Letting $T \to \infty$ and using Fatou's lemma, then maximizing over τ, we obtain $\varphi(x) \ge T_a g(x)$. Relations (4.14), (4.15) follow from (4.2) and (4.17).

If $a \in (0,\infty)$, then the function $h(x) = x$ on $[0,a]$ is two times continuously differentiable on $(0,a)$ and satisfies $\mathcal{L}_g h(x) \ge c$. Since $h(x) \ge \gamma x$ for each $0 \le x \le a$, we can do the same computation as above for the function h instead of φ and obtain (4.16).

The remainder of the proof is the construction of b and φ. For $b > e^2$, define the positive function $\eta(b) \triangleq (1-\gamma)/\left(\frac{1}{2}\log b + \frac{1}{\sqrt{b}} - 1\right)$. Consider the function

$$k(b) \triangleq c\,[\gamma - \eta(b) - 1] + \frac{1}{2}\delta\sqrt{b}(1-\gamma)\left(\gamma - \eta(b)\right) - \frac{1}{2}\sigma^2\eta(b)\sqrt{b}.$$

Since $\lim_{b\to\infty}\eta(b) = 0$, we have $\lim_{b\to\infty}k(b) = \infty$. We fix a value $b > e^2$ for which $k(b) > 0$, $\eta(b) < \gamma$. For any $g \in \overline{\mathcal{G}}_a$ we know that $\lim_{x\to\infty}[g(x) - \gamma x] = 0$, so for b

sufficiently large, we also have

$$x - g(x) \geq \frac{1}{2}(1-\gamma)\sqrt{b} \quad \forall x \in [\sqrt{b}, \infty), \tag{4.20}$$

$$\delta(x - g(x)) \geq \frac{(1-\gamma)c}{\gamma} \quad \forall x \in [b, \infty). \tag{4.21}$$

With b chosen to satisfy all the above properties, we set

$$\varphi(x) = \begin{cases} x, & 0 \leq x \leq \sqrt{b}, \\ \gamma x + \eta(b)\sqrt{b}\left(\frac{x}{b} - \log\frac{x}{b} - 1\right), & \sqrt{b} < x < b, \\ \gamma x, & x \geq b. \end{cases} \tag{4.22}$$

A straightforward computation verifies that φ has the desired properties.

COROLLARY 4.5. *The function $T_a g$ is continuous on $\mathcal{D}_a$.*

Proof. Recall from the proof of Lemma 4.3 that for each $y \geq 0$, the stopping time τ_y^x is a continuous function of x. The complement of the closed set $\mathcal{S}_g$,

$$\mathcal{C}_g \triangleq \{x \in \mathcal{D}_a; T_a g(x) > \gamma x\},$$

is a countable union of disjoint open intervals, and on each of these intervals (α, β), we have $\tau_*^x = \tau_\alpha^x \wedge \tau_\beta^x$, which is a continuous function of $x \in [\alpha, \beta]$. On the set $\mathcal{S}_g$, $\tau_*^x \equiv 0$. Hence, τ_*^x is continuous on both $\mathcal{S}_g$ and its complement $\mathcal{C}_g$. To show that τ_*^x is continuous on $\mathcal{D}_a = \mathcal{C}_g \cup \mathcal{S}_g$, it remains only to show that if $\{x_n\}_{n=1}^\infty$ is a sequence in $\mathcal{C}_g$ converging to $x \in \mathcal{S}_g$, then $\tau_*^{x_n} \to \tau_*^x = 0$. But $\tau_*^{x_n} \leq \tau_x^{x_n}$ and $\tau_x^{x_n} \to \tau_x^x = 0$, almost surely (Lemma 4.2), so the desired result holds.

For $a < \infty$ we have $0 \leq X^x(t \wedge \tau_*^x) \leq a$. For $a = \infty$, Lemma 4.4 implies there exists $b > 0$ such that $[b, \infty) \subset \mathcal{S}_g$. In this case, $0 \leq X^x(t \wedge \tau_*^x) \leq \max\{x, b\}$. The continuity of $T_a g$ follows from the representation (4.12), the continuity of τ_*^x, the joint continuity of $Y^x(t)$ on $[0, \infty] \times \mathcal{D}_a$, and the dominated convergence theorem. □

PROPOSITION 4.6. *The function $T_a g$ is twice continuously differentiable on $\mathcal{C}_g$ and satisfies the equation*

$$\mathcal{L}_g T_a g = c \text{ on } \mathcal{C}_g. \tag{4.23}$$

If $g \in \mathcal{G}_a$, then $T_a g$ is three times continuously differentiable on $\mathcal{C}_g$.

Proof. Let $x \in \mathcal{C}_g$ be given, and choose $0 < \alpha < x < \beta$ such that $(\alpha, \beta) \subset \mathcal{C}_g$. Consider the linear, second-order ordinary differential equation

$$\mathcal{L}_g h(x) = c \quad \forall x \in (\alpha, \beta), \tag{4.24}$$

with the boundary conditions $h(\alpha) = T_a g(\alpha)$, $h(\beta) = T_a g(\beta)$. Because the coefficients of (4.24) are continuous, the equation has a twice continuously differentiable solution h satisfying these boundary conditions. If $g \in \mathcal{G}_a$, so that the coefficients of (4.24) are continuously differentiable, then h is three times continuously differentiable. Itô's formula implies that

$$\begin{aligned} d\left[e^{-rt} h(X^x(t))\right] &= e^{-rt}\left[-\mathcal{L}_g h(X^x(t))\,dt + \sigma X^x(t) h'(X^x(t))\,dW(t)\right] \\ &= -e^{-rt} c\,dt + e^{-rt}\sigma X^x(t) h'(X^x(t))\,dW(t). \end{aligned}$$

Integrating this equation from $t=0$ to $t=\tau_\alpha^x\wedge\tau_\beta^x$ and taking expectations, we obtain

$$\begin{aligned} h(x) &= \mathbb{E}\left[\int_0^{\tau_\alpha^x\wedge\tau_\beta^x} e^{-rt}c\,dt + e^{-r(\tau_\alpha^x\wedge\tau_\beta^x)}h(X^x(\tau_\alpha^x\wedge\tau_\beta^x))\right] \\ &= \mathbb{E}\left[\int_0^{\tau_\alpha^x\wedge\tau_\beta^x} e^{-rt}c\,dt + e^{-r(\tau_\alpha^x\wedge\tau_\beta^x)}T_a g(X^x(\tau_\alpha^x\wedge\tau_\beta^x))\right] \\ &= \mathbb{E}Z^x(\tau_\alpha^x\wedge\tau_\beta^x) \;=\; Z^x(0) \;=\; T_a g(x), \end{aligned}$$

where we have used the fact that $Z^x(t\wedge\tau_\alpha^x\wedge\tau_\beta^x)$ is a bounded martingale, since $\tau_\alpha^x\wedge\tau_\beta^x\le\tau_*^x$. ◻

Remark 4.7. Let us denote by $D^\pm T_a g$ the derivatives from the right and left of $T_a g$, when these one-sided derivatives exist. We likewise denote by $DT_a g$ the derivative of $T_a g$, when the derivative exists. Because it is open, the set $\mathcal{C}_g$ is a countable union of disjoint open intervals, which we call the *components* of $\mathcal{C}_g$. Let (α,β) be one of these components. The second-order differential operator $\mathcal{L}_g$ does not degenerate to a first-order operator at any point in $[\alpha,\beta]$, except at α when $\alpha=0$. Therefore, the function h in the proof of Proposition 4.6 is twice continuously differentiable at the endpoint β and also at α provided that $\alpha>0$. We conclude that $D^- T_a g(\beta)=\lim_{x\uparrow\beta} DT_a g(x)$ exists. If $\alpha>0$, then $D^+T_a g(\alpha)=\lim_{x\downarrow\alpha} DT_a g(x)$ also exists.

5. The invariance property of T_a.

As in the previous section, let $a\in(0,\infty]$ and $g\in\overline{\mathcal{G}}_a$ be given, and define $T_a g$ by (3.5). In this section we show that T_a maps $\mathcal{G}_a$ into itself. For this we use the theory of viscosity solutions of Hamilton–Jacobi–Bellman equations developed by Crandall and Lions (see [7], [9]). The proof of Proposition 5.1 below is standard, so we omit it. See [19] for a similar proof.

PROPOSITION 5.1. *The function $T_a g$ is a viscosity solution of the Hamilton–Jacobi–Bellman equation*

$$\min\{\mathcal{L}_g h(x)-c, h(x)-\gamma x\}=0 \quad \forall x\in\mathcal{D}_a. \tag{5.1}$$

We use Proposition 5.1 to deduce other information about $T_a g$.

COROLLARY 5.2. *Given any $b\in(0,a)$, the set $\mathcal{C}_g\cap(0,b)$ is nonempty.*

Proof. Suppose $T_a g(x)=\gamma x$ for all $x\in[0,b]$. Then

$$\mathcal{L}_g T_a g(x)-c=(\gamma-1)c+\delta\gamma(x-g(x)),$$

which is strictly negative for $x>0$ sufficiently small. This violates the viscosity supersolution property of $T_a g$. ◻

LEMMA 5.3. *If $(0,a)\cap\mathcal{S}_g$ contains a point b, then $[b,\infty)\cap\mathcal{D}_a\subset\mathcal{S}_g$.*

Proof. Assume $b\in(0,a)\cap\mathcal{S}_g$ and denote $\overline{\varphi}(x)=\gamma x$. Because $T_a g(b)=\overline{\varphi}(b)$ and $T_a g\ge\overline{\varphi}$, the viscosity supersolution property for $T_a g$ implies

$$c\le\mathcal{L}_g\overline{\varphi}(b)=c\gamma+\delta\gamma(b-g(b)).$$

But the function $x\to x-g(x)$ is nondecreasing on $\mathcal{D}_a$. Therefore

$$c\le c\gamma+\delta\gamma(x-g(x)) \quad \forall x\in[b,\infty)\cap\mathcal{D}_a.$$

We must show that $T_a g(x)\le\overline{\varphi}(x)$ for all $x\in[b,\infty)\cap\mathcal{D}_a$. Assume on the contrary that $\eta\triangleq\sup\{T_a g(x)-\overline{\varphi}(x); x\in[b,\infty)\cap\mathcal{D}_a\}$ is positive and let x_0 attain the supremum in the definition of η. (The supremum is attained because both $T_a g$ and $\overline{\varphi}$ are continuous, and if $a = \infty$, then $T_a g(x) = \overline{\varphi}(x)$ for all sufficiently large x.)

We take $\varphi(x) = \overline{\varphi}(x) + \eta$ for $x \in [b,\infty) \cap \mathcal{D}_a$, so that $\varphi(x) \geq T_a g(x)$ for $x \in [b,\infty) \cap \mathcal{D}_a$ and $\varphi(x_0) = T_a g(x_0)$. We have $\varphi(b) > T_a g(b)$ and can choose φ on $(0,b)$ so that it is twice continuously differentiable and dominates $T_a g$ on all of $(0,a)$. Because $T_a g$ is a viscosity subsolution of (5.1) and $\varphi(x_0) = T_a g(x_0) > \gamma x_0$, we obtain

$$\mathcal{L}_g \varphi(x_0) = r T_a g(x_0) - \gamma(r x_0 - c) + \delta\gamma(x_0 - g(x_0)) \leq c \leq \gamma c + \delta\gamma(x_0 - g(x_0)),$$

and hence $T_a g(x_0) \leq \gamma x_0$, a contradiction to the choice of x_0. We conclude that $T_a g(x) \leq \overline{\varphi}(x)$ for $x \in [b,\infty) \cap \mathcal{D}_a$. ◻

From Corollary 5.2 and Lemmas 5.3 and 4.4, we have the following.

PROPOSITION 5.4. *If a is finite, then $\mathcal{C}_g = (0,b)$ for some $b \in (0,a]$ and $\mathcal{S}_g = \{0\} \cup [b,a]$. If $a = \infty$, then $\mathcal{C}_g = (0,b)$ for some $b \in (0,\infty)$ and $\mathcal{S}_g = \{0\} \cup [b,\infty)$.*

Let b be as in Proposition 5.4. We have already seen that $T_a g$ is twice continuously differentiable on $\mathcal{C}_g = (0,b)$ with a one-sided derivative $D^- T_a g(b)$ at b. Since $T_a g(x) = \gamma x$ on $\mathcal{S}_g$, this function is clearly differentiable on the set (b,a) if $b < a$, with one-sided derivative $D^+ T_a g(b) = \gamma$. It remains to examine the differentiability of $T_a g$ at the point b.

PROPOSITION 5.5 (smooth pasting). *The function $T_a g$ is continuously differentiable on $(0,a)$.*

Proof. It suffices to show in the case that $b < a$ that $D^- T_a g(b) = D^+ T_a g(b)$. Because $T_a g(x) \geq \gamma x$ for all $x \in \mathcal{D}_a$ and $T_a g(b) = \gamma b$, we must have $D^- T_a g(b) \leq \gamma$. If $D^- T_a g(b) < \gamma$, we choose $m \in (D^- T_a g(b), D^+ T_a g(b))$, $k > 0$, and define $\varphi(x) = \gamma b + m(x-b) + k(x-b)^2$ for x in an open interval containing b. Note that $\varphi(b) = T_a g(b)$ and $\varphi'(b) = m$. Therefore, $\varphi(x) < T_a g(x)$ for $x \neq b$ in a sufficiently small neighborhood of b (whose size depends on k). We construct φ outside this neighborhood so that φ is twice continuously differentiable on $(0,a)$ and $\varphi(x) \leq T_a g(x)$ for all $x \in (0,a)$. Because $T_a g$ is a viscosity supersolution of (5.1), the inequality

$$0 \leq \mathcal{L}_g \varphi(b) - c = r\gamma b - (rb - c)m + \delta(b - g(b))m - \sigma^2 b^2 k - c$$

must hold. Since $k > 0$ is arbitrary, this is impossible. ◻

We have proved so far the following properties of the value function $T_a g$: for any $g \in \overline{\mathcal{G}}_a$, $T_a g$ is a continuous function on $\mathcal{D}_a$ and it has a continuous derivative on $(0,a)$, $T_a g(0) = 0$ and $T_a g(x) \geq \gamma x$ for all $x \in \mathcal{D}_a$. If a is finite, then $T_a g(x) = \gamma a$; if $a = \infty$, then $T_\infty g(x) = \gamma x$ for x sufficiently large.

We now need to prove an invariance property for the operator T_a. Up to this point, we have taken g to be an arbitrary function in $\overline{\mathcal{G}}_a$. For the next proposition, we must restrict our attention to $g \in \mathcal{G}_a$.

PROPOSITION 5.6. *Let $a \in (0,\infty]$ be given. Then T_a maps $\mathcal{G}_a$ into $\mathcal{G}_a$.*

Proof. Assume that $g \in \mathcal{G}_a$. By the above remark, it remains only to show that $-M_a \leq DT_a g < 1$ on $(0,a)$ if a is finite and $0 \leq DT_a g < 1$ if $a = \infty$.

First we claim that the function $\psi = DT_a g$ (defined on $(0,a)$) cannot attain a positive local maximum or a negative local minimum in $\mathcal{C}_g$. By Proposition 4.6, ψ is C^2 on $\mathcal{C}_g$. Assume that ψ has a positive local maximum at $x_* \in \mathcal{C}_g$. Thus, we have $\psi'(x_*) = 0$. In particular,

$$\frac{d}{dx}(T_a g(x) - x\psi(x))\Big|_{x=x_*} = -x_*\psi'(x_*) = 0.$$

Equation (4.23) implies for $x \in \mathcal{C}_g$ that

$$c = \mathcal{L}_g T_a g(x)$$
$$= r(T_a g(x) - x\psi(x)) + c\psi(x) + \delta(x - g(x))\psi(x) - \frac{1}{2}\sigma^2 x^2 \psi'(x),$$

and thus

$$0 = \frac{d}{dx}\mathcal{L}_g T_a g(x)\Big|_{x=x_*} = \delta(1 - g'(x_*))\psi(x_*) - \frac{1}{2}\sigma^2 x_*^2 \psi''(x_*).$$

Because ψ has a local maximum at x_*, $\psi''(x_*) \leq 0$. But $1 - g'(x_*)$ is positive, and $\psi(x_*) > 0$. We have a contradiction, and hence ψ cannot have a positive local maximum in $\mathcal{C}_g$. If ψ has a negative local minimum at x_*, we likewise have a contradiction.

We consider now the case that $a = \infty$. For $x < y$ we have $X^x(t) \leq X^y(t)$ almost surely and $\tau_0^x \leq \tau_0^y$ almost surely. It follows from (3.5) that $T_\infty g$ is nondecreasing. The lower bound $DT_\infty g \geq 0$ is established. For the upper bound, $DT_\infty g(x) < 1$, we recall that $\mathcal{C}_g = (0, b)$ for some $b \in (0, \infty)$. Assume there were a point $x_0 \in (0, b)$ where $DT_\infty g(x_0) \geq 1$. We know that $DT_\infty g(b) = \gamma < 1$. Now consider a point $x_1 \in (0, x_0)$. If $DT_\infty g(x_1) < 1$, then $DT_\infty g$ would have a positive local maximum in the interval (x_1, b), which is impossible. We conclude that $DT_\infty g(x_1) \geq 1$. In other words, if there were a point $x_0 \in (0, b)$ where $DT_\infty g(x_0) \geq 1$, then $DT_\infty g \geq 1$ on the whole interval $(0, x_0)$. The upper bound in (4.14) would immediately imply that $T_\infty g(x) = x$ for $0 \leq x \leq x_0$, and once again $DT_\infty g$ would have a positive local maximum in $(0, b)$. We conclude that $DT_\infty g(x_0) < 1$ for all $x_0 \in (0, b)$.

If a is finite, we can modify the above argument, using (4.16) in place of (4.14) and $D^- T_a g(b) \leq \gamma$ (in case $\mathcal{C}_g = (0, a)$), to obtain the upper bound $DT_a g < 1$ on $(0, a)$.

The proof of the lower bound $DT_a g(x) \geq -M_a$ for the case $a < \infty$ is more involved. Again using the notation $\mathcal{C}_g = (0, b)$, we assume there is $x_0 \in (0, b)$ such that $DT_a g(x_0) < 0$. Let $x_1 \in (0, x_0)$. The continuous function $DT_a g$ attains its minimum on $[x_1, b]$ at x_1 or b, since it cannot attain a negative interior minimum. In case the minimum is attained at x_1, this means that $DT_a g(x_1) < DT_a(x_0) < 0$. For any $0 < x_2 < x_1$, DT_g cannot attain a negative interior minimum on $[x_2, x_0]$, so we can conclude that $DT_a g(x_2) < DT_a g(x_1) < 0$. This should hold for any $0 < x_2 < x_1$, which is in contradiction to $T_a g(0) = 0$, $T_a g(x) \geq \gamma x$. So if $DT_a g(x_0) < 0$, then for any $x_1 \in (0, x_0)$, $DT_a g$ attains its negative minimum on $[x_1, b]$ at b. This means that

$$D^- T_a g(b) \leq \inf_{0 < x \leq b} D_a T g(x) < 0. \tag{5.2}$$

In other words, the derivative $D_a T g$ either is nonnegative or, if it has negative values, is bounded below by $DT_a^- g(b)$. Of course, the latter case can only happen for $b = a$. The first case satisfies the conclusion, so we assume

$$D^- T_a g(a) = \min_{x \in (0, a]} DT_a g(x) < 0. \tag{5.3}$$

This means that $\mathcal{C}_g = (0, a)$ and hence $\mathcal{L}_g T_a g(x) = c$ for all $x \in (0, a)$. Let h satisfy $\mathcal{L}_g h(x) = c$ for $x \in (\gamma a, a)$ and $h(\gamma a) = \gamma a$, $h(a) = \gamma a$. Since $T_a g(\gamma a) \leq \gamma a = h(\gamma a)$, $T_a g(a) = \gamma a = h(a)$ and $\mathcal{L}_g T_a g(x) = \mathcal{L}_g h(x)$ for all $x \in (0, a)$, the usual comparison argument based on the maximum principle yields $T_a g(x) \leq h(x)$ for all $x \in [\gamma a, a]$. But $T_a g(a) = h(a)$, and this implies

$$D^- T_a g(a) \geq D^- h(a). \tag{5.4}$$

It suffices to find a lower bound on $D^-h(a)$. We have $0 \le \gamma x \le T_a g(x) \le h(x)$ for $x \in [\gamma a, a]$. In order to find an upper bound on h, we let $x^* \in [\gamma a, a]$ be such that $h(x^*) = \max_{x\in[\gamma a,a]} h(x)$. If x^* is an interior point of $[\gamma a, a]$, then $h'(x^*) = 0$ and $h''(x^*) \le 0$. But $\mathcal{L}_g h(x^*) = c$ from which we conclude that $\max_{x\in[\gamma a,a]} h(x) = h(x^*) \le \frac{c}{r}$. If x^* is not an interior point of $[\gamma a, a]$, then $\max_{x\in[\gamma a,a]} h(x) \le h(\gamma a) = h(a) = \gamma a$. In either case, we have

$$0 \le h(x) \le \max\left\{\gamma a, \frac{c}{r}\right\} \quad \forall x \in [\gamma a, a]. \tag{5.5}$$

We know that

$$0 \le g(x) \le a \quad \forall\, x \in [\gamma a, a]. \tag{5.6}$$

Neither (5.5) nor (5.6) depends on the lower bound $-M_a \le g'(x)$ satisfied by functions g in $\mathcal{C}_a$ when a is finite.

Since $h(\gamma a) = h(a)$, there exists $x_0 \in (\gamma a, a)$ such that $h'(x_0) = 0$. We solve the equation $\mathcal{L}_g h = c$ on $(\gamma a, a)$ for h'' and then integrate to obtain

$$h'(x) = \int_{x_0}^{x} \frac{2}{\sigma^2 y^2}\big[rh(y) - (ry - c)h'(y) + \delta(y - g(y))h'(y) - c\big]\,dy \tag{5.7}$$

for all $x \in [x_0, a]$. Taking into account the bounds (5.5) and (5.6), we may use Gronwall's inequality to obtain $|h'(a)| \le M_a$ for some constant M_a depending only on the bounds $\max\{\gamma a, \frac{c}{r}\}$ and a appearing in (5.5) and (5.6) and also depending on the interval $[\gamma a, a]$. From (5.3), (5.4) we conclude that $DT_a g(x) \ge -M_a$ for all $x \in (0, a)$. ∎

Remark 5.7. M_a is bounded in a as long as a is bounded away from 0.

6. The fixed point property.

For $a = \infty$ we recall that $\overline{\mathcal{G}}_\infty$ is a closed subset of the complete metric space (C_γ, d) (see Definition 3.1). For $a < \infty$, the set $\overline{\mathcal{G}}_a$ is a closed convex subset of the Banach space $C[0, a]$ endowed with the supremum norm. We denote by $d(f, g)$ the metric associated with the supremum norm. We have proved that $T_\infty(\overline{\mathcal{G}}_\infty) \subset C_\gamma$ and $T_a(\overline{\mathcal{G}}_a) \subset C[0, a]$ for $a < \infty$. We also know (in both cases $a = \infty$ and $a < \infty$) that $T_a(\mathcal{G}_a) \subset \mathcal{G}_a$. In this section we prove that $T_a(\overline{\mathcal{G}}_a) \subset \overline{\mathcal{G}}_a$ and the operator T_a has a unique fixed point in $\overline{\mathcal{G}}_a$. Many of the arguments in the rest of the paper are based on the following lemma.

LEMMA 6.1 (comparison). *Let $0 \le \alpha < \beta$ and $f, g \in C(\alpha, \beta)$ be given. Consider $\varphi \in C^1(\alpha, \beta)$ a viscosity subsolution of $\mathcal{L}_f\varphi(x) \le c$ on (α, β) and $\psi \in C^1(\alpha, \beta)$ a viscosity supersolution of $\mathcal{L}_g\psi(x) \ge c$ on (α, β). Assume that at least one of the functions is a classical ($C^2(\alpha, \beta)$) solution of the corresponding differential inequality and that the function $\varphi - \psi$ attains a local maximum at $x_* \in (\alpha, \beta)$. Then*

$$r(\varphi(x_*) - \psi(x_*)) \le \delta(f(x_*) - g(x_*))\varphi'(x_*) = \delta(f(x_*) - g(x_*))\psi'(x_*).$$

Proof. Let us assume that $\varphi \in C^2(\alpha, \beta)$ is a classical solution of $\mathcal{L}_f\varphi(x) \le c$. (The argument in the other case is identical.) This means that

$$r\varphi(x_*) - (rx_* - c)\varphi'(x_*) + \delta(x_* - f(x_*))\varphi'(x_*) - \frac{1}{2}\sigma^2 x_*^2\varphi''(x_*) \le c.$$

The function $\psi - \varphi$ attains a local minimum at x_*, and since φ is C^2 in a neighborhood of x_*, we can consider φ as a test function when we apply the definition of the viscosity supersolution ψ. We obtain the inequality

$$r\psi(x_*) - (rx_* - c)\varphi'(x_*) + \delta(x_* - g(x_*))\varphi'(x_*) - \frac{1}{2}\sigma^2 x_*^2\varphi''(x_*) \ge c.$$

Comparing the above results, we conclude that

$$r(\varphi(x_*) - \psi(x_*)) \le \delta(f(x_*) - g(x_*))\varphi'(x_*).$$

Since x_* is a point of interior maximum for $\varphi - \psi$, and both φ and ψ have continuous derivatives on $(0, \alpha)$, we have that $\varphi'(x_*) = \psi'(x_*)$. ∎

PROPOSITION 6.2. *For $0 < a \le \infty$, we have $T_a(\overline{\mathcal{G}}_a) \subset \overline{\mathcal{G}}_a$, and the mapping T_a has a unique fixed point in $\overline{\mathcal{G}}_a$.*

Proof. Let $f, g \in \overline{\mathcal{G}}_a$ be given. We denote $\varphi = T_a f$ and $\psi = T_a g$. Since $\varphi(0) = \psi(0) = 0$, we know that $\sup_{x\in\mathcal{D}_a}(\varphi(x) - \psi(x)) \ge 0$. We recall that φ, ψ are continuous on $[0, a]$ for finite a (or they are continuous on $[0, \infty)$ and equal to γx for x large enough if $a = \infty$). Thus there exists x_* such that $\varphi(x_*) - \psi(x_*) = \max_{x\in[0,a]}(\varphi(x) - \psi(x))$. If $\varphi(x_*) - \psi(x_*) = 0$, then

$$\sup_{x\in\mathcal{D}_a}(\varphi(x) - \psi(x)) \le \varphi(x_*) - \psi(x_*) = 0 \le \frac{\delta}{r}\max\{M_a, 1\}d(f, g).$$

Assume that $\varphi(x_*) - \psi(x_*) > 0$. Since $\varphi(0) = \psi(0)$ and $\varphi(a) = \psi(a)$ (or $\varphi(x) = \psi(x)$ for all x large enough if $a = \infty$), we see that $0 < x_* < a$. Moreover, since $\varphi(x_*) > \psi(x_*) \ge \gamma x_*$, we know $x_* \in \mathcal{C}_f = \{x : \varphi(x) > \gamma x\}$.

We remember that φ is a C^2 function on the open set $\mathcal{C}_f$, it is a classical solution of $\mathcal{L}_f\varphi = c$ on $\mathcal{C}_f$, and ψ is a viscosity supersolution of $\mathcal{L}_g\psi \ge c$. Lemma 6.1 implies $r(\varphi(x_*) - \psi(x_*)) \le \delta(f(x_*) - g(x_*))\varphi'(x_*)$. Therefore,

$$\sup_{x\in[0,a]}\big(\varphi(x) - \psi(x)\big) \le \varphi(x_*) - \psi(x_*) \le \frac{\delta}{r}|f(x_*) - g(x_*)||\varphi'(x_*)|.$$

Since $\varphi'(x_*) = \psi'(x_*)$, it is enough to assume that at least one of the functions f and g is an element of $\mathcal{G}_a$ to conclude that $|\varphi'(x_*)| \le \max\{M_a, 1\}$, where $M_a = 0$ for $a = \infty$. Consequently, we obtain

$$\sup_{x\in[0,a]}\big(\varphi(x) - \psi(x)\big) \le \frac{\delta}{r}\max\{M_a, 1\}d(f, g).$$

We can switch φ and ψ in the argument above and obtain a similar inequality for $\psi - \varphi$. In other words, we have proved that

$$d(T_a f, T_a g) \le \frac{\delta}{r}\max\{M_a, 1\}d(f, g), \tag{6.1}$$

provided that at least one of the functions f and g is an element of $\mathcal{G}_a$.

We now choose $f \in \overline{\mathcal{G}}_a$, and let $f_n \in \mathcal{G}_a$ be such that $d(f_n, f) \to 0$ as $n \to \infty$. Using (6.1) we immediately obtain $d(T_a f_n, T_a f) \to 0$ as $n \to \infty$, and since $T_a f_n \in \mathcal{G}_a$ for all n, we conclude that $T_a f \in \overline{\mathcal{G}}_a$.

A similar approximation argument ($f_n \to f,\ g_n \to g,\ f_n, g_n \in \mathcal{G}_a$), together with

$$d(T_a f, T_a g) \le d(T_a f, T_a f_n) + d(T_a f_n, T_a g_n) + d(T_a g_n, T_a g)$$

yields

$$d(T_a f, T_a g) \le \frac{\delta}{r}\max\{M_a, 1\}d(f, g) \quad \forall f, g \in \overline{\mathcal{G}}_a.$$

We consider separately the two cases $a = \infty$ and $a < \infty$. If $a = \infty$, then $M_\infty = 0$. Since $\delta < r$, T_a is a contraction on the complete metric space $(\overline{\mathcal{G}}_\infty, d)$. Applying the Banach fixed point theorem, we conclude that T_a has a unique fixed point in $\overline{\mathcal{G}}_\infty$.

If $a < \infty$, the Arzela–Ascoli theorem implies that $\overline{\mathcal{G}}_a$ is a convex and compact subset of the Banach space $C[0,a]$. Since $T_a : \overline{\mathcal{G}}_a \to \overline{\mathcal{G}}_a$ is a continuous mapping with respect to the norm of $C[0,a]$, Schauder's fixed point theorem implies that there exists a fixed point of T_a in $\overline{\mathcal{G}}_a$. Suppose there were two fixed points of T_a, namely f and g. Assume without loss of generality that

$$f(x_*) - g(x_*) = \max_{x \in [0,a]} \big(f(x) - g(x)\big) > 0,$$

so $x_* \in \mathcal{C}_f$. We apply Lemma 6.1 to conclude

$$r(f(x_*) - g(x_*)) \leq \delta f'(x_*)(f(x_*) - g(x_*)),$$

which is impossible since $f(x_*) - g(x_*) > 0$, $\delta < r$, and $f'(x_*) \leq 1$. (We use here the fact that f has a continuous derivative on $(0,a)$ and $f \in \overline{\mathcal{G}}_a$ to conclude $f'(x_*) \leq 1$.) This means that $f \leq g$ on $[0,a]$. Interchanging f and g, we obtain $f = g$, so the fixed point is unique. □

We denote by f_a the unique fixed point of T_a in $\overline{\mathcal{G}}_a$. The function f_a is continuous on $\mathcal{D}_a$ and continuously differentiable on $(0,a)$. Associated with the function f_a is a number $b_a \in (0,a]$ such that

$$\mathcal{L}_{f_a} f_a(x) = c, \quad f_a(x) > \gamma x, \quad 0 < x < b_a, \tag{6.2}$$

$$\mathcal{L}_{f_a} f_a(x) \geq c, \quad f_a(x) = \gamma x, \quad b_a < x \leq a. \tag{6.3}$$

Even if $a = \infty$, b_∞ is finite.

Proposition 6.3. *The number b_a is given by*

$$b_a = \begin{cases} a & \text{if } a \leq b_\infty, \\ b_\infty & \text{if } a \geq b_\infty. \end{cases} \tag{6.4}$$

Proof. The proof is based on the same comparison argument for viscosity solutions that allowed us to conclude that the fixed point is unique in Proposition 6.2, namely an application of Lemma 6.1.

Consider first the case $a \leq b_\infty$, and suppose $b_a < a$. The function f_a is defined only on $[0,a]$, but we may extend it by the formula

$$\overline{f}_a(x) = \begin{cases} f_a(x) & \text{if } 0 \leq x \leq a, \\ \gamma x & \text{if } x \geq a. \end{cases} \tag{6.5}$$

It is apparent from (6.3) that $\overline{f}_a$ is continuous on $[0,\infty)$ and continuously differentiable on $(0,\infty)$. Furthermore, for $x \geq a$ we have

$$c \leq \mathcal{L}_{\overline{f}_a} \overline{f}_a(a) = c\gamma + \delta\gamma(1-\gamma)a \leq c\gamma + \delta\gamma(1-\gamma)x = \mathcal{L}_{\overline{f}_a} \overline{f}_a(x). \tag{6.6}$$

Using (6.6) we conclude that $\overline{f}_a$ is a viscosity solution of the equation $\min\{\mathcal{L}_f f(x) - c, f(x) - \gamma x\} = 0$ on $(0,\infty)$. Furthermore, $\overline{f}_a$ has a continuous derivative on $(0,\infty)$ and $\overline{f}_a(x) = \gamma x$ for large x. We can now compare $\overline{f}_a$ and f_∞. We know that either

$$\sup_{x \in [0,\infty)} (\overline{f}_a(x) - f_\infty(x)) \leq 0$$

or there exists $x_* \in (0, b_\infty)$ such that

$$\overline{f}_a(x_*) - f_\infty(x_*) = \sup_{x\in[0,\infty)} (\overline{f}_a(x) - f_\infty(x)) > 0.$$

In the latter case, $x_* \in \mathcal{C}_{\overline{f}_a}$ and Lemma 6.1 implies

$$r(\overline{f}_a(x_*) - f_\infty(x_*)) \leq \delta f'_\infty(x_*)(\overline{f}_a(x_*) - f_\infty(x_*)),$$

which is impossible since $r < \delta$ and $f'_\infty(x_*) \leq 1$. This means that the only possibility is $\overline{f}_a \leq f_\infty$. In the same way we prove that $f_\infty \leq \overline{f}_a$, so $\overline{f}_a = f_\infty$. This implies that $b_a = b_\infty$, which contradicts the hypothesis $b_a < a \leq b_\infty$.

The case $a > b_\infty$ is similar since $f_a(a) = \gamma a = f_\infty(a)$, and the restriction of f_∞ to $[0, a]$ is a viscosity solution of (3.1) on $(0, a)$. We can use the same comparison argument to conclude that $f_\infty|_{[0,a]} = f_a$, which implies $b_a = b_\infty$.

COROLLARY 6.4. *For $0 < a \leq \infty$, the function f_a is in $\mathcal{G}_a$.*

Proof. We have already seen that f_a is continuously differentiable, and since $f_a \in \overline{\mathcal{G}}_a$, we conclude that $-M_a \leq f'_a(x) \leq 1$ for $0 < x < a$. It remains only to prove that the derivative f'_a cannot attain the value 1.

Assume, by contradiction, that $f'_a(x_0) = 1$ for some $x_0 \in (0, a)$. This means f'_a has a maximum at x_0 and $x_0 \in \mathcal{C}_{f_a}$, where f_a is two times continuously differentiable. Hence, $f''_a(x_0) = 0$. Moreover, $\mathcal{L}_{f_a} f_a(x_0) = c$, so $(r - \delta)(x_0 - f_a(x_0)) = 0$. Since $r - \delta > 0$ we see that $f_a(x_0) = x_0$. The function f_a is thus a solution of the ordinary differential equation $\mathcal{L}_f f(x) = c$ with initial conditions $f(x_0) = x_0,\ f'(x_0) = 1$ on the interval $[x_0, b_a]$. However, the only such solution to this equation is $f(x) = x$, and we conclude that $f_a(x) = x$ for $x_0 \leq x \leq b_a$. This contradicts the fact that $f_a(b_a) = \gamma b_a < b_a$.

COROLLARY 6.5. *For every $0 < a \leq \infty$, the function f_a is concave for small values of x, it has a right derivative at $x = 0$, and $D^+ f_a(0) \leq 1$.*

Proof. Since $f_a = T_a f_a$ and we just proved that $f_a \in \mathcal{G}_a$, we know from the first part of the proof of Proposition 5.6 that the derivative $f'_a = DT_a f_a$ cannot attain a positive local maximum in $(0, b_a)$. Since $f_a(0) = 0$, $f_a(b_a) = \gamma b_a$, and f_a is differentiable on $(0, b_a)$, we can conclude from the mean-value theorem that there exists $x_\gamma \in (0, b_a)$ with $f'_a(x_\gamma) = \gamma$. Since $D^- f_a(b_a) \leq \gamma$, we can argue that for any $x_1 < x_2 \leq x_\gamma$ we have $f'_a(x_1) > f'_a(x_2)$. To do this, we first use the fact that f'_a cannot attain a positive interior maximum on $[x_2, b_a]$ to conclude that $f'_a(x_2) > f'_a(x_\gamma) = \gamma$ and then use the fact that f'_a cannot attain a positive interior maximum on $[x_1, x_\gamma]$ to further conclude that $f'_a(x_1) > f'_a(x_2)$. In other words, the derivative f'_a is strictly decreasing on $(0, x_\gamma)$. This means that the function f_a is concave on $[0, x_\gamma]$ and

$$D^+ f_a(0) \triangleq \lim_{x\to 0} \frac{f_a(x) - 0}{x - 0} = \lim_{x\to 0} f'_a(x) \tag{6.7}$$

is well defined. It is obvious that $D^+ f_a(0) \leq 1$.

7. Proofs of Theorems 2.2, 2.4, and 2.5. For each call price K we construct a function f^* so that $f^*(x)$ is the value of the convertible bond when the value of the firm is x. For small values of x, the function $f^*(x)$ agrees with $f_a(x)$ for an appropriately chosen a, depending on K. In order to proceed, we must first understand the dependence of f_a on the parameter a. For this purpose, we define $m\colon (0, \infty) \to (0, \infty)$ by

$$m(a) = \max_{x\in[0,a]} f_a(x). \tag{7.1}$$

Because $f_a = f_\infty|_{[0,a]}$ for $a \geq b_\infty$ and f_∞ is nondecreasing by virtue of its membership in $\overline{\mathcal{G}}_\infty$, we have

$$m(a) = \gamma a \quad \forall a \geq b_\infty. \tag{7.2}$$

For $a < b_\infty$ and $x \in (0, a)$, we have $f_a(x) > \gamma x$ (Proposition 6.3 and the inequality in (6.2)), and so it is possible that $m(a) > \gamma a$ for $0 < a < b_\infty$. We shall in fact discover that there is a number $b_0 \in [0, b_\infty)$ such that $m(a) > \gamma a$ for $0 < a < b_0$, whereas $m(a) = \gamma a$ for $a \geq b_0$ (see Remark 7.3).

LEMMA 7.1. *The function* $m\colon (0,\infty) \to (0,\infty)$ *is strictly increasing and continuous and satisfies* $\lim_{a\downarrow 0} m(a) = 0$.

Proof. It is clear from (7.2) that m is strictly increasing on $[b_\infty, \infty)$. We first show that m is nondecreasing on $(0, b_\infty]$. Let $0 < a_1 < a_2 \leq b_\infty$ be given. Since $f_{a_1}(0) = 0 = f_{a_2}(0)$ and $f_{a_1}(a_1) = \gamma a_1 < f_{a_2}(a_1)$, if the function $f_{a_1} - f_{a_2}$ attains a positive maximum over $[0, a_1]$ it must be at an interior point $x_* \in (0, a_1)$. But $\mathcal{L}_{f_{a_1}} f_{a_1}(x) = c = \mathcal{L}_{f_{a_2}} f_{a_2}(x)$ for $0 < x < a_1$, and $x \in \mathcal{C}_{f_{a_1}}$, where f_{a_1} is C^2. Lemma 6.1 implies that

$$r\big(f_{a_1}(x_*) - f_{a_2}(x_*)\big) \leq \delta\big(f_{a_1}(x_*) - f_{a_2}(x_*)\big) f'_{a_1}(x_*),$$

which is impossible because $\delta < r$ and $f'_{a_1}(x_*) \leq 1$. We conclude that $f_{a_1}(x) \leq f_{a_2}(x)$ for all $x \in [0, a_1]$. Therefore m is nondecreasing on $(0, b_\infty]$.

By the same comparison argument, the function $f_{a_2} - f_{a_1}$ cannot attain a positive maximum in $(0, a_1)$, so $f_{a_2}(x) - f_{a_1}(x) \leq f_{a_2}(a_1) - \gamma a_1$ for $0 \leq x \leq a_1$. It follows that

$$\begin{aligned} m(a_2) - m(a_1) &= \max\Big\{ \max_{x\in[0,a_1]} f_{a_2}(x), \max_{x\in[a_1,a_2]} f_{a_2}(x)\Big\} - m(a_1) \\ &\leq \max\Big\{ \max_{x\in[0,a_1]} \big(f_{a_2}(x) - f_{a_1}(x)\big), \max_{x\in[a_1,a_2]} \big(f_{a_2}(x) - \gamma a_1\big)\Big\} \\ &= \max\Big\{ f_{a_2}(a_1) - \gamma a_1, \max_{x\in[a_1,a_2]} \big(f_{a_2}(x) - \gamma a_1\big)\Big\} \\ &= -\gamma a_1 + \max_{x\in[a_1,a_2]} f_{a_2}(x). \end{aligned} \tag{7.3}$$

By virtue of its membership in $\overline{\mathcal{G}}_{a_2}$ and Remark 5.7, the function f_{a_2} satisfies $f'_{a_2}(x) \geq -C$ for all $x \in (0, a_2)$ and some positive constant C which is bounded away from zero so long as a_2 is bounded away from zero. Thus, for $x \in [a_1, a_2]$,

$$f_{a_2}(x) = f_{a_2}(a_2) - \int_x^{a_2} f'_{a_2}(y)\, dy \leq \gamma a_2 + C(a_2 - x) \leq \gamma a_2 + C(a_2 - a_1).$$

Substituting this into (7.3), we conclude that

$$0 \leq m(a_2) - m(a_1) \leq (C + \gamma)(a_2 - a_1), \tag{7.4}$$

so long as a_2 is bounded away from zero. The function m is thus continuous.

We now prove that $m(a_1) < m(a_2)$. Assume, by contradiction, that $m(a_1) = m(a_2)$. Let $x_0 \in [0, a_1]$ be such that $f_{a_1}(x_0) = m(a_1)$. We must actually have $x_0 \in (0, a_1)$ because $m(a_1) = m(a_2) \geq \gamma a_2 > \gamma a_1 = f_{a_1}(a_1) > 0 = f_{a_1}(0)$. We have already shown that f_{a_2} dominates f_{a_1} on $[0, a_1]$, and hence we must have $f_{a_1}(x_0) = f_{a_2}(x_0)$. The comparison argument using Lemma 6.1 shows that neither $f_{a_2} - f_{a_1}$ nor $f_{a_1} - f_{a_2}$ can have a positive maximum in the open interval $(0, x_0)$; we conclude that

$$f_{a_1}(x) = f_{a_2}(x) \quad \forall x \in [0, x_0]. \tag{7.5}$$

Both f_{a_1} and f_{a_2} are solutions of the ordinary differential equation $\mathcal{L}_f f(x) = c$ on $[x_0, a_1]$ and have the same initial conditions $f_{a_1}(x_0) = f_{a_2}(x_0)$, $f'_{a_1}(x_0) = f'_{a_2}(x_0)$. It follows that

$$f_{a_1}(x) = f_{a_2}(x) \quad \forall x \in [x_0, a_1]. \tag{7.6}$$

This implies that $f_{a_2}(a_1) = f_{a_1}(a_1) = \gamma a_1$, which contradicts Proposition 6.3. We conclude that m is strictly increasing on $(0, b_\infty]$.

Finally, since $f_a(x) \leq x$ for $0 \leq x \leq a$, we see that $0 \leq m(a) \leq a$, and consequently $\lim_{a\downarrow 0} m(a) = 0$. $\square$

LEMMA 7.2.

(i) *Assume* $m(\overline{a}) > \gamma\overline{a}$ *for some* $\overline{a} > 0$. *Then* $\overline{a} < \frac{c}{r\gamma}$ *and* $m(a) > \gamma a$ *for all* $a \in (0, \overline{a})$.

(ii) *If* $m(a) > \gamma a$, *the function* f_a *attains its maximum over* $[0, a]$ *at a unique point* $x_a \in (0, a)$.

Proof. (i) If $a \geq \frac{c}{r\gamma}$, we define $h(x) = \gamma a \geq \frac{c}{r}$ for $x \in [0, a]$. Then $\mathcal{L}_{f_a} h(x) \geq c$ for $0 < x < a$. Lemma 6.1 shows that $f_a - h$ cannot have a positive maximum in $(0, a)$, and since $f_a(0) = 0 \leq h(0)$ and $f_a(a) = \gamma a = h(x)$, we conclude that $f_a(x) \leq h(a)$ for all $0 \leq x \leq a$. Consequently, the maximum of f_a is $m(a) = \gamma a$.

Assume now that $m(\overline{a}) > \gamma\overline{a}$ for some $\overline{a} > 0$. We have just seen that $\overline{a} < \frac{c}{r\gamma}$. Let $a \in (0, \overline{a})$ be given. Define $\ell = a/\overline{a} < 1$ and rescale the function $f_{\overline{a}}$ by setting $f(x) = \ell f_{\overline{a}}(\frac{x}{\ell})$ for all $x \in [0, a]$. We compute $f'(x) = f'_{\overline{a}}(\frac{x}{\ell})$ and $f''(x) = \frac{1}{\ell} f''_{\overline{a}}(\frac{x}{l})$, from which we conclude that

$$\mathcal{L}_f f(x) = \ell \mathcal{L}_{f_{\overline{a}}} f_{\overline{a}}\left(\frac{x}{l}\right) + c(1-l) f'_{\overline{a}}\left(\frac{x}{l}\right) \leq \ell c + c(1-\ell) = c \quad \forall x \in (0, a).$$

Lemma 6.1 shows that $f - f_a$ cannot have a positive maximum over $[0, a]$ at a point in $(0, a)$. But $f(0) = f_a(0) = 0$ and $f(a) = f_a(a) = \gamma a$, and therefore $f_a(x) \geq f(x)$ for all $x \in [0, a]$. In particular,

$$m(a) = \max_{x\in[0,a]} f_a(x) \geq \max_{x\in[0,a]} f(x) = \ell m(\overline{a}) > \ell\gamma\overline{a} = \gamma a. \tag{7.7}$$

(ii) Let us assume now that $m(a) > \gamma a$ and there exist $0 < x_0 < y_0 < a$ such that $f_a(x_0) = f_a(y_0) = m(a)$. Since $f_a(x) \leq m(a)$ for $x_0 \leq x \leq y_0$ we see that f_a has a local minimum at some point $x_1 \in (x_0, y_0)$. Then $f'_a(x_1) = 0$, $f''_a(x_1) \geq 0$, and we may use the equation $\mathcal{L}_{f_a} f_a(x_1) = c$ to obtain $r f_a(x_1) \geq c$. This is impossible because $f_a(x_1) \leq m(a) < m(\frac{c}{\gamma r}) = \frac{c}{r}$.

Remark 7.3. We define $b_0 \triangleq \sup\{a > 0, m(a) > \gamma a\}$, where we set $b_0 = 0$ if $m(a) = \gamma a$ for all $a > 0$. Lemma 7.2 shows that $m(a) > \gamma a$ for all $a \in (0, b_0)$. This lemma further shows that $b_0 \leq \frac{c}{\gamma r}$. Since for $x \geq b_\infty$ we have $f_\infty(x) = \gamma x$ and $\mathcal{L}_{f_\infty} f_\infty(x) \geq c$, we conclude that

$$r\gamma b_\infty - (r b_\infty - c)\gamma + \delta(b_\infty - \gamma b_\infty)\gamma \geq c,$$

which implies $\delta(1-\gamma) b_\infty \gamma \geq c(1-\gamma)$, and consequently $b_\infty \geq \frac{c}{\gamma\delta}$. In summary,

$$0 \leq b_0 \leq \frac{c}{\gamma r} < \frac{c}{\gamma\delta} \leq b_\infty. \tag{7.8}$$

LEMMA 7.4. *If* $0 < \gamma < \frac{1}{2}$, *then* $b_0 > 0$.

Proof. For small values of a, we construct a quadratic subsolution of

$$\begin{cases} \mathcal{L}_g g(x) \le c \text{ for } \ 0 < x < a, \\ g(0) = 0, \ g(a) = \gamma a, \end{cases} \tag{7.9}$$

which satisfies $\max_{x \in [0,a]} g(x) > \gamma a$. According to Lemma 6.1, $g - f_a$ cannot have a positive maximum over $[0, a]$ in $(0, a)$, and since $g(0) = f_a(0) = 0$, $g(a) = f_a(a) = \gamma a$, we see that $f_a \ge g$ on $[0, a]$. It follows that $m(a) > \gamma a$.

The remainder of the proof is the construction of g. We define

$$g(x) = -\frac{x^2}{2a} + \left(\gamma + \frac{1}{2}\right) x,$$

so that $g(0) = 0$ and $g(a) = \gamma a$. Direct computation results in

$$\begin{aligned} &\mathcal{L}_g g(x) \\ &= \frac{rx^2}{2a} - \frac{cx}{a} + \left(\gamma + \frac{1}{2}\right) c - \frac{\delta x^3}{2a^2} + \frac{3\delta\gamma x^2}{2a} - \frac{\delta x^2}{4a} - \delta\gamma^2 x + \frac{\delta x}{4} + \frac{\sigma^2 x^2}{2a} \\ &\le \frac{ra}{2} + \left(\gamma + \frac{1}{2}\right) c + \frac{3\delta\gamma a}{2} + \frac{\delta a}{4} + \frac{\sigma^2 a}{2} \quad \forall x \in [0, a]. \end{aligned}$$

Since $\left(\gamma + \frac{1}{2}\right) c < c$, we have $\sup_{x \in [0,a]} \mathcal{L}_g g(x) \le c$ for sufficiently small a. ☐

We summarize what has so far been established.

(a) For $a > b_\infty$ we have $f_a = f_\infty|_{[0,a]}$ and the maximum $m(a) = \gamma a$ of f_a over $[0, a]$ is attained at the right endpoint a. We have $f_a(x) > \gamma x$ for $x \in (0, b_\infty)$ and $f_a(x) = \gamma x$ for $x \in [b_\infty, a]$.

(b) For $b_0 \le a \le b_\infty$, the maximum $m(a) = \gamma a$ of f_a over $[0, a]$ is attained at the right endpoint a and $f_a(x) > \gamma x$ for all $x \in (0, a)$.

(c) If $b_0 > 0$ (a sufficient condition for this is $0 < \gamma < \frac{1}{2}$), then for $0 < a < b_0$, we have $f_a(x) > \gamma x$ for all $x \in (0, a)$ and the maximum $m(a) > \gamma a$ of f_a over $[0, a]$ is attained at a unique point $x_a \in (0, a)$.

For a fixed call price K we want to define $f^*(x)$ to be $f_a(x)$ for small values of x, where a is the unique parameter such that $m(a) = K$. Denoting

$$K_1 = \gamma b_0, \quad K_2 = \gamma b_\infty,$$

we have the following three situations corresponding to the three cases of Theorem 2.5.

(i) If $K > K_2$, we set $a = \frac{K}{\gamma}$. We define

$$f_*(x) = \begin{cases} f_a(x) = f_\infty(x) & \text{if } 0 \le x \le a, \\ \gamma x & \text{if } x \ge a. \end{cases} \tag{7.10}$$

We see that $f_*(x) = f(x, C_a^*, C_o^*)$ for $C_a^* = \frac{K}{\gamma}$ and $C_o^* = b_\infty < C_a^*$.

(ii) If $K_1 \le K \le K_2$, then again we set $a = \frac{K}{\gamma}$. We define

$$f_*(x) = \begin{cases} f_a(x) & \text{if } 0 \le x \le a, \\ \gamma x & \text{if } x \ge a. \end{cases} \tag{7.11}$$

In this case, $f_*(x) = f(x, C_a^*, C_o^*)$ for $C_a^* = C_o^* = \frac{K}{\gamma}$.

(iii) Assume $K_1 > 0$ and $0 < K < K_1$. Because $m(b_0) = K_1$, there exists a unique $a = m^{-1}(K) < b_0$ such that $m(a) = K$. Since $K < K_1$, Lemma 7.2 implies that $K = m(a) > \gamma a$ and there exists a unique $x_a \in (0, a)$ such that $f_a(x_a) = m(a)$. Since $f'_a < 1$, we obtain that $K = m(a) = f_a(x_a) < x_a$, so $K < x_a < a < \frac{K}{\gamma}$. We now take $C^*_a = x_a$, $C_o = \frac{K}{\gamma}$ and define

$$(7.12) \qquad f^*(x) = \begin{cases} f_a(x) & \text{for } 0 \le x \le C^*_a, \\ K & \text{for } C^*_a \le x \le C^*_o, \\ \gamma x & \text{for } x \ge C^*_o. \end{cases}$$

Again we have $f_*(x) = f(x, C^*_a, C^*_o)$. Since $f'_a(C^*_a) = 0$, f_* is a C^1 function on $(0, \frac{K}{\gamma})$.

It is apparent that the function f_* and the numbers C^*_o, C^*_a just defined have all the properties set forth in Theorem 2.5. The uniqueness of solutions to $\mathcal{N} f = c$ in that theorem follows from Lemma 6.1. We now accept Theorem 2.2, whose proof will be given later in this section, and show that the function f_* defined by (7.10)–(7.12) is indeed the function f_* given by (2.19), and the numbers C^*_o and C^*_a defined above satisfy (2.18). Using f_*, C^*_o, and C^*_a just defined in this way means that the proof of Theorem 2.4 given below also completes the proof of Theorem 2.5.

Proof of Theorem 2.4. We need to prove that

$$(7.13) \qquad f(x, C^*_a, C^*_o) \le f(x, C_a, C^*_o) \text{ for each } C_a \ge K,\ x \in (0, \infty),$$

$$(7.14) \qquad f(x, C^*_a, C^*_o) \ge f(x, C^*_a, C_o) \text{ for each } C_o > 0,\ x \in (0, \infty).$$

Case (i). $K > K_2 = \gamma b_\infty$.

If $C_a \ge C^*_a = \frac{K}{\gamma}$, then clearly $f(x, C^*_a, C^*_o) = f(x, C_a, C^*_o)$ for $x \in (0, \infty)$.

If $C^*_o < C_a \le \frac{K}{\gamma}$, according to Definition 2.3(i) we have $f(x, C_a, C^*_o) = f(x, C^*_a, C^*_o)$ for $0 \le x < C_a$, and $f(x, C_a, C^*_o) = K \ge f(x, C^*_a, C^*_a)$ for $C_a \le x \le \frac{K}{\gamma}$. For $x \ge \frac{K}{\gamma}$, we have $f(x, C^*_a, C^*_o) = f(x, C_a, C^*_o) = \gamma x$.

Finally, consider the case $K \le C_a \le C^*_o = b_\infty$. Using the Case (i) assumption, we have $K > K_2 = \gamma b_\infty = \gamma C^*_o \ge \gamma C_a$. From Definition 2.3(ii),

$$f(C_a, C_a, C^*_o) = \max\{K, \gamma C_a\} = K \ge \gamma C_a = f_*(C_a).$$

Since $f(\cdot) = f(\cdot, C_a, C^*_o)$ satisfies $\mathcal{L}_f f(x) = c$ on $(0, C_a)$ and $\mathcal{L}_{f_*} f_*(x) = c$ on $(0, C_a)$, an application of Lemma 6.1 yields

$$f(x, C^*_a, C^*_o) = f_*(x) \le f(x, C_a, C^*_o) \text{ for } 0 \le x \le C_a.$$

For $C_a \le x \le \frac{K}{\gamma}$, we have

$$f(x, C_a, C^*_o) = \max\{K, \gamma x\} = K = f_*\Big(\frac{K}{\gamma}\Big) \ge f_*(x) = f(x, C^*_a, C^*_o).$$

For $x > \frac{K}{\gamma}$, we have $f(x, C^*_a, C^*_o) = \gamma x = f(x, C_a, C^*_o)$. This completes the proof of (7.13) in Case (i).

To establish (7.14), we let $C_o > 0$ be given. If $C_o \le C^*_o$, then $f(C_o, C^*_a, C_o) = \gamma C_o \le f_*(C_o)$. Applying Lemma 6.1, we get

$$(7.15) \qquad f(x, C^*_a, C_o) \le f_*(x) = f(x, C^*_a, C^*_o) \text{ for } 0 \le x \le C_o.$$

The same inequality is easily verified for $C_o \le x < \infty$.

The case $C_o^* < C_o \le C_a^*$ is the most interesting. We know that the function $f(\cdot) = f(\cdot, C_a^*, C_o)$ satisfies

$$\begin{cases} \mathcal{L}_f f(x) = c \text{ for } 0 < x < C_o, \\ f(0) = 0,\ f(C_o) = \gamma C_o = f^*(C_o) \text{ (since } C_o > C_o^* = b_\infty). \end{cases}$$

We recall that f_* is a C^1 viscosity supersolution of $\mathcal{L}_{f_*} f_*(x) = c$ on $(0, C_o)$, so Lemma 6.1 can be again used to obtain

$$f(x, C_a^*, C_o) = f(x) \le f^*(x) = f(x, C_a^*, C_o^*) \text{ for } 0 \le x \le C_o.$$

For $C_o \le x$, we have

$$f(x, C_a^*, C_o) = \gamma x = f^*(x) = f(x, C_a^*, C_o^*).$$

If $C_o \ge C_a^* = \frac{K}{\gamma}$, we just observe that $f(x, C_a^*, C_o) = f(x, C_a^*, \frac{K}{\gamma})$, so we can reduce this case to the case $C_o = C_a^*$ already considered. This completes the proof of (7.14) in Case (i).

Case (ii). $\gamma b_0 = K_1 \le K \le K_2 = \gamma b_\infty$.

This is the simplest case, all proofs being based on comparison arguments for C^2 solutions of the equation $\mathcal{L}_f f(x) = c$. The details are left to the reader.

Case (iii). $0 < K < K_1 = \gamma b_o$.

If $C_a^* < C_o < \infty$, there is no change:

$$f(x, C_a^*, C_o) = f(x, C_a^*, C_o^*) \text{ for } 0 \le x < \infty.$$

If $0 < C_o \le C_a^*$, then $f(C_o, C_a^*, C_o) = \gamma C_o \le f_*(C_o)$. The Comparison Lemma 6.1 implies

$$f(x, C_a^*, C_o) \le f(x, C_a^*, C_o^*) \text{ for } 0 \le x \le C_o.$$

For $C_o < x < C_a^*$, we have $f(x, C_a^*, C_o) = \gamma x \le f(x, C_a^*, C_o^*)$, and for $C_a^* \le x$ we know that $f(x, C_a^*, C_o) = f(x, C_a^*, C_o^*) = \max\{K, \gamma x\}$. This completes the proof of (7.14) in Case (iii).

We consider (7.13). If $K \le C_a \le C_a^*$, then $f(C_a, C_a, C_o^*) = K \ge f(C_a, C_a^*, C_o^*)$. The Comparison Lemma 6.1 implies $f(x, C_a, C_o^*) \ge f(x, C_a^*, C_o^*)$ for $0 \le x \le C_a$. For $x \ge C_a$, we have $f(x, C_a, C_o^*) = \max\{K, \gamma x\} \ge f(x, C_a^*, C_o^*)$.

The case $C_a \ge C_o^*$ can be reduced to the case $C_a = C_o^*$ since $f(x, C_a, C_o^*) = f(x, C_a^*, C_o^*)$ for all $x \ge 0$ if $C_a \ge C_o^*$. We do that case now.

Assume $C_a^* < C_a \le C_o^*$. First we claim that $f_*(\cdot) = f(\cdot, C_a^*, C_o^*)$ is a C^1 viscosity subsolution of

$$\mathcal{L}_{f_*} f_*(x) \le c \text{ on } \left(0, \frac{K}{\gamma}\right), \tag{7.16}$$

and then we use the Comparison Lemma 6.1 (the difference $f_*(\cdot) - f(\cdot, C_a, C_o^*)$ cannot have a positive maximum in $(0, C_a)$) to conclude that

$$f_*(x) \le f(x, C_a, C_o^*) \text{ for } 0 \le x \le C_a.$$

In the comparison argument we also use the fact that $f(\cdot) = f(\cdot, C_a, C_o^*)$ satisfies $\mathcal{L}_f f(x) = c$ for $0 < x < C_a$, and

$$f_*(0) = 0 = f(0, C_a, C_o^*), \quad f_*(C_a) = K = f(C_a, C_a, C_o^*).$$

For $C_a \le x \le \frac{K}{\gamma}$, we have $f_*(x) = K = f(x, C_a, C_o^*)$, and for $x > \frac{K}{\gamma}$ we know that $f_*(x) = \gamma x = f(x, C_a, C_o^*)$.

This means that the proof of (7.14) is complete, provided we can show that f_* is a viscosity subsolution of (7.16). We know that $\mathcal{L}_{f_*} f_*(x) = c$ for $0 < x < C_a^*$, f_* being a C^2 function on $(0, C_a^*)$. From (7.8) and the Case (iii) assumption, we see that $rK \le c$. Furthermore, $f_*(x) = K$ for $C_o^* \le x \le \frac{K}{\gamma}$. We conclude that $\mathcal{L}_{f_*} f_*(x) \le c$ on $(C_a^*, \frac{K}{\gamma})$.

It remains to show that if $\psi \in C^2(0, C_o^*)$ dominates f^* on $(0, C_o^*)$ and agrees with f^* at C_a^*, then

$$r\psi(C_a^*) - (rC_a^* - c)\psi'(C_a^*) + \delta(C_a^* - \psi(C_a^*))\psi'(C_a^*) - \frac{1}{2}\sigma^2 (C_a^*)^2 \psi''(C_a^*) \le c. \tag{7.17}$$

Since $f_* \in C^1(0, \frac{K}{\gamma})$ and $f_*'(C_a^*) = 0$, we have $\psi'(C_a^*) = 0$. Since $0 < C_a^* < a$, we know $\mathcal{L}_{f_a} f_a(C_a^*) = c$, and since $f_a(C_a^*) = K$ and $f_a'(C_a^*) = 0$, we obtain

$$rK - \frac{1}{2}\sigma^2 (C_a^*)^2 f_a''(C_a^*) = c. \tag{7.18}$$

However, since $\psi(C_a^*) = f_a(C_a^*)$, $\psi'(C_a^*) = f_a'(C_a^*) = 0$, and ψ dominates f_a on $[0, C_a^*]$ (because $f_*(x) = f_a(x)$ on $[0, C_a^*]$), we conclude that

$$f_a''(C_a^*) \le \psi''(C_a^*). \tag{7.19}$$

Substituting this into (7.18), we obtain (7.17). □

Remark 7.5. The proof of the last claim is based on the elementary observation that for a C^2 function, a one-sided maximum is enough to conclude that the second derivative is not positive, provided that the first derivative vanishes. Furthermore, we have proved that f_* is a viscosity solution of the variational inequality $\max\{\mathcal{N} f_*(x) - c, f_*(x) - K\} = 0$ on $(0, \frac{K}{\gamma})$.

Proof of Theorem 2.2. For $y_1 = x_1$, it is easily verified that $f(x) = x$ is a solution of (2.16), and the Comparison Lemma 6.1 establishes uniqueness.

For $0 < y_1 < x_1$, uniqueness again follows from Lemma 6.1 once we have a solution satisfying $f' \le 1$ on $(0, x_1)$. The proof of existence is based on a fixed point argument similar to the proof of Proposition 6.2 with $a < \infty$. In fact, the argument here is simpler, since we deal only with C^2 solutions of the differential equation $\mathcal{L}_g f(x) = c$ rather then viscosity solutions of the variational inequality $\min\{\mathcal{L}_g f(x) - c, f(x) - \gamma x\} = 0$.

For $0 < y_1 < x_1$, we set $A = x_1 - y_1$ and define $\mathcal{G}$ to be the set of all functions $g \in C[0, x_1] \cap C^2(0, x_1)$ such that $g(0) = 0$, $g(x_1) = y_1$, and

$$g(x) \ge \max\{x - A, 0\}, \quad -M(x_1, y_1) \le g'(x) < 1 \quad \forall x \in (0, x_1),$$

where $M(x_1, y_1)$ is a constant to be determined later but depending on only x_1 and y_1. We further define $\overline{\mathcal{G}}$ to be the closure of $\mathcal{G}$ in $C[0, x_1]$ with respect to the supremum norm $\|\cdot\|$. For $g \in \overline{\mathcal{G}}$, we set

$$Tg(x) = \mathbb{E}\left[\int_0^{\tau_0^x \wedge \tau_{x_1}^x} ce^{-ru} du + \mathbb{I}_{\{\tau_{x_1}^x < \tau_0^x\}} e^{-r(\tau_0^x \wedge \tau_{x_1}^x)} y_1\right], \tag{7.20}$$

where $X^x(t)$ is given by (3.4) with $X^x(0) = x$. It is clear from its definition that $Tg \ge 0$ for every $g \in \overline{\mathcal{G}}$. We use the argument in the proof of Proposition 4.6 to conclude that for $g \in \overline{\mathcal{G}}$ the function Tg is of class C^2 on $(0, x_1)$ and $\mathcal{L}_g Tg(x) = c$ for $0 < x < x_1$. The continuity of Tg at 0 and x_1 follows from Lemma 4.2. The functions $\max\{x - A, 0\}$ and x are respective sub- and supersolutions of $\mathcal{L}_g f = c$ which lie respectively below and above Tg at the endpoints 0 and x_1. Lemma 6.1 implies that for all $g, h \in \overline{\mathcal{G}}$,

$$\max\{x - A, 0\} \leq Tg(x) \leq x \text{ for } 0 \leq x \leq x_1, \tag{7.21}$$

$$\|Tg - Th\| \leq \sup_{0<x<x_1} |DTg(x)| \|g - h\|. \tag{7.22}$$

We now prove that $T(\mathcal{G}) \subset \mathcal{G}$, the analogue of Proposition 5.6. For $g \in \mathcal{G}$, the first part of the proof of Theorem 5.6 shows that DTg cannot attain a positive local maximum nor a negative local minimum in $(0, x_1)$. This implies that either DTg is nonnegative on $(0, x_1)$ or else $D^-Tg(x_1) \leq DTg(x)$ for $0 < x < x_1$. To show that $DTg(x) \geq -M(x_1, y_1)$, it suffices to find a lower bound on $D^-Tg(x_1)$ which may depend on x_1 and y_1 but not on g. For this purpose, we let h be the solution on $[y_1, x_1]$ of the equation $\mathcal{L}_g h = c$ with boundary conditions $h(y_1) = h(x_1) = y_1$. Lemma 6.1 shows that h is nonnegative and dominates Tg on $[y_1, x_1]$, and hence $D^-h(x_1) \leq D^-Tg(x_1)$. If h attains a maximum at some point $x_* \in (y_1, x_1)$, the equation $\mathcal{L}_g h(x_*) = c$ implies $h(x_*) \leq \frac{c}{r}$. If h does not attain a maximum in (y_1, x_1), then h is dominated by its value y_1 at the endpoints of this interval. In either case, we obtain a bound on $|h|$ which is independent of g. Furthermore, there must be some point $x_0 \in (y_1, x_1)$ where h' vanishes. We solve the equation $\mathcal{L}_g h = c$ for h'' and integrate from x_0 to obtain (5.7). We then use Gronwall's inequality to obtain a bound on $|h'|$ independent of g.

We need also to obtain the upper bound $DTg < 1$. We observe first that since $Tg(x) \geq \max\{x - A, 0\}$ and these two functions agree at $x = x_1$, we must have $D^-Tg(x_1) \leq 1$. We use the same arguments used to prove $DT_a g < 1$ if $g \in \mathcal{G}_a$ to conclude that $DTg < 1$ on $(0, x_1)$. This completes the proof that $T(\mathcal{G}) \subset \mathcal{G}$. A relation similar to (6.1) shows that the operator T is continuous on $\overline{\mathcal{G}}$, and hence $T(\overline{\mathcal{G}}) \subset \overline{\mathcal{G}}$. Schauder's fixed point theorem implies the existence of a function $f \in \overline{\mathcal{G}}$ satisfying $Tf = f$. This means, in particular, that $f \in C[0, x_1] \cap C^2(0, x_1)$ and $\mathcal{L}_f f = c$, so f is a solution of (2.16). Since f is differentiable and $f \in \overline{\mathcal{G}}$, we know that $f' \leq 1$ on $(0, x_1)$. In fact, $f' < 1$. The proof is identical to the proof of $f_a' < 1$. ◻

Acknowledgment. The third author thanks John Noddings for introducing him to convertible bonds.

REFERENCES

[1] P. Asquith, *Convertible bonds are not called late*, J. Finance, 50 (1995), pp. 1275–1289.

[2] P. Asquith and D. Mullins, Jr., *Convertible debt: Corporate call policy and voluntary conversion*, J. Finance, 46 (1991), pp. 1273–1289.

[3] A. Bensoussan, M. Crouhy, and D. Galai, *Stochastic equity volatility and the capital structure of the firm*, Philos. Trans. Roy. Soc. London Ser. A, 347 (1994), pp. 531–541.

[4] F. Black and M. Scholes, *The pricing of options and corporate liabilities*, J. Political Economy, 81 (1973), pp. 637–659.

[5] M. Brennan and E. Schwartz, *Convertible bonds: Valuation and optimal strategies for call and conversion*, J. Finance, 32 (1977), pp. 1699–1715.

[6] M. Brennan and E. Schwartz, *Analyzing convertible bonds*, J. Financial Quantitative Analysis, 15 (1980), pp. 907–929.

[7] M. Crandall, H. Ishii, and P. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. (N.S.), 27 (1992), pp. 1–67.

[8] K. Dunn and K. Eades, *Voluntary conversion of convertible securities and the optimal call strategy*, J. Financial Econom., 23 (1984), pp. 273–301.
[9] W. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*, Springer-Verlag, New York, 1993.
[10] M. Harris and A. Raviv, *A sequential model of convertible debt call policy*, J. Finance, 40 (1985), pp. 1263–1282.
[11] J. E. Ingersoll, *A contingent-claims valuation of convertible securities*, J. Financial Econom., 4 (1977), pp. 289–322.
[12] J. E. Ingersoll, *An examination of corporate call policies on convertible securities*, J. Finance, 32 (1977), pp. 463–478.
[13] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, 1991.
[14] I. Karatzas and S. Shreve, *Methods of Mathematical Finance*, Springer-Verlag, New York, 1998.
[15] H. Kunita, *Stochastic Flows and Stochastic Differential Equations*, Cambridge University Press, Cambridge, UK, 1990.
[16] R. C. Merton, *Theory of rational option pricing*, Bell J. Econom. Manag. Sci., 4 (1973), pp. 141–183.
[17] M. Miller and F. Modigliani, *The cost of capital, corporation finance, and the theory of investment*, Amer. Econ. Rev., 48 (1958), pp. 261–297.
[18] M. Miller and F. Modigliani, *Dividend policy, growth and the valuation of shares*, J. Business, 34 (1961), pp. 411–433.
[19] B. Øksendal and K. Reikvam, *Viscosity solutions of optimal stopping problems*, Stochastics Stochastics Rep., 62 (1998), pp. 285–301.
[20] A. N. Shiryayev, *Optimal Stopping Rules*, Springer-Verlag, Berlin, 1977.