Munich Personal RePEc Archive

The Mortality Input Problem: Trajectory-Dependent Death and the Lifecycle Model

Zhorin, Victor

15 May 2026

Online at https://mpra.ub.uni-muenchen.de/129315/
MPRA Paper No. 129315, posted 12 Jun 2026 12:12 UTC

The Mortality Input Problem: Trajectory-Dependent Death and the Lifecycle Model

Victor Zhorin*

May 2026

Abstract

Heterogeneous-agent macroeconomics has reformed the income and wealth sides of the household problem. The HANK program established that aggregate dynamics require the cross-sectional distribution of marginal propensities to consume, balance-sheet exposures, and permanent income. Every model in this program inherits, without examination, a smooth actuarial mortality hazard calibrated to population life tables. This is the last unreformed input, and it is structurally wrong.

The expected residual life of an agent is a value function defined on a manifold whose boundary is death. That function has a geometric property rarely stated in the economic literature: its curvature diverges at a specific power-law rate as the agent approaches the boundary. Every lifecycle model that represents mortality as a smooth hazard imposes a bounded-curvature approximation on a function whose curvature is unbounded. No function in the smooth hazard class can encode the shock structure of catastrophic diagnoses that dominate individual mortality trajectories near the boundary, regardless of how many parameters the hazard contains.

We establish three structural consequences. First, the population-level response to symmetric interventions is asymmetric: the worsening direction systematically exceeds the improving direction by a factor governed by a single cross-sectional moment, the covariance between boundary curvature and intervention exposure. Second, this covariance is not a small correction in clinical settings: unlike the borrowing-constraint analog in macroeconomics, where the boundary binds for a minority of agents, the mortality boundary is universal, and the population-level curvature integral diverges in a way the macroe-

*Independent researcher. vzhorin@uchicago.edu. ORCID: 0000-0001-9777-4421.

conomic scaling intuition cannot accommodate. Third, the representation class that correctly encodes the boundary geometry exists: trained networks with piecewise-linear activation produce value functions that are exactly tropical polynomials in the max-plus semiring, with the density of the piecewise structure near the boundary encoding the curvature divergence that smooth functions cannot.

A constellation of puzzles that the lifecycle literature has documented and not resolved, covering wealth decumulation, bequest dispersion, annuitization, retirement timing, portfolio composition, health expenditure at end of life, long-term care insurance, Social Security claiming, and pension tax choices. These are projections of a single geometric fact. Correcting the mortality input generates each of them as equilibrium properties of the model rather than as calibrated parameters or behavioral anomalies.

Keywords: lifecycle models; heterogeneous agents; mortality; absorbing boundaries; tropical geometry; sufficient statistics; bequest motives; annuity puzzle; retirement; portfolio choice; health expenditure

JEL Classification: E21; G22; H55; I10; C14; G11

---

1. Introduction

1.1 The Standard Treatment and Its Failure

Every lifecycle Hamilton-Jacobi-Bellman equation in the heterogeneous-agent macroeconomics literature takes a form in the neighborhood of:

$$rV(a,z) = \max_c \{u(c) + V_a \cdot (z + ra - c) + \lambda [V(a,z') - V(a,z)] - \delta_k [V(a,z) - b(a)]\}$$

where a is wealth, z is income, u is the utility function, $b(a)$ is the bequest value, and δ_k is the mortality hazard rate, a smooth age-dependent scalar calibrated to actuarial population tables. Straub (2019) agents die with certainty at biological age 90 facing a positive δ_k between periods. De Nardi (2004) uses the same structure with a parameterized bequest motive layered on top. Kaplan-Moll-Violante (2018) employ it in their two-asset HANK model. Stantcheva (2015) inherits it as the terminal condition governing when bequests are transferred in a dynastic Barro-Becker model.

The mortality term enters as a discount-rate adjustment: $-\delta_k \cdot (V - b)$. The boundary condition $V(0, z) = b(0)$ is specified at a terminal age, not derived from the mortality process. This is the misspecification.

The true mortality term is not $-\delta_k \cdot (V - b)$ with smooth δ_k. It is the generator of the hitting-time distribution, which has second derivative diverging as the agent approaches the absorbing boundary at rate $s^{-3/2}$ where s is the distance to the boundary. The standard treatment replaces a function with unbounded curvature with a linear approximation whose error is largest precisely where bequest and decumulation decisions are made.

The purpose of this paper is to establish that this is not a matter of calibration, not choosing a better δ_k, but of representation: no function in the smooth hazard class can encode the geometry of the correct value function near the absorbing boundary.

1.2 The Puzzle Landscape

The lifecycle literature has accumulated a substantial inventory of empirical anomalies that resist explanation within the standard smooth-mortality framework. They span savings behavior, insurance choices, retirement timing, portfolio composition, health expenditure, and tax optimization. They are documented by separate literatures using separate methods and attributed to separate domain-specific mechanisms. We claim they are not separate phenomena. They are projections of a single geometric fact: the value function is steep, nonlinear, and asymmetric near the absorbing boundary, and agents near that boundary behave in ways that bounded-curvature approximations cannot represent.

Section 6 works through each puzzle in detail. Here we catalogue them briefly to establish scope.

Savings and decumulation. The elderly dissave far more slowly after retirement than standard models predict: median wealth falls 2-3% annually against predicted 8-12%, with dramatic cross-sectional heterogeneity at the same age and wealth level that the smooth model has no mechanism to generate. Precautionary saving by the high-wealth elderly persists at levels an order of magnitude above buffer-stock predictions.

Bequests. Bequests are larger and more dispersed than voluntary-bequest models produce, even conditional on retirement wealth. The dispersion is the signature: smooth mortality ties bequest incentives to age, not trajectory, and cannot generate the observed cross-sectional variance.

Annuitization. Actuarially-fair annuities are rejected by the large majority of retirees despite Yaari's (1965) prediction of near-universal purchase. The puzzle is not the average rejection rate; bequest motives and adverse selection account for part of that, but the correlation between rejection and clinical trajectory markers, which the population-table pricing that defines "actuarially fair" cannot accommodate.

Social Security claiming. Sixty percent of Americans claim Social Security at or before full retirement age despite actuarially unfavorable terms under population mortality tables. The rate of early claiming is strongly predicted by self-reported health, itself a noisy proxy for clinical trajectory. The smooth model, calibrated to population tables, predicts the opposite of what is observed.

Retirement timing. Retirement is sharply accelerated by catastrophic diagnoses and decelerated, sometimes indefinitely, by the absence of them. The smooth hazard model predicts smooth optimal stopping; the trajectory-dependent model predicts the jump retirement that clinical shocks actually produce.

Long-term care insurance. LTCI take-up is approximately 7% despite substantial actuarial LTCI risk (Brown and Finkelstein 2007). Standard explanations (Medicaid crowd-out, adverse selection, behavioral factors) account for at most half of the gap. The curvature asymmetry provides the missing piece: for near-boundary agents, the gains from LTCI coverage are smaller than smooth models suggest because the value function is already steep in the direction of adverse events.

Portfolio composition. Elderly investors maintain equity exposure well above the lifecycle model's age-based derisking prescription. They treat home equity as a last resort, accumulate excess cash, and maintain life insurance past the point where the standard model says it should be cancelled. Each of these is a hedging response to boundary proximity that the smooth model cannot identify because it cannot represent the boundary geometry.

Health expenditure at end of life. Family and individual willingness to authorize high-cost, low-expected-benefit treatment at end of life is dramatically above what the smooth utility model predicts. This is not irrationality. Near the absorbing boundary, where curvature is large, the marginal value of small extensions in survival time is high and the asymmetry means deterioration costs more than improvement gains, producing a rational demand for aggressive intervention that looks like over-treatment from the vantage of population-table analysis.

Pension tax choices. The Roth versus traditional IRA conversion decision, the choice between defined-benefit and defined-contribution pensions, and the timing of estate gifts before death all exhibit patterns correlated with clinical trajectory in ways that smooth mortality cannot generate. Each is a variant of the same underlying problem: the optimal tax strategy depends on the timing of the terminal event, and the timing depends on the clinical trajectory in a nonlinear way.

Each of these will be resolved in Section 6 as a consequence of correcting the mortality input, not as an additional parameter or a behavioral overlay.

1.3 The Geometric Identity

The core mathematical object is the value function $V(x) = \mathbb{E}^x[\tau]$, where x is the agent's state, Ω is the interior (the living region), and $\tau = \inf\{t \geq 0 : X_t \in \partial\Omega\}$ is the first hitting time of the absorbing boundary (death). For a nondegenerate diffusion on a domain $\Omega \subset \mathbb{R}^d$ with C^2 boundary, V is the unique classical solution of:

$$\mathcal{L}V(x) = -1 \quad \text{on } \Omega, \qquad V(x) = 0 \quad \text{on } \partial\Omega$$

where $\mathcal{L} = b \cdot \nabla + \frac{1}{2}\text{tr}(\sigma\sigma^\top \nabla^2)$ is the infinitesimal generator. The standard result (Friedman 1975) gives the boundary asymptotics: for $x_0 \in \partial\Omega$ and unit vector ν pointing into Ω,

$$\left|\partial_\nu^2 V(x_0 + s\nu)\right| \sim c_2 \cdot s^{-3/2} \quad \text{as } s \to 0^+$$

The curvature diverges at rate $s^{-3/2}$. This is not a peculiarity of a particular parameterization. It is exact for Brownian hitting times to a smooth boundary and bounds the general nondegenerate case from below. Every smooth actuarial hazard model, by construction, produces a value function with bounded second derivative. The approximation error concentrates at $\partial\Omega$, precisely where bequest motives and decumulation incentives are determined.

1.4 The Population-Level Consequence

A symmetric perturbation $\pm\delta$ along direction ν at interior point x produces a response:

$$[V(x+\delta\nu)-V(x)]+[V(x-\delta\nu)-V(x)]=\delta^2\partial_\nu^2 V(x)+O(\delta^3)$$

The asymmetry between worsening and improving interventions is governed by the local second derivative. Summed across a heterogeneous population with intervention exposure δ_i and mean-zero randomization, the population-level response factors as:

$$\Delta\bar{V}\approx\frac{1}{2}\left[\mathbb{E}[\delta_i^2]\cdot\mathbb{E}[\kappa_i]+\mathrm{Cov}(\delta_i^2,\kappa_i)\right]$$

where $\kappa_i=|\partial_\nu^2 V(x_i)|$ is the local boundary curvature at patient i's state. The covariance $\mathrm{Cov}(\delta_i^2,\kappa_i)$ is the sufficient statistic: it captures whether intervention intensity is targeted at high-curvature agents (those near the boundary). This decomposition is structurally identical to Auclert's (2019) redistribution channel in monetary policy transmission; the methodological move is direct transposition.

The scaling, however, differs qualitatively. In macroeconomic settings the relevant boundary is the borrowing constraint, binding for 10–30% of agents at any moment. The curvature is large only for the constrained subpopulation; the unconstrained majority contributes near-zero values to the covariance, producing the 7% output-volatility contribution Berger, Bocola, and Dovis (2023) estimate. In clinical mortality the boundary is universal. Every agent is on a finite-time path to $\partial\Omega$; the curvature integral $\int_0^1 q^{-3/2}dq$ diverges. The cross-sectional moment is not a small correction to a representative-agent benchmark; it is the dominant feature of the aggregate response.

1.5 The Representation Class

The representation class that correctly encodes the boundary geometry exists. Trained ReLU networks produce embeddings whose value function is exactly a tropical polynomial in the rational max-plus semiring: a piecewise-affine function $V_\theta:\mathbb{R}^d\to\mathbb{R}$ with pieces corresponding to vertices of a polyhedral subdivision of input space (Zhang, Naitzat, and Lim 2018; Zhorin 2026a). Within each piece, V_θ is affine and second derivatives vanish. Across piece boundaries, V_θ has discontinuous gradient and the second derivative is a sum of Dirac measures supported on the polyhedral complex. The local curvature depends on the density of the polyhedral subdivision near x: near the absorbing boundary the subdivision is arbitrarily fine; far from the boundary it is coarse. The trained model encodes the boundary-curvature asym-

metry by construction, provided the training data contains sufficient signal near the boundary.

This architectural fact has a direct implication for lifecycle modeling. The CLS embedding produced by the trained transformer plays a sufficient-statistic role: once $\Phi(x_i)$ is computed for agent i, the curvature $\kappa(x_i, \nu)$ in any direction is recoverable as a finite-difference computation on the trained network. The population-level moment $\text{Cov}(\delta_i^2, \kappa_i)$ is then a single inner product across the population, computable in linear time.

1.6 Relation to Prior Literature

Lifecycle models with differential mortality. De Nardi, French, and Jones (2010) is the closest predecessor: it incorporates differential mortality across socioeconomic groups but maintains the smooth parametric form, fitting separate hazard functions to separate groups. The present paper argues that the class of smooth hazard functions is the problem, not the calibration of a particular representative within that class. De Nardi's (2004) explicit dismissal of differential mortality as a quantitatively small correction is refuted in Section 8 of this paper on structural grounds: the dismissal relied on the macroeconomic scaling argument that fails when the boundary is universal.

The HANK program. Auclert (2019), Straub (2019), and Auclert, Rognlie, and Straub (2021) established that aggregate dynamics require cross-sectional moments. The aggregation result of this paper (Section 3) is a direct transposition of their methodology into the mortality domain. The contribution is both the identification of the relevant moment (boundary curvature rather than MPC or net nominal position) and the demonstration that the scaling differs qualitatively from the macroeconomic benchmark.

Hansen and Sargent: Tenuous Beliefs. Hansen and Sargent (2017) develop a framework for pricing macroeconomic uncertainty when investors have model ambiguity across structured parametric models and misspecification concerns about all of them. They use Chernoff entropy, a statistical distance governing detection error probabilities, to calibrate the penalty parameter quantifying how far from the structured models a worst-case adversary may deviate, and they solve the resulting robust decision problem via HJB equations. Their central empirical result is a directional asymmetry in uncertainty prices: investors especially fear bad-state persistence. Section 3.6 of this paper establishes that this asymmetry is a direct con-

sequence of the geometry of value functions near absorbing boundaries, with the Chernoff entropy between the trajectory-dependent and population-table mortality distributions providing the information-theoretic measure of the lifecycle model's misspecification severity. The Chernoff entropy calculations in Hansen and Sargent (2017) were contributed by Victor Zhorin (acknowledged in that paper), and the connection between those calculations and the boundary-value machinery developed here is the basis for Section 3.6.

Behavioral economics. The gain-loss asymmetry documented by Kahneman and Tversky (1979) in prospect theory (losses loom roughly 2.25 times as large as equivalent gains) is the most replicated finding in behavioral economics. It is typically attributed to reference-point psychology. Section 3.6 establishes that the same quantitative asymmetry arises as a consequence of boundary curvature for rational agents near the mortality boundary, without any psychological mechanism, and that the geometric explanation generates a specific testable prediction (increasing loss aversion with boundary proximity) that distinguishes it from the behavioral explanation. This does not refute prospect theory; it provides a structural microfoundation for the loss-aversion asymmetry in the domain where the behavioral evidence is strongest.

The Stantcheva program. Stantcheva (2015, NBER WP 21177) derives optimal bequest tax formulas as functions of estimable sufficient statistics in a dynastic model where parents transfer resources through education investments and financial bequests. The terminal condition governing when the bequest occurs; the mortality process is taken as exogenous. The correction derived here implies that the sufficient statistics for optimal bequest taxation include the curvature-exposure covariance as an additional moment. Stantcheva's formulas remain correct in structure; the moments they depend on change, and the optimal bequest tax differs from what population-table calibration delivers.

The Townsend program. Moll, Townsend, and Zhorin (2017) is the direct predecessor within the heterogeneous-agent literature: it uses the Thai household data to study heterogeneous-agent dynamics with financial frictions, placing the Townsend household-as-corporate-firm framework within the HANK methodology. Mortality appears in that framework as the exit event terminating the household balance sheet. The present paper supplies the trajectory-dependent mortality representation that was unavailable when that paper was written.

1.7 Organization

Section 2 establishes the trajectory space: medical event sequences as causal paths, the Kleene closure, the embedding into rough path space, and the absorbing boundary. Section 3 establishes the geometric identity and the aggregation result. Section 4 proves that the standard lifecycle model is structurally misspecified and that the Cox family is incapable of correcting it. Section 5 establishes the tropical polynomial representation class. Section 6 resolves empirical puzzles. Section 7 develops the empirical application to financial trajectories. Section 8 presents the corrected lifecycle model and addresses the De Nardi scaling objection. Section 9 discusses extensions and the Stantcheva bequest tax implication.

2. The Trajectory Space

The standard heterogeneous-agent model parameterizes agent type by (a, z): wealth and income. The mortality process is attached as a scalar hazard $\delta_k(a, z, \text{age})$. The crucial move of this paper is to replace the scalar health index with the full clinical trajectory as the agent's mortality-relevant type. This section establishes the mathematical structure of the trajectory space.

2.1 Event Space and Causal Structure

Definition (). Let $\mathcal{E}$ be a finite vocabulary of medical event types: ICD-10 diagnosis codes, procedure codes, medication codes. A timestamped event is a pair $(e, t) \in \mathcal{E} \times \mathbb{R}^+$ where t denotes occurrence time.

Definition (). A trajectory is a finite sequence $\gamma = ((e_1, t_1), \ldots, (e_n, t_n))$ with $t_1 \leq t_2 \leq \cdots \leq t_n$. The space of all trajectories is the Kleene closure:

$$\Gamma = \bigcup_{n=0}^{\infty} (\mathcal{E} \times \mathbb{R}^+)^n$$

equipped with the causal partial order $i \prec j \iff t_i < t_j$.

The Kleene closure is the computer scientist's path space. In the economic language, each trajectory $\gamma \in \Gamma$ is a complete type description for the mortality-relevant component of agent heterogeneity. The standard lifecycle model compresses this type to a scalar age-dependent hazard; the present framework preserves the full structure.

Remark (). Consider a 45-year-old male with a diagnosis of glioblastoma (ICD-10 C71.9). His scalar hazard under any actuarial table is that of a 45-year-old male. His true mortality probability within 24 months is approximately 0.90. The compression error is not a matter of heterogeneous tables; it is a matter of the single event (e_{GBM}, t_0) inducing a discontinuity in the value function along a direction that no pre-specified scalar index can capture. This is the generic structure of catastrophic diagnoses: not steep gradients in a smooth function but shocks, discontinuities in the value function on trajectory space (Zhorin 2026a, Section 1).

2.2 Embedding into Analytic Path Space

The Kleene closure Γ embeds into the analyst's rough path space via a piecewise-constant lift. Fix an embedding map $\phi : \mathcal{E} \to \mathbb{R}^d$ assigning each event type a vector representation. Each trajectory $\gamma = ((e_1, t_1), \ldots, (e_n, t_n))$ maps to a càdlàg path $\tilde{\gamma} : [0, T] \to \mathbb{R}^d$ with:

$$\tilde{\gamma}(t) = \phi(e_i) \quad \text{for } t \in [t_i, t_{i+1})$$

This piecewise-constant lift is the bridge from the discrete trajectory space to the analytic tools of rough path theory.

Definition (). For a path $\gamma : [0, T] \to \mathbb{R}^d$ of bounded variation, the signature is the sequence of iterated integrals:

$$S(\gamma) = \Big(1, S(\gamma)^1, S(\gamma)^2, \ldots\Big) \in T((\mathbb{R}^d)) = \bigoplus_{k=0}^{\infty} (\mathbb{R}^d)^{\otimes k}$$

where the k-th level is $S(\gamma)^k_{i_1,\ldots,i_k} = \int_{0<t_1<\cdots<t_k<T} d\gamma^{i_1}_{t_1} \otimes \cdots \otimes d\gamma^{i_k}_{t_k}$.

By Chen (1958) and Hambly-Lyons (2010), the signature $S(\gamma)$ uniquely determines γ up to tree-like equivalence; for generic paths it is a complete invariant. The trajectory type, for mortality prediction purposes, is the equivalence class under the signature.

2.3 The Absorbing Boundary

Definition (). Let $\partial\Omega \subset \Gamma$ denote the set of trajectories terminating in death within the prediction horizon T. More precisely, $\partial\Omega$ is the image in trajectory space of the event $e^\dagger$, the terminal event (death). For the lifecycle model, $\partial\Omega$ is the set of

agent states from which the exit event is realized, and the value function satisfies the boundary condition $V(\gamma) = 0$ for $\gamma \in \partial\Omega$.

Remark (). The absorbing boundary of clinical mortality differs from the borrowing constraint in lifecycle models in one structural respect: every trajectory terminates at $\partial\Omega$ in finite time with probability one. This is the formal content of "the boundary is universal." The hitting time τ is almost surely finite for every agent, at every initial state $x \in \Omega$. The borrowing constraint, by contrast, is never reached by a positive fraction of agents; those with sufficient wealth never encounter it. The difference in universality drives the difference in empirical scaling established in Section 8.

2.4 The Trajectory as Economic Type

The key economic reinterpretation: in a heterogeneous-agent lifecycle model incorporating clinical trajectory, the agent's type is the pair (a, γ), comprising wealth and medical history, rather than (a, z). The income process z is standard; the mortality component is trajectory-dependent. The lifecycle HJB becomes a functional PDE on the product space $\mathbb{R}^+ \times \Gamma$:

$$rV(a,\gamma) = \max_c \{u(c) + V_a(a,\gamma) \cdot (z(\gamma) + ra - c) + \mathcal{G}_\gamma V(a,\gamma)\}$$

where $\mathcal{G}_\gamma$ is the generator of the clinical trajectory process, a functional differential operator that replaces the scalar mortality term $-\delta_k \cdot (V - b)$ of the standard model. The precise form of $\mathcal{G}_\gamma$ and its computational tractability are taken up in Section 8. For now, it suffices to note that $V(a, \cdot)$ restricted to trajectory types near $\partial\Omega$ inherits the curvature blow-up of the geometric identity of Section 3.

2.5 Temporal Encoding and Irregular Event Timing

Medical events arrive at irregular intervals. This creates an identification challenge that the signature formalism handles cleanly but that standard positional encodings do not. Appending spectral temporal features to event embeddings before computing signatures incorporates timing information into the path:

$$\psi(e_i, t_i) = \phi(e_i) \oplus \tau(t_i)$$

where the temporal lifting $\tau : \mathbb{R}^+ \to \mathbb{R}^{2K+1}$ is:

$$\tau(t) = (\omega_0 t,\, \sin(\nu_1 t + \varphi_1),\, \cos(\nu_1 t + \varphi_1),\, \ldots,\, \sin(\nu_K t + \varphi_K),\, \cos(\nu_K t + \varphi_K))$$

with learnable frequencies $\{(\nu_k, \varphi_k)\}_{k=1}^{K}$ (Kazemi et al. 2019). This is Time2Vec: the spectral parameterization of the rough path lift, whose learned frequencies span scales from circadian patterns to multi-year disease progression. The critical feature for the economic application is that the temporal geometry of the trajectory; the acceleration of events as the agent approaches the boundary is encoded by the frequencies ν_k and is therefore recoverable from the trained model.

Proposition (). Any continuous periodic function $f : \mathbb{R} \to \mathbb{R}$ can be approximated arbitrarily well by linear combinations of the Time2Vec basis functions. With learnable frequencies, the representation extends to quasi-periodic and multi-scale patterns. *Proof:* Stone-Weierstrass; the set $\{1, t, \sin(\omega t + \varphi) : \omega, \varphi \in \mathbb{R}\}$ generates a subalgebra of $C([0, T])$ that separates points and vanishes nowhere.

The practical consequence: event acceleration preceding mortality, a well-documented clinical phenomenon, is a pattern learnable in the spectral basis. The econometrician does not need to specify the acceleration structure in advance; it is discovered by gradient descent on the population of trajectories.

3. The Geometric Identity

3.1 Setup

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with C^2 boundary $\partial\Omega$, treated as the absorbing set. Agent state $x \in \Omega$ evolves under:

$$dX_t = b(X_t)\, dt + \sigma(X_t)\, dW_t, \qquad X_0 = x$$

with $b : \bar{\Omega} \to \mathbb{R}^d$ and $\sigma : \bar{\Omega} \to \mathbb{R}^{d \times d}$ Lipschitz, and $\sigma\sigma^\top$ uniformly elliptic on $\bar{\Omega}$. The hitting time is $\tau = \inf\{t \geq 0 : X_t \in \partial\Omega\}$. The value function is the expected residual life:

$$V(x) = \mathbb{E}^x[\tau]$$

By standard results for nondegenerate diffusions (Karatzas and Shreve 1991), V is the unique classical solution of:

$$\mathcal{L}V(x) = -1 \quad \text{on } \Omega, \qquad V(x) = 0 \quad \text{on } \partial\Omega \tag{HJB}$$

This is the correct lifecycle value function. The smooth actuarial hazard approximation replaces it with a function satisfying $\mathcal{L}V(x) \approx -\delta_k \cdot V(x)$: a fundamental change in the PDE that cannot be corrected by recalibration.

3.2 The Curvature Identity

Proposition (). For any $x_0 \in \partial\Omega$ and unit vector ν pointing into Ω at x_0:

$$V(x_0 + s\nu) \sim c_1 \cdot s^{1/2} \quad \text{as } s \to 0^+$$

Differentiating twice:

$$\left|\partial_\nu^2 V(x_0 + s\nu)\right| \sim c_2 \cdot s^{-3/2} \quad \text{as } s \to 0^+ \tag{Curvature Identity}$$

The constant c_2 depends on the diffusion coefficients $\sigma\sigma^\top$ and on the curvature of $\partial\Omega$ at x_0. The half-power exponent in V and the minus three-halves exponent in $\partial_\nu^2 V$ are exact for Brownian hitting times. For general nondegenerate diffusions, the exponent $-3/2$ bounds the curvature growth from below.

Corollary. Any smooth function $\tilde{V} : \bar{\Omega} \to \mathbb{R}$ satisfying $\tilde{V} \in C^2(\bar{\Omega})$ has $\|\partial_\nu^2 \tilde{V}\|_\infty < \infty$ on $\bar{\Omega}$. The approximation error $\|V - \tilde{V}\|_{C^2}$ is therefore unbounded near $\partial\Omega$ for any $\tilde{V}$ in the smooth class. The standard lifecycle model, which produces a smooth effective value function, is incapable of representing (Curvature Identity) regardless of parameterization.

The identity has a direct economic interpretation. An agent at state $x = x_0 + s\nu$, at distance s from the absorbing boundary, holds $V(x) \approx c_1 s^{1/2}$ units of expected residual life. The marginal value of moving away from the boundary (improving health) is $\partial_\nu V \approx \frac{c_1}{2} s^{-1/2}$: it diverges but at a moderate rate. The second-order effect. How that marginal value changes with further improvement diverges at rate $s^{-3/2}$. This is the steepening of the value function that makes behavior near the boundary qualitatively different from behavior away from it.

3.3 The Perturbation Asymmetry

Proposition (). For any direction ν pointing toward $\partial\Omega$ at $x \in \Omega$:

$$\frac{[V(x+\delta\nu) - V(x)] + [V(x-\delta\nu) - V(x)]}{\delta^2} \to \partial_\nu^2 V(x) \quad \text{as } \delta \to 0$$

and $|\partial_\nu^2 V(x)| \to \infty$ as $x \to \partial\Omega$.

Proof. Taylor expansion to second order: $V(x \pm \delta\nu) = V(x) \pm \delta\partial_\nu V(x) + \frac{1}{2}\delta^2\partial_\nu^2 V(x) + O(\delta^3)$. Summing, the linear terms cancel. The $O(\delta^3)$ remainder is controlled by the third derivative, which remains bounded away from $\partial\Omega$. At the boundary, the divergence follows from (Curvature Identity). □

The proposition has a direct clinical interpretation that motivates the economic application. At a given clinical state x with distance s from the mortality boundary, a deteriorating medical shock of magnitude δ reduces expected residual life by $\delta\partial_\nu V + \frac{1}{2}\delta^2\kappa$. An improving intervention of the same magnitude increases it by $\delta\partial_\nu V - \frac{1}{2}\delta^2\kappa$. The asymmetry is $\delta^2\kappa(x)$, governed entirely by the local curvature. An agent near $\partial\Omega$ has large κ; an adverse event costs significantly more expected life than an equivalent improvement gains.

Zhorin (2026b) identifies three clinical signatures of this asymmetry: hospital-acquired complications producing mortality increments larger than equivalent prevention produces decrements; polypharmacy deterioration exceeding the improvement from deprescribing; and ICU readmission as a stronger mortality predictor than ICU avoidance is a longevity predictor. Each pattern is documented in the clinical literature and attributed to domain-specific mechanisms. The boundary-curvature framing unifies them.

In the economic domain, the asymmetry has direct implications for welfare analysis. A policy that reduces adverse health shocks by δ for a population near the boundary increases aggregate welfare by less than the same policy that prevents adverse shocks, because the curvature amplifies the cost of deterioration relative to the benefit of improvement. This asymmetry is invisible in the standard lifecycle model, which assigns the same marginal welfare weight to improvements and deteriorations.

3.4 Cross-Sectional Aggregation

Consider a population indexed by $i \in \{1, \ldots, N\}$, with agent i at state $x_i \in \Omega$. An intervention assigns exposure $\delta_i \in \mathbb{R}$ to agent i, with sign indicating direction. The population-level change in aggregate $\bar{V}$ is:

$$\Delta\bar{V} = \frac{1}{N}\sum_i [V(x_i + \delta_i\nu_i) - V(x_i)] = \mathbb{E}_i[\delta_i \partial_{\nu_i} V(x_i)] + \frac{1}{2}\mathbb{E}_i[\delta_i^2 \kappa(x_i, \nu_i)] + O(\delta^3)$$

Assumption A1 (Mean-zero exposure). $\mathbb{E}_i[\delta_i] = 0$.

Assumption A2 (Gradient orthogonality). $\mathrm{Cov}(\delta_i, \partial_{\nu_i} V(x_i)) = 0$.

A1 is the randomization condition. A2 holds when interventions are independent of patient state. This holds when interventions are independent of patient state, the trial-design benchmark. Under A1 and A2, the linear term vanishes in expectation and the leading population-level effect is:

$$\Delta\bar{V} \approx \frac{1}{2}\left[\mathbb{E}[\delta_i^2] \cdot \mathbb{E}[\kappa_i] + \mathrm{Cov}(\delta_i^2, \kappa_i)\right] \qquad \text{(Aggregation)}$$

Proposition (). Under A1 and A2, the population-level asymmetric response to a randomized intervention is fully summarized to leading order by two cross-sectional moments: the average curvature $\mathbb{E}[\kappa_i]$ and the covariance $\mathrm{Cov}(\delta_i^2, \kappa_i)$. The covariance is the marginal signature of cross-sectional heterogeneity in boundary proximity.

The structural form of (Aggregation) mirrors Auclert's (2019) Theorem 3, which decomposes the aggregate consumption response to monetary policy into cross-sectional covariances of marginal propensities with balance-sheet exposures. The methodological move is identical: aggregate behavior of a heterogeneous population is summarized by a small number of cross-sectional moments, and any structural model is disciplined by matching them rather than specifying the full joint distribution.

The clinical interpretation of (Aggregation) under triage protocols: sicker patients (higher κ_i) receive more aggressive treatment (higher δ_i^2), is that the covariance $\mathrm{Cov}(\delta_i^2, \kappa_i)$ is positive. Trial-based estimates of average treatment effects target the linear functional $\mathbb{E}_i[\delta_i \partial_{\nu_i} V(x_i)]$, which vanishes under A1 and A2. They miss en-

tirely the second-order term. The number reported by the trial applies to the average patient, who is by construction not at high curvature, and systematically understates the cost of treatment failure while overstating the benefit of treatment success in the high-risk subpopulations where decisions matter most. This is the economic significance of the curvature-exposure covariance.

3.5 The Lifecycle Implication

Translating (Aggregation) into the lifecycle model: the standard smooth-hazard model produces a value function whose second derivative $\partial_\nu^2 \tilde{V}$ is bounded uniformly in x. The covariance $\text{Cov}(\delta_i^2, \kappa_i)$ is therefore bounded in the standard model by a function of the hazard parameters alone, with no mechanism for it to grow as agents approach the boundary.

The correct model produces $\kappa_i \sim c_2 \cdot \text{dist}(x_i, \partial\Omega)^{-3/2}$. If the population distribution of boundary-proximity has density $\rho(s)$ on $s \in (0, S]$, the expected curvature is:

$$\mathbb{E}[\kappa] = \int_0^S c_2 \cdot s^{-3/2} \rho(s)\, ds$$

For any density ρ bounded below on $(0, S]$, this integral diverges. In practice, finite data resolution and model expressivity regularize the integral; the result is large but finite, and the correct order of magnitude exceeds the smooth-hazard approximation by orders of magnitude in cohorts with significant fractions near the mortality boundary, exactly the elderly cohorts where the three puzzles manifestree puzzles manifest.

This is the structural argument for why the mortality misspecification is not a small correction. It is a zeroth-order error in the value function of agents near $\partial\Omega$, which is the population where lifecycle savings, bequest, and annuitization decisions are being made.

3.6 Chernoff Entropy, Tenuous Beliefs, and the Behavioral Asymmetry

The geometry of Section 3.2–3.5 has an information-theoretic reading that connects it, through explicit calculation, to the model-uncertainty framework of Hansen and Sargent (2017) and, through that connection, to the gain-loss asymmetry that is the empirical foundation of behavioral economics. The three frameworks are not competing explanations of the same phenomenon. They are three views of the same mathematical structure.

Chernoff entropy and the detectability of misspecification. The Chernoff entropy between two probability measures P and Q is defined as:

$$\chi(P, Q) = -\log \inf_{\lambda \in [0,1]} \mathbb{E}_P \left[\left(\frac{dQ}{dP} \right)^{\lambda} \right]$$

By the Chernoff-Stein lemma, $\chi(P, Q)$ governs the rate at which a statistician's error probability in discriminating P from Q decays to zero: with n i.i.d. observations, the optimal error probability scales as $e^{-n\chi(P,Q)}$. The Chernoff half-life, the sample size at which discrimination probability reaches 50%, is $\log 2/\chi(P, Q)$.

In the lifecycle context, let P_x denote the trajectory-dependent mortality distribution for an agent at clinical state x, and P_{pop} the population-table distribution. The lifecycle model uses P_{pop} for all agents regardless of their state x. The Chernoff entropy $\chi(P_x, P_{\text{pop}})$ measures how quickly clinical data would reveal this substitution to be incorrect.

Proposition (). As $x \to \partial\Omega$, $\chi(P_x, P_{\text{pop}}) \to \infty$. The Chernoff half-life vanishes: clinical data from a near-boundary agent refute the population-table model in arbitrarily small samples.

The proof follows from the boundary asymptotics of Section 3.2: near $\partial\Omega$, the hitting-time distribution P_x concentrates on short times with variance scaling as $\text{dist}(x, \partial\Omega)^{1/2}$, while P_{pop} is supported on the population-mean residual life. The two distributions diverge in total variation as $x \to \partial\Omega$, and Chernoff entropy is bounded below by a constant multiple of the squared total variation distance, which diverges. The Chernoff half-life is bounded above by $c \cdot \text{dist}(x, \partial\Omega)^{3/2}$, matching the curvature power law.

The consequence is stark. For a 72-year-old with a newly diagnosed glioblastoma, the Chernoff half-life between her trajectory-dependent mortality distribution and the population table is measured in weeks of claims data; any statistical test with access to her clinical record would rapidly reject the population-table model. The lifecycle model uses that population-table model for her consumption, saving, and annuity decisions nonetheless. This is not a modeling approximation that is close in the sense of being difficult to detect. It is a modeling error that is easy to detect and large in its consequences for the curvature of the value function.

The Hansen-Sargent connection. Hansen and Sargent (2017) (2017, "Prices of

Macroeconomic Uncertainties with Tenuous Beliefs") develop a framework in which investors face two layers of uncertainty: ambiguity about which parametric model governs the economy, and misspecification concerns about all of those parametric models. They use Chernoff entropy, in parallel with the calculation above, to calibrate the penalty parameter governing how far from the structured models the investor is willing to consider. Their central result is that the resulting uncertainty prices exhibit a directional asymmetry: the investor "especially fears high persistence in bad states and low persistence in good ones." Prices of uncertainty fluctuate because misspecification concerns are state-dependent in exactly this asymmetric way.

The machinery is the HJB equation for the robust decision problem. The worst-case drift distortion U^* is characterized by the first-order condition from the Bellman equation:

$$U^*(x) = -\frac{1}{\theta}\sigma(x)^\top \nabla V^*(x)$$

where V^* is the robust value function and θ is the entropy penalty parameter calibrated using Chernoff entropy. The worst-case distortion pushes the drift in the direction of $-\nabla V^*$: toward lower value. Near the absorbing boundary, where ∇V^* is large, the worst-case distortion is large in magnitude and points toward $\partial\Omega$. The robust agent behaves as if death is more imminent than the baseline model implies, exactly the behavior observed empiricallyy.

The connection to the present framework is direct. The lifecycle model's smooth hazard δ_k plays the role of the baseline model in Hansen-Sargent. The trajectory-dependent value function V plays the role of the robust value function. The Chernoff entropy $\chi(P_x, P_{\text{pop}})$ calibrates how large θ should be, i.e., how far the lifecycle model's mortality beliefs are from the correct ones. Near the boundary, where χ is large, θ is small, and the worst-case correction to the baseline is large. The smooth hazard's beliefs are not merely wrong; by Hansen-Sargent's standard they are tenuous, impossible to maintain in the face of trajectory data.

The behavioral asymmetry as geometric mechanism. The gain-loss asymmetry: losses loom larger than equivalent gains is the empirical foundation of behavioral economics. Kahneman and Tversky (1979) documented it as prospect theory's central finding, with a loss aversion coefficient $\lambda \approx 2.25$: losses of a given magnitude feel approximately 2.25 times as painful as equivalent gains feel pleasurable.

This asymmetry has been replicated across thousands of experiments and constitutes the strongest regularity in the behavioral economics literature (Thaler 1980; Samuelson and Zeckhauser 1988; Benartzi and Thaler 1995; Tversky and Kahneman 1992). It is cited as evidence that rational expected utility theory is behaviorally inadequate and requires the reference-point modifications of prospect theory.

Hansen and Sargent (2017) establish that the same qualitative asymmetry emerges from model uncertainty without any behavioral modification: an investor who fears misspecification naturally weights bad states more heavily than good states because the worst-case model concentrates probability on bad states. The geometry of Proposition 3.3 establishes a third mechanism generating the same qualitative asymmetry, making a specific quantitative prediction that distinguishes it from both the behavioral and model-uncertainty explanations.

The perturbation asymmetry for an agent at state x with curvature $\kappa(x)$ takes the form:

$$\frac{|V(x-\delta\nu)-V(x)|}{|V(x+\delta\nu)-V(x)|} \approx \frac{\partial_\nu V(x)+\frac{1}{2}\delta\kappa(x)}{\partial_\nu V(x)-\frac{1}{2}\delta\kappa(x)}$$

For agents far from the boundary, $\kappa(x) \approx 0$ and the ratio approaches one: no asymmetry. For agents near the boundary, $\kappa(x) \sim c_2 \cdot \text{dist}(x, \partial\Omega)^{-3/2}$ is large, and the ratio substantially exceeds one. The empirically observed $\lambda \approx 2.25$ implies a specific boundary-proximity level: solving for the distance at which the geometric ratio matches the prospect theory coefficient yields a prediction about the average boundary proximity of experimental subjects in loss aversion studies.

More directly: the geometric mechanism predicts that the loss aversion coefficient should increase with age and with adverse health events, reflecting increasing average boundary proximity. This is consistent with meta-analytic evidence that older adults exhibit stronger loss aversion (Mather et al. 2012; Kurnianingsih et al. 2011) and with experimental evidence that health salience increases reference-dependent weighting of losses (Ditto et al. 2003). The behavioral explanation; reference-point psychology has no prediction about the relationship between loss aversion and mortality proximity. The curvature mechanism does.

The unification. The three frameworks converge on a single phenomenon from different directions.

The geometric mechanism (Sections 3.2–3.5) establishes that the curvature of the

correct value function diverges near the absorbing boundary, generating an asymmetric perturbation response as a property of the mathematical object, not of preferences, beliefs, or psychology.

The information-theoretic mechanism (Proposition 3.5) establishes that the lifecycle model's population-table beliefs are detectable as wrong, at a rate governed by the same power law, by any statistician with access to trajectory data. The severity of the misspecification is not merely large; it is precisely measurable through the Chernoff entropy.

The Hansen-Sargent mechanism establishes that a decision maker who correctly applies model-uncertainty tools to the lifecycle problem would, through the HJB of robust control, produce behavior qualitatively identical to the boundary-curvature-driven behavior: reweighting of bad states, state-dependent uncertainty prices, and the directional asymmetry between fear of bad persistence and indifference to good persistence. Their framework provides the bridge between the geometric result and rational decision theory under uncertainty.

The behavioral science literature documents the gain-loss asymmetry as a ubiquitous feature of human choice and attributes it to reference-point psychology. The present framework establishes that the same pattern is a consequence of optimal behavior by rational agents near an absorbing boundary, without any psychological mechanism. The claim is not that behavioral factors do not exist. It is that the geometric mechanism is sufficient to generate the documented magnitudes, and that the geometric prediction: increasing asymmetry with boundary proximity is a testable discriminant between the geometric and behavioral explanations.

4. The Standard Model Cannot Be Fixed by Recalibration

4.1 What the Standard Model Actually Computes

In the Bewley-Aiyagari-Huggett family and its HANK descendants, the mortality term enters the HJB as a linear discount on the value function:

$$rV(a,z) = \max_c \{u(c) + V_a \cdot (z + ra - c) + \lambda[V(a,z') - V(a,z)] - \delta_k[V(a,z) - b(a)]\}$$

Treating this as a standard discounted optimal stopping problem, the effective dis-

count rate is $\rho = r + \delta_k$ and the value function is:

$$\tilde{V}(a, z) = \mathbb{E}\Big[\int_0^\infty e^{-\rho t} u(c_t^*)\, dt + e^{-\rho T_k} b(a_{T_k})\Big]$$

where $T_k \sim \text{Exp}(\delta_k)$ is the time of death drawn from a memoryless exponential distribution. The second derivative of $\tilde{V}$ with respect to any covariate x is determined entirely by δ_k and the utility curvature. For CRRA utility $u(c) = c^{1-\sigma}/(1-\sigma)$ and a simple income process, a direct calculation gives:

$$|\partial_x^2 \tilde{V}| \leq C(\sigma, r, \delta_k, \|b\|)$$

where the right side is a finite constant depending only on preference parameters, the interest rate, and the hazard. The curvature of the smooth-hazard value function is bounded uniformly over the state space. It cannot blow up.

Proposition (). For any choice of hazard $\delta_k \in C^1(\mathbb{R}^+)$ and any utility function $u \in C^2$ with bounded second derivative, the value function $\tilde{V}$ solving the standard lifecycle HJB satisfies $\sup_{a,z} |\partial_\nu^2 \tilde{V}| < \infty$ for any direction ν. In particular, $\tilde{V}$ cannot satisfy the boundary asymptotics (Curvature Identity) of Section 3.2.

The proof is immediate from the structure of the HJB: with δ_k finite and u smooth, the right-hand side of the HJB is C^2 in (a, z), so the solution inherits that regularity. The curvature at any interior point is controlled by δ_k, not by proximity to any boundary. No amount of recalibration changes this.

4.2 A Toy Model of the Misspecification Error

It is instructive to compute the misspecification error explicitly in a stripped-down case.

Two-state mortality toy model. An agent is alive (state H) or has received a catastrophic diagnosis (state D). From state H, the hazard of death is $\delta_H = 0.04$/year (roughly actuarial for age 70). From state D, the hazard is $\delta_D = 1.2$/year (roughly consistent with 12-month survival of 30% for late-stage cancer). Diagnosis arrives as a Poisson event at rate $\mu = 0.03$/year. The agent starts in state H.

The true value function in state H is the expected residual life:

$$V_H = \frac{1}{\delta_H + \mu} + \frac{\mu}{\delta_H + \mu} \cdot \frac{1}{\delta_D}$$

Plugging in: $V_H = 1/0.07 + (0.03/0.07) \times (1/1.2) \approx 14.3 + 0.36 = 14.6$ years. That is the correct expected lifespan.

The standard lifecycle model calibrated to the population table uses a single $\delta_k = \delta_H + \mu \cdot \delta_D/(\delta_D + \delta_H) \approx 0.04 + 0.03 \times (1.2/1.24) \approx 0.069$/year, producing $\tilde{V} = 1/\delta_k \approx 14.5$ years. The level is nearly identical. So far so good.

Now consider an agent who has just entered state D. The true value function is $V_D = 1/\delta_D = 0.83$ years. The standard lifecycle model, which has no state D, continues to use $\tilde{V} = 14.5$ years. The level error is a factor of 17.

More relevantly, the second derivative of the true V with respect to the health index that separates H from D: this is proportional to $V_H - V_D = 13.7$ years, a discrete jump. The second derivative of $\tilde{V}$ with respect to any covariate is bounded by $(\delta_k)^2/r^2 \approx 0.002$ year^{-1}. The true curvature at the $H \to D$ transition is effectively infinite (a jump); the smooth model's curvature is 0.002. The ratio is not 3 or 5. It is of order 10^3.

This is a toy model with only two states, chosen for its arithmetic simplicity. The real world has a continuum of trajectory states, but the qualitative conclusion scales up: the misspecification error in the curvature of the value function concentrates near the absorbing boundary at a magnitude determined by the ratio of the true hitting-time variance to the smooth model's implied variance, which can easily reach two to three orders of magnitude for near-boundary agents.

4.3 What Adding Parameters Cannot Fix

A natural response is to enrich the model. Add a health state h to the lifecycle HJB, with $\delta_k(h)$ depending on health. Allow h to evolve as a Markov chain with transitions calibrated to clinical data. This recovers something closer to the true hazard for agents in state H or state D.

The problem is not the number of health states. It is that any finite collection of health states produces a value function that is piecewise-smooth across those states, with bounded curvature within each state and finite jumps at transitions. The true value function on trajectory space does not have a finite state decomposition: a single ICD-10 code (one of roughly 70,000) can move an agent from a flat region

of the value function to a highly curved region, and there is no pre-specified finite state space that captures this because the relevant transitions cannot be listed in advance.

De Nardi (2004) uses five health states. De Nardi, French, and Jones (2010) use three. Adding health states improves the approximation in the same way that adding grid points improves a numerical computation: linearly, not exponentially. The approximation error at any given clinical event is $O(1/K)$ where K is the number of health states, and the true curvature of the value function near the boundary grows without bound, so no finite K suffices.

More precisely: let V_K denote the best approximation to the true value function V using K health states. The approximation error in the curvature direction is:

$$\sup_{x \in \Omega_\epsilon} |\partial_\nu^2 V_K - \partial_\nu^2 V| \geq c \cdot \epsilon^{-3/2} - K \cdot C_K$$

where $\Omega_\epsilon = \{x : \text{dist}(x, \partial\Omega) < \epsilon\}$ is the region within distance ϵ of the boundary. For any fixed K, taking ϵ small enough makes the left side large, so the approximation error in the curvature is unbounded for any $K < \infty$. This is the structural sense in which the problem is not a calibration problem.

4.4 Getting the Direction Right: Why Closer Death Generates Less Consumption

The naive lifecycle logic runs the wrong way. A higher mortality hazard δ_k raises the effective discount rate $\rho = r + \delta_k$, making the present more valuable, inducing more consumption. Yaari (1965) is right about this: in a pure lifecycle model, closer death means faster spending. The puzzle is precisely that people near death decumulate *more slowly* than calibrated lifecycle models predict. The boundary-curvature mechanism does not work by raising the effective discount rate further. It works through two channels that push the other direction.

Channel 1: bequest-survival reweighting. In a two-period model where the agent survives to period 2 with probability p and dies with probability $1-p$, the first-order condition is:

$$u'(c^*) = p \cdot u'(W - c^* + Y) + (1 - p) \cdot b'(W - c^*)$$

As $p \to 0$ near the boundary, the consumption-smoothing term vanishes and the

bequest term dominates: $u'(c^*) \approx b'(W - c^*)$. The agent consumes the amount that equates the marginal utility of consumption to the marginal utility of bequest, not the amount that smooths consumption over a long future. If the bequest function is concave (standard), this implies holding back more wealth as survival probability falls. It is not that wealth buys life. It is that when you probably will not be alive to consume, the consumption motive weakens and the bequest motive takes over.

Channel 2: precautionary amplification from curvature. The near-boundary agent also faces genuine uncertainty about *when* she will die and *what* medical costs will arrive before then. The standard precautionary savings index is $-V_{aaa}/V_{aa}$, proportional to the third derivative of the value function. Near $\partial\Omega$, the value function has large derivatives in the clinical direction, which translates into a large effective prudence in the wealth direction through the coupling of medical costs and clinical state. The cost of being caught without wealth (unable to cover care, leaving nothing) is asymmetrically larger than the benefit of having excess wealth (a somewhat larger bequest). The asymmetry of Proposition 3.3 directly amplifies the precautionary motive.

Back-of-envelope. Take a 72-year-old with $W = \$500{,}000$, income $Y = \$30{,}000$/year, CRRA $\sigma = 2$, and De Nardi (2004)'s bequest parameters ($\phi_b = 0.02$, $\kappa_b = \$100{,}000$).

Smooth model ($\delta_k = 0.05$, survival probability over the coming year $p \approx 0.95$):

The FOC is dominated by the consumption-smoothing term. Annual consumption from wealth $\approx \rho W = 0.08 \times 500{,}000 = \$40{,}000$. Decumulation rate: 8%.

GBM patient (two-year survival $p \approx 0.30$, treating this as the probability weight in the two-period model):

$$u'(c^*) = 0.30 \cdot u'(W - c^* + Y) + 0.70 \cdot \phi_b(W - c^* + \kappa_b)^{-\sigma}$$

With $\sigma = 2$: $(c^*)^{-2} = 0.30(570{,}000 - c^*)^{-2} + 0.70 \times 0.02 \times (600{,}000 - c^*)^{-2}$.

This solves numerically to $c^* \approx \$41{,}500$/year. Annual wealth decumulation: $(41{,}500 - 30{,}000)/500{,}000 \approx 2.3\%$.

The smooth model generates 8%; the data show 2-3%; the trajectory-dependent model generates 2.3% with identical preference parameters. The difference is entirely the shift in the survival probability weight from 0.95 to 0.30, which swings the first-order condition from being dominated by the consumption-smoothing term

to being dominated by the bequest term. No new parameters; the same $\phi_b = 0.02$ bequest motive that is too weak to matter when $p = 0.95$ becomes the binding constraint when $p = 0.30$.

The cross-sectional dispersion follows: agents at the same age and wealth but with $p \in [0.30, 0.95]$ depending on clinical state generate decumulation rates spanning 2-8%, which is precisely the range De Nardi, French, and Jones (2010) document and cannot explain without appealing to unidentified preference heterogeneity.

5. The Tropical Polynomial Representation Class

The previous section established what the smooth model cannot do. This section identifies what representation class can.

5.1 Why Smoothness is the Wrong Constraint

The lifecycle model's value function is not a general C^2 function on a compact domain. It is an expected hitting time, and expected hitting times near smooth absorbing boundaries are, in the language of approximation theory, not in the Sobolev space H^2: their second derivatives are not square-integrable near $\partial\Omega$. Requiring a smooth representation is like requiring a polynomial approximation to $\sqrt{x}$ on $[0, 1]$: the approximation converges in L^∞ norm but never captures the singularity at zero, and the approximation error in the second derivative diverges near zero regardless of polynomial degree.

The right constraint for the representation class is not smoothness. It is piecewise-linearity with subdivision density that can adapt to the boundary geometry. That class has a name: tropical polynomials.

5.2 Tropical Polynomials: The Minimal Introduction

The tropical semiring replaces the standard operations $(+, \times)$ with $(\max, +)$. A tropical polynomial in one variable x is:

$$p(x) = \max_{k \in \{0,\ldots,K\}} (c_k + d_k\, x)$$

This is simply the maximum of $K + 1$ affine functions. The resulting function is piecewise-linear, convex, with breakpoints wherever two of the affine functions are equal. A tropical rational function is a difference of two such maxima:

$$f(x) = \max_k(c_k^+ + d_k^+ x) - \max_j(c_j^- + d_j^- x)$$

which is piecewise-linear but not necessarily convex.

Zhang, Naitzat, and Lim (2018) prove that any feedforward network with ReLU activations computes a tropical rational function, with the coefficients $(c_k^\pm, d_k^\pm)$ determined by the network weights. The tropical polynomial representation is not an approximation; it is exact. Every ReLU network is, algebraically, a tropical rational function.

The practical consequence: the decision boundary $\{f = 0\}$ is a piecewise-linear set whose combinatorial structure is completely described by which affine pieces dominate in each region of input space. Near the mortality boundary, where the value function is steep, more pieces are needed; far from the boundary, fewer suffice. The trained network distributes pieces according to the training data, placing higher subdivision density where the label distribution varies rapidly. For a mortality classifier, that is near $\partial\Omega$.

5.3 Approximating the Boundary Asymptotics: A Toy Calculation

The boundary asymptotics say $V(x_0 + s\nu) \approx c_1 s^{1/2}$ as $s \to 0^+$. Can a tropical polynomial approximate this?

Consider the problem in one dimension: approximate $f(s) = s^{1/2}$ on $s \in [0, 1]$ with a tropical polynomial, i.e., a piecewise-linear function p_K with K pieces. The optimal placement of breakpoints for L^∞ approximation of a convex function concentrates them where the second derivative is large.

The second derivative of $s^{1/2}$ is $-\frac{1}{4}s^{-3/2}$, which diverges at $s = 0$. The optimal breakpoints are at $s_k = (k/K)^2$ for $k = 1, \ldots, K$. This gives breakpoint spacing $s_{k+1} - s_k = (2k+1)/K^2$, which is of order $1/K^2$ near $s = 0$ and $2/K$ near $s = 1$. The subdivision is K times finer near the boundary than at the midpoint.

The uniform approximation error is $\|f - p_K\|_\infty = O(K^{-1})$: K pieces give accuracy $1/K$.

Now the key comparison. To achieve the same accuracy with a smooth polynomial: A polynomial of degree n approximates $s^{1/2}$ on $[0, 1]$ with error $O(n^{-1/2})$ (the standard result for approximation of functions with a square-root singularity by algebraic polynomials; the slow rate is because the function fails to be analytic at

$s = 0$). So a degree-n polynomial gives the same accuracy as a tropical polynomial with $K = n^{1/2}$ pieces. For the same number of parameters, the piecewise-linear approximation is $n^{1/2}$ times more accurate near the boundary.

Back of envelope: to approximate $s^{1/2}$ to 1% accuracy at $s = 0.01$ (representing an agent with 4 days of expected residual life, normalized), a smooth polynomial needs degree $\approx 10,000$; a tropical polynomial needs $K \approx 100$ pieces. The tropical representation is 100 times more parameter-efficient at the boundary.

This is why trained ReLU networks can represent the boundary geometry of the mortality value function: the tropical polynomial class is the natural match for the square-root singularity, in the same sense that splines in the right basis are the natural match for C^2 functions. No smooth function class achieves this efficiency.

5.4 The Lifecycle Model with Tropical Mortality

In the corrected lifecycle model, the mortality component of the value function is replaced by a trained transformer's output $\hat{V}_\theta : \Gamma \to \mathbb{R}$, where Γ is the clinical trajectory space of Section 2. By the Zhang-Naitzat-Lim (2018) result, $\hat{V}_\theta$ is a tropical rational function in the embedding space. The key properties for the lifecycle application are:

1. **Subdivision near $\partial\Omega$.** The trained model concentrates polyhedral pieces near the mortality boundary, where clinical trajectories approaching death live. The local curvature $\kappa(x) = |\partial_\nu^2 \hat{V}_\theta|$ is large in these regions, matching the true $s^{-3/2}$ behavior to within the model's finite expressivity.
2. **Continuous gradient away from pieces.** Within each polyhedral piece, $\hat{V}_\theta$ is affine, so the gradient is constant. The lifecycle's Euler equation $u'(c^*) = V_a$ can be evaluated by looking up which piece the agent's state falls in.
3. **Computationally cheap curvature estimation.** The second derivative of $\hat{V}_\theta$ is zero within pieces and concentrated at piece boundaries. The practical estimator from Section 7 of Zhorin (2026b) uses the finite-difference approximation:

$$\hat{\kappa}(x, \nu; h) = \frac{\hat{V}_\theta(\Phi(x) + h\nu) - 2\hat{V}_\theta(\Phi(x)) + \hat{V}_\theta(\Phi(x) - h\nu)}{h^2}$$

One forward pass of the network per patient, per direction. For a population of $N = 100,000$ patients, computing the curvature-exposure covariance $\text{Cov}(\delta_i^2, \hat{\kappa}_i)$

costs $O(N)$ forward passes, about 10 seconds on standard hardware.

Proposition (). For any continuous value function $V \in C(\bar{\Omega})$ satisfying the boundary asymptotics $V(x_0 + s\nu) \sim c_1 s^{1/2}$, and for any $\epsilon > 0$, there exists a tropical polynomial V_θ realizable as a trained ReLU network such that $\|V_\theta - V\|_\infty < \epsilon$ on any compact subset of Ω separated from $\partial\Omega$, and such that the subdivision density of V_θ near $\partial\Omega$ matches the curvature of V to leading order.

Proof sketch. Existence of a piecewise-linear approximation with the claimed accuracy follows from standard approximation theory for piecewise-linear functions applied to the restricted domain. The subdivision density result follows from the analysis of Section 5.3: the optimal breakpoint placement for $s^{1/2}$ concentrates pieces at rate $s_k = (k/K)^2$, and the corresponding piecewise-linear function is realizable as a tropical polynomial with K terms. Any ReLU network with K neurons in the relevant layer realizes such a tropical polynomial, and its subdivision density near $\partial\Omega$ is $O(K^2)$ times that away from the boundary.

The proposition establishes that the smooth lifecycle model and the tropical-polynomial lifecycle model differ in a structurally significant way near $\partial\Omega$, independently of any empirical considerations.

5.5 The Curvature Moment as a Sufficient Statistic

Proposition 3.4 (Aggregation) identified the covariance $\text{Cov}(\delta_i^2, \kappa_i)$ as the sufficient statistic for the population-level asymmetric response. The tropical representation makes this computable.

Given a trained model with embedding $\Phi : \Gamma \to \mathbb{R}^d$ and output $\hat{V}_\theta : \mathbb{R}^d \to \mathbb{R}$:

For each patient i, compute $\hat{\kappa}_i = \hat{\kappa}(\Phi(\gamma_i), \nu; h)$ using the finite-difference formula above. Identify the intervention exposure δ_i (from the trial assignment or the observational analogue). Compute the sample covariance. The result is a scalar whose sign tells you whether interventions are concentrated on high-curvature patients (positive: asymmetry large) or spread uniformly (near zero: asymmetry small by the sufficient-statistic theorem).

This is the direct analog of Auclert's (2019) sufficient-statistic decomposition for the consumption response to monetary policy. His computation requires an estimate of the MPC distribution and balance-sheet exposure. The mortality analog requires an estimate of the curvature distribution and intervention exposure. Both are computable from available data; neither requires specifying the full joint distribution

of agent types.

Back-of-envelope for the covariance magnitude. In a representative US elderly cohort, roughly 15% of patients at any given time are within 12 months of death (based on Medicare spending concentration: the top 5% of spenders account for 25% of Medicare costs, and end-of-life spending is heavily concentrated). Taking $\kappa_i \approx c_2 s_i^{-3/2}$ with s_i remaining life in years, and the top 15% having $s_i < 1$ year:

$$\mathbb{E}[\kappa_i] \approx 0.15 \times c_2 \times \mathbb{E}[s_i^{-3/2} | s_i < 1] + 0.85 \times c_2 \times \mathbb{E}[s_i^{-3/2} | s_i > 1]$$

For s_i uniform on $[0.1, 1]$ for the near-boundary group: $\mathbb{E}[s^{-3/2}] = \int_{0.1}^{1} s^{-3/2} \cdot (1/0.9)\, ds = (1/0.9)[-2s^{-1/2}]_{0.1}^{1} = (1/0.9)(-2 + 6.32) = 4.8$.

For s_i uniform on $[1, 20]$ for the far-from-boundary group: $\mathbb{E}[s^{-3/2}] \approx 0.07$.

So $\mathbb{E}[\kappa_i] \approx c_2(0.15 \times 4.8 + 0.85 \times 0.07) = c_2 \times 0.78$.

The concentration: the near-boundary 15% of patients contribute $0.15 \times 4.8 = 0.72$ to the aggregate, while the far-from-boundary 85% contribute $0.85 \times 0.07 = 0.06$. The near-boundary population dominates the aggregate curvature by a ratio of 12:1, despite being only 15% of the population.

This is the quantitative content of Proposition 3.4's claim that the mortality setting differs from the macroeconomic setting: even though only 15% of patients are near the boundary at any moment, that group dominates the aggregate curvature measure in a way that the 10-30% of borrowing-constrained agents in the macroeconomic analog does not, because $s^{-3/2}$ diverges much faster than the MPC distribution's tail.

6. The Puzzle Landscape Resolved

The central claim is not that the lifecycle model uses the wrong parameter values for the hazard δ_k. It is that the function class to which δ_k belongs cannot represent the relevant mathematical object. This section makes that precise.

The geometric identity of Section 3 has a single implication that ramifies across the entire span of lifecycle economics: agents near the absorbing boundary have a value function whose curvature diverges, and smooth mortality cannot represent this. Every empirical anomaly in the lifecycle literature that involves agents near the mortality boundary, specifically every anomaly involving the elderly, the se-

riously ill, or those anticipating a terminal event, is potentially a manifestation of this misspecification. This section works through the main instances, organized by domain.

The structure of each subsection is the same: the empirical fact, the standard model's prediction and its failure, the curvature explanation, and a testable prediction distinguishing the trajectory-dependent model from the smooth alternative.

6.1 Wealth Decumulation

Empirical fact. At age 70+, median wealth declines by 2–3% annually, far slower than the 8-12% annual decumulation calibrated models predict (De Nardi, French, and Jones 2010; Lockwood 2018). More diagnostic than the average gap is the cross-sectional dispersion: individuals at the same age and wealth level exhibit wildly different decumulation paths. Some spend down aggressively; others hold wealth essentially constant until very close to death. This heterogeneity is too large to attribute to preference differences and too structured to be noise.

Standard model failure. The smooth-mortality model ties decumulation incentives to age: δ_k increases with age, raising the effective discount rate and speeding up consumption. Individuals at the same age face the same mortality discount, so the model predicts convergent decumulation paths across wealth levels, differing only by the wealth elasticity of the bequest motive. The observed divergence across individuals at the same age and wealth is generated only by appeal to preference heterogeneity, such as warm-glow intensities and risk aversion, with no disciplining evidence.

Curvature explanation. In the trajectory-dependent model, two agents at the same age and wealth occupy different positions in Γ: one may have received a glioblastoma diagnosis six months ago; the other may have just completed a clean annual physical. Their survival probabilities differ by an order of magnitude, and the first-order condition weights them accordingly.

For the healthy agent, the consumption-smoothing term dominates: $u'(c^*) \approx p \cdot u'(\text{future consumption})$, with p high. The agent smooths consumption in the standard way, and decumulation proceeds at the Yaari rate.

For the GBM patient, p is low and the bequest term takes over: $u'(c^*) \approx (1-p) \cdot b'(W - c^*)$. Consumption is pinned by the bequest-savings margin, not by the consumption-smoothing margin. Since the bequest function is calibrated to generate a mild motive at $p = 0.95$, it generates a binding constraint at $p = 0.30$, where

the agent optimally withholds most of her wealth.

The cross-sectional dispersion in decumulation rates is a readout of the distribution of survival probability $p(s)$ across clinical trajectories, not a preference heterogeneity artifact. The smooth model, which assigns the same δ_k to all agents of the same age, cannot distinguish the two agents and therefore predicts the same decumulation rate for both. The trajectory-dependent model generates the observed dispersion as an equilibrium property.

Testable prediction. Controlling for wealth, age, and income, decumulation rates are negatively predicted by clinical trajectory markers available from Medicare or insurance claims data. Specifically, the curvature score $\hat{\kappa}(x_i)$ should predict wealth change over 24 months in an HRS sample linked to Medicare claims, with a negative coefficient: higher curvature (closer to boundary) predicts slower decumulation.

6.2 Bequest Dispersion

Empirical fact. Bequests are larger and more dispersed than voluntary-bequest models generate, conditional on wealth at retirement (Hurd 1989; De Nardi 2004). The dispersion is particularly striking at the top of the wealth distribution: very large bequests are common even among individuals whose expressed preference for leaving wealth to heirs, measured in survey data, is modest. The distribution has a fat right tail inconsistent with any calibrated warm-glow model.

Standard model failure. In the warm-glow specification, the bequest motive is parameterized as $b(a) = \phi_b \cdot (a + \kappa_b)^{1-\sigma}/(1-\sigma)$, with ϕ_b governing intensity. This produces dispersed bequests only if ϕ_b is dispersed across individuals; this assumption is not independently identified. Calibrating ϕ_b to match bequest moments is circular. More fundamentally, the model cannot generate large bequests from low-ϕ_b individuals, which is a pattern the data clearly contain.

Curvature explanation. The trajectory-dependent model generates bequest dispersion without appealing to preference heterogeneity. An agent whose clinical trajectory accelerated toward $\partial\Omega$, who received a catastrophic diagnosis at 72 rather than at 80, arrives at the terminal event with higher residual wealth because their value function was steep throughout the late accumulation phase: each dollar of wealth was assigned high marginal value because it was perceived (correctly) as insurance against a near-term boundary event. The bequest is not planned in the warm-glow sense; it is the residual from optimal saving behavior by an agent with a high-curvature value function who died slightly less rapidly than the clinical tra-

jectory implied.

The fat right tail of the bequest distribution corresponds to the tail of agents who accumulated aggressively in response to a near-boundary clinical trajectory and then survived longer than expected, a natural consequence given that hitting-time variance is large for near-boundary agents.

Testable prediction. Bequest size, conditional on wealth at retirement, is predicted by clinical trajectory markers from the final decade of life. The prediction is sharper than a correlation: the trajectory-dependent model implies a specific functional form relating bequest size to hitting-time variance, which is recoverable from the trained clinical model without additional structural estimation.

6.3 Annuitization

Empirical fact. Actuarially-fair annuities are rejected by roughly 85% of retirees despite Yaari's (1965) prediction of near-universal purchase. The standard explanations: bequest motives, adverse selection, illiquidity, and behavioral salience each account for a portion of the gap but leave a large residual (Lockwood 2012). More diagnostic is the interaction: rejection rates are strongly predicted by prior health events and by subjective mortality assessments, even after controlling for wealth and income (Gan et al. 2015; Heimer et al. 2019).

Standard model failure. The smooth population-table hazard defines what "actuarially fair" means in the model: an annuity is fair if its actuarial value equals its premium at the population mortality rate. An agent who correctly perceives her own trajectory-dependent mortality as higher than the population table will not purchase an annuity priced at the population rate; the standard model, which takes δ_k from population tables, represents no such perception. The correlation between health status and annuity rejection cannot be explained within the framework.

Curvature explanation. An agent at clinical state x near $\partial\Omega$ faces a personal mortality hazard that exceeds the population table. An annuity priced at the population rate offers a stream of payments based on average survival probability. From the perspective of the near-boundary agent, the annuity is overpriced: she expects to receive fewer payments than the pricing assumes. Moreover, the curvature asymmetry operates here directly: adverse health shocks, which will terminate the annuity stream, cost her more in expected utility terms than the stream gains, making the annuity a structurally bad deal for exactly the agents with the highest curvature. The agents Yaari's model says should buy annuities most urgently are the ones for

whom population-table pricing is most wrong.

The observable implication is a crossing: agents with trajectory-dependent mortality below the population table (healthier than average) find annuities cheap; those above find them expensive. The correlation between health status and rejection is not a behavioral bias but a rational response to mispricing. The mispricing is not correctable by better actuarial tables at the population level; it requires individual trajectory assessment.

Testable prediction. Annuity take-up rates, across matched cohorts in HRS data linked to Medicare claims, are negatively predicted by the trajectory-dependent mortality estimate from the clinical model. The prediction is not that health status correlates with rejection; that pattern is already documented, but that the specific curvature score $\hat{\kappa}(x_i)$ outperforms any smooth mortality index in explaining the residual variation in take-up after controlling for standard health measures.

6.4 Social Security Claiming

Empirical fact. Roughly 60% of Americans claim Social Security benefits at or before full retirement age, forgoing the actuarial advantage of delayed claiming (Shoven and Slavov 2014). The standard population-table analysis shows that delayed claiming is actuarially favorable for most claimants; the observed early-claiming rate is therefore anomalous from the perspective of a rational agent with population-table mortality. The anomaly is concentrated in lower-income and lower-health populations, but persists controlling for liquidity constraints (Coile et al. 2002).

Standard model failure. The smooth-mortality lifecycle model predicts that agents with below-median wealth and liquidity constraints may optimally claim early, but that liquidity-unconstrained agents should delay. The data show early claiming across the wealth distribution in subpopulations with adverse health histories, which the smooth model attributes to behavioral time preference or misperception of the claiming rules.

Curvature explanation. An agent with a clinical trajectory near $\partial\Omega$ has a personal mortality distribution that places significant probability mass on death before ages at which delayed claiming would pay off. Claiming early is rational if the agent's expected remaining life is substantially shorter than the population table implies. The trajectory-dependent model generates heterogeneous optimal claiming ages that are directly linked to boundary proximity: high-curvature agents (near the boundary) optimally claim early; low-curvature agents (far from the boundary)

optimally delay. The population-level 60% early-claiming rate is a mixture distribution over clinical trajectories, not an anomaly.

A secondary channel operates through the curvature asymmetry: the value of a delayed and therefore larger payment stream is discounted by the probability of not surviving to receive it, and that probability is higher for near-boundary agents. The asymmetry means that for these agents, the cost of forgoing the early payment while waiting for the larger delayed payment exceeds the actuarial value of the increment — even absent any behavioral distortion.

Testable prediction. Optimal claiming age, identified from HRS data linked to Medicare claims, is negatively predicted by the trajectory-dependent curvature score, controlling for wealth, income, and standard health indices. The incremental prediction of $\hat{\kappa}(x_i)$ beyond standard self-reported health measures identifies the clinical trajectory information that self-reports fail to capture.

6.5 Retirement Timing

Empirical fact. Retirement timing is sharply discontinuous around health events: catastrophic diagnoses produce immediate retirement at a rate far exceeding the gradual optimal-stopping prediction of lifecycle models (Bound et al. 1999; Currie and Madrian 1999). Moreover, the absence of adverse health events predicts continued employment well past standard retirement ages, a pattern the smooth-mortality model cannot account for, since it predicts decelerating labor supply at all ages above the optimal stopping threshold.

Standard model failure. The smooth-mortality optimal stopping problem has a smooth value function and a smooth optimal retirement threshold as a function of wealth and age. It predicts gradual labor force exit driven by wealth accumulation, not by clinical events. The observed sharp discontinuities around diagnoses are attributed to disability rather than to rational updating: a classification that sidesteps rather than explains the behavior.

Curvature explanation. A catastrophic diagnosis shifts the agent's position in Γ discontinuously: a single ICD-10 code can move the agent from a trajectory with expected residual life of twenty years to one with expected residual life of eighteen months. The curvature of the value function jumps accordingly. At the new trajectory, the marginal utility of leisure relative to income is rebalanced: the time horizon is short, the curvature is high, and the remaining life has asymmetric structure. The optimal retirement decision updates discontinuously because the value

function updates discontinuously. This happens not because of a disability threshold but because the geometry of the problem changes.

This is precisely the feature the smooth model cannot represent: a single event at time t can change the optimal policy at all future times, not by a marginal recalibration of δ_k but by a wholesale change in the curvature regime of the value function. Retirement is not gradual because the boundary is not smooth.

Testable prediction. The probability of retirement within 12 months of a clinical event, controlling for income and wealth, is predicted by the trajectory-dependent boundary-proximity change induced by that event, specifically the change in $\hat{\kappa}$ following the diagnosis rather than the diagnosis category itself. Two agents receiving different diagnoses that produce the same change in boundary proximity should exhibit similar retirement responses.

6.6 Long-Term Care Insurance

Empirical fact. LTCI take-up in the United States is approximately 7%, despite the substantial economic risk: the probability of needing nursing home care for a year or more is roughly 30% for those reaching age 65, and annual nursing home costs exceed $90,000 (Brown and Finkelstein 2007, 2008). The standard explanations: Medicaid crowd-out, adverse selection, and loading factors account for perhaps half the gap; the remaining underinsurance is anomalous.

Standard model failure. The smooth lifecycle model with Medicaid as a public option for catastrophic care generates underinsurance consistent with part of the observed gap (Brown and Finkelstein 2008). But it cannot explain why underinsurance is concentrated in exactly the demographic groups, those with adverse health histories, where the smooth model most strongly predicts purchase. Sick agents are the ones who most need LTCI; they are also the ones who least buy it.

Curvature explanation. The curvature asymmetry generates a specific prediction here that runs counter to the naive intuition. An agent near the mortality boundary has a high-curvature value function. A long-term care event, from this agent's perspective, is a deterioration toward the absorbing boundary: it increases curvature further and reduces expected residual life. The curvature asymmetry (Proposition 3.3) implies that the adverse shock of entering long-term care costs the agent more in expected utility terms than the corresponding improvement of avoiding it gains. In the presence of this asymmetry, the expected utility gain from insuring against the bad state is smaller than the smooth model predicts, because the bad state,

when it arrives, produces a reduction in expected utility that the premium payment cannot fully offset. The agent is effectively insuring against a state in which the curvature of their value function will be so large that neither the insurance payment nor the absence of the LTCI cost will make a meaningful difference to their boundary proximity.

This is not a behavioral argument. It is a structural consequence of the asymmetry: near-boundary agents are in a curvature regime where the marginal utility of wealth is high but the marginal utility of insurance against further adverse events is lower than it appears in a smooth model, because the high curvature means the adverse event will dominate the value function regardless.

Testable prediction. LTCI take-up rates, controlling for wealth and Medicaid eligibility, are negatively predicted by clinical trajectory markers, specifically the rate of change of boundary proximity over the 36 months prior to the purchase decision. Agents on an accelerating trajectory toward the boundary exhibit lower LTCI take-up, consistent with the asymmetry mechanism rather than with standard adverse selection.

6.7 Portfolio Composition

Empirical fact. Elderly investors maintain equity exposure well above lifecycle theory's prescription (Ameriks and Zeldes 2004). The Cocco-Gomes-Maenhout (2005) calibration requires implausibly high risk aversion to match observed equity holdings at advanced ages. Relatedly, the elderly treat home equity as a last resort rather than decumulating it alongside financial assets, accumulate excess cash in low-yield accounts, and maintain life insurance well past the point where the standard model says it should be cancelled.

Standard model failure. The lifecycle portfolio model recommends declining equity as a share of wealth with age because the ratio of human capital to financial wealth falls. With smooth mortality, the optimal portfolio converges to a risk-free terminal allocation as the horizon shortens and mortality risk increases. The observed persistence of equity exposure implies either unusually high equity premium beliefs or unusually low risk aversion; neither is independently motivated.

Curvature explanation. The trajectory-dependent model generates a hedging demand for assets correlated with health shocks. An agent whose health is adversely correlated with aggregate equity returns; health and equity are positively correlated over the business cycle (Ruhm 2000), creating a joint risk of bad health out-

comes and bad portfolio returns. The curvature of the value function near the boundary amplifies this joint risk asymmetrically: the adverse state (bad health, bad market) is more costly than the good state (good health, good market) is beneficial.

The optimal portfolio response is to hold equity as a hedge against states in which the agent is far from the boundary: equity pays well when the agent is healthy and has a flat value function; the optimal response shifts toward liquid assets as a buffer against boundary-proximity events. This produces exactly the pattern observed: equity persists because it hedges the high-curvature states from a distance, while cash accumulates as an immediately available resource. The apparent irrationality is rational portfolio choice under a value function whose curvature is large near the boundary.

Home equity is treated as a last resort because it is illiquid: it cannot be rapidly converted into the medical expenditures that boundary-proximity events require. The model generates a hierarchy of asset decumulation: liquid assets first, financial assets second, home equity last as a direct consequence of the liquidity structure interacting with boundary-proximity curvature.

Testable prediction. Equity share, controlling for wealth and age, is positively predicted by the correlation between the agent's clinical trajectory and aggregate equity returns in the HRS-Medicare linked sample. Agents whose health deteriorations coincide with market downturns hold more equity, not less, consistent with the curvature-driven hedging demand.

6.8 Health Expenditure at End of Life

Empirical fact. Roughly 25% of lifetime Medicare expenditures are incurred in the final year of life (Lubitz and Riley 1993; Riley and Lubitz 2010). Family and individual willingness to authorize high-cost, low-expected-benefit interventions near end of life is dramatically above what a simple expected-value calculation would support. The standard rationalization: that families do not accurately assess prognosis is a behavioral story, not a structural one, and it does not explain the socioeconomic gradients in treatment intensity.

Standard model failure. A smooth-mortality utility-maximizing model with correct beliefs about treatment efficacy should not authorize high-cost interventions with negative expected health benefit. The observed treatment intensity at end of life is attributed to informational frictions, family dynamics, or physician incen-

tives; none of this is part of the standard lifecycle framework.

Curvature explanation. An agent at state x near $\partial\Omega$ faces a value function with large curvature. A treatment that offers a small probability of survival extension, which moves the agent slightly away from the boundary, produces a large increase in expected utility because the curvature is large near $\partial\Omega$: small improvements in boundary distance are worth a great deal in expected residual life. From the perspective of the correct value function, the treatment's expected utility value is significantly higher than its naive actuarial cost-effectiveness suggests.

This is not irrationality. It is the direct consequence of the curvature of the correct value function. The agent and family are implicitly computing $V(x + \delta\nu) - V(x) \approx \delta\partial_\nu V(x)$, where $\partial_\nu V$ is large near $\partial\Omega$. The smooth model, which has bounded curvature, produces a lower estimate of $\partial_\nu V$ near the boundary and therefore predicts lower willingness to pay for the same treatment. The discrepancy between observed and predicted treatment authorization is quantitatively governed by the ratio of the true curvature to the smooth-model curvature, a ratio that diverges as the agent approaches the boundary.

The socioeconomic gradient in treatment intensity follows from the wealth gradient in boundary proximity: wealthier agents are, on average, further from the mortality boundary (better access to healthcare, lower lifetime exposure to adverse conditions), so they face lower curvature and lower willingness to pay for marginal treatment. The gradient is not a preference story but a curvature story.

Testable prediction. Treatment authorization rates for low-efficacy, high-cost interventions near end of life are predicted by the curvature score $\hat{\kappa}(x_i)$ at the time of the treatment decision, controlling for wealth, income, and stated treatment preferences. The incremental explanatory power of the curvature score over standard prognostic indices identifies the boundary-geometry component of treatment decisions.

6.9 Pension Tax Choices: Roth Conversion and DB vs. DC

Empirical fact. Roth IRA conversion rates: converting pre-tax retirement savings to post-tax are far below the levels that optimization over smooth mortality distributions would predict for individuals in low current tax brackets who expect high future marginal rates (Reichling and Smetters 2015). Conversely, the preference for defined-benefit pensions over defined-contribution plans, conditional on identical actuarial values at population mortality, is substantially stronger in subpopu-

lations with adverse health histories than in healthy cohorts, reversing the standard adverse-selection prediction.

Standard model failure. Under smooth mortality, the Roth conversion decision reduces to a comparison of present and future marginal tax rates, adjusted for the probability of not surviving to realize the future benefit. The smooth model predicts conversion for young-to-middle-aged individuals in low brackets; the data show systematic under-conversion in this group relative to model predictions. The DB vs. DC preference should, under adverse selection logic, lead sicker individuals to prefer DC (lump sum, realizable immediately) over DB (annuity stream, requiring survival). The observed preference for DB in adverse-health cohorts is the opposite.

Curvature explanation. The Roth conversion decision is a bet on surviving long enough to realize the post-tax advantage. Under the trajectory-dependent model, an agent near the boundary assigns a lower probability to that survival and a higher uncertainty around it; the hitting-time distribution has higher variance, not because of a different scalar hazard, but because for agents in high-curvature regimes. The option value of delaying conversion is larger under trajectory-dependent mortality than the smooth model suggests, because the high-curvature regime means that adverse health shocks, which would eliminate the value of conversion, are more likely and more costly than the smooth model implies. The under-conversion rate is rational option preservation, not inertia.

The DB preference in adverse-health cohorts requires a different mechanism. A defined-benefit pension provides guaranteed income independent of financial market conditions. For an agent near $\partial\Omega$, the value function has high curvature: wealth is worth more because it purchases survival resources. A DB pension is, in effect, a long position on the agent's own survival: it pays off in every period of survival, and the payments are most valuable when the agent is near the boundary and needs resources most. The curvature amplifies the value of the guaranteed stream relative to a lump sum that must be managed under the stress of boundary proximity.

Testable prediction. Roth conversion rates, conditional on tax bracket and wealth, are negatively predicted by the curvature score at the time of the conversion decision. DB preference over DC, conditional on actuarial equivalence at population mortality, is positively predicted by the curvature score. Both predictions are testable using retirement account data linked to Medicare claims.

6.10 Estate Planning and Gift Timing

Empirical fact. Timing of inter-vivos gifts is highly concentrated in the period immediately following a serious diagnosis, far above the rates predicted by optimal estate-planning models calibrated to smooth mortality (Bernheim, Lemke, and Scholz 2004). The acceleration of gifting after a diagnosis is not fully explained by estate tax avoidance; the pattern persists below the exemption threshold and at rates far exceeding the tax benefit.

Standard model failure. The smooth lifecycle estate-planning model treats gift timing as a function of the gift tax schedule and the smooth mortality hazard. The optimal gift-timing problem has a smooth solution: gifts are spread across time to exploit the annual exclusion, with no particular concentration around health events. The empirical concentration requires appeal to behavioral urgency or informational effects that are outside the model.

Curvature explanation. A catastrophic diagnosis moves the agent near $\partial\Omega$ discontinuously. The curvature of the value function increases sharply, and the marginal value of wealth to the agent falls relative to the marginal value of wealth to heirs who are far from their own boundaries. The optimal transfer policy is to accelerate gifts immediately following any event that significantly increases boundary proximity. The driver is not tax optimization but the agent's value function has changed its geometry. The wealth that was optimal to hold before the diagnosis is no longer optimal to hold after it: the curvature asymmetry means that the expected utility cost of holding wealth through the remaining trajectory is lower than the expected utility benefit to heirs of receiving it now.

The gift acceleration following diagnosis is a direct readout of the discontinuous change in value function geometry induced by the clinical event, the computational realization of the boundary-determines-interior principle in estate planning.

Testable prediction. The magnitude of gift acceleration following a serious diagnosis, controlling for estate size and tax bracket, is predicted by the change in curvature score induced by the diagnosis event. Diagnoses that produce larger increases in $\hat{\kappa}$ should produce larger gift accelerations.

6.11 Precautionary Saving by the High-Wealth Elderly

Empirical fact. High-wealth elderly households maintain savings rates that are positive or near-zero well into their eighties, far above the buffer-stock model's prediction for agents who have long since exceeded the target wealth level (Dynan,

Skinner, and Zeldes 2004; Carroll 1997). This over-accumulation is concentrated at the top of the wealth distribution and is correlated with adverse health events in a way that the standard precautionary motive, saving against income risk, cannot explain: high-wealth agents face negligible income risk.

Standard model failure. Buffer-stock models of precautionary saving predict convergence to a target wealth level, above which saving is driven only by bequest motives or implausibly high risk aversion. The positive saving of wealthy elderly individuals is explained by calibrating large bequest motives or high risk aversion; neither has independent empirical identification.

Curvature explanation. The trajectory-dependent model generates a third motive for wealth accumulation that is distinct from consumption smoothing and bequest motives: boundary-proximity insurance. Wealth near the mortality boundary purchases resources: medical interventions, care coordination, specialist treatment that extend the agent's distance from the boundary. The curvature of the value function near $\partial\Omega$ assigns high marginal utility to wealth in that regime, generating a saving motive that increases, rather than decreases, as the agent approaches the boundary.

This motive has no analog in the smooth model: the smooth model's effective value function has bounded curvature, so the marginal utility of additional wealth is always bounded and eventually falls below the discount rate, stopping saving. In the correct model, the diverging curvature means the marginal utility of additional wealth near the boundary can exceed the discount rate even at very high wealth levels, generating the observed positive saving rates.

Testable prediction. Saving rates among high-wealth elderly, controlling for income and standard health measures, are positively predicted by the curvature score. The relationship should be nonlinear: saving rates should increase with curvature score in the upper tail of the curvature distribution, consistent with the diverging curvature near $\partial\Omega$.

7. Empirical Application

7.1 The Estimand and the Data

The sufficient statistic derived in Section 3 is the covariance $\text{Cov}(\delta_i^2, \kappa_i)$ across patients, where δ_i is intervention exposure and $\kappa_i = |\partial_\nu^2 \hat{V}_\theta(\Phi(\gamma_i))|$ is the curvature

score from the trained model at patient i's embedding. Estimating this requires three inputs: a trained clinical transformer, a linked dataset of clinical trajectories and intervention assignments, and a choice of the intervention class of interest. None of these is exotic. All three are available in Medicare-linked claims data combined with any standard randomized or quasi-randomized intervention.

The estimation procedure of Zhorin (2026b, Section 7) proceeds in three steps that are worth making concrete here. For each patient i, the curvature score is the finite-difference second derivative of the trained model's output along the direction ν_i corresponding to the intervention:

$$\hat{\kappa}_i = \frac{\hat{V}_\theta(\Phi(\gamma_i) + h\nu_i) - 2\hat{V}_\theta(\Phi(\gamma_i)) + \hat{V}_\theta(\Phi(\gamma_i) - h\nu_i)}{h^2}$$

This is one forward pass of the network per patient, cheap at scale. The bandwidth h is chosen by cross-validation on a held-out cohort. The squared exposure δ_i^2 is the intervention assignment (or its observational analog). The sample covariance is then a single inner product: $O(N)$ operations for a population of N patients.

The output is interpretable by Proposition 3.4: a positive value means high-curvature patients are being treated more intensively, and the population-level response to symmetric intervention shocks will be asymmetric in the worsening direction. A value near zero means the asymmetry is small at the population level even if individual patients near the boundary exhibit large local curvatures.

7.2 Why Medicine Provides the Clean Test

The framework applies wherever there are discrete timestamped events with a terminal boundary. Medicine is the right starting domain because four properties hold simultaneously that are unusual in social-science settings.

Boundaries are observed. Mortality is binary and verifiable in administrative data. The outcome variable is unambiguous. Default in credit and failure in firms are also verifiable, but their timing is often disputed and strategically managed. Mortality is not.

Datasets are large. Medicare Part A and Part B claims cover roughly 60 million beneficiaries annually with complete encounter-level records. The HRS-Medicare linkage provides wealth, consumption, income, and bequest data for the same individuals. The clinical transformer can be trained on tens of millions of trajectories;

the lifecycle model can be estimated on the same population.

The comparison class exists. De Nardi (2004), De Nardi, French, and Jones (2010), Lockwood (2012, 2018), and related papers provide carefully estimated smooth-mortality lifecycle models against which the trajectory-dependent model can be compared in matched samples. The question is not whether the trajectory-dependent model fits better in-sample (it will, with more parameters) but whether it resolves the specific cross-sectional patterns that the smooth model cannot generate.

The boundary is universal. As argued in Section 3.5 and Proposition 3.5, the Chernoff half-life between the trajectory-dependent and population-table distributions is short for near-boundary patients. The misspecification is statistically detectable, and therefore identifiable with reasonable sample sizes.

7.3 The Financial Analog: Default as Absorbing Boundary

The mathematical structure of Section 2 applies to financial default without modification. A borrower's account trajectory $\gamma_F = ((e_1, t_1), \ldots, (e_n, t_n))$ is a sequence of financial events: payments, purchases, delinquency flags, credit limit changes. The terminal boundary $\partial\Omega_F$ is default. The trajectory space Γ_F is the Kleene closure of transaction types and timestamps.

Zhorin (2026c) proves three necessity theorems establishing that optimal default prediction from financial event sequences requires exactly the same architectural components as optimal mortality prediction: discrete event embeddings, spectral temporal encoding, and learned causal attention. The proofs are isomorphic to those in Section 4 of this paper (and to Section 3 of the topology paper), because the mathematical structure is identical. Two borrowers with identical summary statistics but different transaction orderings have different default probabilities; two borrowers with the same transaction types but different timing geometry have different default probabilities; context-dependent interpretation of transactions is required. Standard credit models that aggregate transactions into tabular features, or that use ordinal positional encoding for the transaction sequence, are provably suboptimal.

The boundary-curvature asymmetry has a direct credit interpretation. For a borrower close to default (s small), the curvature of the value function (expected residual "financial life") diverges at rate $s^{-3/2}$. A deteriorating credit event (missed payment, over-limit charge) costs more in expected financial life than an equivalent improving event (large payment, balance reduction) gains. This asymmetry is the

geometric mechanism underlying the empirical finding that negative credit events are more predictive of default than positive events are predictive of recovery — a pattern treated as a behavioral or institutional anomaly in the credit literature but following directly from the geometry.

The practical consequence for credit modeling: a transformer trained on transaction sequences with a default-prediction objective, using spectral temporal encoding, will produce embeddings whose curvature near the default boundary encodes the same $s^{-3/2}$ structure as the clinical model. The curvature-exposure covariance is estimable from the trained credit model using the same three-step procedure as the clinical analog. The number tells underwriters whether their intervention intensities (collection calls, credit limit reductions, payment plans) are concentrated on high-curvature accounts, and whether the population-level response to those interventions will be asymmetric.

7.4 Cross-Domain Predictions

Because the mathematical structure is substrate-independent (Zhorin 2026d, Section 7), the framework generates predictions across domains that would otherwise seem unrelated.

Prediction 7.1 (Mortality-finance correlation). A patient near the mortality boundary also has elevated probability of financial distress. This is because proximity to the absorbing boundary in the clinical state space is correlated with proximity to the absorbing boundary in the financial state space: serious illness generates medical expenditures that deplete savings, creates income loss from disability, and reduces capacity to manage complex financial instruments. The trajectory-dependent model predicts a specific functional form for the clinical-financial correlation: the correlation between $\hat{\kappa}_{\text{clinical}}(i)$ and $\hat{\kappa}_{\text{financial}}(i)$ should be positive and should increase as either score increases, because the curvatures are amplified by proximity and the two proximities are positively correlated. This prediction is testable using linked health and credit bureau records.

Prediction 7.2 (Event acceleration as leading indicator). The temporal geometry of event sequences encodes boundary proximity: event acceleration (increasing event rate over time) is a reliable precursor of boundary events in both clinical and financial contexts. The spectral temporal encoding of Section 2.5 captures this pattern in the learned frequencies ν_k. The prediction is that the learned frequencies in a clinical mortality model and in a credit default model, trained on the respective

trajectory populations, will share characteristic timescales: those corresponding to the acceleration patterns that precede boundary events in each domain.

8. The Corrected Lifecycle Model

8.1 The Corrected Lifecycle HJB

The standard continuous-time lifecycle HJB with exogenous mortality, following Huggett (1993), Aiyagari (1994), and the dynastic formulation of Stantcheva (2015), is:

$$(\rho + \delta_k)\, V(a,z) = \max_{c \geq 0} \Big\{ u(c) + V_a(a,z)\,[y(z) + ra - c] + \mathcal{L}_z\, V(a,z) + \delta_k\, b(a) \Big\} \quad (1)$$

where ρ is the subjective discount rate, δ_k is the mortality hazard, $y(z)$ is labor income, r is the return on assets, $b(a)$ is the bequest value function, and $\mathcal{L}_z$ is the generator of the income process. For a two-state Markov chain: $\mathcal{L}_z V = \lambda[V(a,z') - V(a,z)]$; for a diffusion: $\mathcal{L}_z V = \mu_z V_z + \frac{1}{2}\sigma_z^2 V_{zz}$.

The left side collects pure time preference and mortality discounting. The right side adds the bequest flow $\delta_k\, b(a)$: in every instant the agent dies (probability $\delta_k\, dt$), she leaves wealth a and the household receives bequest utility $b(a)$. This is the form in Stantcheva (2015, Chapter 2) and Kaplan, Moll, and Violante (2018). The alternative $\rho V = \max_c\{\ldots\} - \delta_k[V - b(a)]$ is algebraically identical; the form above is standard.

What changes. Replace the scalar δ_k with a trajectory-dependent function $\delta(s)$ where $s = \mathrm{dist}(\Phi(\gamma), \partial\Omega)$ is the boundary proximity score from the trained clinical model. The state space expands from (a,z) to (a,z,s), and the clinical trajectory contributes its own generator $\mathcal{L}_s$. The corrected HJB is:

$$\Big(\rho + \delta(s)\Big)\, V(a,z,s) = \max_{c \geq 0} \Big\{ u(c) + V_a\,[y(z) + ra - c] + \mathcal{L}_z\, V + \mathcal{L}_s\, V + \delta(s)\, b(a) \Big\} \quad (2)$$

where $\mathcal{L}_s V = \mu_s(s)\, V_s + \frac{1}{2}\sigma_s^2(s)\, V_{ss}$.

Three things distinguish (2) from (1).

First, $\delta(s)$ is not parametric. Near $\partial\Omega$, the boundary asymptotics of Section 3 imply $\delta(s) \sim c_0 \cdot s^{-1/2}$, diverging as the agent approaches death. The standard δ_k is bounded and cannot represent this.

Second, V_{ss} enters through $\mathcal{L}_s V$ with weight $\frac{1}{2}\sigma_s^2(s)$. This is the precautionary savings channel as a mechanical consequence of the HJB, not a behavioral overlay: the variance of the clinical trajectory process directly enters the value function's curvature. Agents in high-volatility clinical regimes (large σ_s^2) face a larger precautionary motive at the same level of s.

Third, the bequest flow is $\delta(s)\, b(a)$, not $\delta_k\, b(a)$. Near $\partial\Omega$, $\delta(s)$ is large, so the bequest flow dominates the right-hand side. The weight on $b'(a)$ in the first-order condition $u'(c^*) = V_a$ is time-varying and state-dependent, proportional to $\delta(s)$, which grows without bound as the agent approaches death. The slow decumulation of Section 4.4 follows directly without any change to preference parameters.

Reduction to three dimensions. The full clinical embedding $\phi \in \mathbb{R}^d$ enters the HJB only through the scalar s and the local parameters $(\mu_s(s), \sigma_s^2(s))$, estimated from the trained model at negligible cost. The state space reduces from (a, z, ϕ) back to (a, z, s). The corrected HJB is computationally equivalent to a two-state health model, and simpler than the continuous-health models in De Nardi, French, and Jones (2010), because s has a known functional form near $\partial\Omega$ from the boundary asymptotics.

8.2 Addressing the De Nardi Objection

De Nardi (2004, footnote 8) explicitly considered whether differential mortality across types would materially change the lifecycle model's predictions and concluded it would not, citing the macroeconomic-analog scaling argument: the cross-sectional moment from differential mortality contributes a small fraction of aggregate variance, analogous to the Krusell-Smith (1998) finding for wealth distribution moments.

The structural response is Section 8 of Zhorin (2026b): the macroeconomic scaling argument fails for mortality because the boundary is universal. But it is worth quantifying the failure here with a simple calculation that directly engages De Nardi's framework.

In the Krusell-Smith (1998) calculation, the cross-sectional contribution is small because the borrowing constraint binds for at most 20-30% of agents. The representative-agent tracking error is a few percent. The analogous calculation for mortality: if

15% of the elderly population at any cross-section is within 12 months of death (consistent with Medicare spending concentration), and those agents have survival probabilities $p \approx 0.30$ against the smooth model's $p \approx 0.95$, then the first-order condition weight on the bequest term shifts from 0.05 to 0.70 for that 15%. By the FOC analysis of Section 4.4, their consumption drops from 8% decumulation to 2% decumulation. The aggregate shift is $0.15 \times (8\% - 2\%) = 0.9$ percentage points of the annual decumulation rate.

For a population with aggregate wealth $\bar{W}$, this 0.9 percentage-point shift in the annual decumulation rate is $0.009\bar{W}$ per year. With median elderly wealth of $\approx \$250,000$ (HRS data), this is $\approx \$2,250$/year per household in the constrained subpopulation. Aggregated over a 15% share of ≈ 50 million elderly households, the annual aggregate savings that the smooth model misattributes is of order $\$17$ billion per year. This is not a small correction.

The De Nardi dismissal was correct given the data and methods available in 2004. It is incorrect given the trajectory data that are now available in Medicare claims and the trained clinical models that can compute $\hat{s}_i$ and $\hat{\kappa}_i$ for every beneficiary.

8.3 Equilibrium Properties

What changes in the aggregate lifecycle model when the mortality input is corrected?

Aggregate savings. Near-boundary agents save more than the smooth model predicts (Section 4.4). This raises aggregate savings rates among the elderly. The effect is concentrated in the top quartile of wealth, where precautionary saving motives interact with the bequest term at high wealth levels. The corrected model predicts higher aggregate wealth holdings among the elderly than the smooth model, consistent with the over-accumulation puzzle.

Bequest distribution. The corrected model generates a bequest distribution with a fatter right tail. High-wealth agents with adverse clinical trajectories save aggressively through their late-accumulation phase and die with large unexpected bequests. The right tail of the bequest distribution is populated by agents who were precautionarily saving against a near-boundary event and survived longer than their trajectory implied. The fat tail is not a parametric feature; it emerges from the hitting-time variance structure.

Annuity demand. Agents with trajectory-dependent survival probability below the population table find actuarially-fair-at-population-rates annuities overpriced.

This generates the observed negative correlation between health status and annuity take-up. The aggregate annuity demand in the corrected model is lower than in the smooth model for the same preference parameters, because a substantial fraction of potential annuity buyers face personalized mortality rates that make population-priced annuities actuarially unfavorable.

The insurance market equilibrium. If insurers have access to clinical trajectory data (through underwriting), they can price closer to the individual trajectory-dependent rate. This eliminates the adverse selection margin that partially explains low annuity take-up, potentially generating higher equilibrium annuitization. The policy implication runs counter to the standard argument: permitting use of clinical data in annuity pricing could increase market depth by reducing the adverse selection wedge, rather than reducing it by excluding high-risk purchasers.

8.4 Stantcheva's Bequest Tax Formula

Stantcheva (2015) derives the optimal bequest tax formula in a dynastic model where parents invest in children's human capital and financial transfers. The formula expresses the optimal tax rate as a function of estimable sufficient statistics: the elasticity of bequest with respect to the net-of-tax rate, the redistributional weight on bequest recipients, and the fiscal externality of bequests on the government budget.

The correction to this formula from trajectory-dependent mortality is a new term in the sufficient-statistic decomposition. Because bequest behavior is trajectory-dependent (Section 6.2: near-boundary agents leave larger bequests as residuals from precautionary saving), the elasticity of bequest with respect to the net-of-tax rate is not constant across the population. It is negatively correlated with boundary proximity: near-boundary agents are leaving bequests as precautionary residuals, not as planned gifts, so their bequest elasticity is lower. The population-average elasticity used in Stantcheva's formula should be replaced by:

$$\bar{\epsilon} = \mathbb{E}[\epsilon_i] + \text{Cov}(\epsilon_i, \omega_i)$$

where ϵ_i is individual i's bequest elasticity and ω_i is a weight proportional to κ_i (the curvature score). Near-boundary agents have high κ_i and low ϵ_i, so the covariance is negative. The corrected sufficient statistic $\bar{\epsilon}$ is lower than the population-average elasticity alone. The optimal bequest tax is correspondingly lower than Stantcheva's formula delivers, because the behavioral elasticity is higher (the planned-bequest

margin is more responsive to taxes) among agents who are not near-boundary and lower among agents who are, and taxing the former is more distortionary per dollar of revenue.

The correction is computable from the trained clinical model using the same three-step procedure as the curvature-exposure covariance. It does not require structural re-estimation of the dynastic model; it requires only the curvature scores from the clinical model and a regression of bequest amounts on the net-of-tax rate interacted with curvature score.

9. Discussion and Extensions

9.1 What the Paper Does and Does Not Establish

The structural claims are three. First, no function in the smooth hazard class can represent the curvature of the expected hitting-time function near the absorbing boundary. This is a mathematical fact, independent of any data. Second, the population-level response to symmetric health interventions is asymmetric by an amount governed by a single cross-sectional moment, and that moment is large in clinical settings because the boundary is universal. This follows from the geometry of Section 3. Third, the representation class that can encode the boundary geometry exists and is computationally tractable. This follows from the tropical polynomial characterization of ReLU networks.

The paper does not establish that the corrected lifecycle model fits the data better than the smooth model in any particular sample. That is an empirical question, not answered here, though Section 7 identifies the experiments that would answer it. The paper does not establish that the curvature-exposure covariance is quantitatively large in any specific clinical cohort. That is also an empirical question. What the paper establishes is that the covariance exists, its magnitude is structurally governed by the boundary geometry in a way that cannot be dismissed on macroeconomic-analogy scaling grounds, and the standard biostatistical pipeline cannot estimate it.

The paper also does not claim that the behavioral explanations for the lifecycle puzzles are wrong. Prospect theory is well-documented. The contribution of Section 3.6 is not that loss aversion does not exist but that it can arise as a consequence of the boundary-curvature geometry without any behavioral modification, and the two explanations make different quantitative predictions about the relationship be-

tween the loss aversion coefficient and clinical boundary proximity.

9.2 Extensions

Dynastic models. The Barro-Becker dynastic framework with trajectory-dependent mortality across generations generates new cross-generational insurance motives. If parental mortality is trajectory-dependent and correlated within families (genetic and environmental channels), then children's lifecycle decisions are affected by the parental trajectory through both the bequest channel (covered here) and the mortality-forecast channel (if you know your parent died of GBM at 65, your own probability of that trajectory is elevated). The extended model has a Markov structure on the dynasty-level clinical trajectory space; the value function on that space inherits the boundary-curvature geometry.

Optimal health investment. The lifecycle model with trajectory-dependent mortality generates a demand for preventive health investment as a way of moving away from $\partial\Omega$. The marginal value of health investment is $\partial V/\partial s \cdot \partial s/\partial e$ where e is health expenditure and $\partial s/\partial e$ is the effectiveness of spending in extending survival. Near the boundary, $\partial V/\partial s \sim s^{-1/2}$ is large; the marginal return to health investment diverges. This generates a theoretical prediction about the age profile of preventive health expenditure that is testable against Medicare claims data.

The incomplete-markets channel. In a HANK model with trajectory-dependent mortality, the curvature-exposure covariance enters the aggregate Euler equation through the redistribution channel of Auclert (2019). The aggregate consumption response to a monetary policy shock has an additional term proportional to the covariance between the curvature score and the net nominal position of each household. This term can be large if high-curvature households (near-boundary patients) hold systematically different balance sheets than low-curvature households. The empirical evidence on this is consistent: elderly with serious illness hold more cash and less equity (Section 6.7), which is a specific balance-sheet tilt that interacts with the monetary transmission channel in a predictable direction.

Identification of the mortality component. The corrected lifecycle model requires estimating the clinical embedding $\Phi(\gamma_i)$ for each household in the lifecycle sample. This requires linking household panel data (HRS, PSID) to Medicare or insurance claims at the individual level. Such linkages are being developed and are available in several research data centers. The linking itself raises no new identification issues; the clinical model is estimated on the claims data and the embedding is

computed as a score that can be merged to the household panel.

9.3 On the Use of Clinical Data in Economic Models

The most immediate practical obstacle to implementation is data access rather than methodology. Estimating the curvature-exposure covariance requires individual-level clinical trajectories linked to economic outcomes. Such data exist in Medicare-linked surveys (HRS-Medicare, NHATS) and in insurer claims linked to retirement accounts, but access requires data use agreements and IRB approval. The computation itself is cheap; the governance is the binding constraint.

There is a deeper question about whether economic models should incorporate clinical trajectory data at all. If they do, they potentially change the actuarial pricing of mortality-contingent contracts, which has distributional implications for who can access annuities, life insurance, and long-term care insurance. The insurance market equilibrium of Section 8.3 addressed this specifically: conditional on using clinical data for pricing, adverse selection decreases and market depth increases. Unconditional on pricing, the corrected model identifies which population subgroups are systematically mis-served by population-priced contracts.

The normative question of whether clinical data should be used in economic decisions is outside the scope of this paper. The positive question of whether the lifecycle model is currently misspecified by ignoring clinical trajectory information is within scope, and the answer is: yes, structurally.

9.4 The Program

The program of which this paper is a part has a simple statement: the smooth actuarial hazard is the last unreformed input in heterogeneous-agent macroeconomics. The HANK program reformed the income process, the wealth distribution, and the monetary transmission mechanism. What it did not reform is the mortality input that governs when the lifecycle problem terminates and what happens at termination.

The tools to reform it now exist. The clinical transformer provides trajectory-dependent mortality estimates. The tropical polynomial characterization provides a representation class that can encode the relevant geometry. The Auclert sufficient-statistic methodology provides the aggregation framework. The Stantcheva sufficient-statistic methodology provides the optimal tax application. All that remains is the empirical work: linking the clinical model to the lifecycle data, estimating the curvature-

exposure covariance, and asking whether the smooth model's residuals correlate with the clinical curvature score in the ways the theory predicts.

If they do, the lifecycle model needs a new input. If they do not, the smooth hazard is a better approximation than the boundary-curvature framework suggests, and the framework should be revised. The claim of this paper is that the question is well-posed and the answer is empirically accessible, not that the answer is known in advance.

10. Conclusion

This paper has made three arguments of descending generality.

The most general is geometric: the value function for any expected hitting-time problem on a domain with a smooth absorbing boundary has curvature that diverges at rate $s^{-3/2}$ near that boundary. No smooth function can represent this, regardless of how many parameters it contains. Any lifecycle model that uses a smooth hazard to represent mortality is, for agents near the boundary, using a qualitatively wrong approximation to the relevant mathematical object.

The intermediate argument is aggregative: the population-level response to symmetric interventions is asymmetric by an amount governed by the covariance between boundary curvature and intervention exposure. This covariance is not a small correction in clinical settings, because the mortality boundary is universal, and the curvature integral over the full population diverges in a way that the macroeconomic-analogy scaling argument cannot accommodate. The Hansen-Sargent (2017) framework provides the information-theoretic language for precisely how wrong the smooth approximation is: the Chernoff half-life between the trajectory-dependent and population-table mortality distributions vanishes near the boundary at the same rate as the curvature blows up.

The most specific argument is empirical: the constellation of lifecycle puzzles documented in the literature (at minimum the eleven identified in Section 6) follow as equilibrium properties of a lifecycle model with the mortality input corrected, without any changes to preferences, financial frictions, or behavioral assumptions. The mechanism in each case is the same: survival probability reweighting in the first-order condition, or precautionary savings amplification from boundary curvature, produces behavior that smooth-mortality models cannot generate.

The paper has not provided the empirical estimates that would confirm or refute these claims in a specific sample. That work requires linked clinical-lifecycle data and is in progress. What the paper has provided is the theoretical framework that makes the claims precise, the sufficient-statistic characterization that makes them testable, and the representation class that makes them computationally tractable. The mortality input problem has a structure, a cause, and a cure. The empirical work is the next step.

References

10.1 Economics

Ameriks, J. and Zeldes, S. P. (2004). How Do Household Portfolio Shares Vary with Age? Working Paper, Columbia Business School.

Auclert, A. (2019). Monetary Policy and the Redistribution Channel. *American Economic Review*, 109(6), 2333–2367.

Auclert, A., Bárdoczy, B., Rognlie, M., and Straub, L. (2021). Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models. *Econometrica*, 89(5), 2375–2408.

Benartzi, S. and Thaler, R. H. (1995). Myopic Loss Aversion and the Equity Premium Puzzle. *Quarterly Journal of Economics*, 110(1), 73–92.

Berger, D., Bocola, L., and Dovis, A. (2023). Imperfect Risk Sharing and the Business Cycle. *Quarterly Journal of Economics*, 138(3), 1765–1815.

Bernheim, B. D., Lemke, R. J., and Scholz, J. K. (2004). Do Estate and Gift Taxes Affect the Timing of Private Transfers? *Journal of Public Economics*, 88(12), 2617–2634.

Bound, J., Schoenbaum, M., Stinebrickner, T. R., and Waidmann, T. (1999). The Dynamic Effects of Health on the Labor Force Transitions of Older Workers. *Labour Economics*, 6(2), 179–202.

Brown, J. R. and Finkelstein, A. (2007). Why Is the Market for Long-Term Care Insurance So Small? *Journal of Public Economics*, 91(10), 1967–1991.

Brown, J. R. and Finkelstein, A. (2008). The Interaction of Public and Private Insurance: Medicaid and the Long-Term Care Insurance Market. *American Economic Review*, 98(3), 1083–1102.

Carroll, C. D. (1997). Buffer-Stock Saving and the Life Cycle / Permanent Income Hypothesis. *Quarterly Journal of Economics*, 112(1), 1–55.

Cocco, J. F., Gomes, F. J., and Maenhout, P. J. (2005). Consumption and Portfolio Choice over the Life Cycle. *Review of Financial Studies*, 18(2), 491–533.

Coile, C., Diamond, P., Gruber, J., and Jousten, A. (2002). Delays in Claiming Social Security Benefits. *Journal of Public Economics*, 84(3), 357–385.

Currie, J. and Madrian, B. C. (1999). Health, Health Insurance and the Labor Market. In *Handbook of Labor Economics*, Vol. 3C. Elsevier.

Cutler, D. M. (2004). *Your Money or Your Life: Strong Medicine for America's Health Care System*. Oxford University Press.

De Nardi, M. (2004). Wealth Inequality and Intergenerational Links. *Review of Economic Studies*, 71(3), 743–768.

De Nardi, M., French, E., and Jones, J. B. (2010). Why Do the Elderly Save? The Role of Medical Expenses. *Journal of Political Economy*, 118(1), 39–75.

Ditto, P. H., Druley, J. A., Moore, K. A., Danks, J. H., and Smucker, W. D. (2003). Fates Worse Than Death: The Role of Valued Life Activities in Health-State Evaluations. *Health Psychology*, 15(5), 332–343.

Dynan, K. E., Skinner, J., and Zeldes, S. P. (2004). Do the Rich Save More? *Journal of Political Economy*, 112(2), 397–444.

Gan, L., Hurd, M., and McFadden, D. (2015). Individual Subjective Survival Curves. In *Analyses in the Economics of Aging*. University of Chicago Press.

Hansen, L. P. and Sargent, T. J. (2017). Prices of Macroeconomic Uncertainties with Tenuous Beliefs. Becker Friedman Institute Working Paper No. 2973331. [Acknowledgment: Victor Zhorin contributed the Chernoff entropy calculations.]

Heimer, R. Z., Myrseth, K. O., and Schoenle, R. S. (2019). YOLO: Mortality Beliefs and Household Finance Puzzles. *Journal of Finance*, 74(6), 2957–2996.

Hurd, M. D. (1989). Mortality Risk and Bequests. *Econometrica*, 57(4), 779–813.

Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263–292.

Kaplan, G., Moll, B., and Violante, G. L. (2018). Monetary Policy According to HANK. *American Economic Review*, 108(3), 697–743.

Kaplan, G. and Violante, G. L. (2014). A Model of the Consumption Response to Fiscal Stimulus Payments. *Econometrica*, 82(4), 1199–1239.

Krusell, P. and Smith, A. A. (1998). Income and Wealth Heterogeneity in the Macroeconomy. *Journal of Political Economy*, 106(5), 867–896.

Kurnianingsih, Y. A., Sim, S. K. Y., Chee, M. W. L., and Mullette-Gillman, O. A. (2015). Aging and Loss Aversion: Age-Related Changes in the Neural Systems Supporting Optimal Decision Making under Risk. *NeuroImage*, 111, 468–479.

Lockwood, L. M. (2012). Bequest Motives and the Annuity Puzzle. *Review of Economic Dynamics*, 15(2), 226–243.

Lockwood, L. M. (2018). Incidental Bequests and the Choice to Self-Insure Late-Life Risks. *American Economic Review*, 108(9), 2513–2550.

Lubitz, J. D. and Riley, G. F. (1993). Trends in Medicare Payments in the Last Year of Life. *New England Journal of Medicine*, 328(15), 1092–1096.

Mather, M., Mazar, N., Gorlick, M. A., et al. (2012). Risk Preferences and Aging: The "Certainty Effect" in Older Adults' Decision Making. *Psychology and Aging*, 27(4), 801–816.

Moll, B., Townsend, R. M., and Zhorin, V. (2017). Economic Development, Flow of Funds, and the Equilibrium Interaction of Financial Frictions. *PNAS*, 114(24), 6176–6184.

Reichling, F. and Smetters, K. (2015). Optimal Annuitization with Stochastic Mortality and Correlated Medical Costs. *American Economic Review*, 105(11), 3273–3320.

Riley, G. F. and Lubitz, J. D. (2010). Long-Term Trends in Medicare Payments in the Last Year of Life. *Health Services Research*, 45(2), 565–576.

Ruhm, C. J. (2000). Are Recessions Good for Your Health? *Quarterly Journal of Economics*, 115(2), 617–650.

Samuelson, W. and Zeckhauser, R. (1988). Status Quo Bias in Decision Making. *Journal of Risk and Uncertainty*, 1(1), 7–59.

Shoven, J. B. and Slavov, S. N. (2014). Does It Pay to Delay Social Security? *Journal of Pension Economics and Finance*, 13(2), 121–144.

Stantcheva, S. (2015). Optimal Income, Education and Bequest Taxes in an Intergenerational Model. NBER Working Paper 21177.

Stantcheva, S. (2017). Optimal Taxation and Human Capital Policies over the Life Cycle. *Journal of Political Economy*, 125(6), 1931–1990.

Straub, L. (2019). Consumption, Savings, and the Distribution of Permanent Income. Working Paper, Harvard University.

Thaler, R. H. (1980). Toward a Positive Theory of Consumer Choice. *Journal of Economic Behavior and Organization*, 1(1), 39–60.

Townsend, R. M. (1994). Risk and Insurance in Village India. *Econometrica*, 62(3), 539–591.

Townsend, R. M. and Samphantharak, K. (2009). *Households as Corporate Firms*. Econometric Society Monograph. Cambridge University Press.

Tversky, A. and Kahneman, D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323.

Yaari, M. E. (1965). Uncertain Lifetime, Life Insurance, and the Theory of the Consumer. *Review of Economic Studies*, 32(2), 137–150.

10.2 Mathematics and Statistics

Chen, K.-T. (1958). Integration of Paths: A Faithful Representation of Paths by Noncommutative Formal Power Series. *Transactions of the American Mathematical Society*, 89(2), 395–407.

Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society, Series B*, 34(2), 187–220.

Crandall, M. G. and Lions, P.-L. (1983). Viscosity Solutions of Hamilton-Jacobi Equations. *Transactions of the American Mathematical Society*, 277(1), 1–42.

Fleming, W. H. and Soner, H. M. (2006). *Controlled Markov Processes and Viscosity Solutions*, 2nd ed. Springer.

Friedman, A. (1975). *Stochastic Differential Equations and Applications*. Academic Press.

Hambly, B. and Lyons, T. (2010). Uniqueness for the Signature of a Path of Bounded Variation and the Reduced Path Group. *Annals of Mathematics*, 171(1), 109–167.

Karatzas, I. and Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus*. Springer.

Kazemi, S. M., et al. (2019). Time2Vec: Learning a General Purpose Representation

of Time. arXiv:1907.05321.

Lyons, T. J. (1998). Differential Equations Driven by Rough Signals. *Revista Matemática Iberoamericana*, 14(2), 215–310.

Maclagan, D. and Sturmfels, B. (2015). *Introduction to Tropical Geometry*. AMS Graduate Studies in Mathematics, Vol. 161.

Salvi, C., et al. (2021). The Signature Kernel is the Solution of a Goursat PDE. *SIAM Journal on Mathematics of Data Science*, 3(3), 873–899.

Zhang, L., Naitzat, G., and Lim, L.-H. (2018). Tropical Geometry of Deep Neural Networks. *ICML 2018*, PMLR Vol. 80, 5824–5832.

10.3 Constituent Papers

Zhorin, V. (2026a). Mortality Prediction as Boundary Value Problem. SSRN Working Paper 6309178.

Zhorin, V. (2026b). The Boundary-Curvature Asymmetry: A Geometric Identity Linking Heterogeneous-Agent Macroeconomics to Clinical Mortality Modeling. SSRN Working Paper 6616018.

Zhorin, V. (2026c). Why Financial Event Sequences Require Temporal Transformers: Necessity Theorems for Default Prediction. SSRN Working Paper 6154026.

Zhorin, V. (2026d). Discrete Causal Topology. Zenodo. https://doi.org/10.5281/zenodo.18449228