GUARANTEED MINIMUM WITHDRAWAL BENEFIT IN VARIABLE ANNUITIES

Min Dai

Department of Mathematics, National University of Singapore

Yue Kuen Kwok

Department of Mathematics, Hong Kong University of Science and Technology

Jianping Zong

Department of Mathematics, National University of Singapore

Abstract

We develop a singular stochastic control model for pricing variable annuities with the guaranteed minimum withdrawal benefit. This benefit promises to return the entire initial investment, with withdrawals spread over the term of the contract, irrespective of the market performance of the underlying asset portfolio. A contractual withdrawal rate is set and no penalty is imposed when the policyholder chooses to withdraw at or below this rate. Subject to a penalty fee, the policyholder is allowed to withdraw at a rate higher than the contractual withdrawal rate or surrender the policy instantaneously. We explore the optimal withdrawal strategy adopted by the rational policyholder that maximizes the expected discounted value of the cash flows generated from holding this variable annuity policy. An efficient finite difference algorithm using the penalty approximation approach is proposed for solving the singular stochastic control model. Optimal withdrawal policies of the holders of the variable annuities with the guaranteed minimum withdrawal benefit are explored. We also construct discrete pricing formulation that models withdrawals on discrete dates. Our numerical tests show that the solution values from the discrete model converge to those of the continuous model.

Keywords: guaranteed minimum withdrawal benefit, variable annuities, singular stochastic control model, penalty approximation, optimal withdrawal policies

1   Introduction

A variable annuities policy is a financial contract between a policyholder and an insurance company which promises a stream of annuities cash flows. At the initiation of the contract, the policyholder pays a single lump sum premium to the issuer. The trusted fund is then invested in a well diversified reference portfolio of a specific class of assets. Under the policy, the insurer promises to make variable periodic payments to the policyholder on preset future dates. The variable payments would depend on the performance of the reference portfolio, thus the policyholders are provided with the equity participation. Variable annuities are attractive to investors not only because of the tax-deferred feature. In addition, they also offer different types of benefits, such as guaranteed minimum death benefit, guaranteed minimum accumulation benefit, guaranteed minimum income benefit.

In recent years, variable annuities with the guaranteed minimum withdrawal benefits (GMWBs) have attracted significant attention and sales. These benefits allow the policyholders to withdraw funds on an annual or semi-annual basis. There is a contractual withdrawal rate such that the policyholder is allowed to withdraw at or below this rate without a penalty. The GMWB promises to return the entire initial

investment, thus the guarantee can be viewed as an insurance option. More precisely, even when the personal account (investment net of withdrawal and proportional insurance fees) of the policyholder falls to zero prior to the policy maturity date, the insurer continues to provide the guaranteed withdrawal amount until the entire original premium is paid out. If the account stays positive at maturity, the whole remaining balance in the account is paid to the policyholder at maturity. Therefore, the total sum of cash flows received by the policyholder is guaranteed to be the same or above the original premium deposit (not accounting for the time value of the cash flows). Under the dynamic setting of the policy, the policyholder is allowed to withdraw at a rate higher or lower than the contractual rate or in a finite amount or even surrender instantaneously, according to his best economic advantage. The annuity contract may include the following clause that serves to discourage excessive withdrawal. When the policyholder withdraws at a higher rate than the contractual withdrawal rate, the guarantee level is reset to the minimum of the prevailing guarantee level and the account value. For example, suppose the policyholder decides to withdraw $10, which is higher than the contractual withdrawal amount $7. Suppose the current guarantee level is $80 while the personal account is $60, then the guarantee level drops to $\min(\$80, \$60) - \$10 = \$50$ after the withdrawal of $10. In addition, there is a percentage penalty charge applied on the excessive portion of the withdrawal amount.

There has been much research devoted to the pricing and hedging of variable annuities and insurance policies with various forms of embedded options. For hedging strategies, Coleman *et al.* (2006) suggest risk minimization hedging for variable annuities under both equity and interest rate risks. Milevsky and Posner (2001) use risk neutral option pricing theory to value the guaranteed minimum death benefit in variable annuities. Chu and Kwok (2004) and Siu (2005) analyze the withdrawal and surrender options in various equity-linked insurance products. Milevsky and Salisbury (2006) develop the pricing model of variable annuities with GMWB under both static and dynamic withdrawal policies. Under the static withdrawal policies, the policyholders are assumed to behave passively with withdrawal rate kept fixed at the contractual rate and to hold the annuity to maturity. In their dynamic model, policyholders are assumed to follow an optimal withdrawal policy seeking to maximize the annuity value by lapsing the product at an optimal time. Since the withdrawal is allowed to be at a finite rate or in discrete amount (infinite withdrawal rate), the pricing model leads to a singular stochastic control problem with the withdrawal rate as the control variable.

In this paper, we would like to study the nature of GMWB in variable annuities beyond the results reported by Milevsky and Salisbury (2006). We provide a rigorous derivation of the singular stochastic control model for pricing variable annuities with GMWB using the Hamilton-Jacobi-Bellman equation. Both cases of continuous and discrete withdrawal of funds are considered. The valuation of the variable annuities product is performed under the risk neutral framework, assuming the underlying equity portfolio is tradeable or the holder is a risk neutral investor. Our pricing models do not include mortality factor since mortality risk is not quite crucial in guaranteed minimum withdrawal benefit riders. Also, we have assumed deterministic interest rate structure since interest rate plays its influence mainly on discount factors in pricing the guaranteed minimum withdrawal benefit. This is different from equity fluctuation, where it has much more profound impact on the optimal withdrawal policy. We assume the policyholder to be fully rational in the sense that he chooses the optimal dynamic withdrawal strategy so as to maximize the expected discounted value of the cash flows generated from holding the annuity policy. In our pricing formulation, we manage to obtain a set of parabolic variational inequalities that govern the fair value of the variable annuity policy with the GMWB. The constraint inequalities are seen to involve the gradient of the value function. By extending the penalty method in the solution of optimal stopping problems as proposed by Forsyth and Vetzal (2002) and Dai *et al.* (2007), we propose an efficient finite difference scheme following the penalty approximation approach to solve for the fair value of the annuities. The

numerical procedure of using the penalty approximation approach represents a nice contribution to the family of numerical methods for solving singular stochastic control problems (Kumar and Muthuraman, 2004; Forsyth and Labahn, 2006). In addition, we design the finite difference scheme that allows for discrete jumps across discrete withdrawal dates for solving the discrete time withdrawal model.

The paper is organized as follows. In the next section, we derive the singular stochastic control model that incorporates the GMWB into the variable annuities pricing model. We start with the formulation that assumes continuous withdrawal, then generalize the model to allow for a discrete withdrawal on specified dates. We outline the numerical approach using the finite difference scheme with penalty approximation for solving the set of variational inequalities of the pricing formulation. Numerical tests were performed that serve to illustrate the robustness of the proposed numerical schemes for both the continuous and discrete models. In Section 3, we analyze the optimal withdrawal behaviors of the policyholders. We also examine the impact of various parameters in the singular stochastic control pricing model on the fair insurance fee to be charged by the insurer for provision of the guarantee. A summary and concluding remarks are presented in the last section.

2 Model formulation

Mathematically, it is more convenient to construct the pricing model of the annuity policy that assumes continuous withdrawal. In actual practice, withdrawal of discrete amount occurs at discrete time instants during the life of the policy. In this section, we start with the construction of the continuous model by assuming continuous withdrawal. The more realistic scenario of discrete withdrawal will be considered afterwards. In our singular stochastic control model for pricing the GMWB, the discretionary withdrawal rate is the control variable. Some of the techniques used in the derivation of our pricing model are similar to those used in the singular stochastic control model proposed by Davis and Norman (1990) in the analysis of portfolio selection with transaction costs.

2.1 Continuous withdrawal model

Let S_t denote the value of the reference portfolio of assets underlying the variable annuity policy, before the deduction of any proportional fees. Taking the usual assumption on the price dynamics of equity in option pricing theory, the evolution of S_t under the risk neutral measure is assumed to follow

$$dS_t = rS_t \, dt + \sigma S_t \, dB_t, \tag{2.1}$$

where B_t represents the standard Brownian motion, σ is the volatility and r is the riskfree interest rate. Let $\mathcal{F}_t$ be the natural filtration generated by the Brownian process B_t and α be the proportional annual insurance fee paid by the policyholder. The most important feature of the GMWB is the guarantee on the return of premium via withdrawal, where the accumulated sum of all withdrawals throughout the policy's life is the premium w_0 paid up front (not accounting for the time value of the cash flows).

We let A_t denote the account balance of the guarantee, where A_t is right-continuous with left limit, non-negative and non-increasing $\{\mathcal{F}_t\}_{t \geq 0}$-adaptive process. At initiation, A_0 equals w_0; and the withdrawal guarantee becomes insignificant when A_t hits 0. As withdrawal continues, A_t decreases over the life of the policy until it hits the zero value. By the maturity date T, A_t must become zero. To derive the continuous time pricing model, we first consider a restricted class of withdrawal policies in which A_t

is constrained to be absolutely continuous with bounded derivatives, that is

$$A_t = A_0 - \int_0^t \gamma_s \, ds, \quad 0 \leq \gamma_s \leq \lambda. \tag{2.2}$$

Penalty charges are incurred when the withdrawal rate γ exceeds the contractual withdrawal rate G. Supposing a proportional penalty charge k is applied on the portion of γ above G, then the net amount received by the policyholder is $G + (1 - k)(\gamma - G)$ when $\gamma > G$. Let g denote the percentage withdrawal rate, say, $g = 7\%$ means 7% of premium is withdrawn per annum. We then have $G = gw_0$.

Let W_t denote the value of the personal variable annuity account, then its dynamics follows

$$dW_t = (r - \alpha)W_t \, dt + \sigma W_t \, dB_t + dA_t, \text{ for } W_t > 0. \tag{2.3}$$

Once W_t hits the value 0, it stays at this value thereafter. Let w_0 be the initial account value of the policy, which is the same as the premium paid up front. When the personal account value stays positive at maturity T, the remaining balance is paid back to the policyholder at T.

Let $f(\gamma)$ denote the rate of cash flow received by the policyholder as resulted from the continuous withdrawal process, we then have

$$f(\gamma) = \begin{cases} \gamma & \text{if } 0 \leq \gamma \leq G \\ G + (1 - k)(\gamma - G) & \text{if } \gamma > G \end{cases}. \tag{2.4}$$

The policyholder receives the continuous withdrawal cash flow $f(\gamma_u)$ over the life of the policy and the remaining balance of the personal account at maturity. Based on the assumption of rational behavior of the policyholder that he chooses the optimal withdrawal policy dynamically so as to maximize the present value of cash flows generated from holding the variable annuity policy and under the restricted class of withdrawal policies as specified by Eq. (2.2), the no-arbitrage value $\overline{V}$ of the variable annuity with GMWB is given by

$$\overline{V}(W, A, t) = \max_\gamma \mathrm{E}_t \left[e^{-r(T-t)} \max(W_T, 0) + \int_t^T e^{-r(u-t)} f(\gamma_u) \, du \right], \tag{2.5}$$

where T is the maturity date of the policy and expectation E_t is taken under the risk neutral measure conditional on $W_t = W$ and $A_t = A$. Here, γ is the control variable that is chosen to maximize the expected value of discounted cash flows. Using the standard procedure of deriving the Hamilton-Jacobi-Bellman (HJB) equation in stochastic control problems (Yong and Zhou, 1999), the governing equation for $\overline{V}$ is found to be

$$\frac{\partial \overline{V}}{\partial t} + \mathcal{L}\overline{V} + \max_\gamma h(\gamma) = 0 \tag{2.6}$$

where

$$\mathcal{L}\overline{V} = \frac{\sigma^2}{2} W^2 \frac{\partial^2 \overline{V}}{\partial W^2} + (r - \alpha)W \frac{\partial \overline{V}}{\partial W} - r\overline{V}$$

and

$$\begin{aligned} h(\gamma) &= f(\gamma) - \gamma \frac{\partial \overline{V}}{\partial W} - \gamma \frac{\partial \overline{V}}{\partial A} \\ &= \begin{cases} \left(1 - \dfrac{\partial \overline{V}}{\partial W} - \dfrac{\partial \overline{V}}{\partial A} \right) \gamma & \text{if } 0 \leq \gamma < G \\ kG + \left(1 - k - \dfrac{\partial \overline{V}}{\partial W} - \dfrac{\partial \overline{V}}{\partial A} \right) \gamma & \text{if } \gamma \geq G \end{cases}. \end{aligned}$$

4

The function $h(\gamma)$ is piecewise linear, so its maximum value is achieved at either $\gamma = 0, \gamma = G$ or $\gamma = \lambda$. Recall that we place a sufficiently large upper bound λ for γ, namely, $0 \leq \gamma \leq \lambda$. It is easily seen that

$$\max_{\gamma} h(\gamma) = \begin{cases} kG + \lambda\left(1 - k - \dfrac{\partial \overline{V}}{\partial W} - \dfrac{\partial \overline{V}}{\partial A}\right) & \text{if } 1 - \dfrac{\partial \overline{V}}{\partial W} - \dfrac{\partial \overline{V}}{\partial A} \geq k \\[2mm] \left(1 - \dfrac{\partial \overline{V}}{\partial W} - \dfrac{\partial \overline{V}}{\partial A}\right) G & \text{if } 0 < 1 - \dfrac{\partial \overline{V}}{\partial W} - \dfrac{\partial \overline{V}}{\partial A} < k \\[2mm] 0 & \text{if } 1 - \dfrac{\partial \overline{V}}{\partial W} - \dfrac{\partial \overline{V}}{\partial A} \leq 0 \end{cases} \quad . \tag{2.7}$$

Substituting into Eq. (2.6), we obtain the following equation for $\overline{V}$:

$$\frac{\partial \overline{V}}{\partial t} + \mathcal{L}V + \min\left[\max\left(1 - \frac{\partial \overline{V}}{\partial W} - \frac{\partial \overline{V}}{\partial A}, 0\right), k\right] G$$

$$+ \quad \lambda \max\left(1 - k - \frac{\partial \overline{V}}{\partial W} - \frac{\partial \overline{V}}{\partial A}, 0\right) = 0. \tag{2.8}$$

For the general case where A_t is allowed to be discontinuous (instantaneous withdrawal of finite amount), the no-arbitrage value V of the variable annuity with GMWB is given by

$$V(W, A, t) = \max_{A} \mathrm{E}_t\left[e^{-r(T-t)} \max(W_T, 0) + \int_t^T e^{-r(u-t)} F(-dA_u)\right], \tag{2.9}$$

where

$$F(-dA_u) = \begin{cases} -dA_u & \text{if } -dA_u \leq G \, dt \\ kG \, dt - (1-k)dA_u & \text{if } -dA_u > G \, dt \end{cases} \quad .$$

To obtain $V(W, A, t)$ from $\overline{V}(W, A, t)$, we allow the upper bound λ on γ to be infinite. It is well known that Eq. (2.8) is a penalty approximation to Eq. (2.9) (Friedman, 1982). Taking the limit $\lambda \to \infty$ in Eq. (2.8), we obtain the following linear complementarity formulation of the value function $V(W, A, t)$:

$$\min\left[-\frac{\partial V}{\partial t} - \mathcal{L}V - \max\left(1 - \frac{\partial V}{\partial W} - \frac{\partial V}{\partial A}, 0\right) G, \frac{\partial V}{\partial W} + \frac{\partial V}{\partial A} - (1-k)\right] = 0,$$

$$W > 0, \quad 0 < A < w_0, \quad t > 0. \tag{2.10}$$

One can follow a similar argument presented in Zhu (1992) to show that the value function $V(W, A, t)$ defined in Eq. (2.9) is indeed the generalized solution to the HJB equation (2.10) subject to the auxiliary conditions presented below. To complete the formulation of the pricing model, it is necessary to prescribe the terminal condition at time T and boundary conditions along the boundaries: $W = 0, W \to \infty$ and $A = 0$. Note that it is not necessary to prescribe the boundary condition at $A = w_0$ due to the hyperbolic nature of the variable A in the governing equation (2.10).

- At maturity, the policyholder takes the maximum between the remaining guarantee withdrawal net of penalty charge and the remaining balance of the personal account.

- When either $A = 0$ or $W \to \infty$, the withdrawal guarantee becomes insignificant. The value of the annuity becomes $We^{-\alpha(T-t)}$. The discount factor $e^{-\alpha(T-t)}$ arises due to discounting at the rate α as a proportional fee at the rate α is paid during the remaining life of the annuity.

- When $W = 0$, the equity participation of the policy vanishes. The pricing formulation reduces to a simpler optimal control model with no dependence on W. Let $V_0(A, t)$ be the value function of the annuity when $W = 0$, which is the solution to the following linear complementarity formulation [considered as a reduced version of Eq. (2.10) with no dependence on W]:

$$\min \left[-\frac{\partial V_0}{\partial t} + rV_0 - \max \left(1 - \frac{\partial V_0}{\partial A}, 0 \right) G, \frac{\partial V_0}{\partial A} - (1 - k) \right] = 0,$$
$$0 < A < A_0, 0 < t < T,$$
$$V_0(A, T) = (1 - k)A \quad \text{and} \quad V_0(0, t) = 0. \tag{2.11}$$

In summary, the auxiliary conditions of the linear complementarity formulation (2.10) are given by

$$V(W, A, T) = \max(W, (1 - k)A)$$
$$V(W, 0, t) = e^{-\alpha(T-t)}W, \quad V(0, A, t) = V_0(A, t),$$
$$V(W, A, t) \to e^{-\alpha(T-t)}W \text{ as } W \to \infty. \tag{2.12}$$

Interestingly, a closed form solution to $V_0(A, t)$ can be found. Defining

$$\tau^* = \min \left(-\frac{\ln(1 - k)}{r}, T - t \right),$$

it can be shown that

$$V_0(A, t) = (1 - k) \max(A - G\tau^*, 0) + \frac{G}{r} \left[1 - e^{-r \min(A/G, \tau^*)} \right]. \tag{2.13}$$

The analytic derivation of $V_0(A, t)$ and its financial interpretation are presented in the Appendix.

As a remark, Milevsky and Salisbury (2006) have derived a similar dynamic control model that allows for dynamic withdrawal rate adopted by the policyholder. However, their formulation is not quite complete since it does not contain time dependency in the value function. Also, there is no full prescription of the auxiliary conditions associated with their pricing formulation.

Construction of finite difference scheme

The numerical solution of the singular stochastic control formulation in Eqs. (2.10) and (2.12) poses a difficult computational problem. Instead of solving the singular stochastic control model directly, we solve for the penalty approximation model (2.8) in which the allowable control is bounded. In our numerical procedure to solve for $\overline{V}(W, A, t)$, we apply the standard finite difference approach to discretize the penalty approximation formulation (2.8). Since the governing equation (2.8) is a degenerate diffusion equation with only the first order derivative of A appearing, upwind discretization must be used to deal with the first order derivative terms in the differential equation. This technique serves to avoid excessive numerical oscillations in the calculations when the penalty parameter λ assumes a large value.

We employ the two-level implicit finite difference scheme and use time to expiry $\tau = T - t$ as the time variable. Let $\overline{V}_{j,k}^n$ denote the numerical approximation to $\overline{V}(j\triangle W, k\triangle A, n\triangle \tau)$, where $\triangle W$ and $\triangle A$ are the spatial step sizes and $\triangle \tau$ is the time step. Let $\mathcal{L}_h \overline{V}_{j,k}^n$ denote the spatial discretization of the differential term $\mathcal{L}V$, where

$$\mathcal{L}_h \overline{V}_{j,k}^n = \frac{\sigma^2}{2} W_j^2 \frac{\overline{V}_{j+1,k}^n - \overline{V}_{j,k}^n + \overline{V}_{j-1,k}^n}{\triangle W^2} + (r - \alpha) W_j \frac{\overline{V}_{j+1,k}^n - \overline{V}_{j-1,k}^n}{2\triangle W} - r\overline{V}_{j,k}^n.$$

6

The general family of two-level implicit schemes for solving Eq. (2.8) is given by

$$
\begin{aligned}
\frac{\overline{V}_{j,k}^{n+1} - \overline{V}_{j,k}^{n}}{\triangle \tau} \;=\; & \theta \mathcal{L}_h \overline{V}_{j,k}^{n+1} + (1-\theta)\mathcal{L}_h \overline{V}_{j,k}^{n} \\
& + \min\Big(\max\Big(1 - \frac{\overline{V}_{j,k}^{n+1} - \overline{V}_{j-1,k}^{n+1}}{\triangle W} - \frac{\overline{V}_{j,k}^{n+1} - \overline{V}_{j,k-1}^{n+1}}{\triangle A}, 0\Big), k\Big)G \\
& + \lambda \; \max\Big(1 - k - \frac{\overline{V}_{j,k}^{n+1} - \overline{V}_{j-1,k}^{n+1}}{\triangle W} - \frac{\overline{V}_{j,k}^{n+1} - \overline{V}_{j,k-1}^{n+1}}{\triangle A}, 0\Big), \qquad (2.14)
\end{aligned}
$$

where θ is a weighting factor, $0 < \theta \le 1$. When $\theta = 1$, we have the fully implicit scheme; while $\theta = \frac{1}{2}$ corresponds to the Crank-Nicholson scheme. Due to the presence of the mildly non-linear penalty term in the differential equation, a non-linear algebraic system of equations has to be solved at each time step. Newton type iterations are applied to solve the non-linear algebraic equations.

When we apply the above numerical scheme (2.14) to obtain the numerical approximation solution to the singular stochastic control model (2.10), there are two sources of errors. One source is the analytic approximation error that arises from the penalty approximation of the singular stochastic control model. This error can be controlled by choosing the penalty parameter to be sufficiently large. The other source of error comes from the numerical discretization of the penalty approximation model (2.8). Since the solution to Eq. (2.8) is expected to have a linear growth at infinity, the strong comparison principle holds in the sense of viscosity solution [Crandal et $al.$ (1992); Barles et $al.$ (1995)]. As a consequence, by virtue of the result established by Barles and Souganidis (1991), one can establish the convergence of the fully implicit scheme (corresponding to $\theta = 1$) to the viscosity solution of Eq. (2.8) when the penalty parameter λ is taken to be sufficiently large and the step sizes in the numerical schemes become vanishingly small. Due to the lack of monotonicity property, the analytic proof of convergence of the Crank-Nicholson scheme cannot be established in a similar manner. We resort to numerical experiments to test for convergence of the Crank-Nicholson scheme.

In Table 1, we list the numerical results obtained from the Crank-Nicholson scheme using varying number of time steps and spatial steps. The values of the model parameters used in the calculations are: $G = 7, \sigma = 0.2, \alpha = 0.036, k = 0.1, r = 0.05, T = 14.28, w_0 = 100$ and $\lambda = 10^6$. Let N_t, N_W and N_A denote the number of time steps and number of spatial steps in W and A, respectively. The apparent convergence of the numerical solution is revealed in Table 1. We expect to have a quadratic rate of convergence of the numerical solution using the Crank-Nicholson scheme such that the numerical error is reduced by a factor of $1/4$ when the number of time steps and number of spatial steps are doubled. Our numerical results show that the actual rate of convergence is slightly slower than the expected rate. This may be attributed to the upwind treatment of the first order derivative terms in the numerical scheme. We also examine the convergence of the numerical solution to the penalty approximation model (2.8) with varying values of λ to the annuity value of the continuous model. The numerical results shown in Table 2 were obtained using the Crank-Nicholson scheme with $N_t = 512, N_W = 1024, N_A = 1024$. We choose two different values of k and all the other model parameters are taken to be the same as those used to generate the numerical results in Table 1. The apparent convergence of the numerical solution to the penalty approximation model is revealed when the penalty parameter increases to a sufficiently high value.

2.2 Discrete withdrawal model

Consider the real life situation where discrete withdrawal amount is only allowed at time $t_i, i = 1, 2, \cdots, N$. Here, t_0 denotes the time of initiation and the last withdrawal date t_N is the maturity date T. Let the discrete withdrawal amount at time t_i be denoted by γ_i. Since the account balance of the withdrawal guarantee A_t remains unchanged within the interval $(t_{i-1}, t_i), i = 1, 2, \cdots, N$, the annuity value function $V(W, A, t)$ satisfies the following differential equation which has no dependence on A:

$$\frac{\partial V}{\partial t} + \mathcal{L}V = 0, \quad t \in (t_{i-1}, t_i), \quad i = 1, 2, \cdots, N. \tag{2.15}$$

The updating of A_t only occurs at the withdrawal dates. Upon withdrawing an amount γ_i at t_i, the annuity account drops from W_t to $\max(W_t - \gamma_i, 0)$, while the guarantee balance drops from A_t to $A_t - \gamma_i$. The jump condition of $V(W, A, t)$ across t_i is given by

$$V(W, A, t_i^-) = \max_{0 \leq \gamma_i \leq A} \{V(\max(W - \gamma_i, 0), A - \gamma_i, t_i^+) + \widehat{f}(\gamma_i)\}. \tag{2.16}$$

Here, $\widehat{f}(\gamma_i)$ represents the actual cash amount received by the policyholder subject to a penalty charge under excessive withdrawal, which can be defined in a similar manner as that for $f(\gamma)$ in Eq. (2.4). The auxiliary conditions for $V(W, A, t)$ remain the same as those stated in Eq. (2.12), except that the boundary value function $V_0(A, t)$ under discrete withdrawal is governed by

$$\begin{aligned}
&\frac{\partial V_0}{\partial t} - rV = 0, && t \neq t_i, \quad i = 1, 2, \cdots, N, \\
&V_0(A, t^-) = \max_{0 \leq \gamma_i \leq A} \{V_0(A - \gamma_i, t^+) + \widehat{f}(\gamma_i)\}, && t = t_i, \quad i = 1, 2, \cdots, N, \\
&V_0(A, T) = \widehat{f}(A) \quad \text{and} \quad V_0(0, t) = 0.
\end{aligned} \tag{2.17}$$

The above formulation resembles that of the pricing models of discretely monitored path dependent options. Here, A serves the role as the path dependent variable, which is updated whenever the calendar time sweeps across a fixing date. To solve for $V(W, A, t)$ under the discrete withdrawal model, we apply standard finite difference technique to discretize Eq. (2.15). The guarantee balance A is updated on those time steps that correspond to fixing dates. In our numerical calculations, we assume a finite set of discrete values that can be taken by γ_i at fixing date t_i. According to Eq. (2.16), we choose γ_i such that $V(\max(W - \gamma_i, 0), A - \gamma_i, t_i)$ is maximized. This is plausible since we know the values of V at all discrete points of (W, A) in the computational domain.

Reset provision on the guarantee level

The GMWB annuity may contain the reset provision on the guarantee level that serves as a disincentive to excessive withdrawals beyond G. After the guarantee balance A_t and account W_t are debited by the withdrawal amount γ_i at time t_i, the guarantee balance is reset to $\min(A_t, W_t) - \gamma_i$ if $\gamma_i > G$. While it is not straightforward to incorporate this reset provision into the continuous withdrawal model, it is relatively easy to modify the jump condition (2.16) to include the provision in the discrete withdrawal model. With the reset provision, the new jump condition becomes

$$V(W, A, t_i^-) = \max_{0 \leq \gamma_i \leq A} \left\{ V(\max(W - \gamma_i, 0), B, t_i^+) + \widehat{f}(\gamma_i) \right\}, \tag{2.18}$$

where

$$B = \begin{cases} \min(A - \gamma_i, \max(W - \gamma_i, 0)) & \text{if } \gamma_i > G \\ A - \gamma_i & \text{if } \gamma_i \leq G \end{cases}. \tag{2.19}$$

The auxiliary conditions remain the same as those of the non-reset case, except that the jump condition used in the calculation of $V_0(A, t)$ has to be modified as follows:

$$V_0(A, t^-) = \max_{0 \leq \gamma_i \leq A} \left\{ V_0((A - \gamma_i)\mathbf{1}_{\{\gamma_i \leq G\}}, t^+) + \widehat{f}(\gamma_i) \right\}, \tag{2.20}$$

where

$$\mathbf{1}_{\{\gamma_i \leq G\}} = \begin{cases} 1 & \text{if } \gamma_i \leq G \\ 0 & \text{otherwise} \end{cases}.$$

We would like to check for consistency between the continuous and discrete withdrawal models. We compute the fair value of the GMWB annuity without the reset provision on the guarantee level under varying values of withdrawal frequency per year. In Table 3, we tabulate the numerical results of annuity values obtained from numerical calculations using the finite difference schemes, where discrete withdrawals can be done monthly (frequency = 12), bimonthly (frequency = 6), etc. The model parameters used in our calculations are the same as those used in Tables 1 and 2. Consistent with obvious financial intuition, the tabulated results reveal that the annuity value increases with higher frequency of withdrawal per year. Also, the annuity value obtained from the continuous withdrawal model using the penalty approximation is seen to be very close to that obtained from the discrete withdrawal model with monthly withdrawal (comparing 93.419 with 93.346). The apparent agreement of annuity values serves to verify the consistency between the continuous and discrete models. The differences in annuity values with and without the reset provision are seen to be small (see Table 3).

3 Pricing behaviors and optimal withdrawal policies

Insurance companies charge proportional insurance fee (denoted by α in our pricing model) to compensate for the provision of the GMWB rider. There have been concerns in the insurance industry that the fee rate has been charged too low due to sales competition. Milevsky and Salisbury (2006) warn that current pricing of products sold in the market is not sustainable. They claim that the GMWB fees will eventually have to increase or product design will have to change. In Table 4, we present the numerical results that show how various model parameters, like GMWB rate g, penalty charge k and equity volatility σ of the account affect the required insurance fee. We chose $r = 0.05$ and used the continuous model in our calculations. The insurance fee α is determined so that the upfront amount invested in the annuity w_0 is set equal to the present value of the future cash flows generated from the annuity contract. We observe that α is an increasing function of the equity volatility σ and the GMWB contractual withdrawal rate g, but a decreasing function of the penalty charge k. Comparing to similar results based on the static withdrawal model as reported in Milevsky and Salisbury (2006), the issuer should charge a substantially higher insurance fee when the policyholder has the flexibility of dynamic withdrawal. For example, the GMWB annuity under the static withdrawal policy which guarantees a 7% withdrawal rate and equity volatility of 20% would demand a fair insurance fee of 73 basis points. However, the fair insurance fee increases to 165 basis points under the dynamic withdrawal policy even a relatively high penalty charge of 5% is imposed.

Also, we would like to examine the optimal dynamic withdrawal policies adopted by the policyholder. Since $h(\gamma)$ apparently achieves its maximum value at either $\gamma = 0, G$ or infinite value ∞, the policyholder

chooses either to withdraw a finite amount (infinite rate of withdrawal), at the contractual rate G or not to withdraw at all. Here, we postulate that the case $\gamma = 0$ should be ruled out. That is, it is always non-optimal not to withdraw. To understand this phenomenon using financial intuition, we note that the non-withdrawal amount is subject to a proportional insurance fee α. Under the risk neutral valuation framework, the drift rate of W_t is $r - \alpha$, which is always less than r for $\alpha > 0$. As a result, withdrawal is more preferable since the withdrawal amount will have a higher return at the rate r as priced under the risk neutral valuation. A mathematical argument is presented as follows. Obviously, we have

$$V(W + \delta, A + \delta, t) \leq V(W, A, t) + \delta \tag{3.1}$$

for any finite amount $\delta > 0$; and from which we infer that

$$\frac{\partial V}{\partial W} + \frac{\partial V}{\partial A} = \lim_{\delta \to 0} \frac{V(W + \delta, A + \delta, t) - V(W, A, t)}{\delta} \leq 1. \tag{3.2}$$

With the positivity of $1 - \dfrac{\partial V}{\partial W} - \dfrac{\partial V}{\partial A}$, Eq. (2.10) is reduced to

$$\max\left[-\frac{\partial V}{\partial t} - \mathcal{L}V - \left(1 - \frac{\partial V}{\partial W} - \frac{\partial V}{\partial A}\right)G, \frac{\partial V}{\partial W} + \frac{\partial V}{\partial A} - (1 - k) \right] = 0, \tag{3.3}$$

further confirming that withdrawal always occurs under optimal dynamic withdrawal strategy. These claims on optimal withdrawal policies are verified in our numerical calculations. In Figure 1, we plot the optimal withdrawal boundary that separates the "$\gamma = G$" region and "$\gamma = \infty$" region. The model parameters used in the calculations are: $G = 7$, $r = 0.05$, $\alpha = 0.01$, $w_0 = 100$, $k = 0.1$, $\sigma = 0.2$ and $T = 14.28$. We can identify two different regions under optimal withdrawal policies in the (W, A)-plane. For a given value of W, we move from the "$\gamma = \infty$" region to the "$\gamma = G$" region as the value of A is decreasing gradually. When (W_t, A_t) lies within the "$\gamma = \infty$" region, the holder should withdraw instantaneously a finite amount until (W_t, A_t) falls to a point on the separating boundary. Corresponding to $W = 0$, the optimal withdrawal boundary in Figure 1 is seen to start from the left end at

$$A = -\frac{G}{r}\ln(1 - k) = -\frac{7}{0.05}\ln(1 - 0.1) = 14.75,$$

agreeing with the result deduced from the closed form solution of $V_0(A, t)$ [see Appendix].

4   Conclusion

As baby boomers are now getting close to retirement, sales of variable annuities have become great success in the life insurance industry in the last decade. Investors like to have the possibility of upside equity participation but they are also concerned about the downside risk. The various forms of guarantees embedded in variable annuities provide competing edge over other investment instruments. These guaranteed minimum benefit riders on variable annuities have complex option like characteristics. The sources of risk associated with these guarantee riders include insurance risk (mortality), market risk (equity and interest rate) and policyholder's behaviors (exercise policies of embedded rights).

Following the well known Hamilton-Jacobi-Bellman approach in stochastic control problems, we have managed to construct singular stochastic control models for pricing variable annuities with guaranteed

minimum withdrawal benefit under both continuous and discrete framework. Here, the withdrawal rate is considered as a control variable.

In our derivation of the continuous model, we apply the penalty approach where an upper bound is placed on the withdrawal rate. We then take the bound to tend to infinity subsequently so as to relax the constraint on the withdrawal rate. Interestingly, this penalty approach leads to an effective numerical approximation methods using the finite difference scheme. On the other hand, we have also constructed the numerical scheme for solving the discrete model, following the standard numerical schemes for pricing discretely monitored path dependent options. Since the discrete and continuous versions of the pricing model are derived using quite different approaches, the apparent agreement of the numerical results from both versions serves to check for consistency of the two pricing approaches.

We have analyzed the impact of various model parameters on the fair insurance fee to be charged by the insurer for the provision of the GMWB. The insurance fee increases with increasing equity volatility level and contractual withdrawal rate but decreases with a higher penalty charge. The insurer should charge a substantially higher insurance fee when the policyholder has the flexibility of *dynamic* withdrawal. Also, we have explored the optimal withdrawal policies of the policyholders. When there is a penalty on withdrawal above the contractual rate, the policyholder either withdraws a finite amount (infinite withdrawal rate) or withdraws at the contractual rate. When it is optimal for the policyholder to choose "withdrawal in a finite amount", he chooses to withdraw an appropriate finite amount instantaneously making the equity value of the personal account and guarantee balance to fall to the level that it becomes optimal for him to withdraw at the contractual rate.

References

[1] Barles, G., Ch. Daher, and M. Romano (1995): Convergence of numerical schemes for parabolic equation arising in finance theory, *Mathematical Models and Methods in Applied Sciences*, 5, 125-143.

[2] Barles, G., and P.E. Souganidis (1991): Convergence of approximation schemes for fully nonlinear second order equations, *Asymptotic Analysis*, 1, 271-283.

[3] Chu, C.C., and Y.K. Kwok (2004): Reset and withdrawal rights in dynamic fund protection, *Insurance: Mathematics and Economics*, 34(2), 273-295.

[4] Coleman, T.F., Y. Li, and M.C. Patron (2006): Hedging guarantees in variable annuities under both equity and interest rate risks, *Insurance: Mathematics and Economics*, 38(2), 215-228.

[5] Crandal, M.G., H. Ishii and P.L. Lions (1992): User's guide to viscosity solutions of second order partial differential equations, *Bull. Amer. Soc.*, 27, 1-67.

[6] Dai, M., Y.K., Kwok, and H. You (2007): Intensity-based framework and penalty formulation of optimal stopping problems, to appear in *Journal of Economic Dynamics and Control*.

[7] Davis, M.H.A., A.R. Norman (1990): Portfolio selection with transaction costs, *Mathematics of Operations Research*, 15, 676-713.

[8] Davis, M.H.A., V.G. Panas, and T. Zariphopoulou (1993): European option pricing with transaction costs, *SIAM Journal of Control and Optimization*, 31(2), 470-493.

[9] Fleming, W.H., and H.M. Soner (2006): *Controlled Markov processes and viscosity solutions*, 2nd edition, Springer-Verlag, New York.

[10] Forsyth, P.A., and K.R. Vetzal (2002): Quadratic convergence for valuing American options using a penalty method, *SIAM Journal on Scientific Computation*, 23, 2096-2123.

[11] Forsyth, P.A. and G. Labahn (2006): Numerical methods for controlled Hamilton-Jacobi-Bellman PDEs in finance, working paper of University of Waterloo.

[12] Friedman, A (1982): *Variational Principles and Free Boundary Problems*, Wiley, New York.

[13] Kumar, S. and K. Muthuraman (2004): A numerical method for solving singular stochastic control problems, *Operations Research*, 52(4), 563-582.

[14] Milevsky, M.A., and S.E. Posner (2001): The Titanic option: Valuation of the guaranteed minimum death benefit in variable annuities and mutual funds, *Journal of Risk and Insurance*, 68(1), 93-128.

[15] Milevsky, M.A., and T.S. Salisbury (2006): Financial valuation of guaranteed minimum withdrawal benefits, *Insurance: Mathematics and Economics*, 38(1), 21-38.

[16] Siu, T.K. (2005): Financial fair valuation of participating policies with surrender options and regime switching, *Insurance: Mathematics and Economics*, 37, 533-552.

[17] Yong, J., and X. Zhou (1999): *Stochastic controls: Hamiltonian systems and HJB equations*, Springer-Verlag, New York.

[18] Zhu, H. (1992): Generalized solution in singular stochastic control: the non-degenerate problem, *Applied Mathematics and Optimization*, 25, 225-245.

Appendix - Derivation of closed form formula of $V_0(A,t)$

First, we consider the solution of $V_0(A,t)$ without the inequality constraint: $\frac{\partial V_0}{\partial A} - (1-k) \geq 0$. Together with the observation that $\frac{\partial V_0}{\partial A} \leq 1$ [see Eq. (3.2)], the governing equation for $V_0(A,t)$ is given by

$$\frac{\partial V_0}{\partial t} - G\frac{\partial V_0}{\partial A} - rV_0 + G = 0, \quad 0 \leq t \leq T, 0 \leq A \leq A_0, \qquad (A.1)$$

with auxiliary conditions: $V_0(A,T) = (1-k)A$ and $V_0(0,t) = 0$. If we define

$$W_0(A,t) = V_0(A,t)e^{r(T-t)} - \frac{G}{r}\big[e^{r(T-t)} - 1\big], \qquad (A.2)$$

then $W_0(A,t)$ satisfies the prototype hyperbolic equation:

$$\frac{\partial W_0}{\partial t} - G\frac{\partial W_0}{\partial A} = 0 \qquad (A.3)$$

with auxiliary conditions: $W_0(A,T) = (1-k)A$ and $W_0(0,t) = -\frac{G}{r}\left[e^{r(T-t)} - 1\right]$. The general solution to $W_0(A,t)$ is of the form

$$W_0(A,t) = F(\xi), \quad \xi = t + \frac{A}{G}, \tag{A.4}$$

where F is some function to be determined by the auxiliary conditions. The characteristics of the hyperbolic equation (A.3) are given by the lines: $\xi = t + \frac{A}{G} = \xi_0$, for varying values of ξ_0 (see Figure 2).

(a) For $\xi_0 \geq T$, given $W_0(A,T) = (1-k)A$, we have

$$W_0(A,T) = F(T + \frac{A}{G}) = (1-k)A \qquad \text{for} \quad t + \frac{A}{G} \geq T.$$

We deduce that

$$F(\xi) = (1-k)G(\xi - T)$$

so that

$$V_0(A,t) = e^{-r(T-t)}(1-k)[A - G(T-t)] + \frac{G}{r}[1 - e^{-r(T-t)}], \tag{A.5a}$$
$$A \geq G(T-t).$$

(b) For $\xi_0 < T$, given $W_0(0,t) = -\frac{G}{r}[e^{-r(T-t)} - 1]$, we have

$$W_0(0,t) = F(t) = -\frac{G}{r}\left[e^{r(T-t)} - 1\right] \quad \text{for} \quad t < T - \frac{A}{G}.$$

We deduce that

$$W_0(A,t) = F(t + \frac{A}{G}) = -\frac{G}{r}\left[e^{r(T-t)-\frac{r}{G}A} - 1\right],$$

so that

$$V_0(A,t) = \frac{G}{r}(1 - e^{-\frac{r}{G}A}), \quad A < G(T-t). \tag{A.5b}$$

In the continuation region, $V_0(A,t)$ satisfies Eq. (A.1) together with the inequality:

$$\frac{\partial A_0}{\partial A} > 1 - k. \tag{A.6}$$

The solution of the form given in Eq.(A.5a) is ruled out since the inequality constraint (A.6) is not satisfied. The solution given in Eq.(A.5b) is feasible only if

$$e^{-\frac{r}{G}A} > 1 - k, \quad \text{that is,} \quad A < -\frac{G}{r}\ln(1-k).$$

Hence, the continuation region is limited to the region:

$$\{(A,t) : A < -\frac{G}{r}\ln(1-k) \text{ and } A < G(T-t)\}$$

as shown in the shaded region in Figure 3. Define

$$\tau^* = \min\left(-\frac{\ln(1-k)}{r}, T - t\right),$$

13

then the solution of $V_0(A, t)$ in the continuation region is given by

$$V_0(A, t) = \frac{G}{r}\left(1 - e^{-\frac{r}{G}A}\right) \quad \text{if} \quad A < G\tau^*. \tag{A.7}$$

In the stopping region, $V_0(A, t)$ satisfies $\frac{\partial V_0}{\partial A} = 1 - k$. The solution takes the form:

$$V_0(A, t) = (1 - k)A + C(t),$$

where $C(t)$ is some arbitrary function. The solution in the stopping region is given by

$$V_0(A, t) = (1 - k)A + \frac{G}{r}(1 - e^{-r\tau^*}) - (1 - k)G\tau^*, \quad A \geq G\tau^*. \tag{A.8}$$

Combining the results, the solution is found to be

$$V_0(A, t) = (1 - k)\max(A - G\tau^*, 0) + \frac{G}{r}\left[1 - e^{-r\min(\tau^*, \frac{A}{G})}\right]. \tag{A.9}$$

The above price formula can be interpreted using the following financial argument. To minimize the penalty charge, the policyholder either withdraws at the rate G or infinite rate (instantaneous withdrawal of finite amount). When the guarantee balance $A(t)$ is sufficiently high, the optimal strategy is to withdraw a certain part of $A(t)$ instantaneously, followed by withdrawing the remaining balance at the rate G. To decide on the optimal withdrawal policy, the policyholder strikes the balance between the penalty charge and the time value of the cash flows. The present value of the sum of cash flows received at the rate G from time t to T_0, where $t < T_0 \leq T$, is given by

$$\int_t^{T_0} e^{-r(u-t)}G \; du = \frac{G}{r}\left[1 - e^{-r(T_0 - t)}\right].$$

If the policyholder chooses to receive the lump sum $G(T_0 - t)$ instantaneously at time t, the actual account net of penalty charge to be received is only $(1 - k)G(T_0 - t)$. We define the difference of these two values by

$$D(T_0) = \frac{G}{r}\left[1 - e^{-r(T_0 - t)}\right] - (1 - k)G(T_0 - t), \quad t < T_0 \leq T.$$

The function $D(T_0)$ is concave in T_0, strictly increasing on $T_0 \in (t, T_0^*)$ and strictly decreasing on $T_0 \in (T_0^*, T)$. The critical point T_0^* is given by

$$T_0^* = t + \min\left(-\frac{\ln(1-k)}{r}, \; T - t\right).$$

The optimal withdrawal policies can be deduced as follows. For $t < T_0^*$ [or $t \geq T_0^*$], when $A(t) \leq G(T - T_0^*)$ [or $A(t) \leq G(T - t)$], the policyholder withdraws at the constant rate G throughout the remaining life. Otherwise, when $A(t) > G(T - T_0^*)$ [or $A(t) > G(T - t)$], the policyholder withdraws the discrete amount $A(t) - G(T - T_0^*)$ [or $A(t) - G(T - t)$] instantaneously, then followed by withdrawing at the rate G throughout the remaining life. The present value of the sum of cash flows received by the policyholder following the above optimal withdrawal policies is then equal to the price formula (A.9).

N_t	N_W	N_A	annuity value	change in value	ratio of change
32	64	64	96.241		
64	128	128	94.720	-1.521	
128	256	256	93.788	-0.932	1.6
256	512	512	93.506	-0.282	3.3
512	1024	1024	93.419	-0.087	3.3

Table 1: Examination of the rate of convergence of the Crank-Nicholson scheme for solving the penalty approximation model.

penalty	$k = 1\%$	$k = 10\%$
parameter λ	annuity value	annuity value
10^1	89.515	87.187
10^2	99.924	92.720
10^3	101.884	93.327
10^4	101.028	93.410
10^5	101.043	93.418
10^6	101.045	93.419
10^7	101.045	93.419
10^8	101.045	93.419

Table 2: Test of convergence of the numerical approximation solution to the annuity value with varying values of the penalty parameter λ and penalty charge k.

	$k = 10\%$	
frequency	no reset provision	reset provision
1	92.172	92.168
2	92.800	92.785
3	92.980	92.955
4	93.111	93.088
5	93.163	93.133
6	93.186	93.159
12	93.346	93.299
∞	93.419	—

Table 3: The dependence of the fair value of the GMWB annuity on the withdrawal frequency per year. The annuity value obtained using the continuous withdrawal model (frequency becomes ∞) is close to that corresponding to monthly withdrawal (frequency equal 12). The differences in annuity values with and without the reset provision are seen to be small.

		$k = 5\%$		$k = 10\%$	
contractual rate, g	maturity, $T = 1/g$	$\sigma = 20\%$	$\sigma = 30\%$	$\sigma = 20\%$	$\sigma = 30\%$
4%	25.00	103 bp	213 bp	56 bp	133 bp
5%	20.00	125 bp	260 bp	69 bp	162 bp
6%	16.67	145 bp	305 bp	83 bp	192 bp
7%	14.29	165 bp	348 bp	97 bp	221 bp
8%	12.50	185 bp	390 bp	111 bp	251 bp
9%	11.11	202 bp	429 bp	124 bp	277 bp
10%	10.00	219 bp	466 bp	137 bp	304 bp
15%	6.67	296 bp	639 bp	198 bp	434 bp

Table 4: Impact of the GMWB contractual rate g, penalty charge k and equity volatility σ of the account on the required insurance fee α (in basis points).
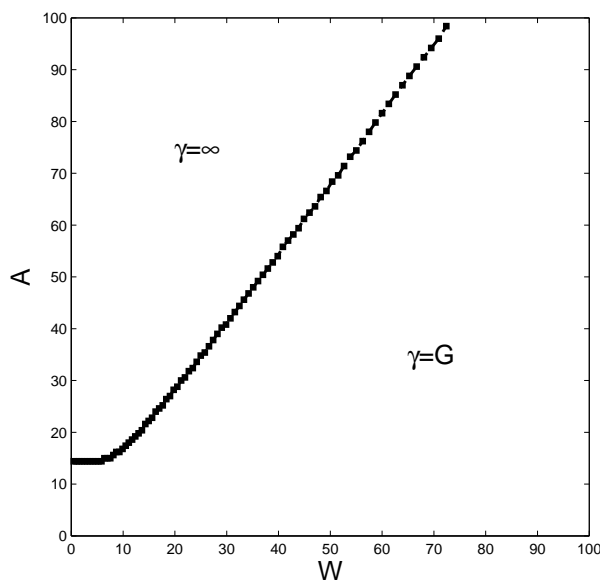

Figure 1: Plot of the optimal withdrawal boundary in the (W, A)-plane, separating the "$\gamma = \infty$" region at the top from the "$\gamma = G$" region at the bottom. The boundary intersects the A-axis at $A = -\frac{G}{r}\ln(1 - k)$.
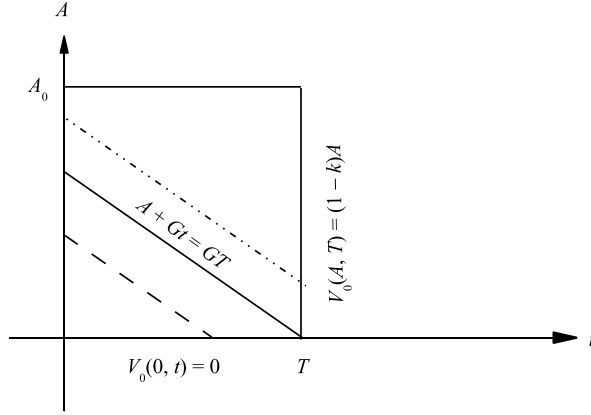
Figure 2: The characteristic lines are given by $t + \dfrac{A}{G} = \xi_0$ for varying values of ξ_0. For $\xi_0 > T$, the characteristic lines intersect the right vertical boundary: $t = T$; and for $\xi_0 \leq T$, the characteristics lines intersect the bottom horizontal boundary: $A = 0$.
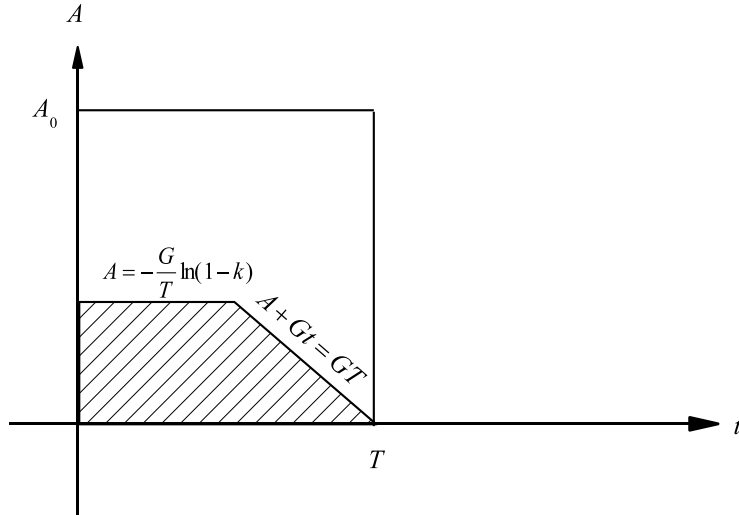

Figure 3: The continuation region lies in the region (shaded part) $\{(t, A) : A \leq -\dfrac{G}{r} \ln(1 - k)$ and $A - G(T - t) \leq 0\}$, with $V_0(t, A) = \dfrac{G}{r}(1 - e^{-\frac{r}{G}A})$.

17