Optimal pension decumulation under age-dependent mortality and lifetime ruin constraints

Tim Prokosch

*AG Finanzmathematik, RPTU Kaiserslautern-Landau, Kaiserslautern, Germany.

Corresponding author(s). E-mail(s): prokosch@rptu.de;

Abstract

We study a pension decumulation problem in which an investor withdraws income at a constant rate and allocates wealth between a risk-free asset and multiple risky assets, while seeking to minimize the probability of reaching a positive impoverishment threshold before an independent time of death. Because realistic age-dependent mortality generally destroys the classical closed-form solution, we reverse the usual modelling order and construct a mortality law that is compatible with analytic solvability. For a finite maximal lifetime, the resulting survival law produces a monomial value function in normalized wealth and an explicit optimal investment strategy. The solution also yields a closed-form sustainable withdrawal rate and a sufficient condition under which the unconstrained strategy automatically satisfies no-short-selling and no-borrowing constraints. We show that the associated hazard rate has a Gompertz–Makeham leading term with an explicit positive remainder and converges to the constant-hazard benchmark as the maximal lifetime tends to infinity. Under the optimal strategy, the distance from the safe wealth boundary follows a geometric Brownian motion, reducing the time of impoverishment to a Brownian first-passage problem with a curved boundary. The constructed family therefore provides a tractable analytic benchmark for evaluating policies and approximations under more general mortality specifications.

Keywords: pension decumulation; self-annuitization; stochastic control; age-dependent mortality; lifetime ruin; Gompertz–Makeham law

1 Introduction

Germany's old-age security system is heavily shaped by the statutory pay-as-you-go pension scheme. Demographic change and increasing life expectancy make this system progressively sensitive to the ratio between contributors and retirees. Whereas the ratio of contributors to retirees was around 6 in 1962, it declined to around 2 in the early 2000s and is projected to decline further to around 1.3 by 2035 [8, 30]. Moreover, the Federal Statistical Office reports that, under all variants of its 16th coordinated population projection, approximately one quarter of Germany's population will be aged 67 or older by 2035, compared with roughly one fifth in 2024 [13]. The German Council of Economic Experts similarly emphasizes that the retirement of the baby-boomer cohorts places renewed pressure on the long-run sustainability of the pay-as-you-go system [16]. OECD pension reports provide an international comparative perspective and repeatedly stress that pension systems have to adapt to aging populations and changing labor-market conditions [27].

For an individual's pension provision, this background motivates a portfolio problem that differs from the usual accumulation or consumption problem, respectively. A retiree is often not primarily interested in maximizing terminal wealth or expected discounted utility of lifetime consumption. The central question is instead more concrete:

> *How much real income can be withdrawn per unit time while keeping the probability of running out of wealth acceptably small?*

In German terminology, this is close to asking for an individually sustainable *Rentenniveau*: a stable withdrawal level that supports an existing, or desired, standard of living.

The practical appeal of this question is reflected in Bengen's historical analysis of retirement withdrawals [4]. For prescribed stock–bond allocations, Bengen examines how large an inflation-adjusted withdrawal could have been sustained over historical retirement periods. His conclusion is commonly summarized as the "4% rule": an initial withdrawal equal to roughly four percent of the portfolio, followed by inflation-adjusted withdrawals, survived the historical 30-year periods considered in the study. This provides an intuitive benchmark, but it is a historical stress-test for a prespecified asset allocation and horizon rather than the solution of a stochastic control problem. In particular, it neither selects a dynamically optimal portfolio nor links the withdrawal rate to an explicitly chosen probability of impoverishment over a random lifetime.

At the theoretical level, classical continuous-time consumption–investment models originating with Samuelson and Merton [21, 22, 29] usually select consumption and investment so as to maximize expected discounted lifetime utility of consumption

$$\mathbb{E}\left[\int_0^\infty \mathrm{e}^{-\delta t} U(C_t)\,\mathrm{d}t\right]$$

for some discount rate $\delta > 0$ and utility function U. Value-preserving portfolio strategies provide a different, non-utility-based perspective. Korn [18] derives this principle from a local mean–variance argument and studies value-preserving trading and consumption strategies in a general semimartingale framework. These criteria are

mathematically elegant, but the output is not always directly interpretable as a pension statement. For one, no single individual knows their own utility function, and the choice of a particular functional form is often more a matter of mathematical convenience than economic realism. Moreover, overly aggressive discounting is often times necessary to ensure that the infinite-horizon integral converges, leading to a focus on short-term consumption at the expense of long-term sustainability. Similarly, the value-preserving principle may require capital injections to maintain a prescribed wealth level, which is not feasible for retirees who have no future labor income and depend on their existing wealth to fund consumption.

For this reason, we instead consider a ruin-based formulation where the investor seeks to maintain a sustainable withdrawal level, while controlling the probability of running out of wealth within a random retirement horizon. That is, they seek to find an investment strategy satisfying

$$\mathbb{P}(\tau_{\text{imp}} < \mathcal{T}) \leq \alpha,$$

where τ_{imp} is the first time wealth reaches an impoverishment level x_{imp} that is deemed unacceptable, $\mathcal{T}$ is the (random) duration of the retirement period, and α is the maximum acceptable probability of impoverishment. Here, $\mathbb{P}$ denotes the probability measure induced by the underlying stochastic model and the chosen investment strategy. One should note that, in the following, we will differentiate the terms 'ruin' and 'impoverishment', as the former is often used in the actuarial literature to refer to the event of wealth reaching zero, whereas the latter can be interpreted more broadly as wealth falling below a certain reservation level that is considered insufficient for maintaining an acceptable standard of living. The details of the model and the precise definition of the optimization problem will be given in the next section. The main takeaway is that the model inputs x_{imp}, $\mathcal{T}$, and α are directly interpretable as a minimum acceptable reservation level, a retirement duration, and a maximum acceptable probability of impoverishment, respectively. Likewise, the output of the optimization problem is then a sustainable withdrawal level C that can be supported under the given constraints.

This problem belongs to the literature on self-annuitization and lifetime ruin. Milevsky and Robinson [24] study retirement ruin under fixed withdrawals, a fixed asset allocation, and Gompertz mortality. Young [34] formulates the corresponding dynamic investment problem when the remaining lifetime is exponentially distributed. The constant hazard rate makes that problem time-homogeneous and leads to a power-form value function and an optimal risky investment that is linear in the distance from the safe wealth level. This solution is an important analytic benchmark, but a constant hazard rate does not capture the usual increase of mortality with age. Age-dependent mortality has already been incorporated into lifetime-ruin models. Moore and Young [26] consider general deterministic mortality laws, including Gompertz' law, and obtain optimal policies numerically. Liang and Young [20] treat a general deterministic hazard rate and analyze a simple time-dependent strategy derived from the constant-hazard solution, together with bounds on its value loss. Portfolio constraints, regularity and duality methods, and ambiguity about mortality have likewise been studied in related settings [2, 3, 35]. Thus, neither age-dependent mortality nor

constrained lifetime-ruin control is new in itself. The remaining analytical gap is that explicit optimal feedbacks are generally unavailable once the hazard rate varies with time.

We approach this gap through an inverse-solvability question: which mortality laws are compatible with an explicitly solvable lifetime-impoverishment HJB equation? Instead of prescribing an arbitrary mortality law and then solving the resulting time-inhomogeneous control problem numerically, we impose tractable structure on the transformed value function and recover compatible survival laws. This deliberately exchanges mortality-model generality for exact solvability.

Our principal results are as follows.

(i) We construct a finite-endpoint, time-dependent mortality law for which the lifetime-impoverishment problem has an explicit solution. Relative to the moving safe boundary, the minimum impoverishment probability is a power of the normalized safe gap, and the optimal feedback is available in closed form for a market with multiple risky assets. As the endpoint tends to infinity, the model recovers the classical constant-hazard power solution.

(ii) The explicit value function can be inverted to obtain a sustainable withdrawal rate for a prescribed risk tolerance. We also derive a sufficient condition under which the unconstrained multi-asset feedback satisfies the no-short-sale and no-borrowing simplex constraint throughout the state space. This result identifies the globally nonbinding region; it does not solve the constrained HJB when the constraint binds.

(iii) The constructed hazard rate admits a Gompertz–Makeham leading term and an explicit positive remainder. The expansion is uniform on intervals bounded away from the finite endpoint and makes precise both the actuarial interpretation of the mortality law and its departure from an exact Gompertz–Makeham specification.

(iv) Under the optimal feedback, the distance between wealth and the moving safe boundary follows a geometric diffusion. This yields a Brownian curved-boundary representation of the impoverishment time. Independence of mortality and financial risk then converts the financial first-passage law into the distribution of impoverishment before death.

The remainder of the paper is organized as follows. Section 2 introduces the market and mortality models, derives the moving safe boundary and HJB equation, and establishes the verification framework. Section 3 develops the inverse solvability conditions. Section 4 presents the compactified-exponential mortality law and its withdrawal and portfolio consequences. Section 4.2 relates its hazard rate to the Gompertz–Makeham form. Section 5 studies the distribution of the impoverishment time, and Section 6 describes the numerical assessment. Sections 7 and 8 discuss the limitations of the construction and conclude.

2 The model and verification framework

This section establishes the model and verification framework used throughout the paper. We first specify the financial market, mortality law, controlled wealth dynamics, and admissible investment strategies. We then identify the safe boundary, normalize wealth by the withdrawal rate, and derive the Hamilton–Jacobi–Bellman equation together with its boundary conditions. The verification result will provide sufficient conditions under which a candidate solution equals the value function and its associated feedback strategy is optimal.

In the setting of this paper, any investor $\mathcal{I}$ is characterized by the tuple

$$\mathcal{I} := (x_0, x_{\mathrm{imp}}, \alpha, S),$$

where $x_0 > x_{\mathrm{imp}} > 0$ denote initial wealth and the impoverishment level, $\alpha \in (0,1)$ is the maximal acceptable lifetime-impoverishment probability, and S describes the investor's mortality law. Given the financial market, she then aims to choose a constant withdrawal rate $C > 0$ and an investment strategy π so as to minimize the probability of running out of wealth during her remaining lifetime. The precise construction of the mortality law, the controlled wealth process, and the class of admissible strategies is provided below.

2.1 Market, mortality, and admissible controls

First, we will define the financial market, the controlled wealth process together with the corresponding admissible strategies, and the investor's mortality model. To this end, let $\mathcal{P}^B := (\Omega^B, \mathcal{F}^B, \mathbb{F}^B, \mathbb{P}^B)$ denote the filtered probability space supporting the financial market, and let $\mathcal{P}^{\mathcal{T}} := (\Omega^{\mathcal{T}}, \mathcal{F}^{\mathcal{T}}, \mathbb{P}^{\mathcal{T}})$ denote the probability space supporting a non-negative random time $\mathcal{T}$. We consider the product probability space $\mathcal{P}^B \otimes \mathcal{P}^{\mathcal{T}} =: (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where $\Omega := \Omega^B \times \Omega^{\mathcal{T}}$, $\mathcal{F} := \mathcal{F}^B \otimes \mathcal{F}^{\mathcal{T}}$, and $\mathbb{P} := \mathbb{P}^B \otimes \mathbb{P}^{\mathcal{T}}$. For $t \geq 0$, define the mortality filtration by $\mathcal{F}_t^{\mathcal{T}} := \sigma(\mathcal{T} \wedge t)$, and equip the product space with the filtration $\mathbb{F} := (\mathcal{F}_t)_{t \geq 0}$, where $\mathcal{F}_t := \mathcal{F}_t^B \otimes \mathcal{F}_t^{\mathcal{T}}$, together with its usual augmentation. The market processes are lifted canonically to Ω by dependence on the first coordinate only, whereas $\mathcal{T}$ depends only on the second coordinate. Consequently, $\mathcal{T}$ is an $\mathbb{F}$-stopping time and is independent of the financial market, in the sense that $\mathcal{F}_\infty^B$ and $\sigma(\mathcal{T})$ are independent under $\mathbb{P}$.

Definition 1 (Financial market) Let B be a d-dimensional standard Brownian motion. The market is modeled via the usual Black–Scholes dynamics, that is, the money-market account S^0 and the risky assets S^i, $i = 1, \ldots, d$, satisfy

$$\begin{aligned} \mathrm{d}S_t^0 &= r S_t^0 \, \mathrm{d}t, \quad S_0^0 = 1, \\ \mathrm{d}S_t^i &= S_t^i \left((r + \theta_i) \, \mathrm{d}t + \sum_{j=1}^{d} \sigma_{ij} \, \mathrm{d}B_t^j \right), \quad S_0^i > 0, \quad i = 1, \ldots, d, \end{aligned} \tag{1}$$

with a risk-free rate $r > 0$, a vector of excess returns[1] $\theta \in \mathbb{R}^d$, and a non-singular volatility matrix $\sigma \in \mathbb{R}^{d\times d}$. For ease of notation we additionally define the derived quantities

$$\Sigma := \sigma\sigma^\top \succ 0, \qquad \mathfrak{m} := \Sigma^{-1}\theta, \qquad \rho := \frac{1}{2}\theta^\top\Sigma^{-1}\theta.$$

Further, we assume that the vector of excess returns θ is non-zero, that is, there exists at least one risky asset with a mean return different from the risk-free rate. This assumption is standard in the literature and ensures that the investor has a non-trivial investment opportunity set. Since Σ is positive definite, this assumption is equivalent to $\rho > 0$. This will be used in subsequent derivations.

The investor's wealth process X is controlled by a risky-fraction strategy π and a constant withdrawal rate $C > 0$. The dynamics of the controlled wealth process are given below.

Definition 2 (Wealth process) For a risky-fraction strategy π and a constant (net[2]) withdrawal rate $C > 0$, wealth evolves according to

$$\mathrm{d}X_t = \left(rX_t - C + X_t\pi_t^\top\theta\right)\mathrm{d}t + X_t\pi_t^\top\sigma\ \mathrm{d}B_t. \tag{2}$$

To ensure the well-posedness of the controlled wealth process, we introduce the following standard definition of admissible strategies.

Definition 3 (Admissible strategies) Let $K \subseteq \mathbb{R}^d$. For an initial state (t, x), an admissible strategy is a $\mathbb{F}^B$-progressively measurable, K-valued process π for which (2) has a unique strong solution and, for every finite $u \geq t$,

$$\mathbb{E}_{t,x}\left[\int_t^u \left|X_s\pi_s^\top\theta\right|\ \mathrm{d}s\right] < \infty, \qquad \mathbb{E}_{t,x}\left[\int_t^u X_s^2\pi_s^\top\Sigma\pi_s\ \mathrm{d}s\right] < \infty.$$

Here $\mathbb{E}_{t,x}$ denotes the conditional expectation given $X_t = x$.[3] The corresponding class is denoted by $\mathcal{A}_K(t, x)$. We consider

$$K_{\text{unc}} := \mathbb{R}^d, \qquad K_{\text{con}} := \left\{u \in \mathbb{R}^d : u_i \geq 0,\ \mathbf{1}^\top u \leq 1\right\}.$$

If the particular choice of $K \in \{K_{\text{unc}}, K_{\text{con}}\}$ is clear from the context, or does not matter, we will simply write K and $\mathcal{A}(t, x)$ for sake of brevity.

One special class of admissible strategies is the class of *feedback controls*. As we will later see, such controls will naturally arise as the optimal controls in our setting.

Definition 4 (Feedback control) A measurable function $\pi : [0, \infty) \times \mathbb{R} \to K$ is called a feedback rule. Given a state (t, x), let X^π denote a solution of the closed-loop wealth equation (2) with initial condition $X_t^\pi = x$, and define the induced control process by

$$\pi_s := \pi(s, X_s^\pi), \quad s \geq t.$$

[1] The excess return of an asset is the difference between its expected return and the risk-free rate, less any potential management fees common in many investment products like ETFs.

[2] If a continuous stream of income is present, the net withdrawal rate is the total withdrawals less the rate of income. To shorten the notation, we will omit the term "net" in the future.

[3] For the corresponding conditional probability measure, we write $\mathbb{P}_{t,x}$.

We call the feedback rule admissible, if the induced control process is admissible, that is, $\pi \in \mathcal{A}_K(t,x)$.

To model the investor's mortality, we make use of classical concepts from survival analysis, which we summarize below.

Definition 5 (Mortality model) As constructed above, let $\mathcal{T}$ be a non-negative random time independent of the Brownian filtration driving the financial market. Its survival function is

$$S(t) := \mathbb{P}(\mathcal{T} > t), \quad t \in [0, \infty),$$

where for $t = \infty$ we set $S(\infty) := 0$, and the deterministic right endpoint of its support is

$$T := \sup \operatorname{supp} S \in (0, \infty].$$

Throughout this paper, we assume $S(0) = 1$, $S(t) \to 0$ as $t \to T$, and $S \in C^1([0,T))$. The corresponding force of mortality, or hazard rate, is

$$\lambda(t) := -\frac{S'(t)}{S(t)}, \quad 0 \le t < T.$$

With the central objects of the model now defined, we can proceed to formalize the investor's objective of minimizing the conditional probability of lifetime impoverishment.

2.2 Impoverishment objective and domain reduction

Having defined the controlled wealth process and the investor's mortality model, we can now formalize the investor's objective of minimizing the conditional probability of lifetime impoverishment. To this end, we introduce the following definition of the impoverishment time and the corresponding minimal risk function.

Definition 6 (Impoverishment time and minimal risk) For a controlled wealth process started from $(t,x) \in [0,T) \times [0,\infty)$, define

$$\tau_{\text{imp}} := \inf\{s \ge t : X_s \le x_{\text{imp}}\}$$

with the convention $\inf \emptyset = \infty$. Then, the conditional minimal lifetime-impoverishment probability is

$$p(t,x;C) := \inf_{\pi \in \mathcal{A}(t,x)} \mathbb{P}_{t,x}\big(\tau_{\text{imp}} < \mathcal{T} \,\big|\, \mathcal{T} > t\big). \tag{3}$$

Unless stated otherwise, we suppress from the notation the dependence of p on the fixed model primitives, including the financial-market parameters, the impoverishment level x_{imp}, the mortality law S, and the constraint set K. Thus, $p(t,x;C)$ displays only the current state and the withdrawal rate.

With this notation, the investor's optimal sustainable withdrawal rate can be expressed as the inverse of the minimal risk function p in the following sense:

$$C^*(x_0;\alpha) := \sup\big\{C > 0 : p(0,x_0;C) \le \alpha\big\}. \tag{4}$$

Consistently with the preceding convention, the dependence of C^* on the fixed model primitives remains suppressed. When x_0 and α are fixed as well, we simply write C^*. Note that, although the investor's risk tolerance is restricted to $\alpha \in (0,1)$, some of the explicit formulas derived below admit a continuous extension to $\alpha = 0$. Whenever this boundary value is invoked, it is understood only in this limiting sense and not as an additional instance of the now degenerate control problem.

This representation already outlines the subsequent steps one needs to follow in order to solve the optimal withdrawal problem:

Step 1: For a given withdrawal rate C, solve the stochastic control problem (3) to obtain the minimal risk function $(t,x) \mapsto p(t,x;C)$.

Step 2: Invert the mapping $C \mapsto p(0,x_0;C)$ to obtain the optimal sustainable withdrawal rate C^* as defined in (4).

We will see that for the setting considered in this paper, the second step is especially simple, as the mapping $C \mapsto p(0,x_0;C)$ is a simple monomial in a transformed variable similar to previous work on the lifetime ruin problem [20, 26, 34]. Hence, the main challenge is to solve the stochastic control problem (3) for a given withdrawal rate C.

An immediate consequence of Definition 6 is the existence of a so-called *safe-level*.

Lemma 1 (Existence of a safe-level) *Assume $C > r x_{\mathrm{imp}}$. For $0 \le t < T$, define*

$$x_{\mathrm{safe}}(t) := x_{\mathrm{safe}}^{\infty} - \left(x_{\mathrm{safe}}^{\infty} - x_{\mathrm{imp}}\right) \mathrm{e}^{-r(T-t)}, \tag{5}$$

with $x_{\mathrm{safe}}^{\infty} := C/r$, and the convention $\exp(-r(\infty - t)) = 0$. Then $x_{\mathrm{safe}}(t)$ is the least wealth from which the strategy $\pi \equiv 0$ can finance withdrawals without crossing x_{imp} before the endpoint T. In particular,

$$p(t, x_{\mathrm{safe}}(t)) = 0.$$

If $T < \infty$, then $\lim_{t \uparrow T} x_{\mathrm{safe}}(t) = x_{\mathrm{imp}}$; when $T = \infty$, $x_{\mathrm{safe}} \equiv C/r$.

Proof Placing all wealth in the riskless asset, the wealth dynamics reduce to the ODE

$$\dot{X} = rX - C.$$

In case $T = \infty$, one easily verifies that at least $x_{\mathrm{safe}}^{\infty} = C/r$ is required to avoid impoverishment indefinitely. For $T < \infty$, we can determine the minimal safe level $x_{\mathrm{safe}}(t)$ by solving the ODE backward from the terminal condition $X_T = x_{\mathrm{imp}}$. The solution is given by (5). The rest of the lemma follows directly from this construction. □

For later use, we also invert the safe-boundary relation at time zero. For fixed initial wealth $x_0 > x_{\mathrm{imp}}$, we define the *safe withdrawal rate* by

$$C_{\mathrm{safe}}(x_0) := r\left(x_{\mathrm{imp}} + \frac{x_0 - x_{\mathrm{imp}}}{1 - \mathrm{e}^{-rT}}\right), \tag{6}$$

again with the convention $\mathrm{e}^{-r\infty} = 0$. This is the maximal constant withdrawal rate that can be financed by riskless investment alone without crossing x_{imp} before T; equivalently, it is the unique withdrawal rate for which $x_{\mathrm{safe}}(0) = x_0$.

The assumption $C > rx_{\text{imp}}$ of Lemma 1 is natural, as otherwise the safe level $x_{\text{safe}}(t)$ would be below the impoverishment threshold x_{imp}, leading to a degenerate risky region. For this reason, we will assume $C > rx_{\text{imp}}$ throughout the remainder of this paper and not state it explicitly in the subsequent results. As the following lemma shows, this is not a restriction.

Lemma 2 *The optimal sustainable withdrawal rate C^* defined in (4) can be equivalently expressed as*

$$C^*(x_0;\alpha) = \sup\big\{C > rx_{\text{imp}} : p(0,x_0;C) \le \alpha\big\}.$$

Proof We only need to show that withdrawal rates above but sufficiently close to rx_{imp} still result in a value function that is less than or equal to α. To this end, let $\varepsilon > 0$ and write $C_\varepsilon := rx_{\text{imp}} + \varepsilon$. Then, by noticing that

$$x_{\text{safe}}(0) = x_{\text{imp}} + \frac{\varepsilon}{r}\left(1 - \mathrm{e}^{-rT}\right),$$

and choosing

$$0 < \varepsilon \le \frac{r(x_0 - x_{\text{imp}})}{1 - \mathrm{e}^{-rT}},$$

we have $x_0 \ge x_{\text{safe}}(0)$ and thus $p(0,x_0;C_\varepsilon) = 0 \le \alpha$. □

One remarkable observation is that, in case $T < \infty$, the safe boundary $x_{\text{safe}}(t)$ is strictly decreasing in t and approaches x_{imp} as t approaches T, leading to a collapse of the risky region at the terminal time. This is in contrast to the infinite-horizon case, where the safe level is constant and independent of time. For this reason, we introduce the following definition of the risky region and a normalized safe-gap coordinate, which will be used in the subsequent analysis.

Definition 7 (Risky region and normalized safe gap) The nontrivial state domain and its normalized safe-gap coordinate are

$$\mathcal{D} := \big\{(t,x) : 0 \le t < T,\ x_{\text{imp}} < x < x_{\text{safe}}(t)\big\}, \qquad z(t,x) := \frac{x_{\text{safe}}(t) - x}{x_{\text{safe}}(t) - x_{\text{imp}}}.$$

Thus, $z = 0$ at the safe boundary and $z = 1$ at the impoverishment boundary. We also set

$$\tau_{\text{safe}} := \inf\{s \ge t : X_s \ge x_{\text{safe}}(s)\}, \qquad \tau_{\mathcal{D}} := \tau_{\text{imp}} \wedge \tau_{\text{safe}}.$$

Additionally, we define the spatial closure of the risky region by

$$\overline{\mathcal{D}}^x := \big\{(t,x) : 0 \le t < T,\ x_{\text{imp}} \le x \le x_{\text{safe}}(t)\big\}.$$

Analogously, we define

$$\mathcal{D}^z := \big\{(t,z) : 0 \le t < T,\ 0 < z < 1\big\}, \qquad \overline{\mathcal{D}}^z := \big\{(t,z) : 0 \le t < T,\ 0 \le z \le 1\big\}.$$

It is straightforward to see that the minimal risk function p is constant outside the risky region, as stated in the following corollary.

Corollary 3 (Minimal risk outside the risky region) *The minimal risk function p satisfies*

$$p(t,x) = 1 \quad \text{for } (t,x) \in [0,T) \times [0, x_{\text{imp}}],$$
$$p(t,x) = 0 \quad \text{for } (t,x) \in [0,T) \times [x_{\text{safe}}(t), \infty).$$

The results of Lemma 1 and Corollary 3 motivate us to work with the normalized safe-gap coordinate z instead of the wealth process X in the subsequent analysis. Despite $t = T$ not being in the domain of interest, it is useful to record what happens at the terminal time T when $T < \infty$. As mentioned above, the risky region collapses to the single point $x_{\text{imp}} = x_{\text{safe}}(T)$, so the normalized safe-gap coordinate z is not defined there. Indeed, for $t < T$,

$$z(t, x_{\text{safe}}(t)) = 0 \quad \text{and} \quad z(t, x_{\text{imp}}) = 1,$$

whereas at the terminal time both expressions formally give the indeterminate form $0/0$. The minimal risk function also has distinct limiting values along these two boundary curves:

$$\lim_{t \uparrow T} p(t, x_{\text{imp}}) = 1 \quad \text{while} \quad \lim_{t \uparrow T} p(t, x_{\text{safe}}(t)) = 0.$$

This, however, causes no difficulty, since all subsequent analysis is carried out for $t < T$.

This will be addressed in more detail in the subsequent sections, where we will derive the Hamilton–Jacobi–Bellman equation and the corresponding boundary conditions for the minimal risk function p. However, we will see that this will cause no issue for the class of survival functions considered in this paper.

2.3 Hamilton–Jacobi–Bellman equation

Having set up the controlled wealth process and the investor's mortality model, we can now derive the Hamilton–Jacobi–Bellman (HJB) equation corresponding to the stochastic control problem (3) and prove the associated verification theorem. As this has been done in the literature before, we will only summarize the main steps and refer to the relevant references for the detailed derivation. To this end we introduce the following two objects.

Definition 8 (Controlled generator) For $u \in K$ and $f \in C^{1,2}(\mathcal{D})$, let

$$\mathcal{L}^u f(t,x) := \big(rx - C + xu^\top \theta\big) f_x(t,x) + \frac{x^2}{2} u^\top \Sigma u \, f_{xx}(t,x).$$

Definition 9 (Normalizing factor) For $0 \le t < T$, let

$$\mathcal{E}(t) := \frac{r}{1 - \mathrm{e}^{-r(T-t)}}$$

where we adopt the convention $\exp(-r(\infty - t)) = 0$ for $t \ge 0$.

Remark 1 Using the normalizing factor $\mathcal{E}(t)$, we can express the boundary of the safe region $x_{\text{safe}}(t)$ equivalently as

$$x_{\text{safe}}(t) = x_{\text{imp}} + \frac{C - r x_{\text{imp}}}{\mathcal{E}(t)}.$$

Similarly, the safe withdrawal rate $C_{\text{safe}}(x_0)$ defined in (6) can be expressed as

$$C_{\text{safe}}(x_0) = r x_{\text{imp}} + \mathcal{E}(0)(x_0 - x_{\text{imp}}).$$

Following the standard approach in stochastic control, we can derive the HJB equation for the minimal risk function p by considering the infinitesimal generator of the controlled wealth process and applying the dynamic programming principle, under the assumption of sufficient regularity [34]. Doing so, we obtain the following system of equations:

$$\begin{aligned} p_t(t,x;C) + \inf_{u\in K} \mathcal{L}^u p(t,x;C) - \lambda(t) p(t,x;C) = 0, \quad & (t,x) \in \mathcal{D}, \\ p(t, x_{\text{imp}}; C) = 1, \quad & 0 \le t < T, \\ p(t, x_{\text{safe}}(t); C) = 0, \quad & 0 \le t < T. \end{aligned} \tag{7}$$

By switching from x to the normalized safe-gap coordinate z, we can equivalently express the HJB equation in terms of z as follows:

$$\begin{aligned} p_t(t,z;C) + \inf_{u\in K} \overline{\mathcal{L}}^u p(t,z;C) - \lambda(t) p(t,z;C) = 0, \quad & (t,z) \in \mathcal{D}^z, \\ p(t,1;C) = 1, \quad & 0 \le t < T, \\ p(t,0;C) = 0, \quad & 0 \le t < T \end{aligned} \tag{8}$$

with $\overline{\mathcal{L}}^u$ given by

$$\overline{\mathcal{L}}^u f(t,z) := \big(\mathcal{E}(t) z - a(t,z) u^\top \theta\big) f_z(t,z) + \frac{1}{2} a(t,z)^2 u^\top \Sigma u \, f_{zz}(t,z),$$

where $a(t,z) := 1 - z + \frac{x_{\text{imp}}}{x_{\text{safe}}(t) - x_{\text{imp}}}$.

Before proceeding to the verification theorem, we will first provide a short analysis of the Hamiltonian terms in the HJB equation (7) for both the unconstrained and constrained cases, that is

$$\mathcal{H}[p](t,x) := \inf_{u\in K} \mathcal{L}^u p(t,x;C) = \inf_{u\in K} \left\{ (rx - C + x u^\top \theta) p_x + \frac{x^2}{2} u^\top \Sigma u \, p_{xx} \right\}.$$

Assuming that $p_x < 0$, we can simplify the Hamiltonian and obtain

$$\mathcal{H}[p](t,x) = (rx - C) p_x + x p_x \sup_{u\in K} \left\{ u^\top \theta + \frac{1}{2} \frac{x p_{xx}}{p_x} u^\top \Sigma u \right\}.$$

Thus, we can restrict our attention to what we call the *reduced Hamiltonian*:

$$\widehat{\mathcal{H}}[p](\kappa) := \sup_{u \in K} \left\{ u^\top \theta + \frac{1}{2} \kappa u^\top \Sigma u \right\}.$$

From this representation, one can immediately derive the following result.

Lemma 4 (Optimal control) *Let $\kappa < 0$. Then,*

$$u^*(\kappa; K) := \arg\max_{u \in K} \left\{ u^\top \theta + \frac{1}{2} \kappa u^\top \Sigma u \right\}$$

is well-defined and unique. Moreover, $u^*(\kappa)$ can be expressed as

$$u^*(\kappa; K) = \arg\min_{u \in K} \|u - u^*_{\mathrm{unc}}(\kappa)\|_\Sigma =: \Pi_K^{\|\cdot\|_\Sigma}(u^*_{\mathrm{unc}}(\kappa)),$$

where $u^*_{\mathrm{unc}}(\kappa) := u^*(\kappa; K_{\mathrm{unc}}) = -\kappa^{-1}\mathfrak{m}$ is the unconstrained optimal control and $\Pi_K^{\|\cdot\|_\Sigma}$ denotes the projection onto K in the covariance norm[4] $\|v\|_\Sigma := (v^\top \Sigma v)^{1/2}$.

Proof The form of the unconstrained optimal control u^*_{unc} follows from the first-order condition of the reduced Hamiltonian when $K = K_{\mathrm{unc}}$. Note that the existence and uniqueness of u^*_{unc} follows from the strict concavity of the reduced Hamiltonian in u for $\kappa < 0$. Expressing the risk premium via $\theta = -\kappa \Sigma u^*_{\mathrm{unc}}(\kappa)$, we can rewrite the objective function of the reduced Hamiltonian as

$$\begin{aligned} u^\top \theta + \frac{1}{2} \kappa u^\top \Sigma u &= \frac{\kappa}{2} \left[(u - u^*_{\mathrm{unc}}(\kappa))^\top \Sigma (u - u^*_{\mathrm{unc}}(\kappa)) - u^*_{\mathrm{unc}}(\kappa)^\top \Sigma u^*_{\mathrm{unc}}(\kappa) \right] \\ &= \frac{\kappa}{2} \left[\|u - u^*_{\mathrm{unc}}(\kappa)\|_\Sigma^2 - \|u^*_{\mathrm{unc}}(\kappa)\|_\Sigma^2 \right] \\ &= -\frac{|\kappa|}{2} \|u - u^*_{\mathrm{unc}}(\kappa)\|_\Sigma^2 + \text{constant}. \end{aligned}$$

Hence, maximizing the reduced Hamiltonian over $u \in K$ is equivalent to minimizing the distance to the unconstrained optimal control $u^*_{\mathrm{unc}}(\kappa)$ in the covariance norm. Since K_{con} is a closed convex set, the projection onto K_{con} in the covariance norm is well-defined and unique, which proves the claim. □

A nice consequence of this result is that, under mild conditions, the optimal control $u^*(\kappa; K_{\mathrm{con}})$ in the constrained case coincides with the unconstrained optimal control $u^*_{\mathrm{unc}}(\kappa)$ for sufficiently negative κ.

Corollary 5 *Suppose that $\mathfrak{m} \geq 0$ componentwise and let $\kappa < 0$. Then*

$$u^*(\kappa; K_{\mathrm{con}}) = u^*_{\mathrm{unc}}(\kappa)$$

if and only if

$$\kappa \leq -\|\mathfrak{m}\|_1.$$

[4] As we assumed Σ to be positive definite, this is a proper norm.

Proof By Lemma 4,

$$u^*(\kappa; K_{\text{con}}) = \Pi_{K_{\text{con}}}^{\|\cdot\|_\Sigma}\left(u^*_{\text{unc}}(\kappa)\right).$$

Hence, the constrained and unconstrained controls coincide if and only if $u^*_{\text{unc}}(\kappa) \in K_{\text{con}}$. Since

$$u^*_{\text{unc}}(\kappa) = -\kappa^{-1}\mathfrak{m} \geq 0,$$

this is equivalent to

$$\mathbf{1}^\top u^*_{\text{unc}}(\kappa) = -\kappa^{-1}\|\mathfrak{m}\|_1 \leq 1,$$

or, equivalently,

$$\kappa \leq -\|\mathfrak{m}\|_1.$$

□

This observation will allow us to derive explicit solutions for the minimal risk function p and the corresponding optimal control u^* in the constrained case, as we will see in the subsequent sections after we have established the verification theorem.

2.4 A verification theorem

In what follows, we will verify that the candidate solution to the HJB equation is indeed the value function of our optimal control problem. In previous work, this has been done for the case of an infinite-horizon problem with a constant safe level [20, 26, 34]. However, in our setting, we allow for a finite-horizon problem with a moving safe level, which requires some additional care in the verification argument.

Before we state the verification theorem, we first provide an alternative representation of the minimal risk function p in terms of the survival function S, which we will use in the proof of the verification theorem.

Lemma 6 *The minimal risk function p can be expressed in terms of the survival function S as*

$$p(t, x; C) = \inf_{\pi \in \mathcal{A}(t,x)} \mathbb{E}_{t,x}\left[\frac{S(\tau_{\text{imp}})}{S(t)}\right].$$

Proof By definition, we have

$$p(t, x; C) = \inf_{\pi \in \mathcal{A}(t,x)} \mathbb{P}_{t,x}\left(\tau_{\text{imp}} < \mathcal{T} \,\middle|\, \mathcal{T} > t\right) = \inf_{\pi \in \mathcal{A}(t,x)} \frac{\mathbb{P}_{t,x}\left(\tau_{\text{imp}} < \mathcal{T}, \mathcal{T} > t\right)}{\mathbb{P}_{t,x}\left(\mathcal{T} > t\right)}.$$

Now, since $\tau_{\text{imp}} \geq t$ almost surely, we have $\{\tau_{\text{imp}} < \mathcal{T}, \mathcal{T} > t\} = \{\tau_{\text{imp}} < \mathcal{T}\}$, and by the definition of the survival function S, it follows that

$$p(t, x; C) = \inf_{\pi \in \mathcal{A}(t,x)} \frac{\mathbb{P}_{t,x}\left(\tau_{\text{imp}} < \mathcal{T}\right)}{S(t)}.$$

Finally, using the independence of τ_{imp} and $\mathcal{T}$, we can express the probability in terms of the expectation of the survival function, which yields

$$\mathbb{P}_{t,x}\left(\tau_{\text{imp}} < \mathcal{T}\right) = \mathbb{E}_{t,x}\left[\mathbb{P}_{t,x}\left(\tau_{\text{imp}} < \mathcal{T} \,\middle|\, \tau_{\text{imp}}\right)\right] = \mathbb{E}_{t,x}\left[S(\tau_{\text{imp}})\right],$$

which completes the proof. □

Now, we are ready to state the verification theorem for the minimal risk function p.

Theorem 7 (A verification theorem for the minimal risk function) *Let $v : \overline{\mathcal{D}}^x \to [0,1]$ be of class $C^{1,2}(\mathcal{D}) \cap C(\overline{\mathcal{D}}^x)$. Additionally, suppose that*

$$v_t(t,x) + \inf_{u\in K} \mathcal{L}^u v(t,x) - \lambda(t)v(t,x) = 0, \quad (t,x) \in \mathcal{D}, \tag{9a}$$

$$v(t, x_{\text{imp}}) = 1, \quad 0 \le t < T, \tag{9b}$$

$$v(t, x_{\text{safe}}(t)) = 0, \quad 0 \le t < T, \tag{9c}$$

that is, v solves the HJB equation (7) with the appropriate boundary conditions, and that there exists a measurable selector $\psi : \mathcal{D} \to K$ such that

$$\psi(t,x) \in \arg\min_{u\in K} \mathcal{L}^u v(t,x)$$

for all $(t,x) \in \mathcal{D}$, and that the corresponding feedback rule

$$\varpi(t,x) := \begin{cases} \psi(t,x), & (t,x) \in \mathcal{D}, \\ 0, & \text{otherwise}, \end{cases}$$

induces an admissible feedback control $\varpi_s := \varpi(s, X_s^{\varpi})$. Then

$$v(t,x) = p(t,x;C), \quad (t,x) \in \mathcal{D},$$

and the above feedback strategy is optimal.

Proof Fix $(t,x) \in \mathcal{D}$ and $\pi \in \mathcal{A}_K(t,x)$, and denote by X^π the corresponding wealth process. Now, let $\varepsilon > 0$ be sufficiently small, and define

$$\mathcal{D}_\varepsilon := \left\{(s,y) : 0 \le s \le \min(T-\varepsilon, \varepsilon^{-1}), x_{\text{imp}} + \varepsilon \le y \le x_{\text{safe}}(s) - \varepsilon\right\}.$$

Then, $\mathcal{D}_\varepsilon$ is a compact subset of $\mathcal{D}$, and we have $\mathcal{D}_\varepsilon \uparrow \mathcal{D}$ as $\varepsilon \downarrow 0$. Moreover, since $v \in C^{1,2}(\mathcal{D})$, we have $v \in C^{1,2}(\overline{\mathcal{D}_\varepsilon})$ for each $\varepsilon > 0$. Thus, with the stopping time

$$\tau_\varepsilon := \inf\{s \ge t : (s, X_s^\pi) \notin \mathcal{D}_\varepsilon\}$$

and the function

$$\tilde{v}(s,x) := \frac{S(s)}{S(t)} v(s,x),$$

we can apply Itô's formula to the process $\tilde{v}(s \wedge \tau_\varepsilon, X^\pi_{s\wedge\tau_\varepsilon})$ for $s \ge t$, which yields

$$\tilde{v}(s\wedge\tau_\varepsilon, X^\pi_{s\wedge\tau_\varepsilon}) - \tilde{v}(t,x) = \int_t^{s\wedge\tau_\varepsilon} \frac{S(u)}{S(t)} \Big(\underbrace{v_t(u,X_u^\pi) + \mathcal{L}^{\pi_u} v(u,X_u^\pi) - \lambda(u)v(u,X_u^\pi)}_{=:\,\mathrm{H}[v](u)} \Big)\, \mathrm{d}u$$

$$+ \int_t^{s\wedge\tau_\varepsilon} \tilde{v}_x(u, X_u^\pi) X_u^\pi \pi_u^\top \sigma \, \mathrm{d}B_u.$$

Since π is admissible, and $\mathcal{D}_\varepsilon$ is compact, we have that

$$\mathbb{E}_{t,x}\left[\int_t^{s\wedge\tau_\varepsilon} \tilde{v}_x(u,X_u^\pi)^2 (X_u^\pi)^2 \pi_u^\top \Sigma \pi_u \, \mathrm{d}u\right] \le \|\tilde{v}_x\|^2_{L^\infty(\overline{\mathcal{D}_\varepsilon})} \mathbb{E}_{t,x}\left[\int_t^{s\wedge\tau_\varepsilon} (X_u^\pi)^2 \pi_u^\top \Sigma \pi_u \, \mathrm{d}u\right] < \infty.$$

Hence, the stopped stochastic integral above is a martingale, and taking expectations on both sides yields

$$\mathbb{E}_{t,x}\left[\tilde{v}(s\wedge\tau_\varepsilon, X^\pi_{s\wedge\tau_\varepsilon})\right] - \tilde{v}(t,x) = \mathbb{E}_{t,x}\left[\int_t^{s\wedge\tau_\varepsilon} \frac{S(u)}{S(t)} \mathrm{H}[v](u)\, \mathrm{d}u\right]. \tag{10}$$

Now, since v satisfies the HJB (9a), we have

$$\mathrm{H}[v](u) = v_t(u,X_u^\pi) + \mathcal{L}^{\pi_u} v(u,X_u^\pi) - \lambda(u)v(u,X_u^\pi) \ge 0, \tag{11}$$

and therefore

$$\tilde{v}(t,x) \leq \mathbb{E}_{t,x}\left[\tilde{v}(s \wedge \tau_\varepsilon, X^\pi_{s\wedge\tau_\varepsilon})\right].$$

Since $X^\pi_{s\wedge\tau_\varepsilon}$ has continuous paths, and $\mathcal{D}_\varepsilon \uparrow \mathcal{D}$ as $\varepsilon \downarrow 0$, we have that $\tau_\varepsilon \uparrow \tau_\mathcal{D}$ as $\varepsilon \downarrow 0$, and, in particular, $X^\pi_{s\wedge\tau_\varepsilon} \to X^\pi_{s\wedge\tau_\mathcal{D}}$ almost surely. Hence, by continuity and the dominated convergence theorem, we can take the limit $\varepsilon \downarrow 0$ to obtain

$$\tilde{v}(t,x) \leq \mathbb{E}_{t,x}\left[\tilde{v}(s \wedge \tau_\mathcal{D}, X^\pi_{s\wedge\tau_\mathcal{D}})\right].$$

Now, we split the expectation into two parts, depending on whether $\tau_\mathcal{D} \leq s$ or $\tau_\mathcal{D} > s$. That is,

$$\begin{aligned}\tilde{v}(t,x) &\leq \mathbb{E}_{t,x}\left[\tilde{v}(s \wedge \tau_\mathcal{D}, X^\pi_{s\wedge\tau_\mathcal{D}})\mathbf{1}_{\{\tau_\mathcal{D}\leq s\}}\right] + \mathbb{E}_{t,x}\left[\tilde{v}(s \wedge \tau_\mathcal{D}, X^\pi_{s\wedge\tau_\mathcal{D}})\mathbf{1}_{\{\tau_\mathcal{D}> s\}}\right] \\ &= \mathbb{E}_{t,x}\left[\tilde{v}(\tau_\mathcal{D}, X^\pi_{\tau_\mathcal{D}})\mathbf{1}_{\{\tau_\mathcal{D}\leq s\}}\right] + \mathbb{E}_{t,x}\left[\tilde{v}(s, X^\pi_s)\mathbf{1}_{\{\tau_\mathcal{D}> s\}}\right].\end{aligned}$$

In the first expectation, we can use the boundary conditions (9b) and (9c) to obtain

$$\mathbb{E}_{t,x}\left[\tilde{v}(\tau_\mathcal{D}, X^\pi_{\tau_\mathcal{D}})\mathbf{1}_{\{\tau_\mathcal{D}\leq s\}}\right] = \mathbb{E}_{t,x}\left[\frac{S(\tau_{\text{imp}})}{S(t)}\mathbf{1}_{\{\tau_{\text{imp}}\leq s, \tau_{\text{imp}}<\tau_{\text{safe}}\}}\right],$$

hence

$$S(t)v(t,x) \leq \mathbb{E}_{t,x}\left[S(\tau_{\text{imp}})\mathbf{1}_{\{\tau_{\text{imp}}\leq s, \tau_{\text{imp}}<\tau_{\text{safe}}\}}\right] + S(s)\mathbb{E}_{t,x}\left[v(s, X^\pi_s)\mathbf{1}_{\{\tau_\mathcal{D}> s\}}\right].$$

Finally, taking the limit $s \uparrow T$, and using the fact that $S(s) \to 0$ as $s \uparrow T$, we conclude that

$$S(t)v(t,x) \leq \mathbb{E}_{t,x}\left[S(\tau_{\text{imp}})\mathbf{1}_{\{\tau_{\text{imp}}<\tau_{\text{safe}}\}}\right] \leq \mathbb{E}_{t,x}\left[S(\tau_{\text{imp}})\right].$$

Since $\pi \in \mathcal{A}_K(t,x)$ was arbitrary, we can take the infimum over all admissible policies and use Lemma 6 to conclude that

$$v(t,x) \leq p(t,x;C), \quad (t,x) \in \mathcal{D}.$$

Now, suppose that there exist ψ and ϖ as in the statement of the theorem. Then, by construction, we have that (11) holds with equality for ϖ, and therefore, by (10) and the same arguments as above, we obtain

$$v(t,x) = \mathbb{E}_{t,x}\left[\frac{S(\tau_{\text{imp}})}{S(t)}\mathbf{1}_{\{\tau_{\text{imp}}<\tau_{\text{safe}}\}}\right].$$

On $\{\tau_{\text{safe}} \leq \tau_{\text{imp}}\}$, the zero-risk continuation of ϖ leads to the wealth process following x_{safe}. Hence, $\tau_{\text{imp}} = T$ when $T < \infty$, and $\tau_{\text{imp}} = \infty$ when $T = \infty$. In either case $S(\tau_{\text{imp}}) = 0$, and therefore

$$\frac{S(\tau_{\text{imp}})}{S(t)}\mathbf{1}_{\{\tau_{\text{imp}}\geq\tau_{\text{safe}}\}} = \frac{S(T)}{S(t)}\mathbf{1}_{\{\tau_{\text{imp}}\geq\tau_{\text{safe}}\}} = 0.$$

Consequently, we can drop the indicator function in the expectation and obtain

$$v(t,x) = \mathbb{E}_{t,x}\left[\frac{S(\tau_{\text{imp}})}{S(t)}\right].$$

Since $\varpi \in \mathcal{A}_K(t,x)$, we can again take the infimum over all admissible policies to conclude that

$$p(t,x;C) \leq v(t,x), \quad (t,x) \in \mathcal{D}.$$

Combining the two inequalities, we have established that

$$v(t,x) = p(t,x;C), \quad (t,x) \in \mathcal{D},$$

with an optimal strategy ϖ. □

Remark 2 Due to the restriction to survival functions that vanish continuously towards the terminal time T, we can avoid the need for a so-called transversality condition in the verification theorem, which is typically required in infinite-horizon problems [34]. If the survival function does not vanish, the necessary transversality condition can be expressed as

$$\forall \pi \in \mathcal{A}_K(t,x) : \lim_{t \leq s \uparrow T} \mathbb{E}_{t,x}\left[v(s, X_s^\pi)\mathbf{1}_{s<\tau_\mathcal{D}}\right] = 0,$$

Using Lemma 4, we can also derive an explicit expression for the optimal feedback strategy in terms of the candidate solution v to the HJB equation.

Corollary 8 *If the candidate solution v, in addition to the assumptions of Theorem 7, satisfies $v_x < 0$ and $v_{xx} > 0$ in $\mathcal{D}$, then the optimal feedback strategy can be expressed in closed form as*

$$\varpi(t,x) := \begin{cases} u^*\left(\frac{x v_{xx}(t,x)}{v_x(t,x)}; K\right), & (t,x) \in \mathcal{D}, \\ 0, & \text{otherwise.} \end{cases}$$

Moreover, if $K = K_{\text{unc}}$, the HJB equation (9a) reduces to the PDE

$$v_t(t,x) + (rx - C)v_x(t,x) - \rho\frac{(v_x(t,x))^2}{v_{xx}(t,x)} - \lambda(t)v(t,x) = 0, \quad (t,x) \in \mathcal{D}, \tag{12}$$

which turns into

$$v_t(t,z) + \mathcal{E}(t)zv_z(t,z) - \rho\frac{(v_z(t,z))^2}{v_{zz}(t,z)} - \lambda(t)v(t,z) = 0, \quad (t,z) \in \mathcal{D}^z, \tag{13}$$

in the normalized safe-gap coordinate z.

Proof The representation of the optimal feedback strategy follows directly from the results established in Lemma 4. The reduced form of the HJB equation in the unconstrained case follows from substituting the unconstrained optimal control $u^*_{\text{unc}}(\kappa) = -\kappa^{-1}\mathfrak{m}$ into (9a) and using the definition of ρ. Thus, it remains to show that the HJB equation in the normalized safe-gap coordinate z has the form (13). To this end, we let $(t,x) \in \mathcal{D}$, and note that

$$v_x(t,x) = v_z(t,z)\frac{\partial z}{\partial x}(t,x) = -\frac{v_z(t,z)}{x_{\text{safe}}(t) - x_{\text{imp}}}$$

and

$$v_{xx}(t,x) = v_{zz}(t,z)\left(\frac{\partial z}{\partial x}(t,x)\right)^2 = \frac{v_{zz}(t,z)}{(x_{\text{safe}}(t) - x_{\text{imp}})^2}.$$

Thus, by denoting $\ell(t) := x_{\text{safe}}(t) - x_{\text{imp}}$ and using $x(t,z) = x_{\text{safe}}(t) - z\ell(t)$, one obtains

$$\frac{x v_{xx}(t,x)}{v_x(t,x)} = -\frac{(x_{\text{safe}}(t) - z\ell(t))v_{zz}(t,z)}{\ell(t)v_z(t,z)}.$$

Hence,

$$a(t,z)\varpi(t,z) = \left(1 - z + \frac{x_{\text{imp}}}{\ell(t)}\right)\frac{\ell(t)v_z(t,z)}{(x_{\text{safe}}(t) - z\ell(t))v_{zz}(t,z)}\mathfrak{m} = \frac{v_z(t,z)}{v_{zz}(t,z)}\mathfrak{m}.$$

Finally, substituting this expression into (8) yields the desired form of the HJB equation in the normalized safe-gap coordinate z. □

With the above results, we can now establish a necessary and sufficient condition under which the constraints on the control variable u are redundant, and the optimal control in the constrained case coincides with the unconstrained optimal control.

Theorem 9 (Redundancy of constraints) *Assume that $\mathfrak{m} \geq 0$ componentwise, and denote by p_{unc} and p_{con} the minimal risk functions corresponding to the unconstrained and constrained cases, respectively. Further, suppose that p_{unc} satisfies the conditions of the verification theorem for $K = K_{\mathrm{unc}}$, and that $\partial_x p_{\mathrm{unc}} < 0$ and $\partial_{xx} p_{\mathrm{unc}} > 0$ in $\mathcal{D}$. Then, if and only if*

$$\frac{x\partial_{xx} p_{\mathrm{unc}}(t,x;C)}{\partial_x p_{\mathrm{unc}}(t,x;C)} \leq -\|\mathfrak{m}\|_1, \tag{14}$$

for all $(t,x) \in \mathcal{D}$, we have $p_{\mathrm{con}} = p_{\mathrm{unc}}$.

Proof Set

$$\kappa(t,x) := \frac{x\partial_{xx} p_{\mathrm{unc}}(t,x;C)}{\partial_x p_{\mathrm{unc}}(t,x;C)} < 0.$$

Suppose first that (14) holds. Corollary 5 then gives

$$u^*\big(\kappa(t,x); K_{\mathrm{con}}\big) = u^*_{\mathrm{unc}}\big(\kappa(t,x)\big)$$

throughout $\mathcal{D}$. Thus, p_{unc} also satisfies the constrained HJB equation, and its optimal feedback is simplex-valued and admissible. Theorem 7 therefore yields

$$p_{\mathrm{con}} = p_{\mathrm{unc}}.$$

Conversely, suppose that

$$p_{\mathrm{con}} = p_{\mathrm{unc}} := p$$

on $\mathcal{D}$. Since the common value function is of class $C^{1,2}(\mathcal{D})$, the dynamic-programming principle implies that it satisfies both the constrained and the unconstrained HJB equations. Hence, their optimized Hamiltonians coincide at every $(t,x) \in \mathcal{D}$.

Since $K_{\mathrm{con}} \subseteq K_{\mathrm{unc}}$ and the corresponding Hamiltonian optimizers are unique by Lemma 4, equality of the optimized Hamiltonians implies

$$u^*\big(\kappa(t,x); K_{\mathrm{con}}\big) = u^*_{\mathrm{unc}}\big(\kappa(t,x)\big).$$

Applying Corollary 5 gives

$$\kappa(t,x) \leq -\|\mathfrak{m}\|_1,$$

which is precisely (14). □

Summarizing the results of this section, we have established a verification theorem for the minimal risk function p and derived an explicit representation of the optimal feedback strategy in terms of the candidate solution to the HJB equation. Moreover, we have provided a necessary and sufficient condition under which the constraints on the control variable u are redundant, and the optimal control in the constrained case coincides with the unconstrained optimal control. In the next section, we will apply these results to derive explicit solutions for the minimal risk function and the corresponding optimal control in both the unconstrained and constrained cases.

3 Inverse-solvable mortality families

In this section, we present a new method for constructing mortality laws that are compatible with closed-form solutions to the optimal pension decumulation problem. Motivated by the results of Theorem 9, we focus on the unconstrained case. To this end, we only consider $K = K_{\text{unc}}$ for the remainder of this section. As Moore and Young [26] have — under the usual smoothness conditions — shown, the value function p must be convex and non-increasing in wealth. For this reason, and in view of the assumptions posed in the verification theorem, we will assume throughout this section that the candidate value function satisfies $v_z > 0$ and $v_{zz} > 0$ on the risky region $\mathcal{D}^z$.

We begin by recalling the normalized HJB PDE (13) for candidates v of the value function p in the moving coordinate z from Corollary 8:

$$\begin{aligned} v_t(t,z) + \mathcal{E}(t) z v_z(t,z) - \rho \frac{(v_z(t,z))^2}{v_{zz}(t,z)} - \lambda(t) v(t,z) = 0, \quad (t,z) \in \mathcal{D}^z, \\ v(t,0) = 0, \quad 0 \le t < T, \\ v(t,1) = 1, \quad 0 \le t < T, \end{aligned}$$

For constant hazard rates $\lambda(t) \equiv \lambda > 0$, Young [34], Bayraktar and Young [2], Moore and Young [26], and Young and Zhang [35] have shown that the above problem admits a closed-form solution, allowing for the derivation of a wide range of explicit results. To later compare with these previous works, we first collect two central results for the constant-hazard case by stating them without proof. To this end, note that when the hazard rate is constant, we have $T = \infty$, and thus $\mathcal{E}(t) \equiv r$ and $x_{\text{safe}}(t) \equiv x^{\infty}_{\text{safe}}$ for all $t \ge 0$.

Theorem 10 *Let $\lambda > 0$, and let $\delta > 1$ be the unique solution to the quadratic equation*

$$r\delta^2 - (\rho + r + \lambda)\delta + \lambda = 0.$$

Then the value function on the risky region is given by

$$p(t,x) = \left(\frac{x^{\infty}_{\text{safe}} - x}{x^{\infty}_{\text{safe}} - x_{\text{imp}}} \right)^{\delta}, \quad (t,x) \in \overline{\mathcal{D}}^x,$$

and the corresponding optimal risky-fraction strategy is given by the feedback rule

$$\pi^*(t,x) = \frac{x^{\infty}_{\text{safe}} - x}{x(\delta - 1)} \mathfrak{m}, \quad (t,x) \in \overline{\mathcal{D}}^x.$$

Theorem 11 *Under the optimal strategy π^* from Theorem 10, the safe wealth level x^{∞}_{safe} is unreachable in finite time, that is,*

$$\mathbb{P}(\tau_{\text{safe}} < \infty) = 0.$$

Moreover, the financial impoverishment time τ_{imp}, given that impoverishment occurs before death, has an inverse-Gaussian distribution,

$$\tau_{\text{imp}} \mid \tau_{\text{imp}} < \mathcal{T} \sim \text{IG}\left(\frac{q_0}{r + \rho/(\delta-1)^2}, \frac{q_0^2 (\delta-1)^2}{2\rho} \right),$$

where $q_0 = \ln(x^{\infty}_{\text{safe}} - x_{\text{imp}}) - \ln(x^{\infty}_{\text{safe}} - x_0)$.

A direct consequence of Theorem 10 is the following corollary, which provides a closed-form expression for the optimal withdrawal rate C^* under the constant-hazard model.

Corollary 12 (Sustainable withdrawal and required initial capital) *Fix $x_0 > x_{\mathrm{imp}}$ and $\alpha \in (0,1)$, and set*

$$\beta := \frac{\alpha^{1/\delta}}{1-\alpha^{1/\delta}}.$$

Then, under the constant-hazard model, the optimal sustainable withdrawal rate defined in (4) is

$$\begin{aligned} C^*(x_0;\alpha) &= C_{\mathrm{safe}}(x_0) + r\beta(x_0 - x_{\mathrm{imp}}) \\ &= r\big(x_{\mathrm{imp}} + (1+\beta)(x_0 - x_{\mathrm{imp}})\big). \end{aligned} \tag{15}$$

Conversely, for any withdrawal rate $C > r x_{\mathrm{imp}}$, the minimal initial wealth required to achieve a lifetime-impoverishment probability of at most α is

$$\begin{aligned} x_0(C;\alpha) &= x^{\infty}_{\mathrm{safe}} - \frac{\beta}{\beta+1}\big(x^{\infty}_{\mathrm{safe}} - x_{\mathrm{imp}}\big) \\ &= x_{\mathrm{imp}} + \frac{1}{\beta+1}\left(\frac{C}{r} - x_{\mathrm{imp}}\right), \end{aligned} \tag{16}$$

where $x^{\infty}_{\mathrm{safe}} = C/r$ is evaluated at the target withdrawal rate C.

Remark 3 Equations (15) and (16) immediately show the trade-off between risk-taking and investment gains. If $\alpha = 0$, then $\beta = 0$ and both expressions reduce to the optimal amounts achievable under riskless-only investment: rx_0 and $x^{\infty}_{\mathrm{safe}}$. As α increases, β also increases, representing the fact that the investor is willing to accept a higher probability of lifetime ruin, which in turn allows for a higher withdrawal rate C^* or a lower required initial capital x_0. Also notice that, due to the exponent $1/\delta$ in the definition of β, close to $\alpha = 0$, C^* increases significantly for small increases in α. In particular, the derivative of C^* with respect to α is unbounded as $\alpha \downarrow 0$.

For general time-dependent hazard rates λ, closed-form solutions to the normalized HJB PDE (13) are not available in general. However, Moore and Young [26] and Liang and Young [20] have shown that the HJB PDE can be transformed into a linear free boundary problem via a Legendre transformation, which allows for the derivation of certain qualitative properties of the value function and the optimal strategy. In particular, they have shown that value functions corresponding to different hazard rates can be ordered according to the reverse ordering of the hazard rates themselves. More precisely, they have shown the following hazard-comparison theorem.

Theorem 13 *Let $\underline{\lambda}, \lambda, \overline{\lambda} : [0,T) \to [0,\infty)$ be three hazard rates such that*

$$\underline{\lambda}(t) \le \lambda(t) \le \overline{\lambda}(t)$$

for all $t \in [0,T)$, and denote by $\underline{p}, p, \overline{p}$ the corresponding value functions. Then, for all $(t,x) \in \overline{\mathcal{D}}^x$, we have

$$\underline{p}(t,x) \ge p(t,x) \ge \overline{p}(t,x).$$

Theorem 13 shows that an explicitly solvable mortality law can also serve as a benchmark for other mortality models. If the target hazard rate can be bounded above and below by hazard rates with known value functions, the comparison principle yields explicit lower and upper bounds on its value function. The classical constant-hazard model therefore provides two-sided bounds only when suitable finite constant envelopes exist. This fails for unbounded hazard rates commonly used in practice: the De Moivre hazard rate diverges as $t \uparrow T < \infty$, whereas the Gompertz–Makeham hazard rate diverges over its unbounded time horizon. In these cases, a constant lower envelope may still provide an upper bound on the value function, but no finite constant upper envelope, and hence no corresponding constant-hazard lower bound, is available.

This one-sided limitation motivates us to enlarge the class of mortality laws for which the HJB admits an analytically tractable solution. However, rather than prescribing the hazard rate and then solving the resulting time-inhomogeneous HJB, we reverse the usual order of construction. We seek candidate value functions of the self-similar form

$$v(t,z) = \frac{\psi(b(t)z)}{\psi(b(t))}, \quad (t,z) \in \mathcal{D}^z,$$

and derive the compatibility conditions linking the self-similar profile ψ, the scaling function b, and the hazard rate λ. Solving these conditions in the inverse direction identifies mortality laws for which both the value function and the optimal feedback admit explicit or semi-explicit representations.

3.1 The compatibility equation

Throughout this section, we will use the notation $\overline{b} := \sup_{t\in[0,T)} b(t)$ for the supremum of the scaling function b, which we assume to be strictly positive and continuously differentiable on $[0,T)$. Further, we suppose that ψ is strictly positive and twice continuously differentiable on $(0,\overline{b})$, and that, if $\overline{b} < \infty$, ψ and ψ' admit finite continuous extensions to the upper boundary of their domain. Some basic properties of ψ follow immediately.

Lemma 14 *It holds that ψ' and ψ'' are strictly positive on $(0,\overline{b})$. Moreover,*

$$\psi(0) := \lim_{y\downarrow 0} \psi(y) = 0,$$

and, if $\overline{b} < \infty$, then

$$\psi(\overline{b}) := \lim_{y\uparrow \overline{b}} \psi(y) > 0, \qquad \psi'(\overline{b}) := \lim_{y\uparrow \overline{b}} \psi'(y) > 0.$$

Proof Recall that we assumed that $v_z, v_{zz}, b, \psi > 0$ on their respective domains. Computing the first two derivatives of v with respect to z yields

$$v_z(t,z) = \frac{b(t)\psi'(b(t)z)}{\psi(b(t))}, \qquad v_{zz}(t,z) = \frac{b(t)^2\psi''(b(t)z)}{\psi(b(t))}.$$

Now, let $y \in (0, \bar{b})$ be arbitrary. Then, there exists $t \in [0, T)$ such that $b(t) > y$. Setting $z = y/b(t) \in (0, 1)$ in the above expressions shows that $\psi'(y), \psi''(y) > 0$. The boundary condition $v(t, 0) = 0$ for all $t \in [0, T)$ implies that

$$\psi(0) = \lim_{y \downarrow 0} \psi(y) = \lim_{z \downarrow 0} v(t, z)\psi(b(t)) = 0.$$

Now assume that $\bar{b} < \infty$. Then, since $\psi' > 0$ on $(0, \bar{b})$, we have $\psi(y) > \psi(0) = 0$ for all $y \in (0, \bar{b})$. Thus,

$$\psi(\bar{b}) = \lim_{y \uparrow \bar{b}} \psi(y) \geq \psi(y_0) > 0$$

for any $y_0 \in (0, \bar{b})$. The same argument shows that $\psi'(\bar{b}) > 0$. □

With the above assumptions in place, we can now derive the compatibility equation. To this end, we note that, since v is strictly positive inside the risky region, we can divide the normalized HJB PDE (13) by v to obtain

$$\frac{v_t(t, z)}{v(t, z)} + \mathcal{E}(t)\frac{z v_z(t, z)}{v(t, z)} - \rho \frac{(v_z(t, z))^2}{v_{zz}(t, z)v(t, z)} - \lambda(t) = 0, \quad (t, z) \in \mathcal{D}^z.$$

Using the self-similar ansatz $v(t, z) = \psi(b(t)z)/\psi(b(t))$, we compute the quotients appearing in the above PDE:

$$\begin{aligned}
\frac{v_t(t, z)}{v(t, z)} &= \frac{b'(t)}{b(t)} \left(\frac{b(t) z \psi'(b(t)z)}{\psi(b(t)z)} - \frac{b(t)\psi'(b(t))}{\psi(b(t))} \right), \\
\frac{z v_z(t, z)}{v(t, z)} &= \frac{b(t) z \psi'(b(t)z)}{\psi(b(t)z)}, \\
\frac{(v_z(t, z))^2}{v_{zz}(t, z)v(t, z)} &= \frac{\psi'(b(t)z)^2}{\psi(b(t)z)\psi''(b(t)z)}.
\end{aligned}$$

Hence, the PDE for v reduces to

$$\left(\frac{b'(t)}{b(t)} + \mathcal{E}(t) \right) \frac{b(t) z \psi'(b(t)z)}{\psi(b(t)z)} - \rho \frac{\psi'(b(t)z)^2}{\psi(b(t)z)\psi''(b(t)z)} = \lambda(t) + \frac{b'(t)}{b(t)} \cdot \frac{b(t)\psi'(b(t))}{\psi(b(t))}. \quad (17)$$

One further simplification is possible by introducing the function

$$\Upsilon(y) := \frac{y\psi'(y)}{\psi(y)}, \quad y \in (0, \bar{b}),$$

which we call the elasticity profile of ψ, or simply its elasticity. As with ψ, we can prove some basic properties of Υ.

Lemma 15 *Let Υ be defined as above. Then, $\Upsilon > 1$ on $(0, \bar{b})$, and, if $\bar{b} < \infty$, then*

$$\Upsilon(\bar{b}) := \lim_{y \uparrow \bar{b}} \Upsilon(y) = \frac{\bar{b}\psi'(\bar{b})}{\psi(\bar{b})} \in (1, \infty).$$

Proof By Lemma 14, we have $\psi(0) = 0$. Moreover, since $\psi'' > 0$ on $(0, \bar{b})$, we have that ψ' is strictly increasing. Now, let $y \in (0, \bar{b})$ be arbitrary. Then,

$$\psi(y) = \int_0^y \psi'(\upsilon)\,\mathrm{d}\upsilon < \int_0^y \psi'(y)\,\mathrm{d}\upsilon = y\psi'(y),$$

which shows that $\Upsilon(y) > 1$. Now assume that $\bar{b} < \infty$. Then, the above argument also works for $y = \bar{b}$, that is,

$$\psi(\bar{b}) = \int_0^{\bar{b}} \psi'(\upsilon)\,\mathrm{d}\upsilon < \int_0^{\bar{b}} \psi'(\bar{b})\,\mathrm{d}\upsilon = \bar{b}\psi'(\bar{b}).$$

Since $\psi(\bar{b}) > 0$ by Lemma 14, we can divide both sides by $\psi(\bar{b})$ to obtain $\Upsilon(\bar{b}) > 1$. □

Using some elementary algebra, we can rewrite the PDE for v in terms of Υ. In particular, by defining the operators

$$\frac{(\psi'(y))^2}{\psi(y)\psi''(y)} = \frac{\Upsilon(y)^2}{y\Upsilon'(y) + \Upsilon(y)^2 - \Upsilon(y)} =: \mathcal{N}[\Upsilon](y) > 0 \tag{18}$$

and

$$\lambda(t) + \frac{b'(t)}{b(t)}\Upsilon(b(t)) =: F[\Upsilon](t), \tag{19}$$

we can write (17) more compactly as

$$\left(\frac{b'(t)}{b(t)} + \mathcal{E}(t)\right)\Upsilon(b(t)z) - \rho\mathcal{N}[\Upsilon](b(t)z) = F[\Upsilon](t). \tag{20}$$

Thus, we have reduced the PDE for v to a compatibility equation for Υ and b, with the right-hand side $F[\Upsilon](t)$ depending only on t. With this, we can now define the notion of a compatible triple.

Definition 10 (Compatible triple) Let b, ψ, and Υ satisfy the standing assumptions and definitions above, and let λ be the hazard rate of a mortality model satisfying Definition 5, with right endpoint T. The triple $(b, \Upsilon; \lambda)$ is called a *compatible triple* if (20) holds for every $(t, z) \in \mathcal{D}^z$.

Using the notion of a compatible triple, we can now show that the candidate value function v can be reconstructed from Υ, without requiring the explicit knowledge of ψ.

Proposition 16 (Reconstruction of the candidate value function) *Let $(b, \Upsilon; \lambda)$ be a compatible triple. Then v can be reconstructed from Υ via the formula*

$$v(t, z) = \exp\left(-\int_z^1 \frac{\Upsilon(b(t)\zeta)}{\zeta}\,\mathrm{d}\zeta\right) \tag{21}$$

for all $(t, z) \in \overline{\mathcal{D}}^z$.

Proof Let $(t,z) \in \mathcal{D}^z$. By the definition of Υ, we have

$$\frac{\psi'(y)}{\psi(y)} = \frac{\Upsilon(y)}{y}, \quad y \in (0,\bar{b}),$$

and integrating both sides from $b(t)z$ to $b(t)$ yields

$$\ln(\psi(b(t))) - \ln(\psi(b(t)z)) = \int_{b(t)z}^{b(t)} \frac{\Upsilon(y)}{y}\,\mathrm{d}y = \int_z^1 \frac{\Upsilon(b(t)\zeta)}{\zeta}\,\mathrm{d}\zeta,$$

where the last equality follows from the change of variables $y \mapsto b(t)\zeta$. Exponentiating both sides and using the definition of v yields the reconstruction formula (21) for $(t,z) \in \mathcal{D}^z$. The equality extends to the spatial closure $\overline{\mathcal{D}}^z$ by using $\Upsilon > 1$, implying the divergence of the integral as $z \downarrow 0$, and the usual convention $\exp(-\infty) = 0$. □

Despite being of much simpler form than the original PDE, the compatibility equation (20) is still a functional equation in two variables. However, as the following theorem shows, the compatibility equation can only hold if b and Υ are of a very specific form. To shorten notation, we define

$$A(t) := \frac{b'(t)}{b(t)} + \mathcal{E}(t), \quad t \in [0,T).$$

Theorem 17 *Let $(b,\Upsilon;\lambda)$ be a compatible triple. Then either there exists $\Upsilon_0 > 1$ such that $\Upsilon \equiv \Upsilon_0$ on $(0,\bar{b})$, or Υ is non-constant and there exist unique constants $A_0, F_0 \in \mathbb{R}$ such that $A \equiv A_0$ and $F[\Upsilon] \equiv F_0$ on $[0,T)$.*

Proof For arbitrary $t_1, t_2 \in [0,T)$ and $0 < y < \min\{b(t_1), b(t_2)\}$, we may evaluate (20) at $(t_i, y/b(t_i))$ for $i = 1,2$. Subtracting the resulting identities gives

$$\big(A(t_1) - A(t_2)\big)\Upsilon(y) = F[\Upsilon](t_1) - F[\Upsilon](t_2). \tag{22}$$

Since y can be chosen arbitrarily in the interval $(0, \min\{b(t_1), b(t_2)\})$, we conclude that either Υ is constant on this interval or $A(t_1) = A(t_2)$ and $F[\Upsilon](t_1) = F[\Upsilon](t_2)$. Motivated by this observation, we define

$$\hat{y} := \sup\left\{ y \in (0,\bar{b}) : \Upsilon \text{ is constant on } (0,y) \right\},$$

with $\sup\emptyset = 0$. We claim that $\hat{y} \in \{0, \bar{b}\}$. Suppose, to the contrary, that $0 < \hat{y} < \bar{b}$. Then there exists $\Upsilon_0 > 1$ such that $\Upsilon(y) = \Upsilon_0$ for all $y \in (0,\hat{y})$. Moreover, since $\hat{y} < \bar{b}$, there exists t_0 such that $b(t_0) > \hat{y}$. Evaluating (20) at $(t_0, y/b(t_0))$ for an arbitrary $y \in (0, b(t_0))$ shows that

$$A(t_0)\Upsilon(y) - \rho\frac{\Upsilon(y)^2}{y\Upsilon'(y) + \Upsilon(y)^2 - \Upsilon(y)} = F[\Upsilon](t_0). \tag{23}$$

Setting $A_0 := A(t_0)$ and $F_0 := F[\Upsilon](t_0)$, this reduces to

$$A_0\Upsilon_0 - F_0 = \rho\frac{\Upsilon_0}{\Upsilon_0 - 1} > 0$$

on $(0,\hat{y}) \subsetneq (0, b(t_0))$. Thus, by continuity, $A_0\Upsilon(y) - F_0 \neq 0$ in a neighborhood of $\hat{y}$. Hence, in such a neighborhood, we can rearrange (23) to obtain

$$\Upsilon'(y) = \underbrace{\frac{\Upsilon(y)}{y}\left(\frac{\rho\Upsilon(y)}{A_0\Upsilon(y) - F_0} - \Upsilon(y) + 1\right)}_{=: f(y,\Upsilon(y))}.$$

Now, since f is smooth and locally Lipschitz in Υ near $(\hat{y}, \Upsilon_0)$, and since $\Upsilon(\hat{y}) = \Upsilon_0$, the Picard–Lindelöf theorem implies that Υ remains constant in a neighborhood to the right of $\hat{y}$. This contradicts the definition of $\hat{y}$. Therefore, we conclude that $\hat{y} \in \{0, \bar{b}\}$.

If $\hat{y} = \bar{b}$, then Υ is constant on $(0, \bar{b})$. If instead $\hat{y} = 0$, then for arbitrary $t_1, t_2 \in [0, T)$, we can choose $y_1, y_2 \in (0, \min\{b(t_1), b(t_2)\})$ such that $\Upsilon(y_1) \neq \Upsilon(y_2)$. Applying (22) at y_1 and y_2 and subtracting gives

$$\big(A(t_1) - A(t_2)\big) \underbrace{\big(\Upsilon(y_1) - \Upsilon(y_2)\big)}_{\neq 0} = 0,$$

hence $A(t_1) = A(t_2)$. Consequently, $A \equiv A_0$ on $[0, T)$, and we must have, again by (22), that $F[\Upsilon] \equiv F_0$ on $[0, T)$ as well, completing the proof. □

Theorem 17 yields a dichotomy among compatible triples. We call a compatible triple *constant-elasticity* if $\Upsilon \equiv \Upsilon_0$ is constant, and *non-constant-elasticity* otherwise. In the non-constant case, there exist unique constants $A_0, F_0 \in \mathbb{R}$ such that $A \equiv A_0$ and $F[\Upsilon] \equiv F_0$. Throughout the remainder of this section, A_0 and F_0 always denote these compatibility constants. The following results characterize the constant and non-constant classes more explicitly. Section 4 further studies the constant-elasticity case and its implications for the value function and optimal strategy.

A direct consequence of Theorem 17 is that, for non-constant-elasticity compatible triples $(b, \Upsilon; \lambda)$, b and Υ must satisfy two uncoupled first-order ODEs, allowing us to separately solve for b and Υ. In particular, the ODE for b can be solved explicitly, yielding a closed-form expression for the scaling function.

Corollary 18 *Let $(b, \Upsilon; \lambda)$ be a non-constant-elasticity compatible triple. Then, there exists a constant $c_b > 0$ such that*

$$b(t) = c_b \frac{\exp\left((A_0 - r)t\right)}{\mathcal{E}(t)}, \quad t \in [0, T). \tag{24}$$

Moreover, the hazard rate λ must then satisfy

$$\lambda(t) = F_0 + \big(\mathcal{E}(t) - A_0\big)\Upsilon(b(t)), \quad t \in [0, T). \tag{25}$$

Proof The first claim follows directly by solving the first-order ODE

$$b'/b + \mathcal{E} = A_0.$$

One easily verifies that the solution b indeed satisfies the assumed positivity and smoothness conditions. The second claim then follows from the definition of $F[\Upsilon]$ and the fact that $F[\Upsilon] \equiv F_0$. □

For every non-constant-elasticity compatible triple, we henceforth call c_b the associated scaling constant. Thus, A_0, F_0, and c_b always denote the uniquely determined constants appearing in (24) and (25). The right endpoint T is inherited from the mortality model associated with λ and will not be displayed separately in the notation for a compatible triple.

Conversely, if $(b, \Upsilon; \lambda)$ is a constant-elasticity compatible triple, then (20) reduces to an explicit expression for the hazard rate λ in terms of $\Upsilon \equiv \Upsilon_0$. Moreover, due to Proposition 16, the value function v is of a particularly simple form.

Corollary 19 *Let $(b, \Upsilon; \lambda)$ be a constant-elasticity compatible triple with $\Upsilon \equiv \Upsilon_0 > 1$. Then, the hazard rate λ must satisfy*

$$\lambda(t) = \left(\mathcal{E}(t) - \frac{\rho}{\Upsilon_0 - 1}\right)\Upsilon_0, \quad t \in [0, T),$$

and the corresponding candidate value function v is given by

$$v(t, z) = z^{\Upsilon_0}, \quad (t, z) \in \overline{\mathcal{D}}^z.$$

Proof If $\Upsilon \equiv \Upsilon_0 > 1$ is constant, then (20) reduces to

$$\left(\frac{b'(t)}{b(t)} + \mathcal{E}(t)\right)\Upsilon_0 - \rho\frac{\Upsilon_0}{\Upsilon_0 - 1} = \lambda(t) + \frac{b'(t)}{b(t)}\Upsilon_0.$$

Rearranging this expression yields the desired formula for λ. The reconstruction formula (21) then gives

$$v(t, z) = \exp\left(-\int_z^1 \frac{\Upsilon_0}{\zeta}\,\mathrm{d}\zeta\right) = z^{\Upsilon_0}, \quad (t, z) \in \overline{\mathcal{D}}^z.$$

□

Note that b drops out of the expression, so that the hazard rate is independent of the scaling function b. Hence, the constant-elasticity case forms a two-parameter family of mortality laws, parameterized by the pair (Υ_0, T).

Similarly, we can derive an explicit expression for the resulting candidate feedback rule ϖ in terms of Υ and b. To this end, recall that, given the value function candidate v, the candidate ϖ is given by

$$\varpi(t, z) = \frac{v_z(t, z)}{a(t, z)v_{zz}(t, z)}\mathfrak{m}, \quad (t, z) \in \mathcal{D}^z,$$

with $a(t, z)$ as defined below (8).

Proposition 20 *Let $(b, \Upsilon; \lambda)$ be a compatible triple and define*

$$\widehat{\Upsilon}(y) := \Upsilon(y) + \frac{y\Upsilon'(y)}{\Upsilon(y)}, \qquad y \in (0, \bar{b}).$$

Then $\widehat{\Upsilon} > 1$, and the corresponding candidate strategy is described by the feedback rule

$$\varpi(t, z) = \frac{z}{a(t, z)\big(\widehat{\Upsilon}(b(t)z) - 1\big)}\mathfrak{m}, \qquad (t, z) \in \mathcal{D}^z.$$

If the compatible triple has non-constant elasticity, then this representation is equivalently given by

$$\varpi(t, z) = \frac{z}{\rho a(t, z)}\left(A_0 - \frac{F_0}{\Upsilon(b(t)z)}\right)\mathfrak{m}.$$

If it has constant elasticity, with $\Upsilon \equiv \Upsilon_0$, then it reduces to

$$\varpi(t, z) = \frac{z}{a(t, z)(\Upsilon_0 - 1)}\mathfrak{m}.$$

Proof Let $(t,z) \in \mathcal{D}^z$ and set $y = b(t)z$. By the definitions of v and Υ,

$$\frac{v_z(t,z)}{v_{zz}(t,z)} = \frac{\psi'(y)}{b(t)\psi''(y)} = \frac{z}{y\psi''(y)/\psi'(y)}.$$

Moreover,

$$0 < \frac{y\psi''(y)}{\psi'(y)} = \Upsilon(y) - 1 + \frac{y\Upsilon'(y)}{\Upsilon(y)} = \widehat{\Upsilon}(y) - 1.$$

Substitution into the feedback rule for ϖ proves the first representation.

If the triple has non-constant elasticity, the compatibility equation gives

$$\widehat{\Upsilon}(y) - 1 = \frac{\rho\Upsilon(y)}{A_0\Upsilon(y) - F_0},$$

which yields the second representation. If the triple has constant elasticity, then $\Upsilon \equiv \Upsilon_0$ and hence $\widehat{\Upsilon} \equiv \Upsilon_0$, proving the final representation. □

Converting back to the original coordinates (t,x), we obtain the following corollary.

Corollary 21 (Candidate feedback in wealth coordinates) *Let $(b, \Upsilon; \lambda)$ be a compatible triple. Then, in the original coordinates,*

$$\varpi(t,x) = \frac{x_{\text{safe}}(t) - x}{x\big(\widehat{\Upsilon}(y(t,x)) - 1\big)}\mathfrak{m}, \qquad (t,x) \in \mathcal{D},$$

where $\widehat{\Upsilon}$ is as defined in Proposition 20 and

$$y(t,x) := b(t)z(t,x) = b(t)\frac{x_{\text{safe}}(t) - x}{x_{\text{safe}}(t) - x_{\text{imp}}}.$$

Proof Apply Proposition 20 with $z = z(t,x)$ and use

$$\frac{z(t,x)}{a\big(t, z(t,x)\big)} = \frac{x_{\text{safe}}(t) - x}{x}.$$

□

So far, we have only provided results on the candidate value function and the corresponding candidate risky-fraction strategy. The following theorem shows that these candidates are indeed the true value function and optimal strategy of the control problem in z-coordinates.

Theorem 22 (Verification of the self-similar candidates) *Let $(b, \Upsilon; \lambda)$ be a compatible triple. Then the candidate value function v defined in (21) is indeed the value function of the control problem (3) in z-coordinates, and the feedback rule ϖ as given in Corollary 21 is admissible and optimal. That is, in x-coordinates, for all $(t,x) \in \overline{\mathcal{D}}^x$, the value function satisfies*

$$p(t,x) = \exp\left(-\int_{z(t,x)}^{1} \frac{\Upsilon(b(t)\zeta)}{\zeta}\,\mathrm{d}\zeta\right),$$

and the optimal risky-fraction strategy is given by the feedback rule

$$\pi^*(t,x) = \frac{x_{\text{safe}}(t) - x}{x\big(\widehat{\Upsilon}(y(t,x)) - 1\big)}\mathfrak{m}, \qquad (t,x) \in \mathcal{D},$$

and zero outside $\mathcal{D}$, where $\widehat{\Upsilon}$ is as defined in Proposition 20 and $y(t,x) = b(t)z(t,x)$.

Proof Define the candidate in the original coordinates by

$$V(t,x) := v\big(t, z(t,x)\big), \qquad (t,x) \in \overline{\mathcal{D}}^x.$$

We verify the assumptions of Theorem 7. For $y = b(t)z$, differentiation of (21) gives

$$v_t(t,z) = \frac{b'(t)}{b(t)} v(t,z)\big(\Upsilon(y) - \Upsilon(b(t))\big), \qquad v_z(t,z) = \frac{v(t,z)\Upsilon(y)}{z},$$

and

$$v_{zz}(t,z) = \frac{v(t,z)}{z^2}\big(y\Upsilon'(y) + \Upsilon(y)^2 - \Upsilon(y)\big).$$

Consequently,

$$V_x(t,x) = -\frac{v_z(t,z(t,x))}{x_{\mathrm{safe}}(t) - x_{\mathrm{imp}}} < 0, \qquad V_{xx}(t,x) = \frac{v_{zz}(t,z(t,x))}{(x_{\mathrm{safe}}(t) - x_{\mathrm{imp}})^2} > 0.$$

The standing regularity assumptions and (18) therefore imply that $V \in C^{1,2}(\mathcal{D})$ is strictly convex in x. Moreover, Lemma 15 gives $\Upsilon > 1$, and hence

$$0 < v(t,z) < \exp\left(-\int_z^1 \frac{1}{\zeta}\,\mathrm{d}\zeta\right) = z, \qquad (t,z) \in \mathcal{D}^z.$$

Together with $v(t,1) = 1$, this shows that $V \in C(\overline{\mathcal{D}}^x)$, $0 \le V \le 1$, and

$$V(t, x_{\mathrm{safe}}(t)) = v(t,0) = 0, \qquad V(t, x_{\mathrm{imp}}) = v(t,1) = 1.$$

Since $(b, \Upsilon; \lambda)$ is compatible, (20) holds and is equivalent to the normalized HJB equation (13) for v. The coordinate transformation from Corollary 8 therefore shows that V solves the HJB equation (12). Its unique minimizing selector is

$$\varpi(t,x) = -\frac{V_x(t,x)}{xV_{xx}(t,x)}\mathfrak{m}, \qquad (t,x) \in \mathcal{D},$$

which agrees with the feedback rule in Corollary 21. Extend this rule by zero outside $\mathcal{D}$. It remains only to prove that the resulting feedback control is admissible.

To this end, write $E(t,x) := x\varpi(t,x)$ for the risky dollar exposure. Then, by Corollary 21,

$$E(t,x) = \frac{x_{\mathrm{safe}}(t) - x}{\widehat{\Upsilon}(y(t,x)) - 1}\mathfrak{m}, \qquad (t,x) \in \mathcal{D}.$$

We will consider the two cases of a compatible triple separately. For a constant-elasticity compatible triple, $\Upsilon \equiv \Upsilon_0 > 1$ is constant, and hence

$$\|E(t,x)\|_1 \le \frac{x^\infty_{\mathrm{safe}} - x_{\mathrm{imp}}}{\Upsilon_0 - 1}\|\mathfrak{m}\|_1 =: M_0, \qquad (t,x) \in \mathcal{D}.$$

On the other hand, if the compatible triple has non-constant elasticity, then

$$\|E(t,x)\|_1 \le \frac{x^\infty_{\mathrm{safe}} - x_{\mathrm{imp}}}{\rho}\big(|A_0| + |F_0|\big)\|\mathfrak{m}\|_1 =: M_1, \qquad (t,x) \in \mathcal{D}.$$

Hence, in either case, there exists a constant $M < \infty$ such that

$$\|E(t,x)\|_1 \le M, \qquad (t,x) \in \mathcal{D}. \tag{26}$$

Now we show that the feedback rule allows for a unique strong solution X^ϖ of the closed-loop wealth equation (2). In the non-constant-elasticity case, the representation of E via

$$E(t,x) = \frac{x_{\mathrm{safe}}(t) - x}{\rho}\left(A_0 - \frac{F_0}{\Upsilon(y(t,x))}\right)\mathfrak{m}, \qquad (t,x) \in \mathcal{D},$$

together with $\Upsilon \in C^1((0,\bar{b}))$ and $\Upsilon > 1$, shows that the closed-loop drift and diffusion coefficients are locally Lipschitz in wealth on $\mathcal{D}$. The same conclusion is immediate in the constant-elasticity case. Consequently, the closed-loop wealth equation has a unique strong solution up to its first exit from $\mathcal{D}$. The bound (26) prevents explosion before that exit. After an exit, continue the wealth process with zero risky investment. At the lower boundary the resulting riskless trajectory remains below x_{imp}, whereas at the upper boundary it follows x_{safe}. If $T < \infty$ is reached without an earlier exit, the collapse of the risky region implies $X_T = x_{\text{imp}}$, after which the same riskless continuation applies. This constructs the unique strong solution induced by the zero-extended feedback rule.

Moreover, for every finite $u \geq t$,

$$\mathbb{E}_{t,x}\left[\int_t^u |X_s \varpi_s^\top \theta| \, \mathrm{d}s\right] \leq M\|\theta\|(u-t) < \infty$$

and

$$\mathbb{E}_{t,x}\left[\int_t^u X_s^2 \varpi_s^\top \Sigma \varpi_s \, \mathrm{d}s\right] \leq M^2\|\Sigma\|(u-t) < \infty.$$

Thus, the integrability conditions of Definition 3 are satisfied as well.

Therefore, the feedback rule is admissible in the sense of Definition 3. Theorem 7 now yields $V = p$ on $\mathcal{D}$ and proves the optimality of ϖ. The equality extends to $\overline{\mathcal{D}}^x$ by the boundary values established above, and Proposition 16 gives the stated representation of p. □

This theorem shows that the candidate value functions we have constructed are indeed the true value functions of the control problem, and that the corresponding feedback rules are admissible and optimal. Thus, by finding compatible triples, we have effectively solved the control problem.

For now, we will focus on the non-constant-elasticity case, which leads to a more general family of mortality laws. In the next subsection, we will see that in this case the problem of finding Υ can be reduced to a first-order autonomous ODE, which allows for a semi-explicit representation.

3.2 Reduction to a first-order autonomous ODE

Based on the results of Theorem 17 and Corollary 18, we can now derive a semi-explicit representation for non-constant Υ. In the current setting, the central object of our analysis is the following first-order ODE for Υ on $(0,\bar{b})$:

$$A_0\Upsilon - \rho\mathcal{N}[\Upsilon] = F_0. \tag{27}$$

Recall that we have $\mathcal{N}[\Upsilon](y) > 0$ for all $y \in (0,\bar{b})$, and consequently

$$A_0\Upsilon - F_0 = \rho\mathcal{N}[\Upsilon] > 0,$$

allowing us to divide by this quantity. Rearranging this ODE as in the proof of Theorem 17, we can thus write it in the form

$$\Upsilon'(y) = \frac{\Upsilon(y)}{y}\left(\frac{\rho\Upsilon(y)}{A_0\Upsilon(y) - F_0} - \Upsilon(y) + 1\right). \tag{28}$$

By changing coordinates to $U(s) := \Upsilon\big(y_0 \exp(-s)\big)$ for some fixed $0 < y_0 < \bar{b}$, the above ODE becomes

$$U' = -U\left(\frac{\rho U}{A_0 U - F_0} - U + 1\right) \tag{29}$$

on the new domain $I_0 := \{s \in \mathbb{R} : y_0 \exp(-s) < \bar{b}\}$. Introducing the quadratic function

$$Q(\upsilon) := (\upsilon - 1)(A_0\upsilon - F_0) - \rho\upsilon = A_0\upsilon^2 - (\rho + F_0 + A_0)\upsilon + F_0, \tag{30}$$

we can also write (29) in the more compact form

$$U' = \frac{UQ(U)}{A_0 U - F_0} =: f(U).$$

Since the mapping $s \mapsto y_0 \exp(-s)$ is a bijection from I_0 onto $(0, \bar{b})$, the functions U and Υ carry the same information. In particular,

$$\Upsilon(y) = U\left(\ln\left(\frac{y_0}{y}\right)\right), \qquad y \in (0, \bar{b}).$$

We use U only as a logarithmic reparametrization of Υ; the notion of a compatible triple remains formulated in terms of $(b, \Upsilon; \lambda)$.

Lemma 23 *Let $(b, \Upsilon; \lambda)$ be a non-constant-elasticity compatible triple. Then*

$$F_0 < A_0 < 0 \qquad \textit{and} \qquad 1 < \Upsilon(y) < \frac{F_0}{A_0} =: \kappa, \quad y \in (0, \bar{b}).$$

Proof Since (29) is autonomous, any non-constant solution U must satisfy $f(U(s)) \neq 0$ for all $s \in I_0$. We can thus separate variables to obtain the integral identity

$$s = s_0 + \int_{U(s_0)}^{U(s)} \frac{\mathrm{d}\upsilon}{f(\upsilon)}, \quad s, s_0 \in I_0. \tag{31}$$

Using this identity, we will show that $A_0 \geq 0$ leads to a contradiction.

To this end, suppose that $A_0 > 0$. Then, in order for $A_0 U - F_0 > 0$ to hold on the entire domain, we must have $U > \kappa$. Now, let $\hat{\kappa} := \max\{1, \kappa\}$. Then, since $Q(\hat{\kappa}) = -\rho\hat{\kappa} < 0$ and $Q(\upsilon) \to \infty$ as $\upsilon \to \infty$, there exists a unique $\upsilon^* > \hat{\kappa}$ such that $Q(\upsilon^*) = 0$, and consequently $f(\upsilon^*) = 0$. We now have two possibilities: either $\hat{\kappa} < U < \upsilon^*$, or $U > \upsilon^*$. In the first case, we have $f(U(s)) < 0$ for all $s \in I_0$, implying that U is strictly decreasing in s. Moreover, since

$$\lim_{\upsilon \downarrow \hat{\kappa}} f(\upsilon) = \begin{cases} -\frac{\rho}{A_0(1-\kappa)} < 0, & \text{if } \kappa < 1, \\ -\infty, & \text{if } \kappa \geq 1, \end{cases}$$

we immediately get $\overline{f} := \sup_{\upsilon \in (\hat{\kappa}, U(s_0))} f(\upsilon) < 0$ for any fixed $s_0 \in I_0$. Hence, (31) yields

$$s = s_0 + \int_{U(s_0)}^{U(s)} \frac{\mathrm{d}\upsilon}{f(\upsilon)} < s_0 + \int_{\hat{\kappa}}^{U(s_0)} \frac{\mathrm{d}\upsilon}{|\overline{f}|} < \infty,$$

which contradicts the fact that s can be arbitrarily large. Similarly, if $U > \upsilon^*$, then $f(U(s)) > 0$ for all $s \in I_0$. Since $f(\upsilon) \sim \upsilon^2$ as $\upsilon \to \infty$, we can again use (31) to obtain a finite upper

bound for s, yielding the same contradiction. Therefore, we conclude that $A_0 > 0$ cannot occur.

Now, suppose that $A_0 = 0$. In this case $F_0 < 0$ must hold, since otherwise $A_0 U - F_0 > 0$ would be violated on the entire domain. Then (29) reduces to

$$U' = -U\,(1 - \tilde{\kappa} U)\,,$$

where $\tilde{\kappa} := 1 + \rho/F_0 < 1$. If $\tilde{\kappa} > 0$, then again, either $1 < U < \tilde{\kappa}^{-1}$, or $U > \tilde{\kappa}^{-1}$. Exactly as in the previous case, both possibilities lead to a finite upper bound for s when using (31), which is a contradiction. If instead $\tilde{\kappa} \leq 0$, then

$$\sup_{\upsilon > 1} \{-\upsilon(1 - \tilde{\kappa}\upsilon)\} < 0.$$

Thus, U is strictly decreasing, and

$$s \leq s_0 + \int_1^{U(s_0)} \frac{\mathrm{d}\upsilon}{\upsilon(1 - \tilde{\kappa}\upsilon)} < \infty,$$

a contradiction. Hence, we conclude that $A_0 = 0$ cannot occur either, and we must have $A_0 < 0$.

Finally, since $A_0 < 0$ and $A_0 U - F_0 > 0$, we must have $1 < U < \kappa$, and consequently $F_0 < A_0$. Switching back to Υ completes the proof. □

Having established $A_0 < 0$, we have

$$\frac{b'(t)}{b(t)} = A_0 - \mathcal{E}(t) < 0.$$

Hence, b is strictly decreasing and $\bar{b} = b(0) < \infty$. Since (29) is autonomous, a translation of the variable s allows us, without loss of generality, to use the normalization

$$U(s) = \Upsilon\big(\bar{b}\exp(-s)\big)\,, \qquad s \in (0, \infty),$$

from now on. With this normalization, we can now analyze the asymptotic behavior of solutions to (29). In particular, we can show that any non-constant solution must converge to the larger zero of the quadratic Q as $s \to \infty$.

Lemma 24 (Equilibria and limiting elasticity) *Let $(b, \Upsilon; \lambda)$ be a non-constant-elasticity compatible triple. Then Q has two, possibly coinciding, zeros $\upsilon_\pm$ satisfying*

$$1 < \upsilon_- \leq \upsilon_+ < \kappa.$$

Moreover, exactly one of the following alternatives holds on $(0, \bar{b})$:

$$\upsilon_- < \Upsilon < \upsilon_+,$$

or

$$\upsilon_+ < \Upsilon < \kappa.$$

In either case,

$$\lim_{y \downarrow 0} \Upsilon(y) = \upsilon_+.$$

Proof By Lemma 23, we know that $A_0 < 0$ and $1 < U < \kappa$. In particular, $Q''(\upsilon) = 2A_0 < 0$, so Q is strictly concave. Moreover,

$$Q(1) = -\rho < 0 \quad \text{and} \quad Q(\kappa) = -\rho\kappa < 0. \tag{32}$$

Since U is non-constant, it cannot attain a zero of f. Hence, $f(U(s)) \neq 0$ for all $s > 0$, and U is strictly monotone. Since U is also bounded, the limit

$$U_\infty := \lim_{s\to\infty} U(s)$$

exists and is finite.

We claim that $Q(U_\infty) = 0$. Suppose otherwise. By continuity, there exist $\delta > 0$ and $s_0 > 0$ such that

$$|Q(U(s))| \geq \delta,$$

for all $s \geq s_0$. Hence,

$$|U'(s)| = \frac{U(s)|Q(U(s))|}{A_0U(s) - F_0} \geq \frac{\delta}{A_0 - F_0} > 0$$

is bounded away from zero for $s \geq s_0$. This contradicts the existence of the finite limit U_∞, and therefore $Q(U_\infty) = 0$. Moreover, due to (32), we have $U_\infty \in (1, \kappa)$.

Thus, Q has a zero in $(1, \kappa)$. However, since Q is a strictly concave quadratic and is negative at both endpoints, it must have two real zeros, counted with multiplicity. Therefore, there exist $1 < \upsilon_- \leq \upsilon_+ < \kappa$ such that $Q(\upsilon_-) = Q(\upsilon_+) = 0$.

Since a non-constant solution cannot attain either equilibrium υ_- or υ_+, its range must be contained in one of

$$(1, \upsilon_-), \qquad (\upsilon_-, \upsilon_+), \qquad (\upsilon_+, \kappa).$$

The first possibility cannot occur. Indeed, $Q < 0$ on $(1, \upsilon_-)$, and hence $f < 0$ there. Thus, U would be decreasing and bounded below by 1. Its limit would belong to $[1, \upsilon_-)$, but this interval contains no zero of f, which contradicts the preceding limit argument. Therefore, either

$$\upsilon_- < U < \upsilon_+, \quad \text{or} \quad \upsilon_+ < U < \kappa.$$

Finally, since U is strictly increasing in the first, and strictly decreasing in the second case, the limit U_∞ must be equal to υ_+ in either case. Again, switching back to Υ completes the proof. □

Using the above lemma, we can infer that

$$\Upsilon(0) := \lim_{y\downarrow 0} \Upsilon(y) = \upsilon_+.$$

In particular, Υ is bounded on $(0, \bar{b})$. Moreover, since Q is a quadratic polynomial, we can explicitly compute its zeros $\upsilon_\pm$ in terms of A_0, F_0, and ρ. Doing so yields

$$\upsilon_\pm = \frac{\rho + F_0 + A_0 \mp \sqrt{(\rho + F_0 + A_0)^2 - 4A_0F_0}}{2A_0}.$$

In order for those zeros to be real, we must have $\Delta := (\rho + F_0 + A_0)^2 - 4A_0F_0 \geq 0$, which is equivalent to requiring that either

$$F_0 \leq A_0 - \rho - 2\sqrt{-A_0\rho}, \quad \text{or} \quad F_0 \geq A_0 - \rho + 2\sqrt{-A_0\rho}.$$

However, since $0 < Q'(1) = A_0 - F_0 - \rho$ due to concavity and the fact that $1 < \upsilon_-, \upsilon_+$, we must have $F_0 < A_0 - \rho$, excluding the second possibility. Hence, we conclude that the parameters A_0 and F_0 must additionally satisfy the following discriminant condition:

$$F_0 \leq A_0 - \rho - 2\sqrt{-A_0\rho}. \tag{33}$$

The case of equality in (33) corresponds to the double-root case $\Delta = 0$, which we will analyze later. If a compatible triple satisfies the strict inequality in (33), we will refer to it as *distinct-root*, otherwise we will call it *double-root*.

3.2.1 The distinct-root case

For now, we will focus on the case where $\Delta > 0$, which corresponds to a strictly concave quadratic Q with two distinct zeros. In this case, we can factor Q as

$$Q(\upsilon) = A_0(\upsilon - \upsilon_-)(\upsilon - \upsilon_+).$$

Using the integral identity (31), we get

$$s = s_0 + \int_{U(s_0)}^{U(s)} \frac{A_0\upsilon - F_0}{A_0\upsilon(\upsilon - \upsilon_-)(\upsilon - \upsilon_+)}\,\mathrm{d}\upsilon.$$

Using partial fraction decomposition, we can rewrite the integrand as

$$\frac{A_0\upsilon - F_0}{A_0\upsilon(\upsilon - \upsilon_-)(\upsilon - \upsilon_+)} = \frac{r_-}{\upsilon - \upsilon_-} - \frac{r_+}{\upsilon - \upsilon_+} - \frac{1}{\upsilon},$$

with

$$r_\pm := \mp\frac{1}{f'(\upsilon_\pm)} = -\frac{\rho}{A_0(\upsilon_\pm - 1)(\upsilon_+ - \upsilon_-)} > 0.$$

Hence, we can integrate to obtain the implicit solution

$$s = s_0 + r_- \ln\left|\frac{U(s) - \upsilon_-}{U(s_0) - \upsilon_-}\right| - r_+ \ln\left|\frac{U(s) - \upsilon_+}{U(s_0) - \upsilon_+}\right| - \ln\left|\frac{U(s)}{U(s_0)}\right|.$$

Converting back to the original coordinates, we have

$$\ln\left(\frac{\bar{b}}{y}\right) = \ln\left(\frac{\bar{b}}{y_0}\right) + r_- \ln\left|\frac{\Upsilon(y) - \upsilon_-}{\Upsilon(y_0) - \upsilon_-}\right| - r_+ \ln\left|\frac{\Upsilon(y) - \upsilon_+}{\Upsilon(y_0) - \upsilon_+}\right| - \ln\left|\frac{\Upsilon(y)}{\Upsilon(y_0)}\right|.$$

Thus, by applying the exponential function to both sides and collecting all constants into a single constant $c_\Upsilon > 0$, we obtain the following implicit representation of Υ:

$$y = c_\Upsilon \Upsilon(y)\,|\Upsilon(y) - \upsilon_-|^{-r_-}\,|\Upsilon(y) - \upsilon_+|^{r_+}. \tag{34}$$

Considering our two possible branches for U, we see that the above representation can be rewritten as

$$y = c_\Upsilon \Upsilon(y) \left(\Upsilon(y) - \upsilon_-\right)^{-r_-} \left(\upsilon_+ - \Upsilon(y)\right)^{r_+} =: g_-(\Upsilon(y)) \tag{35}$$

when $\upsilon_- < \Upsilon(y) < \upsilon_+$, and as

$$y = c_\Upsilon \Upsilon(y) \left(\Upsilon(y) - \upsilon_-\right)^{-r_-} \left(\Upsilon(y) - \upsilon_+\right)^{r_+} =: g_+(\Upsilon(y)) \tag{36}$$

when $\upsilon_+ < \Upsilon(y) < \kappa$. From now on we will denote the two branches of Υ by Υ_- and Υ_+, respectively.

Therefore, we have obtained an implicit representation of $\Upsilon_\pm$ in terms of the functions $g_\pm$. As we will see in the next lemma, the functions $g_\pm$ are strictly monotone on their respective domains and thus invertible, allowing us to express $\Upsilon_\pm$ explicitly in terms of $g_\pm^{-1}$. In what follows, the same notation is used for the maximal extensions of these solution branches; the elasticity profile of a compatible triple is their restriction to $(0, \overline{b})$.

Lemma 25 *Let $g_- : (\upsilon_-, \upsilon_+) \to (0, \infty)$ and $g_+ : (\upsilon_+, \kappa) \to (0, \overline{g})$ with $\overline{g} := \lim_{\upsilon \uparrow \kappa} g_+(\upsilon)$ be defined as in (35) and (36), respectively. Then $g_\pm$ are bijective, hence invertible, and $\Upsilon_\pm$ must satisfy*

$$\Upsilon_-(y) = g_-^{-1}(y), \quad y \in (0, \infty),$$

and

$$\Upsilon_+(y) = g_+^{-1}(y), \quad y \in (0, \overline{g}).$$

Moreover, Υ_+ can be continuously extended to $\overline{g}$ by setting $\Upsilon_+(\overline{g}) := \kappa$.

Proof We will only prove the claim for g_+, as the proof for g_- is analogous. To this end, we first compute the derivative of g_+:

$$g_+'(\upsilon) = g_+(\upsilon) \left(\frac{1}{\upsilon} - \frac{r_-}{\upsilon - \upsilon_-} + \frac{r_+}{\upsilon - \upsilon_+} \right).$$

Since $\upsilon_- < \upsilon_+ < \upsilon < \kappa$, we have $g_+(\upsilon) > 0$, and it remains to show that the second factor is also strictly positive. To this end, we define

$$h(\upsilon) := \frac{1}{\upsilon} - \frac{r_-}{\upsilon - \upsilon_-} + \frac{r_+}{\upsilon - \upsilon_+}.$$

A direct calculation using Vieta's formulas gives

$$\begin{aligned} r_- - r_+ &= 1, \\ \upsilon_- \upsilon_+ &= \kappa, \\ r_- \upsilon_- - r_+ \upsilon_+ &= 1. \end{aligned}$$

Using these relations, we can rewrite h as

$$h(\upsilon) = \frac{\kappa - \upsilon}{\upsilon(\upsilon - \upsilon_-)(\upsilon - \upsilon_+)}.$$

Thus, we have $h(\upsilon) > 0$ on (υ_+, κ), and consequently $g_+'(\upsilon) > 0$ for all $\upsilon \in (\upsilon_+, \kappa)$. Hence, g_+ is strictly increasing. Moreover, since

$$\lim_{\upsilon \downarrow \upsilon_+} g_+(\upsilon) = 0, \qquad \lim_{\upsilon \uparrow \kappa} g_+(\upsilon) = \overline{g} \in (0, \infty),$$

g_+ is bijective on its domain. Thus, g_+ is invertible on (υ_+, κ), and we can express Υ_+ in terms of g_+^{-1}. The claim for g_- follows analogously, and the lemma is proved. □

Using this representation, we can now derive a lower bound for the constant c_Υ in terms of the constant c_b from (24).

Lemma 26 *Let $(b, \Upsilon; \lambda)$ be a distinct-root compatible triple with $\Upsilon = \Upsilon_+$, and let $c_\Upsilon > 0$ be the representation constant appearing in (36). Then*

$$\frac{(\kappa - \upsilon_-)^{r_-}}{\mathcal{E}(0)\kappa(\kappa - \upsilon_+)^{r_+}} \cdot c_b \leq c_\Upsilon.$$

Proof Recall that the scaling function b is given by

$$b(t) = c_b \frac{\exp\left((A_0 - r)t\right)}{\mathcal{E}(t)}, \quad t \in [0, T),$$

for some constant $c_b > 0$. Since $A_0 < 0$, b is strictly decreasing, and we have $\overline{b} = b(0) = c_b/\mathcal{E}(0)$. Hence, the domain of Υ is given by $(0, \overline{b}) = (0, c_b/\mathcal{E}(0))$. Now, since

$$y = c_\Upsilon \Upsilon(y) \left(\Upsilon(y) - \upsilon_-\right)^{-r_-} \left(\Upsilon(y) - \upsilon_+\right)^{r_+},$$

for all $y \in (0, \overline{g})$, and since g_+ is strictly increasing, we must have

$$\frac{c_b}{\mathcal{E}(0)} \leq \overline{g} = c_\Upsilon \kappa \left(\kappa - \upsilon_-\right)^{-r_-} \left(\kappa - \upsilon_+\right)^{r_+},$$

in order for Υ to be well-defined on its entire domain. Rearranging this inequality yields the desired lower bound for c_Υ. □

Remark 4 If $\Upsilon = \Upsilon_-$, then c_Υ may be arbitrarily small since g_-^{-1} has domain $(0, \infty)$. Moreover, it will prove convenient to, instead of c_Υ, work with the constant c_b from (24). To this end, note that if Υ_{c_Υ} denotes the branch associated with c_Υ, then

$$\Upsilon_{c_\Upsilon}(y) = \Upsilon_1(y/c_\Upsilon).$$

Consequently, replacing $(b, \Upsilon_{c_\Upsilon}; \lambda)$ by $(b/c_\Upsilon, \Upsilon_1; \lambda)$ leaves every occurrence of $\Upsilon(b(t)z)$ unchanged. We may therefore normalize $c_\Upsilon = 1$ and absorb the scaling into c_b. With this normalization, Lemma 26 implies that c_b must satisfy

$$0 < c_b \leq \begin{cases} \mathcal{E}(0)\overline{g}, & \text{if } \Upsilon \text{ is given by (36)}, \\ \infty, & \text{if } \Upsilon \text{ is given by (35)}. \end{cases}$$

After normalizing $c_\Upsilon = 1$, every distinct-root compatible triple is represented by a parameter vector

$$(\circ, c_b, A_0, F_0, T), \qquad \circ \in \{-, +\},$$

subject to the corresponding domain restriction. Conversely, every such parameter vector for which the reconstructed function λ satisfies the requirements of Definition 5 defines a compatible triple.

Now suppose that we have a distinct-root compatible triple $(b, \Upsilon; \lambda)$. Then, according to Corollary 18, the resulting hazard rate λ is given by

$$\lambda(t) = F_0 + \big(\mathcal{E}(t) - A_0\big)\Upsilon(b(t)), \quad t \in [0, T).$$

Using this representation, we can show the following result which studies the asymptotic behavior of the hazard rate λ as $t \to T$.

Proposition 27 *Let $(b, \Upsilon; \lambda)$ be a distinct-root compatible triple. Then the following hold.*

(i) If $T < \infty$, then

$$\lim_{t \uparrow T} \frac{\lambda(t)}{\mathcal{E}(t)} = \upsilon_+.$$

Equivalently,

$$\lim_{t \uparrow T} (T - t)\lambda(t) = \upsilon_+.$$

(ii) If $T = \infty$, then

$$\lim_{t \to \infty} \lambda(t) = F_0 + (r - A_0)\upsilon_+ > 0.$$

Proof By Corollary 18, we have

$$b(t) = c_b \frac{\exp((A_0 - r)t)}{\mathcal{E}(t)}$$

and

$$\lambda(t) = F_0 + \big(\mathcal{E}(t) - A_0\big)\Upsilon(b(t)).$$

Since $A_0 < 0$, we have $b(t) \to 0$ as $t \to T$, both when $T < \infty$ and when $T = \infty$. Consequently,

$$\lim_{t \to T} \Upsilon(b(t)) = \upsilon_+.$$

Suppose first that $T < \infty$. Since $\mathcal{E}(t) \to \infty$, we obtain

$$\frac{\lambda(t)}{\mathcal{E}(t)} = \frac{F_0}{\mathcal{E}(t)} + \left(1 - \frac{A_0}{\mathcal{E}(t)}\right) \Upsilon(b(t)) \longrightarrow \upsilon_+.$$

Moreover, $(T - t)\mathcal{E}(t) \to 1$ as $t \uparrow T$, and hence

$$\lim_{t \uparrow T} (T - t)\lambda(t) = \upsilon_+.$$

Now suppose that $T = \infty$. In this case, $\mathcal{E}(t) \equiv r$, and therefore

$$\lim_{t \to \infty} \lambda(t) = F_0 + (r - A_0)\upsilon_+ := L.$$

Since $\lambda \geq 0$, we necessarily have $L \geq 0$. It remains to exclude $L = 0$.

By the representation (34), and since $\Upsilon(y) \to \upsilon_+$ as $y \downarrow 0$, there exist constants $C > 0$ and $y_0 > 0$ such that

$$|\Upsilon(y) - \upsilon_+| \leq C y^{1/r_+}, \qquad 0 < y < y_0. \tag{37}$$

If $L = 0$, then

$$\lambda(t) = (r - A_0)\big(\Upsilon(b(t)) - \upsilon_+\big).$$

Furthermore, because

$$b(t) = \frac{c_b}{r}\exp\big(-(r - A_0)t\big),$$

it follows by (37) that, for all sufficiently large t,

$$|\lambda(t)| \leq C' \exp\left(-\frac{r - A_0}{r_+}t\right)$$

for some $C' > 0$. Consequently,

$$\int_0^\infty \lambda(t)\ \mathrm{d}t < \infty.$$

Since

$$S(t) = \exp\left(-\int_0^t \lambda(s)\ \mathrm{d}s\right),$$

this would imply

$$\lim_{t\to\infty} S(t) > 0,$$

contradicting the assumption in Definition 5. Thus, $L > 0$, completing the proof. □

In view of the results obtained in this subsection, we can reverse engineer the construction of a hazard rate λ in the following way:

(i) Choose constants $A_0 < 0$ and $F_0 < A_0 - \rho - 2\sqrt{-A_0\rho}$.

(ii) Select a branch for Υ given by either (35) or (36).

(iii) Choose a constant $c_b > 0$ satisfying the restriction in Remark 4 and define the scaling function b as in (24).

(iv) Define the function λ as in (25) and check whether it is non-negative on $[0, T)$ and satisfies the integrability condition

$$\int_0^T \lambda(t)\,\mathrm{d}t = \infty.$$

If not, go back to step (i) and choose different constants. Otherwise, λ defines a mortality model and $(b, \Upsilon; \lambda)$ is a compatible triple with the desired asymptotic behavior.

Despite the fact that the above construction is not guaranteed to yield a non-negative hazard rate, it provides a systematic way to generate a variety of compatible mortality laws. In view of the results of Proposition 27, one can do so by first selecting the asymptotic behavior of the hazard rate λ as $t \to T$ and then fine-tuning the constants A_0, F_0, and c_b to obtain the desired small-time behavior of λ as $t \to 0$. However, due to the lack of a general closed-form expression for $g_\pm^{-1}$, it is difficult to analyze the properties of the resulting hazard rate λ in a more explicit manner. As we will see in the next subsection, the double-root case $\Delta = 0$ allows for a more explicit analysis of the resulting hazard.

3.2.2 The double-root case

In this subsection, we will analyze the double-root case $\Delta = 0$, which corresponds to the equality case in (33). Here, the quadratic Q from (30) has a double root, and the ODE (29) can be solved explicitly. In particular, we have

$$\upsilon_* := \upsilon_- = \upsilon_+ = 1 + \sqrt{-\frac{\rho}{A_0}},$$

and

$$F_0 = A_0 - \rho - 2\sqrt{-A_0\rho} = A_0\upsilon_*^2.$$

This allows us to derive a closed-form expression for the resulting hazard rate λ.

As in the preceding subsubsection, we can factor Q as

$$Q(\upsilon) = A_0(\upsilon - \upsilon_*)^2,$$

where υ_* is the double root of Q. Again, by using the integral identity (31), we can write

$$s = s_0 + \int_{U(s_0)}^{U(s)} \frac{A_0\upsilon - F_0}{A_0\upsilon(\upsilon - \upsilon_*)^2}\,\mathrm{d}\upsilon = s_0 + \int_{U(s_0)}^{U(s)} \frac{\upsilon - \upsilon_*^2}{\upsilon(\upsilon - \upsilon_*)^2}\,\mathrm{d}\upsilon.$$

Using the partial fraction decomposition

$$\frac{\upsilon - \upsilon_*^2}{\upsilon(\upsilon - \upsilon_*)^2} = \frac{1}{\upsilon - \upsilon_*} - \frac{\upsilon_* - 1}{(\upsilon - \upsilon_*)^2} - \frac{1}{\upsilon},$$

and following the same steps as in the distinct-root case, we obtain the implicit representation

$$y = c_\Upsilon \frac{\Upsilon(y)}{\Upsilon(y) - \upsilon_*} \exp\left(-\frac{\upsilon_* - 1}{\Upsilon(y) - \upsilon_*}\right), \tag{38}$$

where $c_\Upsilon > 0$ is again a suitable constant. Fortunately, the above equation can be solved explicitly for Υ in terms of the Lambert–W function, which is implicitly defined via the relation $W(x)\exp(W(x)) = x$ and has multiple branches. The principal branch W_0 is defined on $[-\mathrm{e}^{-1}, \infty)$, while the branch W_{-1} is defined on $[-\mathrm{e}^{-1}, 0)$. Again, we study the extension of Υ to the maximal interval $(0, \overline{y})$, and restrict it back to $(0, \overline{b})$ to obtain the elasticity profile of a compatible triple.

Proposition 28 *The function Υ can be expressed in terms of the Lambert–W function as*

$$\Upsilon(y) = \upsilon_* - \frac{\upsilon_* - 1}{W_{-1}\left(-\frac{y}{\mathrm{e}\overline{y}}\right) + \frac{\upsilon_* - 1}{\upsilon_*}}, \quad y \in (0, \overline{y}), \tag{39}$$

where $\overline{y} := c_\Upsilon \upsilon_*/(\upsilon_* - 1) \cdot \exp(-1/\upsilon_*)$.

Proof Rewriting (38) as

$$-\frac{y}{\mathrm{e}\overline{y}} = -\frac{y(\upsilon_* - 1)\exp(1/\upsilon_*)}{c_\Upsilon \mathrm{e}\upsilon_*} = -\frac{\Upsilon(y)(\upsilon_* - 1)}{\upsilon_*(\Upsilon(y) - \upsilon_*)} \exp\left(\frac{1}{\upsilon_*} - \frac{\upsilon_* - 1}{\Upsilon(y) - \upsilon_*} - 1\right),$$

we get, after some algebraic manipulations,

$$-\frac{y}{\mathrm{e}\overline{y}} = -\frac{\Upsilon(y)(\upsilon_* - 1)}{\upsilon_*(\Upsilon(y) - \upsilon_*)} \exp\left(-\frac{\Upsilon(y)(\upsilon_* - 1)}{\upsilon_*(\Upsilon(y) - \upsilon_*)}\right).$$

By Lemma 24, we have $\upsilon_* < \Upsilon(y) < \upsilon_*^2$. Consequently,

$$-\frac{\Upsilon(y)(\upsilon_* - 1)}{\upsilon_*(\Upsilon(y) - \upsilon_*)} < -1,$$

and since $-y/(\mathrm{e}\overline{y}) \in (-\mathrm{e}^{-1}, 0)$ for all $y \in (0, \overline{y})$, we can apply the branch W_{-1} of the Lambert–W function to both sides of the above equation to obtain

$$W_{-1}\left(-\frac{y}{\mathrm{e}\overline{y}}\right) = -\frac{\Upsilon(y)(\upsilon_* - 1)}{\upsilon_*(\Upsilon(y) - \upsilon_*)},$$

which can be solved for $\Upsilon(y)$ to yield the desired expression. □

To shorten notation, we will define the function $D : (0, \overline{y}) \to (0, \infty)$ by

$$D(y) := -W_{-1}\left(-\frac{y}{\mathrm{e}\overline{y}}\right) - \frac{\upsilon_* - 1}{\upsilon_*}.$$

Note that D can be continuously extended to $\overline{y}$ by defining $D(\overline{y}) := 1/\upsilon_*$. With this, we can express Υ more compactly as

$$\Upsilon(y) = \upsilon_* + \frac{\upsilon_* - 1}{D(y)}, \quad y \in (0, \overline{y}).$$

With this explicit representation of Υ, we immediately obtain the following result.

Lemma 29 *Let Υ be given by (39). Then Υ is strictly increasing on $(0, \overline{y})$ and satisfies*

$$\upsilon_* < \Upsilon(y) < \upsilon_*^2,$$

for all $y \in (0, \overline{y})$. Moreover, the left- and right-hand limits of Υ at the endpoints of its domain are given by

$$\lim_{y \downarrow 0} \Upsilon(y) = \upsilon_*, \quad \text{and} \quad \lim_{y \uparrow \overline{y}} \Upsilon(y) = \upsilon_*^2.$$

Consequently, Υ can be continuously extended to $\overline{y}$ by setting $\Upsilon(\overline{y}) := \upsilon_*^2$.

Proof Since $\upsilon_* - 1 > 0$ and $D(y) > 0$ for all $y \in (0, \overline{y})$, we have $\Upsilon(y) > \upsilon_*$ for all $y \in (0, \overline{y})$. Moreover, since $D(y) > 1/\upsilon_*$ for all $y \in (0, \overline{y})$, we also have $\Upsilon(y) < \upsilon_*^2$ for all $y \in (0, \overline{y})$. The limits at the endpoints of the domain follow immediately from

$$\lim_{y \downarrow 0} D(y) = \infty, \quad \text{and} \quad \lim_{y \uparrow \overline{y}} D(y) = \frac{1}{\upsilon_*}.$$

Finally, since W_{-1} is strictly decreasing on $[-\mathrm{e}^{-1}, 0)$, we have that D is strictly decreasing on $(0, \overline{y})$, and hence Υ is strictly increasing on $(0, \overline{y})$. □

As before, we can now use the explicit representation of Υ to derive a lower bound for the constant c_Υ in terms of the constant c_b from (24).

Lemma 30 *Let $(b, \Upsilon; \lambda)$ be a double-root compatible triple with Υ given by (39), and let $c_\Upsilon > 0$ be the representation constant appearing in (38). Then*

$$\frac{(\upsilon_* - 1)\exp(1/\upsilon_*)}{\upsilon_* \mathcal{E}(0)} \cdot c_b \leq c_\Upsilon.$$

Proof The proof is analogous to that of Lemma 26. Since b is strictly decreasing, we have $\overline{b} = b(0) = c_b/\mathcal{E}(0)$. Hence, the domain of Υ is given by $(0, \overline{b}) = (0, c_b/\mathcal{E}(0))$. Now, since

$$\Upsilon(y) = \upsilon_* + \frac{\upsilon_* - 1}{D(y)}, \quad y \in (0, \overline{y}),$$

we must have $c_b/\mathcal{E}(0) \leq \overline{y}$. Hence,

$$\frac{c_b}{\mathcal{E}(0)} \leq \overline{y} = \frac{c_\Upsilon \upsilon_*}{\upsilon_* - 1} \exp\left(-\frac{1}{\upsilon_*}\right),$$

which can be rearranged to yield the desired lower bound for c_Υ. □

As in Remark 4, we will now normalize $c_\Upsilon = 1$ and absorb the scaling into c_b. With this normalization, Lemma 30 implies that c_b must satisfy

$$0 < c_b \leq \frac{\upsilon_* \mathcal{E}(0)}{(\upsilon_* - 1)\exp(1/\upsilon_*)}. \tag{40}$$

Since in this case the only free parameters are A_0 and c_b for a fixed T, we can derive explicit sufficient conditions on A_0 to ensure that the resulting hazard rate λ is non-negative on $[0, T)$.

Proposition 31 *Let*

$$-\mathcal{E}(0)^2 \leq \rho A_0 < 0.$$

Then, for any $c_b > 0$ satisfying

$$0 < c_b \leq \frac{\upsilon_* \mathcal{E}(0)}{(\upsilon_* - 1)\exp(1/\upsilon_*)},$$

the function λ defined by

$$\lambda(t) = \upsilon_*^2 A_0 + \big(\mathcal{E}(t) - A_0\big)\Upsilon(b(t)), \quad t \in [0, T),$$

with Υ given by (39) and b given by (24), is non-negative on $[0, T)$.

Proof Since $A_0 < 0$, we have that $\mathcal{E}(t) - A_0 > 0$ for all $t \in [0, T)$. Moreover, by Lemma 29, we have that $\Upsilon(b(t)) > \upsilon_*$ for all $t \in [0, T)$. Hence, we have

$$\lambda(t) \geq \upsilon_*^2 A_0 + \big(\mathcal{E}(t) - A_0\big)\upsilon_* = \upsilon_*(\mathcal{E}(t) + A_0(\upsilon_* - 1)), \quad t \in [0, T).$$

Now, since $\mathcal{E}(t)$ is non-decreasing, we have

$$\lambda(t) \geq \upsilon_*(\mathcal{E}(0) + A_0(\upsilon_* - 1)), \quad t \in [0, T).$$

Since $\upsilon_* = 1 + \sqrt{-\rho/A_0} > 1$, it is thus sufficient to show that $\mathcal{E}(0) + A_0(\upsilon_* - 1) \geq 0$. Using the definition of υ_*, we have

$$\mathcal{E}(0) + A_0(\upsilon_* - 1) = \mathcal{E}(0) + A_0\sqrt{-\frac{\rho}{A_0}} = \mathcal{E}(0) - \sqrt{-\rho A_0}.$$

Hence, the condition $-\mathcal{E}(0)^2 \leq \rho A_0 < 0$ is sufficient to ensure that $\lambda(t) \geq 0$ for all $t \in [0, T)$. □

The Lambert–W representation of Υ also yields a closed-form expression for the corresponding value function.

Proposition 32 *Let $(b, \Upsilon; \lambda)$ be a double-root compatible triple with Υ given by (39). Then the value function p is given by*

$$p(t,x) = \frac{D(b(t)z(t,x))}{D(b(t))} \exp\left(-\upsilon_*\big(D(b(t)z(t,x)) - D(b(t))\big)\right)$$

for all $(t,x) \in \mathcal{D}$.

Proof Let $y \in (0, \bar{b})$, noting that $\bar{b} \leq \overline{y}$, with $\overline{y}$ as in Proposition 28, and recall that Υ was defined via

$$\Upsilon(y) = \frac{y\psi'(y)}{\psi(y)}.$$

Equivalently, we can write

$$\frac{\Upsilon(y)}{y} = \frac{\psi'(y)}{\psi(y)} = \frac{\mathrm{d}}{\mathrm{d}y}\ln(\psi(y)).$$

Now, by (28), we have

$$\frac{\Upsilon(y)}{y} = -\Upsilon'(y)\frac{\Upsilon(y) - \upsilon_*^2}{(\Upsilon(y) - \upsilon_*)^2},$$

and hence,

$$\frac{\mathrm{d}}{\mathrm{d}y}\ln(\psi(y)) = -\frac{\Upsilon(y) - \upsilon_*^2}{(\Upsilon(y) - \upsilon_*)^2}\Upsilon'(y).$$

Thus, by substituting $\upsilon = \Upsilon(y)$, fixing $y_0 \in (0, \bar{b})$, and using the chain rule, we can write

$$\ln(\psi(y)) - \ln(\psi(y_0)) = \int_{\Upsilon(y_0)}^{\Upsilon(y)} -\frac{\upsilon - \upsilon_*^2}{(\upsilon - \upsilon_*)^2}\,\mathrm{d}\upsilon.$$

Using a partial fraction decomposition and integrating, we obtain

$$\ln(\psi(y)) = c_\psi - \ln|\Upsilon(y) - \upsilon_*| - \frac{\upsilon_*(\upsilon_* - 1)}{\Upsilon(y) - \upsilon_*}$$

for some constant $c_\psi \in \mathbb{R}$. Hence, by exponentiating both sides, and using the fact that $\Upsilon(y) > \upsilon_*$, we have

$$\psi(y) = \widetilde{c_\psi}\frac{1}{\Upsilon(y) - \upsilon_*}\exp\left(-\frac{\upsilon_*(\upsilon_* - 1)}{\Upsilon(y) - \upsilon_*}\right),$$

for some constant $\widetilde{c_\psi} > 0$. Further, notice that

$$\frac{D(y)}{\upsilon_* - 1} = \frac{1}{\Upsilon(y) - \upsilon_*}, \qquad \upsilon_* D(y) = \frac{\upsilon_*(\upsilon_* - 1)}{\Upsilon(y) - \upsilon_*},$$

for all $y \in (0, \bar{b})$. Therefore, we can write

$$\psi(y) = \widetilde{c_\psi}\frac{D(y)}{\upsilon_* - 1}\exp\left(-\upsilon_* D(y)\right).$$

The preceding identity extends to $y = \bar{b}$ by continuity. Hence, substituting $y = b(t)z$ and $y = b(t)$ into the above expression for ψ and applying Theorem 22 yields the stated formula for p. □

The preceding results can be summarized in the following exhaustive classification of compatible triples.

Theorem 33 (Classification of compatible triples) *Let $(b, \Upsilon; \lambda)$ be a compatible triple. Then exactly one of the following cases occurs.*

(i) Constant-elasticity case. *There exists $\Upsilon_0 > 1$ such that $\Upsilon \equiv \Upsilon_0$, and*

$$\lambda(t) = \left(\mathcal{E}(t) - \frac{\rho}{\Upsilon_0 - 1}\right)\Upsilon_0.$$

The scaling function b is non-identifiable in this case, while the induced hazard rate, value function, and optimal feedback are determined by (Υ_0, T).

(ii) Distinct-root case. *The elasticity is non-constant, the compatibility constants satisfy*

$$A_0 < 0, \qquad F_0 < A_0 - \rho - 2\sqrt{-A_0\rho},$$

and hence $\Delta > 0$. The elasticity is one of the two branches

$$\Upsilon = \Upsilon_- = g_-^{-1} \qquad \text{or} \qquad \Upsilon = \Upsilon_+ = g_+^{-1},$$

subject to the corresponding parameter restriction from Remark 4.

(iii) Double-root case. *The elasticity is non-constant and*

$$A_0 < 0, \qquad F_0 = A_0 - \rho - 2\sqrt{-A_0\rho},$$

so that $\Delta = 0$. In this case, Υ is given by the Lambert–W representation (39), subject to the parameter restriction in (40).

In the two non-constant cases, the scaling function and hazard rate are recovered from

$$b(t) = c_b \frac{\exp((A_0 - r)t)}{\mathcal{E}(t)}, \qquad \lambda(t) = F_0 + \big(\mathcal{E}(t) - A_0\big)\Upsilon(b(t)).$$

In all three cases, the value function and optimal feedback are given by Theorem 22.

Proof Theorem 17 yields the dichotomy between constant and non-constant elasticities. The constant-elasticity case is characterized by Corollary 19.

In the non-constant case, Lemma 23 and Lemma 24 imply $A_0 < 0$ and the discriminant condition (33). If the inequality is strict, then $\Delta > 0$, and Lemma 25 gives the two distinct-root branches. If equality holds, then $\Delta = 0$, and Proposition 28 gives the double-root elasticity. These three cases are mutually exclusive and exhaustive.

The recovery formulas for b and λ follow from Corollary 18, while the final assertion follows from Theorem 22. □

The next section examines the constant-elasticity case of Theorem 33 in greater detail: it constructs the associated mortality law, relates it to Gompertz–Makeham mortality, and derives its policy consequences.

4 The constant-elasticity family

Section 3 classifies the self-similar solutions into constant- and non-constant-elasticity families. We now study the constant-elasticity case in greater detail. Motivated by the notation used in Theorem 10, we write

$$\delta := \Upsilon_0 > 1$$

for the constant value of the elasticity profile throughout this section.

Corollary 19 then determines the compatible hazard rate, while Theorem 22 supplies the value function and optimal feedback. Its simple power value function permits an explicit withdrawal inversion and a transparent test of when the simplex constraint is globally nonbinding, allowing us to derive results analogous to those of Young [34]. As we will see in the next subsection, for finite right endpoints $T < \infty$, the compatible hazard rate can be interpreted as a compactification of exponential mortality.

4.1 The compactified-exponential mortality law

We first separate the mortality law generated by the constant-elasticity branch from its financial parametrization. This allows us to study its properties in a generic setting before returning to the restrictions imposed by compatibility with the investment problem.

Definition 11 (Compactified-exponential distribution) Let $\lambda, p, \kappa, \tau > 0$ be constants and set $\eta := p\kappa/(\exp(\kappa\tau) - 1)$. Further define the survival function

$$S_{\mathrm{CE}}(t) := \begin{cases} \mathrm{e}^{(\eta-\lambda)t}\left(\dfrac{1-\mathrm{e}^{-\kappa(\tau-t)}}{1-\mathrm{e}^{-\kappa\tau}}\right)^p, & 0 \le t < \tau, \\ 0, & t \ge \tau. \end{cases} \tag{41}$$

Its hazard rate is given by

$$\lambda_{\mathrm{CE}}(t) := \lambda - \eta + \frac{p\kappa}{\mathrm{e}^{\kappa(\tau-t)} - 1} \ge \lambda, \qquad 0 \le t < \tau. \tag{42}$$

We say that a nonnegative random variable $\mathcal{T}$ has a compactified-exponential distribution with parameters $(\lambda, p, \kappa, \tau)$ if its survival function is given by S_{CE}. In this case, we write $\mathcal{T} \sim \mathrm{Exp}(\lambda; p, \kappa, \tau)$.

The correction term η normalizes the finite-endpoint contribution so that $\lambda_{\mathrm{CE}}(0) = \lambda$. The hazard then increases from this initial level and diverges as $t \uparrow \tau$, forcing death before the finite endpoint τ. As the following corollary shows, moving this endpoint to infinity recovers the baseline exponential law, justifying the name "compactified-exponential" and the notation $\mathrm{Exp}(\lambda; p, \kappa, \tau)$.

Corollary 34 (Exponential limit) *Let $\mathcal{T} \sim \mathrm{Exp}(\lambda; p, \kappa, \tau)$. Then, for every fixed $t \ge 0$,*

$$S_{\mathrm{CE}}(t) \longrightarrow \mathrm{e}^{-\lambda t}, \qquad \lambda_{\mathrm{CE}}(t) \longrightarrow \lambda,$$

as $\tau \to \infty$. Consequently, $\mathrm{Exp}(\lambda; p, \kappa, \tau)$ converges in distribution to the exponential distribution $\mathrm{Exp}(\lambda)$ with rate λ.

Proof The two limits follow directly from the definitions of S_{CE} and λ_{CE} in (41) and (42), respectively, and the fact that $\eta \to 0$ as $\tau \to \infty$. The convergence in distribution follows from the convergence of the survival functions, which are the complementary cumulative distribution functions of the respective random variables. □

Remark 5 When $\lambda = 0$, the same formulas converge pointwise to $S(t) \equiv 1$ and $\lambda(t) \equiv 0$. This is a useful boundary case of the formulas, but it is not an infinite-horizon mortality model under Definition 5, which requires survival to vanish at the right endpoint.

The limiting argument explains the name of the distribution analytically. The next result provides a complementary probabilistic interpretation and an exact simulation scheme under mild conditions on the parameters: the lifetime can be expressed as the minimum of two independent random variables, one of which is exponentially distributed and the other is a transformation of a beta-distributed random variable. To this end, we define the increasing map

$$\phi_\kappa(z) := \frac{1}{\kappa}\log(1+z), \qquad z \geq 0,$$

for $\kappa > 0$. Its inverse is given by $\phi_\kappa^{-1}(t) = \mathrm{e}^{\kappa t} - 1$ for $t \geq 0$.

Proposition 35 (Competing-risks representation) *Let $\lambda, p, \kappa, \tau > 0$, and define η as in Definition 11. Under the assumption $\lambda > \eta$, let $\mathcal{T}_{\mathrm{exp}}$ and U be independent random variables such that*

$$\mathcal{T}_{\mathrm{exp}} \sim \mathrm{Exp}(\lambda - \eta), \qquad U \sim \mathrm{Beta}(1, p),$$

where the exponential distribution is parametrized by its rate. Further, define

$$\mathcal{T}_{\mathrm{fin}} := \phi_\kappa\left(\left(\mathrm{e}^{\kappa\tau} - 1\right)U\right), \qquad \mathcal{T} := \mathcal{T}_{\mathrm{exp}} \wedge \mathcal{T}_{\mathrm{fin}}.$$

Then $\mathcal{T} \sim \mathrm{Exp}(\lambda; p, \kappa, \tau)$.

Proof Let $0 \leq t < \tau$, and recall that the density of a $\mathrm{Beta}(1, p)$ random variable is given by $f_U(u) = p(1-u)^{p-1}\mathbf{1}_{\{0 \leq u \leq 1\}}$. Hence,

$$\mathbb{P}(U > u) = \int_u^1 p(1-v)^{p-1}\, dv = (1-u)^p.$$

Therefore,

$$\mathbb{P}(\mathcal{T}_{\mathrm{fin}} > t) = \mathbb{P}\left(U > \frac{\mathrm{e}^{\kappa t} - 1}{\mathrm{e}^{\kappa\tau} - 1}\right) = \left(1 - \frac{\mathrm{e}^{\kappa t} - 1}{\mathrm{e}^{\kappa\tau} - 1}\right)^p = \left(\frac{1 - \mathrm{e}^{-\kappa(\tau - t)}}{1 - \mathrm{e}^{-\kappa\tau}}\right)^p.$$

Because $\mathcal{T}_{\mathrm{fin}}$ is a function of U, it is independent of $\mathcal{T}_{\mathrm{exp}}$. Thus, for $0 \leq t < \tau$, the survival function of $\mathcal{T}$ is

$$S_{\mathcal{T}}(t) = \mathbb{P}\big(\mathcal{T}_{\mathrm{exp}} > t, \mathcal{T}_{\mathrm{fin}} > t\big) = \mathbb{P}(\mathcal{T}_{\mathrm{exp}} > t)\mathbb{P}(\mathcal{T}_{\mathrm{fin}} > t) = \mathrm{e}^{(\eta - \lambda)t}\left(\frac{1 - \mathrm{e}^{-\kappa(\tau - t)}}{1 - \mathrm{e}^{-\kappa\tau}}\right)^p.$$

Finally, since $\mathcal{T}$ is non-negative and bounded above by $\mathcal{T}_{\mathrm{fin}}$, $S_{\mathcal{T}}(t) = 0$ for $t \geq \tau$. Hence, $S_{\mathcal{T}} = S_{\mathrm{CE}}$, and the result follows. □

Thus far, the compactified-exponential family has been considered as a stand-alone mortality model. We now return to the control problem. The next theorem identifies the member selected by constant elasticity and shows that this choice yields the power value function and feedback rule obtained in Section 3. In particular, we obtain results that are structurally identical to those of Young [34] for the exponential case, but now for a finite horizon.

Theorem 36 (Compatibility of compactified-exponential mortality) *Let $T \in (0,\infty)$, $\delta > 1 + \rho/\mathcal{E}(0)$, and define*

$$\lambda_\infty := \delta\left(\mathcal{E}(0) - \frac{\rho}{\delta - 1}\right) > 0. \tag{43}$$

If $\mathcal{T} \sim \mathrm{Exp}(\lambda_\infty; \delta, r, T)$, then $\Upsilon \equiv \delta$ is compatible with λ_{CE} for any admissible choice of scaling function b in the sense of Definition 10. The corresponding minimal lifetime-impoverishment probability in the unconstrained case is

$$p(t,x;C) = z(t,x)^\delta = \left(\frac{x_{\mathrm{safe}}(t) - x}{x_{\mathrm{safe}}(t) - x_{\mathrm{imp}}}\right)^\delta, \qquad (t,x) \in \overline{\mathcal{D}}^x, \tag{44}$$

and the corresponding optimal risky-fraction strategy is given on $\mathcal{D}$ by

$$\pi^*(t,x) = \frac{x_{\mathrm{safe}}(t) - x}{x(\delta - 1)}\mathfrak{m}, \tag{45}$$

and is extended by zero outside $\mathcal{D}$.

Proof The assumed lower bound on δ implies

$$(\delta - 1)\mathcal{E}(0) > \rho.$$

Since $\delta > 1$, it follows that

$$\mathcal{E}(0) - \frac{\rho}{\delta - 1} > 0,$$

which verifies the stated inequality $\lambda_\infty > 0$.

We next identify the compatible hazard. If $\Upsilon \equiv \delta$, then $\mathcal{N}[\Upsilon] = \delta/(\delta - 1)$, and the compatibility equation (20) becomes

$$\left(\frac{b'(t)}{b(t)} + \mathcal{E}(t)\right)\delta - \rho\frac{\delta}{\delta - 1} = \lambda(t) + \frac{b'(t)}{b(t)}\delta.$$

Cancelling $b'(t)\delta/b(t)$ from both sides and factoring out δ gives

$$\lambda(t) = \delta\left(\mathcal{E}(t) - \frac{\rho}{\delta - 1}\right). \tag{46}$$

Add and subtract $\mathcal{E}(0)$ inside the parentheses. By (43), this gives

$$\begin{aligned}\lambda(t) &= \delta\left(\mathcal{E}(0) - \frac{\rho}{\delta - 1}\right) + \delta\big(\mathcal{E}(t) - \mathcal{E}(0)\big) \\ &= \lambda_\infty + \delta\big(\mathcal{E}(t) - \mathcal{E}(0)\big).\end{aligned} \tag{47}$$

By Definition 9,

$$\mathcal{E}(t) = \frac{r}{1 - \mathrm{e}^{-r(T-t)}} = r + \frac{r}{\mathrm{e}^{r(T-t)} - 1},$$

$$\mathcal{E}(0) = r + \frac{r}{\mathrm{e}^{rT} - 1}.$$

Consequently,

$$\delta\big(\mathcal{E}(t) - \mathcal{E}(0)\big) = -\frac{\delta r}{\mathrm{e}^{rT} - 1} + \frac{\delta r}{\mathrm{e}^{r(T-t)} - 1}, \tag{48}$$

and (47) turns into

$$\lambda(t) = \lambda_\infty - \frac{\delta r}{\mathrm{e}^{rT} - 1} + \frac{\delta r}{\mathrm{e}^{r(T-t)} - 1}. \tag{49}$$

It remains to match these constants with Definition 11. In the present specialization, its four parameters are

$$(\lambda, p, \kappa, \tau) = (\lambda_\infty, \delta, r, T).$$

Hence, the constant $\eta = p\kappa/(\mathrm{e}^{\kappa\tau} - 1)$ appearing in that definition becomes

$$\eta = \frac{\delta r}{\mathrm{e}^{rT} - 1}, \tag{50}$$

which together with (49) yields

$$\lambda(t) = \lambda_\infty - \eta + \frac{\delta r}{\mathrm{e}^{r(T-t)} - 1} = \lambda_{\mathrm{CE}}(t).$$

Thus, (46) holds with the compactified-exponential hazard. Because b dropped out when Υ was set equal to δ, the triple is compatible for every admissible choice of b.

Finally, Corollary 19 and Theorem 22 yield the stated value function and optimal feedback. □

As we will see in the next subsection, the compatible compactified-exponential law can be interpreted as a modified Gompertz–Makeham law plus a non-negative remainder term consisting of an infinite series of Gompertz-type age components. This interpretation is useful for understanding the behavior of the compatible hazard rate and its relation to empirical mortality data.

4.2 Relation to Gompertz–Makeham mortality

A common parametric model of mortality is the Gompertz–Makeham law, which is widely used in actuarial science and demography. It is a three-parameter distribution that generalizes the Gompertz law by adding a constant term to the hazard rate. The following definition formalizes this law.

Definition 12 (Gompertz–Makeham law) Let $\alpha, \beta > 0$ and $\lambda \geq 0$. A nonnegative random variable $\mathcal{T}$ has a Gompertz–Makeham distribution with parameters (α, β, λ) if its survival function is given by

$$S_{\mathrm{GM}}(t) := \exp\left(-\lambda t - \frac{\alpha}{\beta}\big(\mathrm{e}^{\beta t} - 1\big)\right), \qquad t \geq 0, \tag{51}$$

and its hazard rate is given by

$$\lambda_{\mathrm{GM}}(t) := \lambda + \alpha \mathrm{e}^{\beta t}, \qquad t \geq 0. \tag{52}$$

In this case, we write $\mathcal{T} \sim \mathrm{GM}(\alpha, \beta, \lambda)$. The boundary case $\lambda = 0$ is the Gompertz law.

The exponential age dependence in (52) suggests comparing the two families at the level of their hazard rates. As we will see, decomposing the compatible compactified-exponential hazard in (49) into a constant term and a time-dependent term reveals a hierarchy of Gompertz-type exponential components and clarifies both the connection to, and the distinction from, a single Gompertz–Makeham hazard.

For the compatible law in Theorem 36, the normalizing factor $\mathcal{E}$ gives the exact representation

$$\lambda_{\mathrm{CE}}(t) = \delta\left(\mathcal{E}(t) - \frac{\rho}{\delta - 1}\right) = \delta\left(r - \frac{\rho}{\delta - 1}\right) + \frac{\delta r}{\mathrm{e}^{r(T-t)} - 1}.$$

For every $0 \le t < T$, expanding the time-dependent term as a geometric series and extracting the first summand gives the infinite series representation

$$\begin{aligned}\lambda_{\mathrm{CE}}(t) &= \delta\left(r - \frac{\rho}{\delta - 1}\right) + \delta r \mathrm{e}^{-rT}\mathrm{e}^{rt} + \delta r \sum_{n=2}^{\infty} \mathrm{e}^{-nr(T-t)} \\ &= \lambda + \alpha \mathrm{e}^{\beta t} + \widetilde{R}(t), \end{aligned} \tag{53}$$

where $\lambda = \delta\,(r - \rho/(\delta - 1))$, $\alpha = \delta r \mathrm{e}^{-rT}$, $\beta = r$, and

$$\widetilde{R}(t) := \delta r \sum_{n=2}^{\infty} \mathrm{e}^{-nr(T-t)} \ge 0 \tag{54}$$

is the remainder term. Hence, by letting

$$\alpha_n := \delta r \mathrm{e}^{-nrT}, \qquad \beta_n := nr, \qquad n \ge 1,$$

the remainder term can be written as

$$\widetilde{R}(t) = \sum_{n=2}^{\infty} \alpha_n \mathrm{e}^{\beta_n t}.$$

Note that $\alpha = \alpha_1$ and $\beta = \beta_1$ are the parameters of the first exponential term in (53). Hence, each summand in (54) has the exponential form of the age-dependent component of a Gompertz–Makeham hazard, with exponential growth rate $\beta_n = nr$. The compatible hazard is therefore linked to Gompertz-type mortality through an infinite series of exponential age components.

Its expansion parameter $\exp\bigl(-r(T-t)\bigr)$ increases to one as $t \uparrow T$, so no finite truncation is uniform near the endpoint. Indeed, the logarithmic slope of the time-dependent term satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t} \ln\left(\frac{\delta r}{\mathrm{e}^{r(T-t)} - 1}\right) = \mathcal{E}(t).$$

It exceeds r and diverges as $t \uparrow T$. Thus, the higher exponential terms generate the increasing mortality acceleration and the finite right endpoint. However, the constant

term λ in (53) is negative if $\delta < 1 + \rho/r$, so the non-remainder terms in (53) do not form a Gompertz–Makeham hazard in their own right. The following remark resolves this issue by centering the first exponential term.

Remark 6 (Centering the first exponential term) Recall that the compatible hazard of a compactified-exponential law can be written as

$$\lambda_{\mathrm{CE}}(t) = \lambda_\infty - \underbrace{\frac{\delta r}{\mathrm{e}^{rT} - 1}}_{\eta} + \frac{\delta r}{\mathrm{e}^{r(T-t)} - 1}.$$

Both the subtracted term η and the time-dependent term can be expanded as geometric series. Hence, for $0 \le t < T$,

$$\begin{aligned}\lambda_{\mathrm{CE}}(t) &= \lambda_\infty - \delta r \sum_{n=1}^{\infty} \mathrm{e}^{-nrT} + \delta r \sum_{n=1}^{\infty} \mathrm{e}^{-nrT}\mathrm{e}^{nrt} \\ &= \lambda_\infty + \delta r \sum_{n=1}^{\infty} \mathrm{e}^{-nrT}\big(\mathrm{e}^{nrt} - 1\big).\end{aligned}$$

Thus, the term $-\eta$ subtracts the value of the exponential series at $t = 0$. Equivalently, it replaces every exponential e^{nrt} by its centered counterpart $\mathrm{e}^{nrt} - 1$. Each summand therefore vanishes at $t = 0$, and $\lambda_{\mathrm{CE}}(0) = \lambda_\infty$.

Now set

$$\gamma := \delta r \mathrm{e}^{-rT}$$

and isolate the summand corresponding to $n = 1$. Then

$$\lambda_{\mathrm{CE}}(t) = \lambda_\infty + \gamma\big(\mathrm{e}^{rt} - 1\big) + R(t),$$

where

$$R(t) := \delta r \sum_{n=2}^{\infty} \mathrm{e}^{-nrT}\big(\mathrm{e}^{nrt} - 1\big) \ge 0.$$

Notice that $R(0) = 0$ as well. Thus, the compatible hazard is structurally similar to a Gompertz–Makeham hazard plus a non-negative remainder term. In particular,

$$\lambda_\infty + \gamma\big(\mathrm{e}^{rt} - 1\big) = \lambda_\infty - \gamma + \gamma\mathrm{e}^{rt}.$$

Hence, if $\lambda_\infty \ge \gamma$, the displayed term is the hazard of a $\mathrm{GM}(\gamma, r, \lambda_\infty - \gamma)$ distribution, and the compatible hazard can be interpreted as a Gompertz–Makeham hazard plus a non-negative remainder term. The condition $\lambda_\infty \ge \gamma$ is equivalent to

$$\delta \ge 1 + \frac{\rho}{\mathcal{E}(0) - r\mathrm{e}^{-rT}},$$

which is stronger than the assumption $\delta > 1 + \rho/\mathcal{E}(0)$ in Theorem 36. Since $\mathcal{E}(0) - r\mathrm{e}^{-rT} > r$, the simpler bound $\delta > 1 + \rho/r$ is sufficient, but not necessary, for $\lambda_\infty \ge \gamma$.

Because the riskless rate r need not be close to an empirically fitted rate of aging β of a $\mathrm{GM}(\alpha, \beta, \lambda)$ distribution, whether the compatible compactified-exponential law reproduces a Gompertz–Makeham survival profile over the relevant age range is a numerical question. Section 6 provides a numerical comparison of the two laws and shows that the compactified-exponential law is indeed able to capture the Gompertz–Makeham survival profile over the relevant age range.

This completes the mortality-side analysis of the constant-elasticity family. We now exploit the associated power value function on the policy side. Holding the compatible mortality law fixed, the next subsection varies the withdrawal rate and inverts the initial impoverishment probability to obtain an explicit sustainability threshold and optimal withdrawal rate.

4.3 Policy consequences and feasibility under constraints

For the remainder of the section, write

$$\varepsilon_T := 1 - \mathrm{e}^{-rT}.$$

When the withdrawal rate varies, we hold the mortality law, the market parameters, and the impoverishment level x_{imp} fixed. This keeps the survival law in Theorem 36 fixed and changes only the wealth dynamics and the safe boundary.

Corollary 37 (Sustainable withdrawal and required initial capital) *Fix $x_0 > x_{\mathrm{imp}}$ and $\alpha \in (0,1)$, and set*

$$\beta := \frac{\alpha^{1/\delta}}{1-\alpha^{1/\delta}}.$$

Then the optimal sustainable withdrawal rate defined in (4) is

$$\begin{aligned} C^*(x_0;\alpha) &= C_{\mathrm{safe}}(x_0) + \frac{r\beta}{\varepsilon_T}\big(x_0 - x_{\mathrm{imp}}\big) \\ &= r\left(x_{\mathrm{imp}} + \frac{\beta+1}{\varepsilon_T}\big(x_0 - x_{\mathrm{imp}}\big)\right). \end{aligned} \tag{55}$$

Conversely, for any withdrawal rate $C > r x_{\mathrm{imp}}$, the minimal initial wealth required to achieve an impoverishment probability of at most α is

$$\begin{aligned} x_0(C;\alpha) &= x_{\mathrm{safe}}(0) - \frac{\beta}{\beta+1}\big(x_{\mathrm{safe}}(0) - x_{\mathrm{imp}}\big) \\ &= x_{\mathrm{imp}} + \frac{\varepsilon_T}{\beta+1}\left(\frac{C}{r} - x_{\mathrm{imp}}\right), \end{aligned} \tag{56}$$

where $x_{\mathrm{safe}}(0)$ is evaluated at the target withdrawal rate C.

Proof From (5),

$$x_{\mathrm{safe}}(0) - x_{\mathrm{imp}} = \left(\frac{C}{r} - x_{\mathrm{imp}}\right)\varepsilon_T.$$

On the nontrivial part of the state space, (44) therefore gives

$$p(0,x_0;C)^{1/\delta} = \frac{x_{\mathrm{safe}}(0) - x_0}{x_{\mathrm{safe}}(0) - x_{\mathrm{imp}}} = 1 - \frac{x_0 - x_{\mathrm{imp}}}{(C/r - x_{\mathrm{imp}})\varepsilon_T}.$$

By the definition of β,

$$\alpha^{1/\delta} = \frac{\beta}{\beta+1}, \qquad 1-\alpha^{1/\delta} = \frac{1}{\beta+1}.$$

Solving $p(0, x_0; C) = \alpha$ for C therefore gives

$$\left(\frac{C}{r} - x_{\text{imp}}\right) \varepsilon_T = (\beta + 1)(x_0 - x_{\text{imp}}),$$

and hence

$$\begin{aligned} C &= r x_{\text{imp}} + \frac{r(\beta+1)}{\varepsilon_T}(x_0 - x_{\text{imp}}) \\ &= \left(r x_{\text{imp}} + \frac{r}{\varepsilon_T}(x_0 - x_{\text{imp}})\right) + \frac{r\beta}{\varepsilon_T}(x_0 - x_{\text{imp}}) \\ &= C_{\text{safe}}(x_0) + \frac{r\beta}{\varepsilon_T}(x_0 - x_{\text{imp}}), \end{aligned}$$

where the last equality follows by definition of $C_{\text{safe}}(x_0)$. This proves (55). The right-hand side of the preceding expression for $p(0, x_0; C)^{1/\delta}$ is strictly increasing in C whenever $x_0 < x_{\text{safe}}(0)$. Withdrawal rates with $x_0 \geq x_{\text{safe}}(0)$ have zero impoverishment probability, whereas rates above the solution produce a probability larger than α. Thus, the solution is the supremum in (4).

Conversely, solving the same equality for x_0 gives

$$\begin{aligned} x_0 &= x_{\text{safe}}(0) - \alpha^{1/\delta}\big(x_{\text{safe}}(0) - x_{\text{imp}}\big) \\ &= x_{\text{safe}}(0) - \frac{\beta}{\beta+1}\big(x_{\text{safe}}(0) - x_{\text{imp}}\big). \end{aligned}$$

Substitution of $x_{\text{safe}}(0) - x_{\text{imp}} = \varepsilon_T(C/r - x_{\text{imp}})$ yields the second representation in (56). Since $p(0, \cdot; C)$ is strictly decreasing on $(x_{\text{imp}}, x_{\text{safe}}(0))$, this is the minimal initial wealth for which the impoverishment probability is at most α. □

Remark 7 Equation (55) adds a risk-taking bonus to the safe withdrawal rate $C_{\text{safe}}(x_0)$, whereas (56) subtracts the corresponding capital relief from the safe initial capital $x_{\text{safe}}(0)$. As $\alpha \downarrow 0$, one has $\beta \downarrow 0$, so both adjustments vanish.

Moreover, as $T \to \infty$,

$$\varepsilon_T \longrightarrow 1, \qquad C_{\text{safe}}(x_0) \longrightarrow r x_0, \qquad x_{\text{safe}}(0) \longrightarrow x_{\text{safe}}^{\infty} = \frac{C}{r}.$$

Consequently, (55) and (56) converge exactly to (15) and (16), respectively. Thus, the constant-hazard results in Corollary 12 are recovered as the infinite-horizon benchmark.

In Theorem 9, we showed under which conditions the unconstrained feedback is also optimal for the simplex-constrained problem. Due to the explicit form of the feedback in (45), we can now give a simple condition for this coincidence to hold in the compactified-exponential case.

Proposition 38 (Coincidence with the simplex-constrained problem) *Assume that* $\mathfrak{m} \geq 0$ *componentwise, and set*

$$M := \mathbf{1}^{\top}\mathfrak{m} = \|\mathfrak{m}\|_1.$$

The feedback (45) *is optimal for the simplex-constrained problem, and the constrained and unconstrained value functions coincide if and only if*

$$x_{\text{imp}} \geq \frac{C}{r}\,\frac{\varepsilon_T M}{\delta - 1 + \varepsilon_T M} \tag{57}$$

holds. Under this condition, the feedback belongs to K_{con} throughout $\mathcal{D}$. As $T \to \infty$, condition (57) *becomes*

$$x_{\mathrm{imp}} \geq \frac{C}{r}\frac{M}{\delta - 1 + M}.$$

Proof For the first claim, we check the conditions in Theorem 9. As $\mathfrak{m} \geq 0$ componentwise, it remains to check (14). From (44), we have

$$\partial_x p(t, x; C) = -\frac{\delta\big(x_{\mathrm{safe}}(t) - x\big)^{\delta - 1}}{\big(x_{\mathrm{safe}}(t) - x_{\mathrm{imp}}\big)^{\delta}}$$

and

$$\partial_{xx} p(t, x; C) = \frac{\delta(\delta - 1)\big(x_{\mathrm{safe}}(t) - x\big)^{\delta - 2}}{\big(x_{\mathrm{safe}}(t) - x_{\mathrm{imp}}\big)^{\delta}}.$$

In particular,

$$\frac{x\partial_{xx} p(t, x; C)}{\partial_x p(t, x; C)} = -\frac{(\delta - 1)x}{x_{\mathrm{safe}}(t) - x}.$$

Consequently, the inequality in Theorem 9 is equivalent to

$$M\big(x_{\mathrm{safe}}(t) - x\big) \leq (\delta - 1)x$$

for every $(t, x) \in \mathcal{D}$.

The difference

$$(\delta - 1)x - M\big(x_{\mathrm{safe}}(t) - x\big) = (\delta - 1 + M)x - Mx_{\mathrm{safe}}(t)$$

is increasing in x. Moreover, x_{safe} is decreasing in t. Hence, the preceding inequality holds throughout $\mathcal{D}$ if and only if

$$M\big(x_{\mathrm{safe}}(0) - x_{\mathrm{imp}}\big) \leq (\delta - 1)x_{\mathrm{imp}}.$$

Using $x_{\mathrm{safe}}(0) - x_{\mathrm{imp}} = \big(C/r - x_{\mathrm{imp}}\big)\varepsilon_T$, this becomes

$$\varepsilon_T M\left(\frac{C}{r} - x_{\mathrm{imp}}\right) \leq (\delta - 1)x_{\mathrm{imp}},$$

which is equivalent to (57). This proves the equivalence of the constrained and unconstrained value functions.

Finally, letting $T \to \infty$ gives $\varepsilon_T \to 1$ and hence the stated limiting condition. □

One should note that the quantity M as defined in Proposition 38 is strictly positive, if the Merton vector $\mathfrak{m}$ is componentwise non-negative. This is due to our assumption of the existence of at least one risky asset with non-zero risk premium.

Combining Proposition 38 with the sustainable withdrawal formula (55) eliminates the endpoint factor ε_T and yields a condition directly in terms of the investor's wealth and accepted risk level.

For the next results, it is convenient to define the *relative wealth bandwidth* as

$$\mathcal{B}_{\mathrm{rel}} := \frac{x_0 - x_{\mathrm{imp}}}{x_{\mathrm{imp}}} = \frac{x_0}{x_{\mathrm{imp}}} - 1. \tag{58}$$

It measures the relative distance between the initial wealth and the impoverishment level. The following corollary shows that, for a given risk level, the relative wealth bandwidth must be sufficiently small in order for the unconstrained feedback to remain feasible under the simplex constraint.

Corollary 39 (Risk levels compatible with the simplex constraint) *Under the assumptions of Proposition 38, the optimal unconstrained feedback corresponding to the sustainable withdrawal rate $C^*(x_0;\alpha)$ is simplex-valued throughout the risky region if and only if*

$$\mathcal{B}_{\mathrm{rel}} \le \frac{\delta-1}{M}\big(1-\alpha^{1/\delta}\big). \tag{59}$$

Equivalently, if the relative wealth bandwidth is fixed and satisfies

$$\mathcal{B}_{\mathrm{rel}} < \frac{\delta-1}{M},$$

define the maximal risk level compatible with the simplex constraint as

$$\alpha_{\max}(x_0, x_{\mathrm{imp}}) := \left(1 - \frac{M}{\delta-1}\mathcal{B}_{\mathrm{rel}}\right)^{\delta}. \tag{60}$$

Then (59) is equivalent to

$$0 < \alpha \le \alpha_{\max}(x_0, x_{\mathrm{imp}}).$$

If $\mathcal{B}_{\mathrm{rel}} = (\delta-1)/M$, no risk level $\alpha \in (0,1)$ satisfies the simplex-feasibility condition. Feasibility is attained only in the continuously extended boundary case $\alpha = 0$, where the sustainable withdrawal rate reduces to the safe withdrawal rate and is financed by riskless investment alone. If $\mathcal{B}_{\mathrm{rel}} > (\delta-1)/M$, the condition cannot be satisfied even under this extension.

Proof Apply Proposition 38 with $C = C^*(x_0;\alpha)$. Since $\delta - 1 + \varepsilon_T M > 0$, condition (57) is equivalent to

$$\big(\delta - 1 + \varepsilon_T M\big)x_{\mathrm{imp}} \ge \varepsilon_T M \frac{C^*(x_0;\alpha)}{r},$$

$$(\delta-1)x_{\mathrm{imp}} \ge \varepsilon_T M\left(\frac{C^*(x_0;\alpha)}{r} - x_{\mathrm{imp}}\right).$$

By (6),

$$\frac{C_{\mathrm{safe}}(x_0)}{r} - x_{\mathrm{imp}} = \frac{x_0 - x_{\mathrm{imp}}}{\varepsilon_T}.$$

Combining this identity with (55) gives

$$\begin{aligned}\frac{C^*(x_0;\alpha)}{r} - x_{\mathrm{imp}} &= \left(\frac{C_{\mathrm{safe}}(x_0)}{r} - x_{\mathrm{imp}}\right) + \frac{\beta}{\varepsilon_T}(x_0 - x_{\mathrm{imp}}) \\ &= \frac{\beta+1}{\varepsilon_T}(x_0 - x_{\mathrm{imp}}).\end{aligned}$$

Hence, the endpoint factor cancels and the simplex condition becomes

$$(\delta-1)x_{\mathrm{imp}} \ge M(\beta+1)(x_0 - x_{\mathrm{imp}}).$$

Since

$$\beta + 1 = \frac{1}{1-\alpha^{1/\delta}},$$

this is equivalent to

$$M(x_0 - x_{\mathrm{imp}}) \le (\delta-1)\big(1-\alpha^{1/\delta}\big)x_{\mathrm{imp}}.$$

Dividing by $x_{\mathrm{imp}} > 0$ and using the definition of $\mathcal{B}_{\mathrm{rel}}$ proves (59).

It remains to solve this condition for α. Rearranging gives

$$\alpha^{1/\delta} \le 1 - \frac{M}{\delta-1}\mathcal{B}_{\mathrm{rel}}.$$

If $\mathcal{B}_{\text{rel}} < (\delta-1)/M$, the right-hand side belongs to $(0,1]$. Since $z \mapsto z^{\delta}$ is strictly increasing on $[0,\infty)$, the preceding inequality is equivalent to

$$0 < \alpha \le \left(1 - \frac{M}{\delta - 1}\mathcal{B}_{\text{rel}}\right)^{\delta} = \alpha_{\max}(x_0, x_{\text{imp}}).$$

If $\mathcal{B}_{\text{rel}} = (\delta-1)/M$, the right-hand side is zero. Hence, no $\alpha \in (0,1)$ satisfies the inequality. In the continuous boundary limit $\alpha \downarrow 0$, however,

$$\alpha^{1/\delta} \longrightarrow 0, \qquad \beta \longrightarrow 0, \qquad C^*(x_0;\alpha) \longrightarrow C_{\text{safe}}(x_0).$$

Thus, the boundary equality corresponds economically to financing the withdrawals by riskless investment alone. Finally, if $\mathcal{B}_{\text{rel}} > (\delta-1)/M$, the right-hand side is negative, whereas $\alpha^{1/\delta} \ge 0$, so the condition cannot hold even at the continuously extended boundary $\alpha = 0$. □

The same formulas can be read in the inverse direction. The next result chooses the impoverishment level and initial wealth so as to maximize the relative wealth range compatible with a target withdrawal and a globally nonbinding simplex constraint.

Corollary 40 (Maximal relative wealth bandwidth at a target withdrawal) *Retain the assumptions of Corollary 39. Fix a target sustainable withdrawal rate $\overline{C} > 0$ and a risk level $\alpha \in (0,1)$, and consider all pairs $0 < x_{\text{imp}} < x_0$ satisfying*

$$C^*(x_0;\alpha) = \overline{C}$$

for which the unconstrained feedback is simplex-valued throughout the risky region.

Then, the relative wealth bandwidth has the unique maximum

$$\mathcal{B}^*_{\text{rel}}(\alpha) := \frac{\delta-1}{M}\big(1-\alpha^{1/\delta}\big). \tag{61}$$

This maximum is attained at the unique pair

$$x^*_{\text{imp}} := \frac{\overline{C}}{r} \cdot \frac{\varepsilon_T M}{\delta - 1 + \varepsilon_T M} \tag{62}$$

and

$$x^*_0(\alpha) := x^*_{\text{imp}}\big(1 + \mathcal{B}^*_{\text{rel}}(\alpha)\big). \tag{63}$$

At this pair, the simplex-feasibility threshold is attained with equality, and

$$\alpha_{\max}(x^*_0(\alpha), x^*_{\text{imp}}) = \alpha.$$

Proof To shorten notation, set

$$k_\alpha := \varepsilon_T\big(1-\alpha^{1/\delta}\big) \in (0,1), \qquad s := \frac{\varepsilon_T M}{\delta - 1 + \varepsilon_T M} \in [0,1).$$

By (55), the calibration condition $C^*(x_0;\alpha) = \overline{C}$ is equivalent to

$$x_0 = k_\alpha \frac{\overline{C}}{r} + (1-k_\alpha)x_{\text{imp}}.$$

This defines a line in the (x_{imp}, x_0) plane along which the optimal withdrawal rate is constant at $\overline{C}$. By Proposition 38, global simplex feasibility is equivalent along this calibration line to

$$x_{\text{imp}} \ge \frac{s}{r}\overline{C}.$$

Moreover, note that, for every $0 < x_{\mathrm{imp}} < \overline{C}/r$,

$$\mathcal{B}_{\mathrm{rel}} = \frac{x_0}{x_{\mathrm{imp}}} - 1 = k_\alpha \left(\frac{\overline{C}}{r x_{\mathrm{imp}}} - 1 \right). \tag{64}$$

Since $M > 0$, we have $s \in (0,1)$. The non-degeneracy condition $\overline{C} > r x_{\mathrm{imp}}$, requires $x_{\mathrm{imp}} < \overline{C}/r$. Hence, the feasible pairs are parametrized by

$$\frac{s}{r}\overline{C} \le x_{\mathrm{imp}} < \frac{\overline{C}}{r}, \qquad x_0 = k_\alpha \frac{\overline{C}}{r} + (1 - k_\alpha) x_{\mathrm{imp}}.$$

Along this interval (64) is strictly decreasing in x_{imp}. It therefore attains its unique maximum at the lower endpoint, which gives

$$x^*_{\mathrm{imp}} = \frac{\overline{C}}{r} s = \frac{\overline{C}}{r} \frac{\varepsilon_T M}{\delta - 1 + \varepsilon_T M}.$$

At this point,

$$\mathcal{B}^*_{\mathrm{rel}}(\alpha) = k_\alpha \left(\frac{1}{s} - 1 \right) = \varepsilon_T \big(1 - \alpha^{1/\delta}\big) \frac{\delta - 1}{\varepsilon_T M} = \frac{\delta - 1}{M} \big(1 - \alpha^{1/\delta}\big),$$

which proves (61) and (62). The identity $x^*_0(\alpha) = x^*_{\mathrm{imp}}(1 + \mathcal{B}^*_{\mathrm{rel}}(\alpha))$ follows directly from the definition of $\mathcal{B}_{\mathrm{rel}}$, proving (63). Finally,

$$1 - \frac{M}{\delta - 1} \mathcal{B}^*_{\mathrm{rel}}(\alpha) = \alpha^{1/\delta},$$

and substitution into (60) yields $\alpha_{\max}(x^*_0(\alpha), x^*_{\mathrm{imp}}) = \alpha$. □

Motivated by this observation, we introduce the following notation.

Definition 13 (Calibration line) Fix a target sustainable withdrawal rate $\overline{C} > 0$ and a risk level $\alpha \in (0,1)$, and set

$$k_\alpha := \varepsilon_T \big(1 - \alpha^{1/\delta}\big) \in (0,1)$$

as in the proof above. The affine relation

$$x_0 = k_\alpha \frac{\overline{C}}{r} + (1 - k_\alpha) x_{\mathrm{imp}} \tag{65}$$

is called the calibration line associated with $(\overline{C}, \alpha)$.

By (55), an economically admissible pair $0 < x_{\mathrm{imp}} < x_0$ lies on this line if and only if

$$C^*(x_0; \alpha) = \overline{C}.$$

Moreover, the affine relation (65) produces an economically admissible pair if and only if $0 < x_{\mathrm{imp}} < \overline{C}/r$. For such pairs, we will write $(x_{\mathrm{imp}}, x_0) \in \mathcal{C}_{\overline{C},\alpha}$ to indicate that (x_{imp}, x_0) lies on the calibration line associated with $(\overline{C}, \alpha)$.

Corollary 40 identifies the largest globally simplex-feasible relative wealth bandwidth along a calibration line. Exceeding this threshold means that borrowing is possible somewhere in the risky region. As the next corollary shows, however, this possibility is confined to an initial time interval.

Corollary 41 (Finite window of possible borrowing) *Retain the compactified-exponential setting and assume* $\mathfrak{m} \geq 0$ *componentwise with* $M = \mathbf{1}^\top \mathfrak{m} > 0$. *Fix* $\overline{C} > 0$ *and* $\alpha \in (0,1)$, *and let*

$$(x_{\mathrm{imp}}, x_0) \in \mathcal{C}_{\overline{C},\alpha}.$$

Further, define

$$\zeta := \frac{\mathcal{B}_{\mathrm{rel}}}{\mathcal{B}^*_{\mathrm{rel}}(\alpha)} = \frac{\varepsilon_T M}{\delta - 1}\left(\frac{\overline{C}}{r x_{\mathrm{imp}}} - 1\right).$$

If $\zeta \leq 1$, *the feedback is simplex-valued throughout* $\mathcal{D}$. *If* $\zeta > 1$, *set*

$$t^*_{\mathrm{bor}} := T + \frac{1}{r}\ln\left(1 - \frac{\varepsilon_T}{\zeta}\right) \in (0,T). \tag{66}$$

At time $t \in [0,T)$, *there exists a solvent wealth state* $x \in (x_{\mathrm{imp}}, x_{\mathrm{safe}}(t))$ *at which the feedback requires borrowing, that is,* $\mathbf{1}^\top \pi^*(t,x) > 1$, *if and only if* $t < t^*_{\mathrm{bor}}$. *In particular, the feedback is simplex-valued at every solvent wealth state for* $t \in [t^*_{\mathrm{bor}}, T)$.

Proof The second representation of ζ follows directly from (58), (65), and (61).

As shown in the proof of Proposition 38, borrowing at a point $(t,x) \in \mathcal{D}$ occurs exactly when

$$M\big(x_{\mathrm{safe}}(t) - x\big) > (\delta - 1)x.$$

The difference between the two sides decreases strictly with x. Hence, at a fixed time $t < T$, there is a solvent wealth state requiring borrowing if and only if

$$M\big(x_{\mathrm{safe}}(t) - x_{\mathrm{imp}}\big) > (\delta - 1)x_{\mathrm{imp}}.$$

Evaluating (5) at $C = \overline{C}$ and then using the preceding identity for ζ, we obtain

$$\frac{M\big(x_{\mathrm{safe}}(t) - x_{\mathrm{imp}}\big)}{(\delta - 1)x_{\mathrm{imp}}} = \frac{M}{\delta - 1}\left(\frac{\overline{C}}{r x_{\mathrm{imp}}} - 1\right)\big(1 - \mathrm{e}^{-r(T-t)}\big) = \zeta\frac{1 - \mathrm{e}^{-r(T-t)}}{\varepsilon_T}.$$

The fraction on the right decreases strictly from 1 at $t = 0$ to 0 as $t \uparrow T$. Thus, if $\zeta \leq 1$, borrowing is impossible at every time. If $\zeta > 1$, the borrowing condition holds precisely until the unique time at which

$$1 - \mathrm{e}^{-r(T-t)} = \frac{\varepsilon_T}{\zeta}.$$

Solving this equation gives (66). Since $\zeta > 1$, we have

$$\mathrm{e}^{-rT} < 1 - \frac{\varepsilon_T}{\zeta} < 1,$$

so $t^*_{\mathrm{bor}} \in (0,T)$. At the cutoff the strict borrowing inequality fails, and it remains false thereafter. As $\mathfrak{m} \geq 0$ componentwise, there are no short risky positions, so the feedback is simplex-valued at every solvent wealth state from that time onward. □

The preceding corollary identifies when borrowing can occur at some solvent wealth level, but it does not determine whether the initial wealth lies in that region. The next result gives a separate threshold for borrowing at the outset.

Proposition 42 (Initial borrowing) *Retain the compactified-exponential setting and assume* $\mathfrak{m} \geq 0$ *componentwise, with* $M = \mathbf{1}^\top \mathfrak{m} > 0$. *Fix* $\overline{C} > 0$ *and* $\alpha \in (0,1)$, *and let*

$$(x_{\mathrm{imp}}, x_0) \in \mathcal{C}_{\overline{C},\alpha}$$

have relative wealth bandwidth excess $\zeta > 1$.

If

$$\alpha^{1/\delta} > \mathcal{B}^*_{\mathrm{rel}}(\alpha),$$

define

$$\zeta_{\mathrm{init}} := \frac{1}{\alpha^{1/\delta} - \mathcal{B}^*_{\mathrm{rel}}(\alpha)}.$$

Then $\zeta_{\mathrm{init}} > 1$*, and the optimal feedback requires borrowing at* $t = 0$ *if and only if* $\zeta > \zeta_{\mathrm{init}}$*. At* $\zeta = \zeta_{\mathrm{init}}$*, the total risky fraction equals one.*

If, on the other hand,

$$\alpha^{1/\delta} \le \mathcal{B}^*_{\mathrm{rel}}(\alpha),$$

the optimal feedback does not require borrowing at $t = 0$ *for any finite* $\zeta > 1$.

Proof By the calibration condition,

$$\frac{x_{\mathrm{safe}}(0) - x_0}{x_{\mathrm{safe}}(0) - x_{\mathrm{imp}}} = \alpha^{1/\delta}.$$

As in the proof of Corollary 41, borrowing at $t = 0$ occurs if and only if

$$M\big(x_{\mathrm{safe}}(0) - x_0\big) > (\delta - 1)x_0.$$

At equality, the total risky fraction equals one. Using the calibration condition, we have

$$x_{\mathrm{safe}}(0) - x_0 = \alpha^{1/\delta}(x_{\mathrm{safe}}(0) - x_{\mathrm{imp}})$$

and thus

$$x_0 = x_{\mathrm{imp}} + \big(1 - \alpha^{1/\delta}\big)(x_{\mathrm{safe}}(0) - x_{\mathrm{imp}}).$$

The borrowing condition is therefore equivalent to

$$\big(\alpha^{1/\delta}(M + \delta - 1) - (\delta - 1)\big)(x_{\mathrm{safe}}(0) - x_{\mathrm{imp}}) > (\delta - 1)x_{\mathrm{imp}},$$

which simplifies to

$$\left(\alpha^{1/\delta} - \frac{\big(1 - \alpha^{1/\delta}\big)(\delta - 1)}{M}\right) \frac{M(x_{\mathrm{safe}}(0) - x_{\mathrm{imp}})}{(\delta - 1)x_{\mathrm{imp}}} > 1,$$

By the definition of ζ,

$$\zeta = \frac{\varepsilon_T M}{\delta - 1}\left(\frac{\overline{C}}{r x_{\mathrm{imp}}} - 1\right) = \frac{M(x_{\mathrm{safe}}(0) - x_{\mathrm{imp}})}{(\delta - 1)x_{\mathrm{imp}}}.$$

Consequently, by the definition of $\mathcal{B}^*_{\mathrm{rel}}(\alpha)$ in (61), the preceding condition is equivalent to

$$\zeta\big(\alpha^{1/\delta} - \mathcal{B}^*_{\mathrm{rel}}(\alpha)\big) > 1.$$

If the bracket is positive, division yields $\zeta > \zeta_{\mathrm{init}}$. Since $\alpha^{1/\delta} < 1$ and $\mathcal{B}^*_{\mathrm{rel}}(\alpha) > 0$, the bracket is strictly smaller than 1, and hence $\zeta_{\mathrm{init}} > 1$. Equality in this condition corresponds to equality in the borrowing condition, which is equivalent to the total risky fraction being one.

If the bracket is non-positive, the inequality cannot hold for any finite $\zeta > 1$, which proves the final assertion. □

In summary, the constant-elasticity branch yields a finite-endpoint mortality law for which both the minimal lifetime-impoverishment probability and the optimal feedback strategy are available in closed form. This explicit structure also makes the withdrawal problem invertible: the optimal sustainable withdrawal rate is the safe withdrawal rate plus a risk-taking bonus, whereas the required initial capital is the safe capital minus the corresponding capital relief. Moreover, we obtained explicit conditions under which the unconstrained feedback remains feasible under the simplex constraint and identified the maximal relative wealth bandwidth compatible with a prescribed risk level and target withdrawal rate. When $M > 0$, the target-withdrawal condition and the simplex-feasibility threshold are simultaneously binding at the maximizing calibration. For wider bands on the same calibration line, borrowing is possible only before the explicit cutoff t^*_{bor}; thereafter the feedback is simplex-valued at every solvent wealth state. Whether borrowing is required already at the initial wealth is determined separately by the threshold ζ_{init}, when it exists.

These results determine the probability of lifetime impoverishment and the policy that attains it, but they do not describe when impoverishment occurs. Section 5 therefore studies the impoverishment time under the exact feedback. The induced distance from the safe boundary is a geometric diffusion, which reduces impoverishment to a Brownian first-passage problem with a deterministic curved boundary. This representation provides the basis for a Durbin-type approximation.

5 First-passage representation of the impoverishment time

This section studies the timing of financial impoverishment under the optimal feedback rule. We first derive the dynamics of the distance from the safe wealth level and use them to determine which boundary of the risky region can be reached. For the constant-elasticity family, these dynamics reduce to a geometric Brownian motion, which yields a representation of the impoverishment time as the first-passage time of a Brownian motion to a curved boundary. We then apply Durbin's expansion to obtain a tractable density approximation and derive a semi-explicit total-variation bound for the resulting conditional law.

5.1 Gap dynamics and boundary accessibility

The natural state variable for the first-passage analysis is the gap between current wealth and the safe boundary. In Young's constant-hazard benchmark [34], the safe boundary is constant and the optimal risky position is linear in this gap. Consequently, the gap follows a geometric Brownian motion, which ultimately leads to the inverse-Gaussian law in Theorem 11. The next result shows that the same multiplicative structure persists under the general feedback rule, although its coefficients need no longer be constant.

Proposition 43 (Multiplicative safe-gap dynamics) *Assume the general feedback rule from Theorem 22 is used and define*

$$Y_t := x_{\mathrm{safe}}(t) - X_t.$$

Up to the first exit from $\mathcal{D}$, let

$$D_t := \widehat{\Upsilon}\big(y(t, X_t)\big).$$

Then $D_t > 1$, and there is a one-dimensional Brownian motion $\widetilde{B}$ such that

$$\mathrm{d}Y_t = Y_t \left(r - \frac{2\rho}{D_t - 1} \right) \mathrm{d}t + Y_t \frac{\sqrt{2\rho}}{D_t - 1} \, \mathrm{d}\widetilde{B}_t. \tag{67}$$

Thus, the safe gap is a multiplicative diffusion.

Proof The identity established in Proposition 20 gives

$$D_t - 1 = \frac{y(t, X_t)\psi''(y(t, X_t))}{\psi'(y(t, X_t))} > 0.$$

The safe boundary satisfies

$$x'_{\mathrm{safe}}(t) = r x_{\mathrm{safe}}(t) - C. \tag{68}$$

To apply Itô's formula, define $G(t, x) := x_{\mathrm{safe}}(t) - x$. Then $Y_t = G(t, X_t)$, and

$$G_t(t, x) = x'_{\mathrm{safe}}(t), \qquad G_x(t, x) = -1, \qquad G_{xx}(t, x) = 0.$$

Itô's formula therefore yields

$$\begin{aligned} \mathrm{d}Y_t &= G_t(t, X_t)\,\mathrm{d}t + G_x(t, X_t)\,\mathrm{d}X_t + \frac{1}{2} G_{xx}(t, X_t)\,\mathrm{d}\langle X \rangle_t \\ &= \Big(x'_{\mathrm{safe}}(t) - r X_t + C - X_t {\pi_t^*}^\top \theta \Big)\,\mathrm{d}t - X_t {\pi_t^*}^\top \sigma \,\mathrm{d}B_t. \end{aligned}$$

Substituting (68) and using the definition of Y_t gives

$$\mathrm{d}Y_t = \Big(r Y_t - X_t {\pi_t^*}^\top \theta \Big)\,\mathrm{d}t - X_t {\pi_t^*}^\top \sigma \,\mathrm{d}B_t.$$

By Theorem 22,

$$X_t \pi_t^* = \frac{Y_t}{D_t - 1} \mathfrak{m}.$$

Consequently,

$$\mathrm{d}Y_t = Y_t \left(r - \frac{\mathfrak{m}^\top \theta}{D_t - 1} \right) \mathrm{d}t - Y_t \frac{\mathfrak{m}^\top \sigma}{D_t - 1} \,\mathrm{d}B_t.$$

Since $\mathfrak{m} = \Sigma^{-1}\theta$ and $\Sigma = \sigma\sigma^\top$,

$$\mathfrak{m}^\top \theta = 2\rho, \qquad \|\sigma^\top \mathfrak{m}\|^2 = 2\rho.$$

Therefore,

$$\widetilde{B}_t := -\frac{\mathfrak{m}^\top \sigma}{\sqrt{2\rho}} B_t$$

is a one-dimensional Brownian motion, and substitution yields (67). □

In general, the local elasticity D_t makes the drift and volatility of the safe gap state dependent. For a constant-elasticity compatible triple, however, D_t is identically equal to δ. The multiplicative diffusion therefore becomes a geometric Brownian motion with constant coefficients, providing the direct analogue of Young's constant-hazard gap dynamics.

Corollary 44 (Geometric safe-gap dynamics) *Suppose that the compatible triple has constant elasticity* $\Upsilon \equiv \delta > 1 + \rho/\mathcal{E}(0)$. *Then* $D_t \equiv \delta$, *and*

$$\mathrm{d}Y_t = Y_t \left(r - \frac{2\rho}{\delta - 1} \right) \mathrm{d}t + Y_t \frac{\sqrt{2\rho}}{\delta - 1} \, \mathrm{d}\widetilde{B}_t. \tag{69}$$

In particular, with

$$\widetilde{\mu} := r - \frac{2\rho}{\delta - 1}, \qquad \widetilde{\sigma} := \frac{\sqrt{2\rho}}{\delta - 1},$$

up to its first exit from $\mathcal{D}$, *the process satisfies*

$$Y_t = Y_0 \exp\!\left(\left(\widetilde{\mu} - \frac{\widetilde{\sigma}^2}{2} \right) t + \widetilde{\sigma} \widetilde{B}_t \right). \tag{70}$$

Proof If $\Upsilon \equiv \delta$, then the definition of $\widehat{\Upsilon}$ gives $\widehat{\Upsilon} \equiv \delta$. The result therefore follows from Proposition 43 and the explicit solution of a geometric Brownian motion. □

The multiplicative representation also determines the relevant exit boundary. The safe gap cannot vanish while the process remains in the risky region, so the safe wealth level is inaccessible in finite time. When the right endpoint T is finite, the risky region collapses as $t \uparrow T$. Consequently, since exit through the safe boundary is impossible, impoverishment must then occur before T. The following corollary formalizes these observations.

Corollary 45 (Boundary accessibility of the risky region) *Under the assumptions of Proposition 43, suppose that* $x_{\mathrm{imp}} < x_0 < x_{\mathrm{safe}}(0)$. *Then*

$$\mathbb{P}(\tau_{\mathrm{safe}} < T) = 0.$$

If, in addition, $T < \infty$, *then*

$$\mathbb{P}(\tau_{\mathrm{imp}} < T) = 1.$$

Proof In the constant-elasticity case, $D_t \equiv \delta > 1$. In the non-constant-elasticity case, Proposition 20 and Lemma 24 give

$$\widehat{\Upsilon}(y) - 1 = \frac{\rho \Upsilon(y)}{A_0 \Upsilon(y) - F_0}, \qquad \Upsilon(y) \xrightarrow{y \downarrow 0} \upsilon_+ > 1.$$

Since

$$Q(\upsilon) = (\upsilon - 1)(A_0 \upsilon - F_0) - \rho \upsilon,$$

the identity $Q(\upsilon_+) = 0$ implies

$$\frac{\rho \upsilon_+}{A_0 \upsilon_+ - F_0} = \upsilon_+ - 1.$$

Thus, in either case,

$$\liminf_{y \downarrow 0} \big(\widehat{\Upsilon}(y) - 1 \big) > 0. \tag{71}$$

Set

$$\widetilde{\mu}_t := r - \frac{2\rho}{D_t - 1}, \qquad \widetilde{\sigma}_t := \frac{\sqrt{2\rho}}{D_t - 1}.$$

Itô's formula and Proposition 43 yield, before exit,

$$\ln Y_t = \ln Y_0 + \int_0^t \left(\widetilde{\mu}_s - \frac{\widetilde{\sigma}_s^2}{2}\right) \mathrm{d}s + \int_0^t \widetilde{\sigma}_s \, \mathrm{d}\widetilde{B}_s. \tag{72}$$

If $\tau_{\text{safe}} < T$, then $Y_t \to 0$, and hence $y(t, X_t) \to 0$, as $t \uparrow \tau_{\text{safe}}$. Equation (71), together with continuity on the preceding compact part of the path, shows that $D_t - 1$ is bounded away from zero up to τ_{safe}. Both integrals in (72) therefore have finite limits, contradicting $\ln Y_t \to -\infty$. Hence, $\mathbb{P}(\tau_{\text{safe}} < T) = 0$.

Finally, let $T < \infty$. Outside the preceding null event, if $\tau_{\text{imp}} \geq T$, then the process remains in $\mathcal{D}$ and

$$0 < Y_t < x_{\text{safe}}(t) - x_{\text{imp}} \xrightarrow{t \uparrow T} 0.$$

In the non-constant-elasticity case, $0 < y(t, X_t) < b(t) \to 0$ by (24). Thus, (71) and continuity again make both integrals in (72) finite at T, contradicting $Y_t \to 0$. For the constant-elasticity case, the same argument applies. Therefore, $\mathbb{P}(\tau_{\text{imp}} < T) = 1$. □

For $T < \infty$, Corollary 45 shows that impoverishment almost surely occurs before the terminal time. The remaining question is therefore not whether the lower wealth boundary is reached, but when it is reached. In the constant-elasticity case, the explicit geometric-Brownian representation turns this timing question into a standard Brownian first-passage problem.

5.2 Curved-boundary first passage

Throughout this section we assume $T < \infty$ unless stated otherwise. Unlike the linear-boundary problem underlying the exact inverse-Gaussian law in the constant-hazard model, the finite-horizon model leads to a curved deterministic boundary. Although its first-passage distribution is generally not available in closed form, the representation below permits its density to be evaluated by quadrature using Durbin's expansion [10, 11]. It thus provides an efficient alternative to Monte Carlo simulation and, importantly, comes with a computable bound on the approximation error.

Proposition 46 (Brownian curved-boundary representation) *Assume that the compatible triple has constant elasticity* $\Upsilon \equiv \delta > 1 + \rho/\mathcal{E}(0)$*, and let* $\widetilde{\mu}, \widetilde{\sigma}$ *be as in Corollary 44. Define*

$$L(t) := x_{\text{safe}}(t) - x_{\text{imp}} = \frac{C - r x_{\text{imp}}}{\mathcal{E}(t)}.$$

For $x_{\text{imp}} < x_0 < x_{\text{safe}}(0)$*, the impoverishment time admits the pathwise representation*

$$\tau_{\text{imp}} = \inf\{t \in [0, T) : \widetilde{B}_t \geq h(t)\}, \qquad h(t) := \frac{\ln\big(L(t)/Y_0\big) - (\widetilde{\mu} - \widetilde{\sigma}^2/2)t}{\widetilde{\sigma}}. \tag{73}$$

Proof Corollary 44 gives the geometric-Brownian representation of the safe gap. Taking logarithms and rearranging yields

$$\tau_{\text{imp}} = \inf\{t \in [0, T) : \widetilde{B}_t \geq h(t)\},$$

with h as defined in (73). □

Although the boundary in (73) appears to depend separately on C, x_{imp}, and x_0 through L and Y_0, the explicit form of the value function causes this dependence to collapse to the initial minimal impoverishment probability. The following corollary makes this invariance explicit.

Corollary 47 (Impoverishment-time invariance at fixed risk) *Under the assumptions of Proposition 46, let*

$$q := p(0, x_0; C) \in (0, 1).$$

Then the Brownian first-passage boundary in (73) can be written as

$$h(t) = \frac{\ln\big(\mathcal{E}(0)/\mathcal{E}(t)\big) - \frac{1}{\delta}\ln(q) - \big(\widetilde{\mu} - \widetilde{\sigma}^2/2\big)t}{\widetilde{\sigma}}. \tag{74}$$

Consequently, fix the market and mortality parameters and consider any two admissible triples $(C^{(1)}, x_{\mathrm{imp}}^{(1)}, x_0^{(1)})$ and $(C^{(2)}, x_{\mathrm{imp}}^{(2)}, x_0^{(2)})$ that induce the same minimal lifetime-impoverishment probability $q \in (0, 1)$. If their optimal-feedback wealth processes are driven by the same market Brownian motion, then (74) gives both processes the same first-passage boundary h for the same $\widetilde{B}$. Their impoverishment times therefore coincide almost surely. In particular, their impoverishment-time distributions are identical.

The same conclusion applies to any two optimal sustainable-withdrawal policies with prescribed risk level α, since each induces $q = \alpha$.

Proof By (44),

$$q^{1/\delta} = \frac{x_{\mathrm{safe}}(0) - x_0}{x_{\mathrm{safe}}(0) - x_{\mathrm{imp}}} = \frac{Y_0}{L(0)}.$$

Moreover, the definition of L gives

$$\frac{L(t)}{L(0)} = \frac{\mathcal{E}(0)}{\mathcal{E}(t)}.$$

Therefore,

$$\frac{L(t)}{Y_0} = \frac{\mathcal{E}(0)}{q^{1/\delta}\mathcal{E}(t)}.$$

Substitution into (73) proves (74). The final assertion follows from the defining identity

$$p\big(0, x_0; C^*(x_0; \alpha)\big) = \alpha.$$

□

Now, because the safe boundary is inaccessible, the only absorbing boundary is the lower wealth boundary at the impoverishment level. Hence, the exit problem reduces to the single Brownian crossing problem in (73). Durbin's expansion gives a sequence of integral approximations for such curved-boundary first-passage densities. We first state the general result and then apply it to the boundary h.

Theorem 48 (Durbin's first-passage density approximation) *Let $T \in (0, \infty]$, let B be a standard one-dimensional Brownian motion, and let $h \in C^2([0, T))$ satisfy*

$$h(0) > 0 \qquad \text{and} \qquad h(t) - th'(t) > 0, \quad t \in (0, T).$$

Then the first-passage time

$$\tau_h := \inf\{t \in (0, T) : B_t \geq h(t)\}$$

has a (potentially defective) density f_{τ_h} on $(0,T)$. The first term in Durbin's expansion, also called the tangent approximation, is

$$f_1(t) := \left[\frac{h(t)}{t} - h'(t)\right] p_t\big(h(t)\big), \qquad p_u(x) := \frac{1}{\sqrt{2\pi u}} \exp\left(-\frac{x^2}{2u}\right). \tag{75}$$

Thus, the one-term approximation is $f_{\tau_h}(t) \approx f_1(t)$, $t \in (0,T)$.

If, in addition, h is concave on $[0,T)$, then this is a one-sided approximation and its error satisfies

$$0 \le f_1(t) - f_{\tau_h}(t) \le f_2(t), \qquad t \in (0,T),$$

where

$$f_2(t) := \int_0^t \left[\frac{h(s)}{s} - h'(s)\right] \left[\frac{h(t)-h(s)}{t-s} - h'(t)\right] p_s\big(h(s)\big) p_{t-s}\big(h(t)-h(s)\big)\,\mathrm{d}s.$$

Remark 8 In Corollary 45, we have already seen that the safe boundary is inaccessible. The first-passage density $f_{\tau_{\mathrm{imp}}}$ is therefore non-defective. However, note that the functions f_k in Theorem 48 need not be densities themselves. In particular, they may integrate to more than one, in which case they must be rescaled to be interpreted as densities.

Now, to apply Durbin's expansion, we need to verify the relevant shape conditions for the boundary. As the following lemma shows, the representation in (74) makes h strictly concave, and every tangent line has a positive intercept at time zero. These properties ensure that the first term in the expansion is non-negative and, through concavity, that its truncation error has a definite sign.

Lemma 49 *The boundary h in (74) is smooth on $[0,T)$ and $h(t) \to -\infty$ as $t \uparrow T$. Moreover it is strictly concave on $[0,T)$ and satisfies $h(t) - h'(t)t > 0$ for all $t \in [0,T)$.*

Proof From the definition of $\mathcal{E}$,

$$\mathcal{E}'(t) = \frac{r^2 \mathrm{e}^{-r(T-t)}}{\big(1-\mathrm{e}^{-r(T-t)}\big)^2} > 0, \qquad \frac{\mathcal{E}'(t)}{\mathcal{E}(t)} = \mathcal{E}(t) - r.$$

Moreover, $\mathcal{E}(t) \to \infty$ as $t \uparrow T$. Equation (74) therefore shows directly that h is smooth and satisfies $h(t) \to -\infty$ as $t \uparrow T$.

Differentiating (74) and using the definitions of $\widetilde{\mu}$ and $\widetilde{\sigma}$ gives

$$h'(t) = \frac{1}{\widetilde{\sigma}}\left(\frac{\rho(2\delta-1)}{(\delta-1)^2} - \mathcal{E}(t)\right), \qquad h''(t) = -\frac{\mathcal{E}'(t)}{\widetilde{\sigma}} < 0.$$

Hence, h is strictly concave. Moreover,

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(h(t) - th'(t)\big) = -th''(t) > 0, \qquad t \in (0,T).$$

Finally,

$$h(0) = -\frac{\ln(q)}{\delta\widetilde{\sigma}} > 0,$$

because $q \in (0,1)$. Therefore, $h(t) - th'(t) > 0$ for all $t \in [0,T)$. □

Having verified the shape conditions, we can now apply Durbin's expansion to the impoverishment time. The following corollary gives a simple one-term approximation for its density.

Corollary 50 *Under the assumptions of Proposition 46, the density of τ_{imp} is approximated by*

$$f_{\tau_{\text{imp}}}(t) \lesssim f_1(t) = \frac{h(t) - h'(t)t}{\sqrt{2\pi t^3}} \exp\left(-\frac{h(t)^2}{2t}\right), \qquad t \in (0, T). \tag{76}$$

Proof The boundary h in (73) is smooth and concave on $[0, T)$ by Lemma 49. The result therefore follows from Theorem 48. □

The density approximation above concerns the financial impoverishment time itself, whereas the economically relevant object is its law conditional on impoverishment preceding death. Independence weights the first-passage density by the survival function. Since the first Durbin term dominates the true density, its excess mass admits a simple mixture representation. This yields the following total-variation error bound in terms of the true and approximated lifetime-impoverishment probabilities.

Proposition 51 (Total-variation error of the conditional Durbin approximation) *Let ν_{imp} denote the conditional law of the impoverishment time,*

$$\nu_{\text{imp}}(A) := \mathbb{P}\left(\tau_{\text{imp}} \in A \,\middle|\, \tau_{\text{imp}} < \mathcal{T}\right), \qquad A \in \mathcal{B}\big((0, T)\big),$$

define

$$\widehat{q} := \int_0^T S(t) f_1(t)\, \mathrm{d}t, \qquad \widehat{\nu_{\text{imp}}}(A) := \frac{1}{\widehat{q}} \int_A S(t) f_1(t)\, \mathrm{d}t,$$

and again set $q := p(0, x_0; C)$. Then, under the assumptions of Proposition 46,

$$d_{\text{TV}}\big(\widehat{\nu_{\text{imp}}}, \nu_{\text{imp}}\big) \le 1 - \frac{q}{\widehat{q}}, \tag{77}$$

where $d_{\text{TV}}(\mu, \nu) := \sup_A |\mu(A) - \nu(A)|$. Hence, under the optimal sustainable-withdrawal policy at risk level α, one has $q = \alpha$, and consequently

$$d_{\text{TV}}\big(\widehat{\nu_{\text{imp}}}, \nu_{\text{imp}}\big) \le 1 - \frac{\alpha}{\widehat{q}}.$$

Proof Independence of τ_{imp} and $\mathcal{T}$, followed by the verification identity, gives

$$q = \int_0^T S(t) f_{\tau_{\text{imp}}}(t)\, \mathrm{d}t = \mathbb{P}(\tau_{\text{imp}} < \mathcal{T}).$$

Since h is concave, Theorem 48 yields $f_1 - f_{\tau_{\text{imp}}} \ge 0$. Hence, with

$$\Delta := \widehat{q} - q = \int_0^T S(t)\big(f_1(t) - f_{\tau_{\text{imp}}}(t)\big)\, \mathrm{d}t,$$

we have $\Delta \ge 0$. If $\Delta = 0$, then $\widehat{\nu_{\text{imp}}} = \nu_{\text{imp}}$. Otherwise, the measure

$$\varrho(A) := \frac{1}{\Delta} \int_A S(t)\big(f_1(t) - f_{\tau_{\text{imp}}}(t)\big)\, \mathrm{d}t$$

is a probability measure and

$$\widehat{\nu_{\mathrm{imp}}} = \frac{q}{\widehat{q}}\nu_{\mathrm{imp}} + \frac{\Delta}{\widehat{q}}\varrho.$$

Consequently, for every $A \in \mathcal{B}((0,T))$,

$$\widehat{\nu_{\mathrm{imp}}}(A) - \nu_{\mathrm{imp}}(A) = \frac{\Delta}{\widehat{q}}\big(\varrho(A) - \nu_{\mathrm{imp}}(A)\big).$$

Taking absolute values and then the supremum over all such A yields

$$d_{\mathrm{TV}}\big(\widehat{\nu_{\mathrm{imp}}}, \nu_{\mathrm{imp}}\big) = \frac{\Delta}{\widehat{q}} d_{\mathrm{TV}}\big(\varrho, \nu_{\mathrm{imp}}\big).$$

Since ϱ and ν_{imp} are probability measures, $d_{\mathrm{TV}}(\varrho, \nu_{\mathrm{imp}}) \le 1$. Therefore,

$$d_{\mathrm{TV}}\big(\widehat{\nu_{\mathrm{imp}}}, \nu_{\mathrm{imp}}\big) \le \frac{\Delta}{\widehat{q}} = 1 - \frac{q}{\widehat{q}}.$$

This proves (77). □

This upper bound allows us to quantify the error of the conditional Durbin approximation in terms of the true and approximated lifetime-impoverishment probabilities without having to perform a computationally expensive Monte Carlo simulation. In Section 6, we will see that for realistic parameter values, the bound is small, and the approximation is accurate.

6 Numerical assessment

The preceding sections provide explicit policies and first-passage representations for a class of compatible mortality laws. We now examine their behavior under an illustrative market and mortality calibration. We first compare the calibrated survival models, then assess policy performance under a common mortality reference and investigate the effects of portfolio constraints. Finally, we compare the conditional impoverishment-time distribution with its Durbin approximation. These experiments illustrate the analytical results and their sensitivity to modelling choices; they do not constitute an actuarial forecast or identify an optimal policy under general mortality.

The numerical studies also export detailed results and refinement diagnostics as CSV files with accompanying data dictionaries, which can be provided as supplementary material. The files supporting the additional diagnostics below are identified alongside the corresponding results.

6.1 Model calibration

The experiments combine a real-terms financial market with mortality data for German cohorts. We first specify the risk-free rate and the return and covariance assumptions for two risky assets. We then fit an exponential model, a Gompertz–Makeham (GM) model and the compatible compactified-exponential model to the cohort survival targets. Following the notation S_{CE} in Definition 11, we abbreviate the last model by CE in the figures and tables. The market parameters enter its compatibility restriction. Its fit supplies the finite-endpoint law and closed-form policy used below, while the fitted GM law serves as a mortality reference where stated.

6.1.1 Market and financial parameters

We use a transparent long-run calibration for a euro-area investor. All parameters entering the financial model are expressed in annual real units.

We set the annual inflation benchmark to

$$\iota = 0.02,$$

consistent with the medium-term inflation target of the European Central Bank [12]. This inflation rate is used only to translate externally reported nominal return assumptions into the real parameters of the model.

For the risky assets, we consider the MSCI World index and the eb.rexx Government Germany index. Since long-horizon forecasts for the precise eb.rexx maturity segment are not readily available, we use the broader euro-government-bond category as a proxy for its long-run return and volatility. Table 1 summarizes the 30-year nominal annualized return assumptions and annualized volatilities reported by RBC Global Asset Management [28, p. 10]. The reported volatilities are based on historical three-year rolling periods, rather than forward-looking 30-year volatility forecasts.

Asset class	Nominal return	Volatility
MSCI World	0.0675	0.145
Euro government bonds	0.0330	0.040

Table 1 Nominal annualized return assumptions and annualized volatilities used for the two risky assets, from [28, p. 10].

The corresponding return assumptions are also close to the 20-year estimates of 6.7% for developed-market equities and 3.3% for euro government bonds reported by Fidelity International [14, p. 17].

To translate the nominal annualized return assumptions into the real arithmetic drifts required by the Black–Scholes model, let g_i^{N} denote the reported nominal annualized return and let σ_i denote the corresponding annual volatility. For this illustrative conversion, we interpret the annualized return forecast as a compounded growth rate and first define the real continuously compounded growth rate by

$$g_i^{\mathrm{R}} = \ln(1 + g_i^{\mathrm{N}}) - \ln(1 + \iota).$$

Under geometric Brownian motion, the arithmetic drift is related to the continuously compounded growth rate through

$$\mu_i = g_i^{\mathrm{R}} + \frac{\sigma_i^2}{2}.$$

Consequently, the real equity drift is

$$\begin{aligned}\mu_{\mathrm{W}} &= \ln(1.0675) - \ln(1.02) + \frac{1}{2}(0.145)^2 \\ &\approx 0.056029,\end{aligned}$$

whereas the real government-bond drift is

$$\begin{aligned}\mu_{\mathrm{G}} &= \ln(1.0330) - \ln(1.02) + \frac{1}{2}(0.040)^2 \\ &\approx 0.013465.\end{aligned}$$

Thus, in the order MSCI World and German government bonds, we use

$$\mu = \begin{pmatrix} 0.056029 \\ 0.013465 \end{pmatrix}.$$

Expected inflation does not itself determine the real risk-free rate. Under the deterministic inflation benchmark used for the conversion, the effective annual nominal and real rates, denoted by i^{N} and i^{R}, satisfy the multiplicative Fisher relation [15, p. 9]

$$1 + i^{\mathrm{N}} = (1 + i^{\mathrm{R}})(1 + \iota).$$

In the continuous-time model, r is a continuously compounded real rate, so that $1 + i^{\mathrm{R}} = \exp(r)$. Hence the exact conversion used below is

$$1 + i^{\mathrm{N}} = \exp(r)(1 + \iota).$$

The Deutsche Bundesbank reports that longer-term real interest rates in Germany, measured as the ten-year federal-bond yield less expected inflation over the same horizon, were approximately 0.6% in 2025 [9]. ECB estimates of the euro-area natural real rate available at the end of 2024 ranged from -0.5% to 0.5% for the principal term-structure and semi-structural measures [7]. These quantities are not identical to the constant tradable bank-account rate in the model, but they provide a transparent empirical scale for its calibration. We therefore use

$$r = 0.005.$$

Together with $\iota = 0.02$, this corresponds to the effective nominal rate

$$\exp(0.005)(1.02) - 1 \approx 0.025113.$$

As a contemporaneous plausibility check, Trade Republic advertised an annual nominal interest rate of 2.5% on cash in October 2026 [33]. Conversely, combining this rate

with expected inflation of 2% gives the continuously compounded real rate

$$\ln\left(\frac{1.025}{1.02}\right) \approx 0.004890,$$

which is close to the selected value. This comparison is only illustrative: the advertised nominal rate is variable and does not constitute a guaranteed long-run real return.

The corresponding vector of real excess returns is

$$\begin{aligned}\theta &= \mu - r\mathbf{1} \\ &= \begin{pmatrix} 0.056029 - 0.005 \\ 0.013465 - 0.005 \end{pmatrix} \\ &= \begin{pmatrix} 0.051029 \\ 0.008465 \end{pmatrix}.\end{aligned}$$

To complete the covariance specification, we use the RBC volatility assumptions

$$\sigma_{\mathrm{W}} = 0.145, \qquad \sigma_{\mathrm{G}} = 0.040.$$

In its Q3 2025 report, Allianz Global Investors reports a correlation of 0.08 between developed-world equities and euro government bonds [1, p. 2]. We use this value as an approximate long-run cross-asset correlation,

$$\varrho_{\mathrm{WG}} = 0.08.$$

The Allianz estimate is reported from a USD-investor perspective, with the fixed-income position hedged to USD, and is therefore used only as a proxy for the dependence between the two asset classes.

The resulting annual covariance matrix is

$$\begin{aligned}\Sigma &= \begin{pmatrix} \sigma_{\mathrm{W}}^2 & \varrho_{\mathrm{WG}}\sigma_{\mathrm{W}}\sigma_{\mathrm{G}} \\ \varrho_{\mathrm{WG}}\sigma_{\mathrm{W}}\sigma_{\mathrm{G}} & \sigma_{\mathrm{G}}^2 \end{pmatrix} \\ &= \begin{pmatrix} (0.145)^2 & 0.08(0.145)(0.040) \\ 0.08(0.145)(0.040) & (0.040)^2 \end{pmatrix} \\ &= \begin{pmatrix} 0.021025 & 0.000464 \\ 0.000464 & 0.001600 \end{pmatrix}.\end{aligned}$$

The subsequent calculations use the unrounded drifts implied by the formulas above. The resulting Merton vector and optimal mean–variance excess return, as defined in Section 2, are

$$\mathfrak{m} = \Sigma^{-1}\theta \approx \begin{pmatrix} 2.325208 \\ 4.616041 \end{pmatrix},$$

$$\rho = \frac{1}{2}\theta^{\top}\Sigma^{-1}\theta \approx 0.078863.$$

The vector $\mathfrak{m}$ is a model input to the state-dependent feedback rule, not by itself a three-asset portfolio allocation. For any risky-fraction vector π, the bank-account weight is $1-\mathbf{1}^\top\pi$; using $\mathfrak{m}$ directly as the risky-fraction vector would imply substantial borrowing.

In summary, all numerical experiments use annual real parameters and the risky-asset ordering MSCI World and German government bonds. The inflation benchmark and financial-market inputs are

$$\iota = 0.02, \qquad r = 0.005, \qquad \mu = \begin{pmatrix} 0.056029 \\ 0.013465 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} 0.021025 & 0.000464 \\ 0.000464 & 0.001600 \end{pmatrix}.$$

Consequently, the real excess-return vector and the derived market quantities are

$$\theta = \begin{pmatrix} 0.051029 \\ 0.008465 \end{pmatrix}, \qquad \mathfrak{m} = \begin{pmatrix} 2.325208 \\ 4.616041 \end{pmatrix}, \qquad \rho = 0.078863.$$

The risky-fraction strategy has two components in this ordering, while the corresponding bank-account weight is $1 - \mathbf{1}^\top\pi$. All numerical calculations use the unrounded values implied by the preceding calibration formulas.

All monetary quantities are expressed in euros of the calibration year. Accordingly, the constant withdrawal rate C is a real withdrawal rate. For example, $C = 30{,}000$ represents annual consumption with the purchasing power of EUR 30,000 in the calibration year. Under the deterministic inflation benchmark ι, the corresponding nominal withdrawal rate at time t is

$$C^{\mathrm{N}}(t) = C(1+\iota)^t.$$

Inflation is therefore not applied again to C in the real-wealth dynamics.

6.1.2 Mortality data and parametric calibration

We construct the mortality targets from the German cohort life tables supplied by the Federal Statistical Office (Destatis) [31]. These cohort tables combine observed or reconstructed mortality with projected mortality where the cohort's future experience is not yet observed. The baseline uses the medium longevity scenario (variant V2) and valuation year 2026. For every retirement age $a \in \{45, \dots, 75\}$, the corresponding birth cohort is $2026 - a$. Tables 12621-b03 and 12621-b04 provide the sex-specific one-year death probabilities $q_{x,s}$, $s \in \{\mathrm{m}, \mathrm{f}\}$, at attained ages $x = a, \dots, 99$. The open class "100 and above" is not treated as a one-year death probability; the exactly supported horizon therefore ends at attained age 100.

Conditional on survival to age a, the sex-specific survival probability at integer remaining lifetime k is reconstructed as

$$\widehat{S}_{a,s}(k) = \prod_{j=0}^{k-1}(1 - q_{a+j,s}), \qquad k = 0, \dots, h_a, \qquad h_a := 100 - a,$$

where the empty product equals one. For plotting between integer ages, we use a piecewise-constant force of mortality $\mu_{x,s} = -\ln(1 - q_{x,s})$. Thus, for $k \le t < k + 1$,

$$\widehat{S}_{a,s}(t) = \widehat{S}_{a,s}(k) \exp(-\mu_{a+k,s}(t - k)) .$$

This interpolation reproduces every annual $q_{x,s}$ exactly.

Because the financial model describes a generic investor rather than a sex-specific portfolio, the baseline target is the equal-sex survival mixture

$$\widehat{S}_a(t) = \frac{1}{2}\widehat{S}_{a,\mathrm{m}}(t) + \frac{1}{2}\widehat{S}_{a,\mathrm{f}}(t).$$

The mixture is formed after constructing the two conditional survival curves. It represents equal numbers of male and female investors alive at the initial age and is a transparent modeling convention, not an official population-weighted Destatis series.

For every a, we fit three conditional remaining-lifetime laws to this same target. The exponential benchmark has survival function

$$S_{\mathrm{Exp},a}(t) = \exp(-\eta_a t),$$

the Gompertz–Makeham benchmark

$$S_{\mathrm{GM},a}(t) = \exp\left(-\lambda_a t - \frac{\alpha_a}{\beta_a}\big(\exp(\beta_a t) - 1\big)\right),$$

and the CE law is the compatible member of the compactified-exponential family identified in Theorem 36. Here α_a is a Gompertz–Makeham parameter, not the risk tolerance α introduced in Section 2. The GM law provides a smooth parametric mortality benchmark [23], while the exponential law isolates the effect of imposing a constant hazard. Their roles differ from that of the CE law: the latter is fitted subject to the restriction that preserves the explicit value function and feedback rule.

To compare the models on the same footing, we minimize the Kullback–Leibler divergence [19] between the discrete lifetime distributions induced on the annual grid. It compares the probabilities assigned to each year of remaining lifetime, rather than fitting the survival values directly. For the cohort target, define

$$\widehat{p}_{a,k} = \widehat{S}_a(k) - \widehat{S}_a(k + 1), \quad k = 0, \dots, h_a - 1, \qquad \widehat{p}_{a,h_a} = \widehat{S}_a(h_a),$$

where the final category represents survival beyond attained age 100. For a candidate survival function S, define $p_{a,k}(S)$ analogously and minimize

$$\mathcal{L}_a^{\mathrm{KL}}(S) = \sum_{k:\,\widehat{p}_{a,k}>0} \widehat{p}_{a,k} \ln\left(\frac{\widehat{p}_{a,k}}{p_{a,k}(S)}\right).$$

If a model assigns zero mass to a category with positive empirical mass, the loss is infinite. This already implies that the maximal attainable age of the compatible

compactified-exponential law must exceed the attained age 100 of the cohort target in order to avoid an infinite loss.

The exponential rate and the three Gompertz–Makeham parameters are fitted in logarithmic coordinates to better resolve the small-parameter regimes. For the compatible compactified-exponential law, the market parameters $r = 0.005$ and $\rho \approx 0.078863$ remain fixed. Given a candidate remaining-lifetime endpoint T_a, write $\mathcal{E}_a$ for the normalizing factor of Definition 9 with $T = T_a$. At the initial age,

$$\mathcal{E}_a(0) = \frac{r}{1 - \exp(-rT_a)}, \qquad \delta_{\min}(T_a) = 1 + \frac{\rho}{\mathcal{E}_a(0)}.$$

The numerical parametrization enforces

$$T_a > h_a, \qquad \delta_a > \delta_{\min}(T_a),$$

and compatibility then determines the initial hazard as

$$\lambda_{\infty,a} = \delta_a \left(\mathcal{E}_a(0) - \frac{\rho}{\delta_a - 1} \right).$$

As in (43), $\lambda_{\infty,a}$ is the initial CE hazard; the subscript does not denote the fitted endpoint. The first inequality is a support-consistency condition: the cohort target assigns positive probability to survival to age 100. Unlike a common-endpoint calibration, we estimate T_a separately for every initial age and report the implied maximum attained age $\omega_a = a + T_a$. Hence ω_a measures the endpoint sensitivity of the approximation across cohorts; it is not interpreted as an estimate of the biological maximum lifespan. All three models are calibrated by a deterministic multistart search. The compatible fit can approach the lower elasticity bound closely; in that case its initial hazard also approaches zero, so the numerical values should be interpreted with due regard to the imposed parameter restriction.

Figure 1 collects the fitted free parameters. For the compactified law, it also displays the compatibility bound $\delta_{a,\min} = \delta_{\min}(T_a)$ and the fitted remaining-lifetime endpoint T_a. For the baseline age $a = 67$, the CE fit gives approximately $\delta_{67} = 4.205$, $T_{67} = 42.65$ years, and initial hazard $\lambda_{\infty,67} = 0.0060$. These quantities determine the CE-optimal policies considered below; they are held fixed when the investor's risk tolerance is varied at that age.

Figure 2 compares the three initial hazards. Since the exponential and Gompertz–Makeham distributions have unbounded support, their 99.5% attained-age quantiles are shown beside the finite endpoint ω_a of the compatible law. The quantiles are descriptive high-death-age measures and are not interpreted as endpoints. At age 67, the fitted initial hazards are approximately 0.0511 for the exponential law, 0.0134 for GM and 0.0060 for CE. The comparison shows why a similar lifetime distribution need not imply a similar initial hazard: the constant-hazard model must use one rate to represent mortality over the entire remaining lifetime, whereas the other two hazards vary with time.

To show what these parameters mean for survival, Figure 3 compares the fitted curves with the sex-specific data and their equal-weight mixture at retirement ages 60, 67, and 75. The target lies between the two sex-specific curves by construction; the models are fitted to that mixture, not separately to each sex.

Figure 4 summarizes the comparison across all retirement ages through the minimized Kullback–Leibler losses. The GM model closely tracks the cohort targets throughout the age range. The CE model improves substantially on the exponential benchmark, but its compatibility restriction limits its flexibility relative to the unconstrained three-parameter GM fit. At age 67, for example, the losses are approximately 0.3842, 0.0165 and 0.0004 for the exponential, CE and GM models, respectively. These are fit diagnostics for the annual lifetime categories, not differences in lifetime-impoverishment probabilities; the latter also depend on the chosen withdrawal and investment policy.



Fig. 1 Mortality parameters fitted to the equal-weight cohort-survival target over retirement ages $a = 45, \ldots, 75$, using the 2026 medium scenario. The compatible-model panel also shows the lower bound $\delta_{a,\min}$ and remaining-lifetime endpoint T_a. The Gompertz–Makeham and compatible-model panels have logarithmic vertical axes.

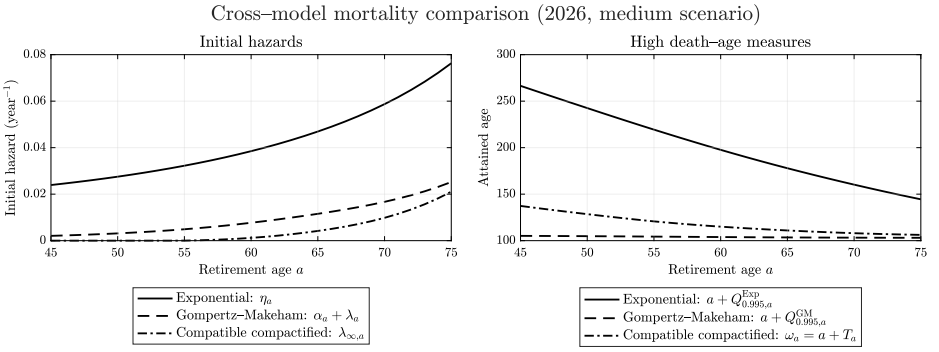


Fig. 2 Initial hazards (left) and high-death-age measures (right) for the three models fitted to the 2026 medium cohort-survival targets. The right panel shows 99.5% attained-age quantiles for the unbounded models and the finite endpoint $\omega_a = a + T_a$ for the compatible compactified-exponential model.

The calibration thus provides two mortality laws for which the preceding theory yields explicit policies, together with the more flexible GM law as a common reference. A closer mortality fit, however, does not by itself reveal how a policy performs when evaluated under a different law. We now turn to that comparison and its implications for sustainable withdrawals and impoverishment risk.

6.2 Policy performance under mortality misspecification

We apply the closed-form withdrawal and feedback rules from Corollaries 12 and 37 and compare their achieved risks under a common mortality law.

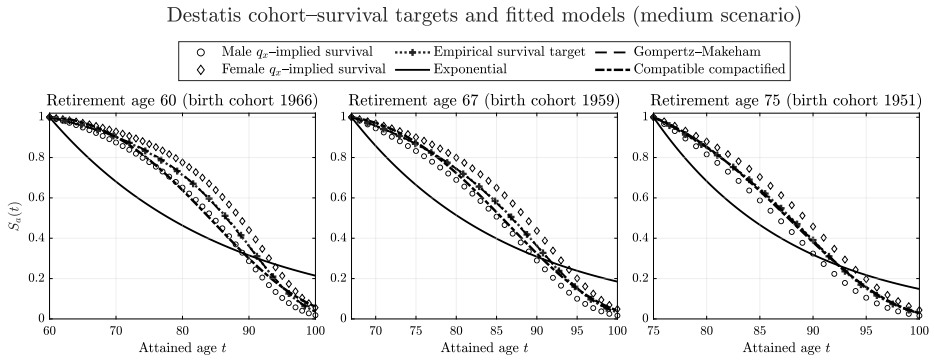


Fig. 3 Sex-specific cohort-survival probabilities, their equal-weight mixture, and the three fitted models at retirement ages 60, 67, and 75. The mixture is the calibration target; the sex-specific curves are shown separately for comparison. Data use the 2026 medium cohort-mortality scenario.

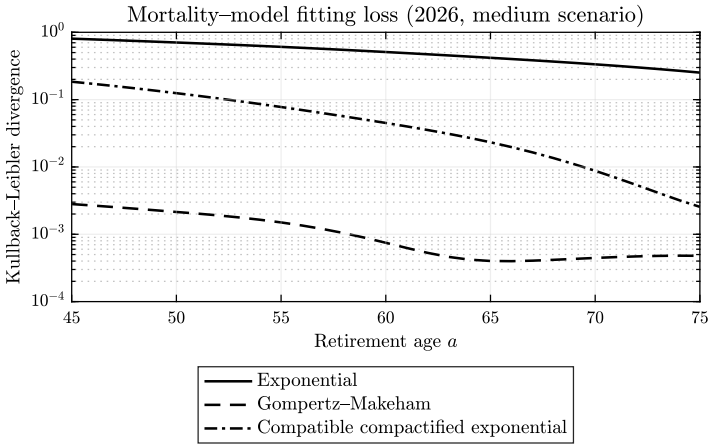


Fig. 4 Minimized Kullback–Leibler divergence between the annual lifetime distributions implied by the equal-weight cohort target and each fitted model, across retirement ages 45 to 75. The vertical axis is logarithmic; the data use the 2026 medium scenario.

For this comparison, we set $a = 67$, $x_0 = 500{,}000$ Euro, and $x_{\mathrm{imp}} = 100{,}000$ Euro. At each target risk level α, the exponential and CE models determine their respective unconstrained optimal sustainable withdrawal rates and feedback policies. As an additional benchmark, we adapt Bengen's four-percent rule [4] to the continuous-time market: the modified Bengen-style policy withdraws $0.04(x_0 - x_{\mathrm{imp}}) = 16{,}000$ real Euro per year and holds the constant risky-fraction allocation $\pi_{\mathrm{B}} = \mathfrak{m}/M$, where $M = \mathbf{1}^\top \mathfrak{m}$. In the present calibration this allocation is long-only and sums to one. It is a model-based comparator inspired by the four-percent rule, not a reproduction of Bengen's historical strategy.

We evaluate all three policies under the fitted GM mortality law. If τ_{imp}^j is the financial impoverishment time under policy $j \in \{\mathrm{Exp}, \mathrm{CE}, \mathrm{B}\}$, its achieved risk is

$$q_{\mathrm{GM}}^j := \mathbb{P}(\tau_{\mathrm{imp}}^j < \mathcal{T}_{\mathrm{GM}}) = \mathbb{E}[S_{\mathrm{GM}}(\tau_{\mathrm{imp}}^j)].$$

The second equality uses independence as in Lemma 6. For the two model-derived policies, $q_{\mathrm{GM}}^j - \alpha$ measures the change in achieved risk when their calibration law is replaced by the GM reference. The Bengen-style policy is not calibrated to a risk target, so its difference from α does not have the same interpretation. In particular, evaluating a policy under GM mortality does not identify a GM-optimal policy.

We simulate financial impoverishment times using a weekly time grid and 100,000 paths arranged in 50,000 independent antithetic pairs; see [17, Ch. 4] for this and related variance reduction techniques. For each simulated time τ_{imp}, the corresponding GM survival probability gives the conditional probability of impoverishment before death. Paths with no observed hit by the simulation horizon contribute zero; when the horizon is the GM survival quantile with tail probability $5 \cdot 10^{-5}$, the omitted continuous-time contribution is bounded by that tail probability. This truncation bound does not include discretization error. Standard errors and the pointwise confidence bands below are computed from the antithetic pair averages.

Monitoring only the grid points would miss some boundary crossings. To detect impoverishment within a step, we use the Brownian-bridge exit-probability construction familiar from barrier-option simulation [25]. In a local distance-to-boundary coordinate, let $d_0 > 0$ and d_1 be the signed endpoint distances, positive on the solvent side, and let $v > 0$ be the frozen variance rate. If $d_1 > 0$, the conditional crossing probability over a step of length Δt is $\exp\bigl(-2d_0 d_1/(v\Delta t)\bigr)$; a nonpositive end distance implies a hit. Conditional on a hit and on both endpoints, we sample the fraction $u \in (0,1)$ of the step elapsed at first passage using [6, Sec. 6]

$$V = \frac{u}{1-u} \sim \mathrm{IG}\left(\frac{d_0}{|d_1|}, \frac{d_0^2}{v\Delta t}\right), \qquad \widehat{\tau}_{\mathrm{hit}} = t + \Delta t \frac{V}{1+V},$$

where the inverse Gaussian is parametrized by its mean and shape. The limiting law is used when $d_1 = 0$. For the two unconstrained model policies, wealth below the safe level is advanced by the exact geometric safe-gap transition; the modified Bengen-style policy uses Euler–Maruyama steps. We use a logarithmic gap-to-boundary coordinate on geometric-gap steps and a frozen wealth-diffusion coordinate on Euler steps and at the compactified endpoint. The sampled time is exact for the approximating bridge,

but the resulting hit detection remains approximate where the wealth diffusion or boundary changes within a step.

Table 2 checks the two closed-form policies against their own mortality laws. It displays the achieved risk minus its prescribed target; the coarse and fine columns provide a first time-step sensitivity check. Coarse and fine paths use coupled market increments and coupled auxiliary bridge draws; they are not a reconstruction of the same continuous path between grid points. Across all 56 main-run own-law estimates, the largest absolute discrepancy from the target is approximately 1.55 Monte Carlo standard errors. These estimates cover 26 risk levels at age 67 and $\alpha = 0.05$ at ages 60 and 75, for each of the two policies; the table displays only selected age-67 cases. The complete results are recorded in `policy_validation.csv`. Neither this comparison nor the table establishes absence of time-step bias.

	Main		Coarse		Fine	
α (%)	Exp	CE	Exp	CE	Exp	CE
0.1	0.000 (0.003)	0.001 (0.004)	0.005 (0.007)	0.001 (0.007)	0.004 (0.007)	0.001 (0.007)
5	−0.007 (0.038)	0.005 (0.050)	0.017 (0.076)	−0.015 (0.099)	0.032 (0.076)	−0.001 (0.099)
20	0.069 (0.068)	0.044 (0.091)	−0.311 (0.136)	−0.330 (0.181)	−0.281 (0.136)	−0.318 (0.181)

Table 2 Own-law validation at retirement age 67. Entries are the simulated impoverishment probability minus α, in percentage points; parentheses give antithetic-pair Monte Carlo standard errors. Main uses 100,000 paths at $\Delta t = 1/52$. Coarse and fine each use 25,000 paths at $\Delta t = 1/52$ and 1/104, respectively.

Figure 5 compares the sustainable withdrawals and the risks attained under the common GM reference. The Bengen-style withdrawal and its achieved risk are independent of α, whereas the two model-derived withdrawals vary with the chosen target. Their distinct safe-withdrawal benchmarks reflect the finite endpoint of the CE law and the infinite endpoint of the exponential law; they should not be interpreted as different risk-free rates.

Table 3 repeats the headline risk level $\alpha = 0.05$ at ages 60, 67, and 75. At age 67, the fitted CE and exponential rules yield annual withdrawals of approximately 20,937 and 19,289 real Euro, respectively; their achieved GM risks are approximately 5.077% and 5.153%. The table also shows how the withdrawals and achieved risks change with retirement age, with a separate mortality fit used at each age. These comparisons concern the stated policies and calibrations, not the unknown optimum under GM mortality.

The preceding comparison allows the model-derived policies to borrow. We next ask when the unconstrained CE feedback would leave the no-borrowing, no-short-selling simplex K_{con} from Definition 3. This connects the global feasibility condition of Proposition 38 with its pathwise consequences: a feedback may fail to be feasible at some state without a given investor ever encountering that state.

We fix retirement age $a = 67$ and the real annual withdrawal $\overline{C} = 30{,}000$ Euro. For each $\alpha \in \{0.01, 0.05, 0.10\}$, we move along $\mathcal{C}_{\overline{C},\alpha}$ while keeping the CE law fixed. With

Age	C_{Exp}	C_{CE}	$q_{\mathrm{GM}}^{\mathrm{Exp}}$	$q_{\mathrm{GM}}^{\mathrm{CE}}$	$q_{\mathrm{GM}}^{\mathrm{B}}$
60	17.65	18.51	5.654 (0.058)	5.083 (0.055)	6.405 (0.040)
67	19.29	20.94	5.153 (0.052)	5.077 (0.051)	1.872 (0.019)
75	22.58	25.55	4.275 (0.042)	5.071 (0.044)	0.182 (0.004)

Table 3 Age sensitivity at target risk $\alpha = 0.05$. Withdrawals are in thousand real Euro per year; GM impoverishment probabilities and their antithetic-pair Monte Carlo standard errors (in parentheses) are percentages. Each age uses 100,000 simulated paths.

the calibrated $\mathfrak{m} \geq 0$ componentwise, let $M = \mathbf{1}^\top \mathfrak{m} > 0$ and recall from (61) that

$$\mathcal{B}_{\mathrm{rel}}^*(\alpha) = \frac{\delta - 1}{M}\big(1 - \alpha^{1/\delta}\big).$$

We vary $\zeta \in [0.75, 4]$ and set $\mathcal{B}_{\mathrm{rel}} = \zeta \mathcal{B}_{\mathrm{rel}}^*(\alpha)$. The wealth pair corresponding to each bandwidth follows by solving the representation of ζ in Corollary 41 for x_{imp} and using $x_0 = x_{\mathrm{imp}}(1 + \mathcal{B}_{\mathrm{rel}})$:

$$x_{\mathrm{imp}}(\zeta) = \frac{\overline{C}}{r} \cdot \frac{\varepsilon_T M}{\varepsilon_T M + \zeta(\delta - 1)}, \tag{78}$$

$$x_0(\zeta, \alpha) = x_{\mathrm{imp}}(\zeta)\big(1 + \zeta \mathcal{B}_{\mathrm{rel}}^*(\alpha)\big). \tag{79}$$

Thus,

$$\big(x_{\mathrm{imp}}(\zeta), x_0(\zeta, \alpha)\big) \in \mathcal{C}_{\overline{C},\alpha},$$

so the CE-implied risk remains α as ζ varies. At a fixed ζ, the impoverishment level is independent of α, whereas initial wealth varies with the risk target. Moreover,

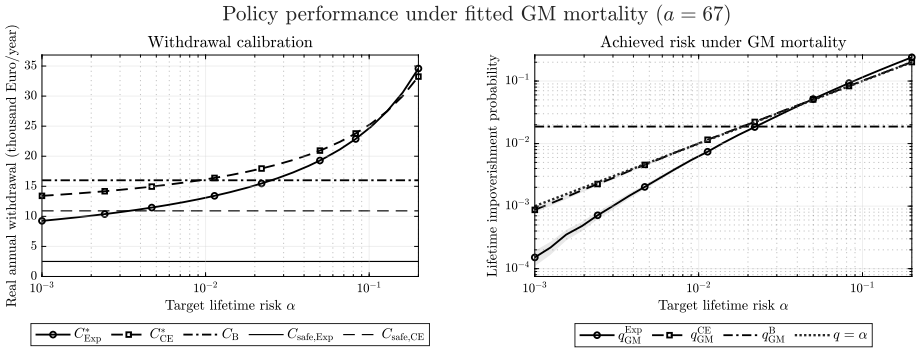


Fig. 5 Policy comparison at retirement age 67. Left: annual real withdrawals, including the two riskless safe-withdrawal benchmarks. Right: lifetime impoverishment probabilities under fitted GM mortality, with pointwise 95% Monte Carlo confidence bands and the reference line $q = \alpha$. The modified Bengen-style policy uses a fixed withdrawal and allocation. The horizontal axes and the right-hand vertical axis are logarithmic. Exp and CE denote the exponential and compatible compactified-exponential models.

Corollary 47 shows that, for each fixed α, the distribution of τ_{imp} is invariant with respect to ζ. The value $\zeta = 1$ recovers the maximizing pair of Corollary 40. The unconstrained feedback is simplex-valued throughout the risky region for $\zeta \leq 1$; for $\zeta > 1$, Corollary 41 confines possible borrowing to times before t^*_{bor}. Proposition 42 further distinguishes this global threshold from ζ_{init}, above which the initial wealth itself lies in the borrowing region, when such a threshold exists.

To measure how often this possibility is encountered along realized wealth paths, write $\pi^*_{\mathrm{tot}}(t,x) := \mathbf{1}^\top \pi^*(t,x)$ for the total risky weight and set $\tau_{\mathrm{bor}} := \inf\{0 \leq t < \tau_{\mathrm{imp}} : \pi^*_{\mathrm{tot}}(t, X_t) > 1\}$, with the infimum of the empty set equal to infinity. Since the calibrated Merton vector is componentwise nonnegative, such an exceedance is precisely a borrowing violation of the simplex constraint. Under independent fitted GM mortality we estimate

$$p^{\mathrm{GM}}_{\mathrm{bor}} = \mathbb{E}\big[S_{\mathrm{GM}}(\tau_{\mathrm{bor}})\mathbf{1}_{\{\tau_{\mathrm{bor}} < \tau_{\mathrm{imp}}\}}\big], \tag{80}$$

$$A^{\mathrm{GM}}_{\mathrm{bor}} = \mathbb{E}\left[\int_0^{\tau_{\mathrm{imp}}} S_{\mathrm{GM}}(t)\mathbf{1}_{\{\pi^*_{\mathrm{tot}}(t,X_t)>1\}}\,\mathrm{d}t\right]. \tag{81}$$

The first quantity is the probability of borrowing before death and impoverishment; the second is the expected time spent alive while borrowing, averaged over all investors, including those who never borrow. Both quantities account for mortality analytically through the GM survival function and vanish for $\zeta \leq 1$. For $\zeta > 1$, the financial paths are simulated only until impoverishment or the last possible borrowing time t^*_{bor}. First borrowing is detected with a separate single-boundary Brownian-bridge approximation, applied only up to the observed financial stopping time. This single-boundary approximation does not account for the conditioning imposed by the impoverishment-bridge outcome within the same time step. In particular, when a step ends at a sampled impoverishment time, consistent conditioning under the local Brownian approximation calls for a first-passage bridge rather than an ordinary Brownian bridge. For a linear boundary, such a bridge can be represented using a three-dimensional Bessel bridge; an alternative representation uses a path transformation of a Brownian bridge [5, Sec. 1 and Thm. 7]. Accounting jointly for both boundaries would nevertheless add considerable complexity. We therefore retain the simpler single-boundary approximation for this supplementary exposure analysis, without asserting that its error is negligible.

For similar reasons, occupation is computed separately from linearly interpolated boundary entries and exits, using a precomputed integral of GM survival to weight each interval. This approximation can miss borrowing excursions between grid points and need not agree pathwise with the bridge-corrected first-borrowing time. We use 100,000 financial paths in 50,000 antithetic pairs on a weekly grid.

Figure 6 displays the calibration geometry and the GM-weighted probability of ever borrowing. The left panel shows parameter pairs, not sample wealth trajectories: the three risk targets share the same x_{imp} at a given ζ, but have different initial wealth. The shaded region in the right panel is analytically feasible. Beyond $\zeta = 1$, global feasibility fails, although the chance of encountering the borrowing region can remain small near the threshold. The right panel also marks the initial-borrowing

thresholds that lie in the displayed range; once a policy starts in the borrowing region, the probability of encountering it is one.

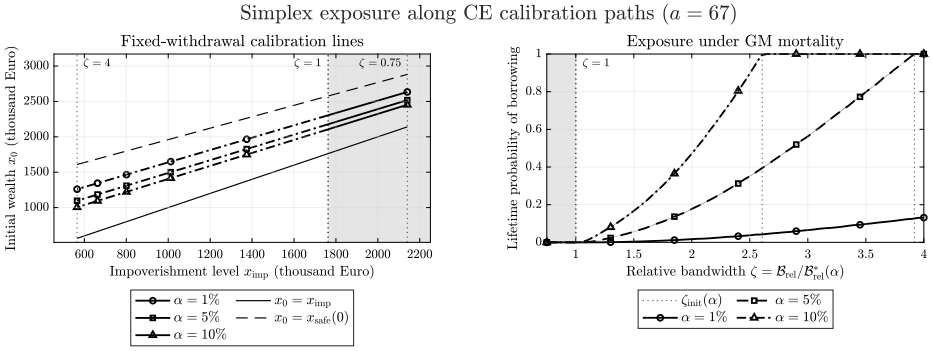


Fig. 6 Constraint exposure along fixed-withdrawal calibration lines at age 67. Left: compatible pairs (x_{imp}, x_0) for target risks of 1%, 5%, and 10%, with $x_0 = x_{\text{imp}}$ and the time-zero safe-wealth locus for reference. Right: GM-weighted probability of borrowing before death and impoverishment, with pointwise 95% confidence bands. Dotted lines mark the displayed initial-borrowing thresholds; the shaded range $\zeta \leq 1$ has zero exposure analytically. Simulated estimates use 100,000 weekly paths.

For example, at $\alpha = 0.05$, the estimated probability of ever borrowing rises from 1.73% at $\zeta = 1.25$ to 17.76% at $\zeta = 2$. The corresponding GM-weighted occupation times rise from 0.896 to 106.735 days (Table 4). At $\alpha = 0.10$ and $\zeta = 2$, the probability reaches 47.10%, with 326.909 expected borrowing days. Thus the loss of a global feasibility guarantee at $\zeta = 1$ need not imply substantial realized exposure just beyond the threshold, although exposure becomes material for wider bands.

α (%)	$\zeta = 1.25$	$\zeta = 1.5$	$\zeta = 2$
1	0.012 (0.003)	0.479 (0.037)	7.271 (0.245)
5	0.896 (0.035)	13.319 (0.250)	106.735 (1.040)
10	4.987 (0.094)	50.332 (0.508)	326.909 (1.643)

Table 4 GM-survival-weighted expected occupation of the borrowing region, in days, at age 67 and a fixed real withdrawal of 30,000 Euro per year. Entries average over all paths and independent death times; they are not durations conditional on borrowing. Parentheses contain antithetic-pair Monte Carlo standard errors. A year is converted to 365.25 days; the omitted $\zeta \leq 1$ entries are exactly zero.

These estimates describe the stated discretized simulation rather than the continuous-time exposure exactly. The borrowing bridge is not conditioned on the sampled impoverishment event and does not reconstruct a joint path between the two boundaries. Moreover, a bridge-detected borrowing excursion may contribute to first borrowing without contributing to the endpoint-linear occupation estimate. At

$\alpha = 0.05$, paired weekly and half-weekly refinements with 25,000 paths each give fine-minus-coarse borrowing probability differences of -0.011 and -0.010 percentage points at $\zeta = 1.25$ and 2.5, respectively, each smaller than its paired Monte Carlo standard error. The corresponding occupation differences are approximately -0.001 and 0.284 days, with paired standard errors of 0.013 and 0.253 days, respectively. At $\zeta = 1.05$, only two paths have an endpoint-interpolated borrowing interval at each resolution. This count need not coincide with the number of bridge-detected first-borrowing events. The full refinement results are recorded in `constraint_refinement.csv`. These comparisons do not bound the bridge-conditioning or remaining time-step bias, and the experiment does not solve the simplex-constrained control problem.

The preceding experiment records when the unconstrained feedback leaves the simplex, but it does not quantify the risk effect of enforcing the constraint. As a preliminary assessment, we therefore project the feedback pointwise onto the simplex in the covariance norm,

$$\widehat{\pi^*}(t,x) = \Pi^{\|\cdot\|_\Sigma}_{K_{\mathrm{con}}}\big(\pi^*(t,x)\big),$$

as suggested by Lemma 4. This produces an admissible policy comparator; it is not asserted to solve the dynamic simplex-constrained problem. In the risky region, the projected policy uses the unconstrained geometric safe-gap transition if projection leaves the feedback unchanged at the start of a step, and an Euler–Maruyama step with projected risky holdings if it changes the feedback. The former transition is exact only while the feedback remains unconstrained; a change in projection status within the step is not resolved.

Let $q^{\mathrm{GM}}_{\mathrm{unc}}(\alpha)$ denote the GM impoverishment probability under the unconstrained feedback and let $q^{\mathrm{GM}}_{\mathrm{proj}}(\zeta,\alpha)$ denote the corresponding probability under its pointwise projection. Along each fixed-withdrawal calibration line, Corollary 47 makes the unconstrained GM risk independent of ζ. On a common simulated market path, define the GM payoffs $X^{\mathrm{proj}} = S_{\mathrm{GM}}(\tau^{\mathrm{proj}}_{\mathrm{imp}})$ and $X^{\mathrm{unc}} = S_{\mathrm{GM}}(\tau^{\mathrm{unc}}_{\mathrm{imp}})$, and construct the control payoff $Y = S_{\mathrm{CE}}(\tau^{\mathrm{unc}}_{\mathrm{imp}})$. Here $S_M(\infty) = 0$. Crucially, Y uses the *same* unconstrained impoverishment time as X^{unc}; only the survival law applied to it changes. The exact CE calibration gives $\mathbb{E}[Y] = \alpha$. Following the control-variate principle [17, Ch. 4], an estimator of the projected GM risk is therefore

$$\widehat{q}^{\mathrm{GM}}_{\mathrm{proj,CV}}(b) = \overline{X}^{\mathrm{proj}} - b(\overline{Y} - \alpha),$$

where bars denote path averages (with antithetic pairing used for standard errors). We take the fixed coefficient $b = 1$. This need not minimize variance for every ζ; the variance-minimizing coefficient is a covariance-to-variance ratio for the antithetic-pair averages of X^{proj} and Y, which may vary with ζ. Nevertheless, $b = 1$ is structurally natural here: if GM mortality were replaced by CE and projection did not bind, $X^{\mathrm{proj}} = Y$ on every path, so their random variation would cancel exactly. It also preserves the following pathwise additive decomposition of the investor-facing risk surplus:

$$q^{\mathrm{GM}}_{\mathrm{proj}}(\zeta,\alpha) - \alpha = \mathbb{E}\big[X^{\mathrm{proj}} - Y\big]$$

$$= \underbrace{\mathbb{E}[X^{\mathrm{unc}} - Y]}_{\text{mortality-model contribution}} + \underbrace{\mathbb{E}\big[X^{\mathrm{proj}} - X^{\mathrm{unc}}\big]}_{\text{projection contribution}}. \tag{82}$$

The unconstrained payoff pair is simulated once per risk target and reused across bandwidths. We estimate both contributions by averages of their paired pathwise differences; in particular,

$$\widehat{q}^{\mathrm{GM}}_{\mathrm{proj,CV}}(1) - \alpha = \overline{(X^{\mathrm{unc}} - Y)} + \overline{(X^{\mathrm{proj}} - X^{\mathrm{unc}})}.$$

Thus their sum is the control-variate total on the same finite sample. In these runs the pairing reduces Monte Carlo noise, particularly in the mortality-model contribution. It does not remove discretization bias: the simulated CE anchor may differ from its exact target on a finite grid. The projection contribution is exactly zero for $\zeta \leq 1$, because the unconstrained feedback is already simplex-valued there.

As a run-specific diagnostic, at $\alpha = 0.05$ the simulated CE anchor falls 0.023 percentage points below its exact target, with a Monte Carlo standard error of 0.049 percentage points. This discrepancy combines Monte Carlo variation and finite-grid error. It is not an additional correction to the control-variate estimate, which already uses the known CE expectation. The anchor diagnostics are recorded in `policy_projection_baseline.csv`.

Figure 7 and Table 5 show the paired estimates from the same run. The dotted baseline isolates the effect of changing mortality for the unconstrained policy; the additional gap to a marked curve is the effect of replacing that policy by its projection. Neither component should be read as the value loss relative to the optimal GM-constrained policy. The mortality-model contributions are approximately -0.010, 0.072, and 0.104 percentage points for target risks of 1%, 5%, and 10%, respectively. The projection contribution is negligible just above $\zeta = 1$, but grows with the relative bandwidth. At $\alpha = 0.05$ it adds approximately 0.029 percentage points at $\zeta = 2$, 0.196 at $\zeta = 3$, and 0.404 at $\zeta = 4$. At $\zeta = 4$, the resulting total surpluses are approximately 0.041, 0.476, and 1.148 percentage points across the three target risks. The estimates indicate a growing effect of projection on wider calibration bands, without identifying the dynamically optimal constrained policy.

These values remain finite-grid estimates. Paired 25,000-path refinements at $\alpha = 0.05$ use weekly and half-weekly steps with coupled market increments and auxiliary bridge draws. At $\zeta = 2$, the control-variate total surpluses are approximately 0.120 and 0.114 percentage points, respectively; their paired difference is -0.006 percentage points with a standard error of 0.016 percentage points; see `policy_projection_refinement.csv` for the refinement results. The bridge check reduces the own-law shortfall associated with endpoint monitoring, but the frozen-coefficient bridge and projected Euler steps still warrant caution when interpreting the figures as continuous-time policy risks.

Taken together, the comparisons distinguish two effects of practical interest: using a tractable mortality model changes the risk attained under the GM reference, while enforcing the simplex constraint can change that risk further. The constraint-exposure

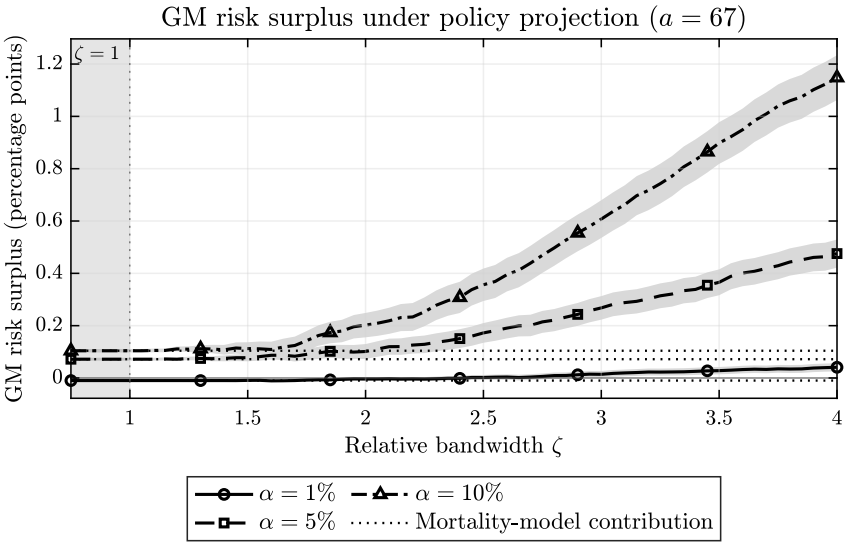


Fig. 7 GM risk surplus under pointwise projection of the compactified-exponential feedback at age 67. Marked curves show the control-variate estimate of $q^{\mathrm{GM}}_{\mathrm{proj}} - \alpha$ with pointwise 95% confidence bands; dotted lines show the mortality-model contribution, whose bands are omitted. Their signed vertical gap is the projection contribution in (82). The shaded range $\zeta \leq 1$ has zero projection contribution. Estimates use 100,000 weekly paths.

α (%)	ζ	Mortality	Projection	Total
1	1	−0.010 (0.002)	0.000 (0.000)	−0.010 (0.002)
1	2	−0.010 (0.002)	0.006 (0.003)	−0.004 (0.003)
1	3	−0.010 (0.002)	0.024 (0.006)	0.014 (0.006)
1	4	−0.010 (0.002)	0.051 (0.008)	0.041 (0.008)
5	1	0.072 (0.003)	0.000 (0.000)	0.072 (0.003)
5	2	0.072 (0.003)	0.029 (0.014)	0.101 (0.015)
5	3	0.072 (0.003)	0.196 (0.023)	0.268 (0.023)
5	4	0.072 (0.003)	0.404 (0.027)	0.476 (0.027)
10	1	0.104 (0.003)	0.000 (0.000)	0.104 (0.003)
10	2	0.104 (0.003)	0.098 (0.023)	0.202 (0.024)
10	3	0.104 (0.003)	0.503 (0.036)	0.607 (0.037)
10	4	0.104 (0.003)	1.044 (0.043)	1.148 (0.044)

Table 5 Paired control-variate GM risk-surplus decomposition under pointwise projection at selected relative bandwidths ζ. The mortality, projection, and total columns report the corresponding quantities in (82), in percentage points; parentheses give antithetic-pair Monte Carlo standard errors. The mortality column repeats the common unconstrained baseline for each target risk, and the projection contribution vanishes at $\zeta = 1$. The run uses 100,000 paths and $\Delta t = 1/52$ year.

experiment complements the latter comparison by showing how likely the unconstrained policy is to require borrowing before death and impoverishment, and the expected time spent alive while borrowing. These findings are specific to the stated calibration and numerical approximations; in particular, the projected feedback is not claimed to be optimal under the constraint. We next turn from lifetime probabilities to the timing of impoverishment itself.

6.3 First-passage approximation error

Before examining approximations to the impoverishment-time distribution, Figure 8 illustrates the underlying financial paths. We use age 67, target risk $\alpha = 0.05$, initial wealth $x_0 = 500{,}000$ Euro and impoverishment level $x_{\text{imp}} = 300{,}000$ Euro. The latter differs from the baseline used in the preceding policy comparison. For each scenario the four policies share the same market increments and the same independently sampled GM death time; financial paths are continued after death as counterfactual trajectories. The figure is illustrative rather than an estimate of event frequencies. It shows how the first time a path reaches the impoverishment level depends on the policy, motivating a closer study of the corresponding first-passage law.

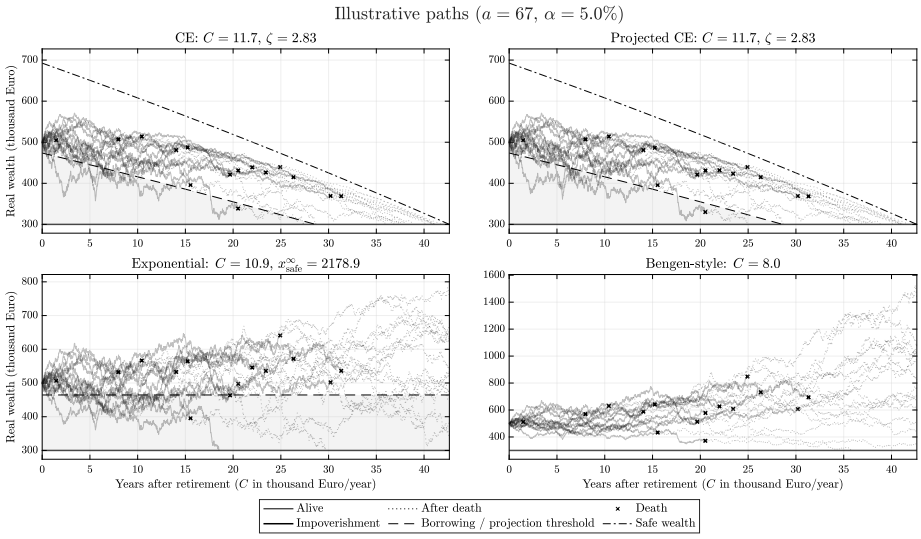


Fig. 8 Fifteen wealth scenarios under the CE, projected CE, exponential and modified Bengen-style policies, using common market increments and GM death times across policies. Paths are solid before death and dotted thereafter; crosses mark deaths before the financial path ends. Dashed lines and shading indicate the borrowing region or, for projected CE, the projection-activation region; dash-dotted lines mark safe wealth. Paths stop at impoverishment or the policy's simulation endpoint. Wealth and safe levels are in thousand real Euro, withdrawals in thousand real Euro per year. The exponential safe level lies above the displayed range; vertical scales differ except between the two CE panels.

We now compare the conditional law of τ_{imp} under compatible compactified-exponential mortality with the survival-weighted first term of Durbin's approximation

[11] from Corollary 50. We retain age 67, $x_0 = 500{,}000$ Euro and $x_{\mathrm{imp}} = 300{,}000$ Euro, and use the CE-optimal withdrawal and feedback at $\alpha \in \{0.01, 0.05, 0.10\}$. For fixed market and mortality parameters, Corollary 47 shows that the resulting financial first-passage law depends on the representative wealth pair only through the prescribed risk level. Thus the selected amounts do not introduce an additional dimension into this comparison.

We simulate 100,000 financial paths in antithetic pairs with weekly steps and the Brownian-bridge impoverishment detection described above. The comparison uses CE mortality, rather than the GM reference used in the preceding policy experiments, so that the conditional law and its bound correspond to Proposition 51. The random time of death is integrated out using CE survival weights $w_i = S_{\mathrm{CE}}(\tau_{\mathrm{imp}}^{(i)})$, with zero weight for a path without a detected hit. Thus the empirical conditional CDF is

$$\widehat{F}_{\mathrm{emp}}(t) = \frac{\sum_i w_i \mathbf{1}_{\{\tau_{\mathrm{imp}}^{(i)} \le t\}}}{\sum_i w_i}.$$

For the corresponding approximation, we integrate the first Durbin term f_1 in (76) against CE survival and normalize by $\widehat{q} = \int_0^T S_{\mathrm{CE}}(t) f_1(t)\, \mathrm{d}t$, as in Proposition 51. Figure 9 compares the resulting cumulative distribution functions. The curves are close at all three risk levels, although the visible discrepancy increases at higher risk levels.

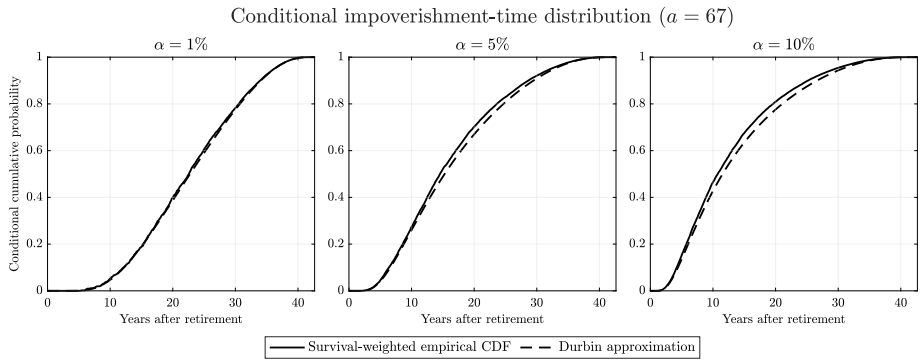


Fig. 9 Conditional impoverishment-time distributions under CE mortality at age 67. The solid curves are survival-weighted empirical CDFs from 100,000 simulated paths; the dashed curves are normalized survival-weighted first-term Durbin approximations. Empirical CDF values are evaluated on the saved time grid and joined by line segments.

Table 6 complements the figure with the continuous total-variation bound from Proposition 51, using the exact calibrated risk $q = \alpha$. The resulting bounds are between 5.64% and 6.59%. Since the event $\{\tau_{\mathrm{imp}} \le t\}$ is one of the sets over which total variation takes its supremum, the same bound controls the difference between the two exact conditional CDFs at every t. The plotted empirical CDF, however, also contains simulation error. To compare the two conditional laws empirically, we also report the TV distance between their probabilities on partitions into 10, 20, and 40 bins.

The bin edges are chosen as quantiles of the normalized, survival-weighted Durbin approximation, so that each bin has the same probability under this approximate conditional law. The corresponding main-run estimates range from 2.17% to 5.36%. For exact bin probabilities, binned TV is bounded above by continuous TV; the displayed Monte Carlo estimates, however, are not certified lower bounds and may themselves exceed the theoretical bound because of sampling and time-discretization error. Finer binning can also amplify sampling noise.

α (%)	$\widehat{q}$ (%)	TV bound (%)	$\widehat{d}_{\mathrm{TV}}^{(10)}$ (%)	$\widehat{d}_{\mathrm{TV}}^{(20)}$ (%)	$\widehat{d}_{\mathrm{TV}}^{(40)}$ (%)
1	1.06	5.64	2.17	3.55	5.36
5	5.34	6.32	3.41	3.80	4.38
10	10.71	6.59	3.88	4.37	4.80

Table 6 Conditional first-passage comparison at age 67. All entries are percentages. $\widehat{q}$ is the survival-weighted Durbin mass, and the TV bound is $1-\alpha/\widehat{q}$. The remaining columns are estimated binned TV discrepancies for the indicated numbers of equal-Durbin-probability bins; they are not certified bounds and no standard error for the TV statistic is reported.

For a time-step check, we simulate 25,000 coupled paths at weekly and half-weekly resolution, using the same construction of coupled market and bridge inputs as in the earlier refinement study. The largest absolute fine-minus-coarse difference between empirical CDF values on the saved grid is at most 0.90 percentage points across the three risk levels. For the 20-bin TV estimates, the fine-minus-coarse changes are approximately 0.08, -0.10, and 0.74 percentage points, respectively. In particular, at $\alpha = 0.01$, the fine-grid estimate of 5.95% slightly exceeds the 5.64% theoretical bound; this is not a contradiction because the estimated binned discrepancy is subject to Monte Carlo and discretization errors. These refinement results are recorded in `first_passage_refinement.csv`.

The refinement also records standard errors for the individual changes in bin probabilities. If $I_{i,b}$ indicates that the simulated hit lies in bin b, the self-normalized estimator is $\widehat{p}_b = \sum_i w_i I_{i,b} / \sum_i w_i$. Both its numerator and denominator are estimated from the same paths. Applying the delta method [32, Ch. 3] to the ratio of their sample means gives estimated pathwise influence values

$$\widehat{\psi}_{i,b} = \frac{w_i\big(I_{i,b} - \widehat{p}_b\big)}{\overline{w}}.$$

The term $-w_i\widehat{p}_b$ accounts for the effect of estimating the normalizing denominator; treating the weights as already normalized fixed probabilities would omit this contribution. We subtract the coarse influence value from the fine value on each coupled path and estimate its standard error from antithetic-pair averages. Across the three risk levels, the largest such binwise standard errors are approximately 0.46, 0.21, and

0.17 percentage points, respectively. These are approximate standard errors for individual bin-probability changes, not for the TV differences or CDF maxima, and they do not quantify time-discretization bias.

Taken together, the experiments illustrate how the closed-form construction translates into calibrated withdrawals, investment policies and impoverishment-time distributions, while also showing the relevance of borrowing constraints and approximation error. The following section discusses the scope of these findings and the modelling assumptions behind them.

7 Discussion and limitations

8 Conclusion

Declarations

Appendix A Technical details

References

[1] Allianz Global Investors (2025) Long-term capital market assumptions: Highlights. Tech. rep., Allianz Global Investors, URL https://ap.allianzgi.com/-/media/allianzgi/ap/ap/ideas-connect/pdfs/web-capitalmarketassumptions-lite-r2025q3-allianzgi-ap-stamped.pdf?hash=987C4302DFE2966615DDCFA6E0F15772&rev=458d3e7924d74830a0ed7728d8c57536, released in Q3 2025; assumptions as at 30 June 2025

[2] Bayraktar E, Young VR (2007) Minimizing the probability of lifetime ruin under borrowing constraints. Insurance: Mathematics and Economics 41(1):196–221. URL https://arxiv.org/abs/math/0703850, arXiv:math/0703850

[3] Bayraktar E, Young VR (2011) Proving regularity of the minimal probability of ruin via a game of stopping and control. Finance and Stochastics 15(4):785–818. URL https://arxiv.org/abs/0704.2244, arXiv:0704.2244

[4] Bengen WP (1994) Determining withdrawal rates using historical data. Journal of Financial planning 7(4):171–180

[5] Bertoin J, Chaumont L, Pitman J (2003) Path transformations of first passage bridges. Electronic Communications in Probability 8:155–166. https://doi.org/10.1214/ECP.v8-1096

[6] Beskos A, Roberts GO (2005) Exact simulation of diffusions. The Annals of Applied Probability 15(4):2422–2444. https://doi.org/10.1214/105051605000000485

[7] Brand C, Lisack N, Mazelis F (2025) Natural rate estimates for the euro area: Insights, uncertainties and shortcomings. ECB Economic Bulletin (1). URL https://www.ecb.europa.eu/press/economic-bulletin/focus/2025/html/ecb.ebbox202501_08~3be5a005f9.en.html, published as part of Issue 1/2025

[8] Bund-Länder – Demographieportal (2024) Beitragszahler je altersrentner in der gesetzlichen rentenversicherung. URL https://www.demografie-portal.de/DE/Fakten/altersrentner-beitragszahler.html

[9] Deutsche Bundesbank (2025) Financial stability review 2025. Tech. rep., Deutsche Bundesbank, URL https://publikationen.bundesbank.de/publikationen-en/reports-studies/financial-stability-reviews/financial-stability-review-2025-969016, see the section "Stability situation in the German financial system"

[10] Durbin J (1985) The first-passage density of a continuous gaussian process to a general boundary. Journal of Applied Probability 22(1):99–122. https://doi.org/10.2307/3213751, URL https://doi.org/10.2307/3213751

[11] Durbin J, Williams D (1992) The first-passage density of the brownian motion process to a curved boundary. Journal of Applied Probability 29(2):291–304. URL http://www.jstor.org/stable/3214567

[12] European Central Bank (2021) Two per cent inflation target. URL https://www.ecb.europa.eu/mopo/strategy/pricestab/html/index.en.html, eCB monetary policy strategy

[13] Federal Statistical Office of Germany (Destatis) (2025) By 2035, one quarter of germany's population will be aged 67 or over. Press release, URL https://www.destatis.de/EN/Press/2025/12/PE25_446_12.html, accessed May 2026

[14] Fidelity International (2026) Capital market assumptions update. H1 2026 capital market assumptions, Fidelity International, URL https://www.fidelityinternational.com/static/master/media/pdf/download-material/capital-market-assumptions-update-h1-2026-fidelity-international.pdf, assumptions based on data as of 31 March 2026

[15] Fisher I (1896) Appreciation and Interest. Macmillan for the American Economic Association, New York, URL https://fraser.stlouisfed.org/files/docs/meltzer/appreciationinte00fish.pdf

[16] German Council of Economic Experts (2020) Demographic change: Sustainable retirement provision. Annual Report 2020/21, Chapter 6, URL https://www.sachverstaendigenrat-wirtschaft.de/fileadmin/dateiablage/gutachten/jg202021/JG202021_Chapter_6.pdf, accessed May 2026

[17] Glasserman P (2003) Monte Carlo Methods in Financial Engineering. Stochastic Modelling and Applied Probability, Springer, New York, https://doi.org/10.1007/978-0-387-21617-1

[18] Korn R (2000) Value preserving strategies and a general framework for local approaches to optimal portfolios. Mathematical Finance 10(2):227–241

[19] Kullback S, Leibler RA (1951) On information and sufficiency. The Annals of Mathematical Statistics 22(1):79–86. https://doi.org/10.1214/aoms/1177729694

[20] Liang X, Young VR (2022) A simple and nearly optimal investment strategy to minimize the probability of lifetime ruin. ASTIN Bulletin 52(2):619–643. https://doi.org/10.1017/asb.2022.3, URL https://doi.org/10.1017/asb.2022.3

[21] Merton RC (1969) Lifetime portfolio selection under uncertainty: The continuous-time case. The Review of Economics and Statistics 51(3):247–257

[22] Merton RC (1971) Optimum consumption and portfolio rules in a continuous-time model. Journal of Economic Theory 3(4):373–413

[23] Milevsky MA (2020) Calibrating Gompertz in reverse: What is your longevity-risk-adjusted global age? Insurance: Mathematics and Economics 92:147–161

[24] Milevsky MA, Robinson C (2000) Self-annuitization and ruin in retirement. North American Actuarial Journal 4(4):112–129

[25] Moon KS (2008) Efficient monte carlo algorithm for pricing barrier options. Communications of the Korean Mathematical Society 23(2):285–294. https://doi.org/10.4134/CKMS.2008.23.2.285

[26] Moore KS, Young VR (2006) Optimal and simple, nearly optimal rules for minimizing the probability of financial ruin in retirement. North American Actuarial Journal 10(4):145–161

[27] OECD (2025) Pensions at a glance 2025: Germany. URL https://www.oecd.org/en/publications/pensions-at-a-glance-2025-country-notes_8a53ef12-en/germany_c480c8fa-en.html, accessed May 2026

[28] RBC Global Asset Management (2026) Developing long-term capital market expectations. Tech. rep., RBC Global Asset Management, URL https://www.rbcgam.com/documents/en/the-big-picture/big-picture-long-term-capital-markets-assumptions.pdf, long-term capital market assumptions as of 30 June 2026

[29] Samuelson PA (1969) Lifetime portfolio selection by dynamic stochastic programming. The Review of Economics and Statistics 51(3):239–246

[30] Statista (2008) Verhältnis der "inaktiven" bevölkerung ab 65 jahre zur gesamten erwerbsbevölkerung in deutschland von 2000 bis 2050. URL https://de.statista.com/statistik/daten/studie/14177/umfrage/deutschland-anteil-rentner-an-erwerbsbevoelkerung/

[31] Statistisches Bundesamt (Destatis) (2026) Lebenserwartung einzelner geburtsjahrg"ange (Kohortensterbetafeln). Cohort life tables and methodological information, URL https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Sterbefaelle-Lebenserwartung/kohortensterbetafeln.html, 2026 update; medium variant V2

[32] van der Vaart AW (1998) Asymptotic Statistics. Cambridge University Press, Cambridge, https://doi.org/10.1017/CBO9780511802256

[33] Verbraucherschutz Sachsen (2026) Trade republic: Zinsen verpasst? URL https://verbraucherschutz-sachsen.de/verfahren/trade-republic-zinsen/, reports the advertised cash rate in October 2026

[34] Young VR (2004) Optimal investment strategy to minimize the probability of lifetime ruin. North American Actuarial Journal 8(4). https://doi.org/10.1080/10920277.2004.10596174, URL https://doi.org/10.1080/10920277.2004.10596174

[35] Young VR, Zhang Y (2016) Lifetime ruin under ambiguous hazard rate. Insurance: Mathematics and Economics 70:125–134. https://doi.org/10.1016/j.insmatheco.2016.06.007, URL https://doi.org/10.1016/j.insmatheco.2016.06.007

Todo list