Dynamic Amortization Governance Model

A Retirement Framework by Luca Fittabile

To my wife, Vivian Olivia Centellas Claros

Abstract

We evaluate retirement withdrawal rules under a criterion aligned with practice: a policy must be executable year after year on a single realized path of wealth and horizon, without foresight of future returns. The **Dynamic Amortization Governance Model (DAGM)** embeds a portfolio-consistent amortization anchor (**DARVA**) and complementary references—Guyton–Klinger, fixed safe-withdrawal, and constant-percentage rules—inside an explicit corridor. A capital-based robustness diagnostic maps the realized portfolio to governance states that expand or compress feasible spending while keeping trade-offs auditable.

Using rolling historical simulations on U.S. (1928–2025) and euro-area (1999–2025) samples, we show that static reference rules encode different economic trade-offs—among average consumption, downside floors, upside participation, and income variability—so they are not interchangeable substitutes. DAGM improves joint outcomes on average and maximum real withdrawals relative to simpler static benchmarks in most configurations, while the comparison with dynamic ARVA is nuanced: ARVA can retain an edge on some floor-like summaries, whereas DAGM more often raises average spending and systematically preserves stronger upside along favorable paths. Additional stability metrics show that DAGM reduces the incidence and cumulative burden of downward year-to-year spending changes versus ARVA across the matrix, reallocating variability rather than eliminating dynamics. Robustness lengthens horizons and adds sleeves (including inflation-linked bonds and gold), and compares matched-terminal-equity glide paths against flat allocations.

The lifecycle extension replaces the fixed plan duration with a dynamically updated demographic horizon derived from official mortality tables at each review, and compares matched lifecycle–fixed twin plans: the lifecycle plan conserves capital for a client still alive at the calendar anchor, whereas the fixed plan depletes toward that date by design.

The paper extends the lifecycle framework to couple households via an *annual-rebalance* architecture: the two members' individual mortality residuals are blended each year into an effective horizon $T_{\text{eff}}(t)$, and the full DAGM engine runs on top. Across seven M65+spouse demographic pairs (USD 1928–2025), the architecture reduces the Phase A spousal coverage cost from up to -10% under a fixed joint-horizon plan to 0–5%, while producing a positive withdrawal jump at first death (+19–44%) as accumulated capital is reallocated from the joint horizon to the survivor's individual lifecycle. Coverage cost, demographic withdrawal jump, and advisor-stated lifestyle step-down are made explicit as three independent, auditable parameters.

A further extension formalizes an **intentional legacy reserve**: a fraction λ of initial capital ring-fenced as a CPI-indexed bequest target, segregated from the spendable capital that drives the withdrawal engine. At each annual review the corridor operates on spendable wealth alone, making the tension between retirement income and bequest intention visible and auditable year by year. The framework distinguishes three structural categories of terminal capital — intentional and quantified (DAGM with $\lambda > 0$), accidental (GK, SWR), and zero by design (ARVA) — and identifies an operating range for λ beyond which depletion risk makes a dedicated inheritance portfolio the more appropriate vehicle.

Disclaimer

This paper is provided for **research and educational purposes only**. It does **not** constitute investment, tax, legal, or personalised financial advice, nor a recommendation to adopt any withdrawal strategy or product. All results are derived from **historical rolling simulations, block-bootstrap resampling, and adversarial stress scenarios** on past market data; they describe how specified rules would have behaved on those paths and **do not** predict future returns, inflation, mortality, or retirement outcomes. Parameter choices, governance rules, and implementation details may differ from any commercial offering. Readers should consult qualified professionals before making financial decisions.

Contents

1 Introduction

Retirement spending is often framed as a forecasting problem: estimate long-run returns, simulate many market paths, and pick a withdrawal rate that "works" with high probability. Retirees do not live in simulation space. They live on one realized path—whatever markets deliver—while liabilities and circumstances evolve. A usable spending model must therefore be real-path implementable: each year it takes the wealth you actually have, incorporates the shocks and life events that actually occur, and maps that state into a transparent spending policy that can be audited and explained.

The DAGM framework follows that premise. It does not predict returns or claim an optimal withdrawal rate under a fixed distributional model. It provides a deterministic mapping from observable inputs at each decision date (wealth, remaining horizon, policy parameters, recorded shocks) into a baseline withdrawal, augmented by a decision-support structure that makes trade-offs explicit. The goal is not to remove judgment, but to replace ad hoc choices with a small set of explicit, disciplined policy levers informed by the best information available at each review.

1.1 Realized paths, sequence risk, and why distributions are not enough

You do not receive historical averages; you receive the order of returns that unfolds during retirement. Two households with identical initial wealth and identical long-run assumptions can diverge sharply if return sequences differ. The practical problem is therefore not only calibration of expected returns, but specification of a policy that updates coherently after good and bad years and can absorb changes in liabilities (pensions, large expenses, windfalls) without breaking the logic of the plan.

Retirement research often evaluates rules with Monte Carlo or historical resampling. These tools generate rich distributions of outcomes and metrics such as depletion probability. They can build intuition—but retirees experience a single path, not a distribution. The relevant criterion is whether a rule can be executed repeatedly on the path that actually occurs: adapting to realized wealth and horizon without requiring foresight of future returns. That distinction motivates this paper.

1.2 Contributions of the Paper

The paper advances the Dynamic Amortization Governance Model along seven dimensions.

(a) It introduces **DARVA** (*Dynamic ARVA*): a portfolio-consistent ARVA conversion anchor whose annuity factor is built from asset-class anchors and portfolio weights instead of a single imposed discount rate. Because amortization speed tracks the implemented mix, the spending baseline aligns with the portfolio's economics and supports glide paths and defensive tilts around retirement—standard responses to sequence-of-returns risk.

(b) It adds a multiplicative policy parameter m on the DARVA reference withdrawal (relative to plain dynamic ARVA at the same state), enabling spending-smile profiles over decumulation and a clean bridge between portfolio withdrawals and public or corporate pension income paid outside the investable portfolio.

(c) It embeds those references in a corridor, so the tension between consumption efficiency and spending stability stays visible and governable.

(d) It models legacy explicitly: a CPI-indexed reserve ring-fenced from the capital amortized through DARVA, so the corridor and governance badges operate on spendable wealth $S_t = W_t - L_t$, not gross portfolio value. Under severe stress, drawing on legacy can be allowed alongside ordinary rules, surfacing the trade-off between honoring the bequest and protecting the spending standard. Empirically (§7), the framework distinguishes three structural categories of terminal capital—intentional and quantified (DAGM with $\lambda > 0$), accidental (GK, SWR), and zero by design (ARVA)—and identifies the operating range of λ: residual spendable capital turns negative on average above $\lambda = 20\%$, and worst-window depletion risk materializes above $\lambda = 40\%$ (USD 1928–2025, 69 rolling 30-year windows).

(e) On top of the standard fixed-ARVA plan duration, DAGM offers an optional mode in which the **plan duration** is updated yearly from official statistical life tables and a client-chosen confidence level, to manage longevity risk effectively.

(f) In the empirical section, it complements classical sustainability-style summaries with path-based stability metrics—the frequency and depth of downward spending adjustments—so performance on realized trajectories is judged with the same granularity the retiree actually experiences.

(g) It extends the lifecycle demographic plan to couple households through an *annual-rebalance* architecture: at each review, the two members' individual $p = 90\%$ mortality residuals are blended into a dynamically updated effective horizon $T_{\text{eff}}(t)$, and the full DAGM governance engine runs on top of that blended anchor. The architecture separates three previously conflated phenomena—spousal coverage cost, demographic withdrawal jump at first death, and advisor-stated lifestyle step-down—into independent, explicitly parameterized plan inputs. Across seven demographic configurations on USD historical data, Phase A coverage cost is reduced to 0–5%, and first-death withdrawal transitions are positive and moderate (+19–44%), reversing the factor-of-two discontinuity of the fixed joint-horizon approach.

2 Taxonomy of Withdrawal Rules

Withdrawal rules can be organized according to the state variables they respond to, the decision cadence they assume, and the extent to which they make trade-offs explicit between sustainability, adequacy, and spending stability. A practical taxonomy is not merely descriptive: it clarifies what each class of methods can and cannot guarantee, and it prevents recommended withdrawals from being interpreted as objective quantities rather than as policy choices.

In broad terms, most retirement spending approaches fall into four families.

The first family consists of fixed real spending rules, which maintain a constant inflation-adjusted withdrawal once retirement begins. These rules prioritize consumption stability but are highly sensitive to the sequence of realized returns.

The second family includes percentage-of-portfolio rules, in which withdrawals are defined as a fixed proportion of current wealth. These rules adapt immediately to changes in portfolio value, eliminating the mechanical risk of depletion but potentially generating significant variability in spending.

The third family comprises conditional decision-rule approaches, often referred to as guardrail methods. These rules maintain a baseline withdrawal path but introduce adjustment mechanisms triggered by observable state variables, such as the withdrawal rate relative to portfolio value. The objective of these frameworks is typically to balance income stability with long-term sustainability by allowing discrete spending adjustments when predefined thresholds are crossed.

The fourth family includes actuarial or amortization-based approaches, in which the portfolio is converted into a spending stream through an annuity factor based on the remaining horizon and a conversion anchor. In these frameworks withdrawals adjust systematically over time as wealth and remaining horizon evolve, effectively distributing portfolio risk across the retirement horizon.

These families differ not only in their mechanics but also in the objectives they implicitly prioritize. Fixed real spending rules emphasize stability, percentage-based rules emphasize automatic sustainability, conditional guardrail systems aim to balance stability and adaptability, and amortization-based methods prioritize consumption efficiency relative to the remaining horizon.

The models examined in this paper are drawn from these families and represent widely used approaches to retirement spending policy. The next section defines the specific reference models used in the analysis.

3 Reference Models

The empirical analysis evaluates four withdrawal frameworks representing the main classes of retirement spending rules:

- Fixed Safe Withdrawal Rate (SWR)

- Constant Percentage Withdrawal (CPW)
- Guyton–Klinger guardrail rule (GK)
- Dynamic ARVA (Actuarial Remaining Value Approach)

The first three models are treated as primary policy references in the empirical analysis alongside DARVA. The CPW rule is included mainly as a conceptual benchmark representing the percentage-of-portfolio class of withdrawal policies.

Each model is briefly described below.

3.1 Fixed Safe Withdrawal Rate

The Safe Withdrawal Rate (SWR) rule represents the traditional fixed real spending approach widely discussed in retirement finance literature. Under this framework the retiree withdraws a constant inflation-adjusted amount determined as a fixed percentage of the initial portfolio value. Once the withdrawal level is established at retirement, it evolves only with inflation and does not respond to subsequent portfolio performance. The rule therefore provides maximum spending stability but does not incorporate any mechanism to adapt to realized market outcomes.

In empirical studies the SWR rule is typically evaluated through historical simulations or Monte Carlo analysis in order to estimate the probability of portfolio depletion before the end of the retirement horizon. Within the empirical analysis of this paper the SWR rule serves as a reference benchmark for fixed real spending policies.

3.2 Constant Percentage Withdrawal

The Constant Percentage Withdrawal (CPW) rule defines annual spending as a fixed proportion of the current portfolio value. Unlike fixed real spending rules, withdrawals therefore adjust automatically to changes in portfolio wealth. Because withdrawals scale directly with portfolio value, the CPW rule eliminates the mechanical possibility of portfolio depletion. However, the resulting spending path can exhibit substantial variability, as withdrawals fluctuate in proportion to market outcomes.

3.3 Guyton-Klinger Guardrail Rule

The Guyton–Klinger (GK) rule is a conditional decision-rule framework designed to balance withdrawal stability with long-term sustainability. The method maintains a baseline spending level that evolves with inflation but introduces adjustment mechanisms triggered by the portfolio withdrawal rate.

Under the GK framework, withdrawals are initially defined as a fixed percentage of the retirement portfolio. In subsequent years the withdrawal amount is adjusted for inflation provided that the implied withdrawal rate remains within predefined guardrails. When the withdrawal rate exceeds the upper guardrail, spending reductions are triggered; when it falls below the lower guardrail, spending increases may be allowed.

The objective of the GK approach is therefore to maintain spending stability under normal market conditions while introducing discrete adjustments when the sustainability of the plan deteriorates or improves significantly.

In this paper the GK rule serves as the stability-oriented reference policy within the DAGM governance framework.

3.4 Dynamic ARVA and DARVA (portfolio-consistent)

Dynamic ARVA is an amortization-based withdrawal rule: at each annual review it maps current capital, remaining horizon, and policy parameters into a baseline withdrawal.

Capital and legacy reserve. Let C denote the total portfolio capital the client records at the annual review decision date—the amount already reflecting any deposits, withdrawals, or windfalls they choose to include in that single figure. Optional capital events CE are labels for non-market explanations (gifts, lump-sum expenses, transfers, etc.): they are stored for audit and traceability but are not numerically added to C for amortization purposes. The economic total used for ARVA/DARVA and the corridor is therefore

$$C_{\text{eff}} = C.$$

A legacy reserve $L \geq 0$ is treated as not available for ordinary decumulation. By default it is set from an initial plan capital C_0 and a share $\lambda \in [0, 1)$,

$$L_{\text{nom}} = \lambda\, C_0,$$

or from an explicit nominal amount; alternatively the policy may specify CPI-indexed legacy from a frozen base amount and base date, in which case the indexed amount is computed and then capped so that $L \leq C_{\text{eff}} = C$. Spendable capital for ARVA is

$$C_{\text{avail}} = \max(0,\ C_{\text{eff}} - L).$$

DARVA conversion anchor. Unlike a single scalar discount rate, the conversion rate is portfolio-consistent: for asset classes $k \in \{\text{Equity}, \text{Bond}, \text{ILBond}, \text{Gold}\}$ with weights w_k (summing to one) and class anchors d_k,

$$d_t = \sum_k w_k\, d_k.$$

Annuity factor and withdrawal. Let T be the remaining horizon in whole years (inclusive). The engine uses the present-value factor for a unit payment each year for T years at rate d_t:

$$A = \begin{cases} T, & |d_t| \approx 0, \\ \dfrac{1 - (1 + d_t)^{-T}}{d_t}, & \text{otherwise.} \end{cases}$$

With policy multiplier m (default $m = 1$), the ARVA reference withdrawal is

$$W_{\text{ARVA}} = \frac{m\, C_{\text{avail}}}{A}.$$

Equivalently, for d_t away from zero, this is the standard amortization of C_{avail} over T periods at rate d_t, applied to capital after legacy is set aside.

This formulation ensures that the baseline tracks allocation (through d_t), horizon (through T and A), and legacy (through C_{avail}). At each review, inputs are refreshed: capital, remaining T, weights, anchors, and—when applicable—indexed legacy, so the rule adapts to the realized path.

ARVA belongs to the class of horizon-explicit decumulation rules. Unlike SWR, GK, and CPW, its withdrawal level is directly tied to the remaining plan horizon T. Longer horizons, all else equal, imply lower supported withdrawals through a larger A.

The Dynamic ARVA / DARVA framework therefore produces a portfolio-adaptive decumulation schedule in which the conversion speed responds to the weighted anchors while legacy is handled explicitly before annuitizing the spendable balance.

4 DAGM Governance Framework

The Dynamic Amortization Governance Model (DAGM) introduces a governance framework designed to translate the reference withdrawal models described in the previous section into a practical decision structure for retirement spending.

Rather than selecting a single withdrawal rule as universally optimal, DAGM treats the reference models as complementary policies reflecting different objectives. In particular, Dynamic ARVA emphasizes consumption efficiency relative to the remaining retirement horizon, while the Guyton–Klinger guardrail rule emphasizes spending stability.

The DAGM framework therefore organizes these models into a governance structure that guides withdrawal decisions along the realized portfolio path.

4.1 Corridor Logic

Under normal portfolio conditions, DAGM defines a withdrawal corridor bounded by the Dynamic ARVA baseline and the Guyton–Klinger guardrail reference.

Dynamic ARVA and the Guyton–Klinger rule represent two complementary responses to realized portfolio dynamics.

Dynamic ARVA converts remaining wealth into consumption over the remaining horizon. As a result, withdrawals respond immediately and mechanically to changes in portfolio value. Under adverse market conditions the ARVA withdrawal can therefore decline significantly, reflecting the deterioration of the available capital base.

The Guyton–Klinger framework instead prioritizes spending stability. The withdrawal level evolves more gradually because adjustments are triggered only when predefined guardrails are breached. This mechanism can temporarily protect spending levels during market downturns, but it may also delay the adjustment of withdrawals to deteriorating portfolio conditions.

The corridor defined by these two rules therefore captures a fundamental trade-off between consumption efficiency and rapid capital adjustment (ARVA) and income stability through gradual policy adjustments (GK).

In this sense the corridor should be interpreted as a normal-state decision range. When the portfolio enters stressed conditions, the feasible withdrawal space may temporarily compress toward more conservative levels as the sustainability constraints of the plan become binding.

4.2 Stress Signals and Conservative References

When the portfolio enters a stressed state, the governance logic of DAGM may require positioning withdrawals below the standard ARVA–GK corridor.

In such situations the framework introduces a more conservative lower reference determined by the sustainability conditions of the plan rather than by a single fixed rule. The appropriate withdrawal level depends on the realized portfolio path, the remaining capital, and the remaining retirement horizon.

As a result, the effective lower bound for withdrawals may vary across different historical scenarios and across different portfolio states. This adaptive mechanism ensures that the governance framework remains responsive to the financial state of the retirement plan while preserving the flexibility required to accommodate client-specific circumstances.

4.3 Parametric calibration of governance parameters

Three design dimensions determine the quantitative behavior of the DAGM badge system: (i) the form of the **expected-capital trajectory** $E(t)$, which defines what "being on track" means at each year of the plan; (ii) the **badge threshold triple** $(\theta_1, \theta_2, \theta_3)$, which maps the robustness ratio $r(t) = C_{\text{eff}}(t)/E(t)$

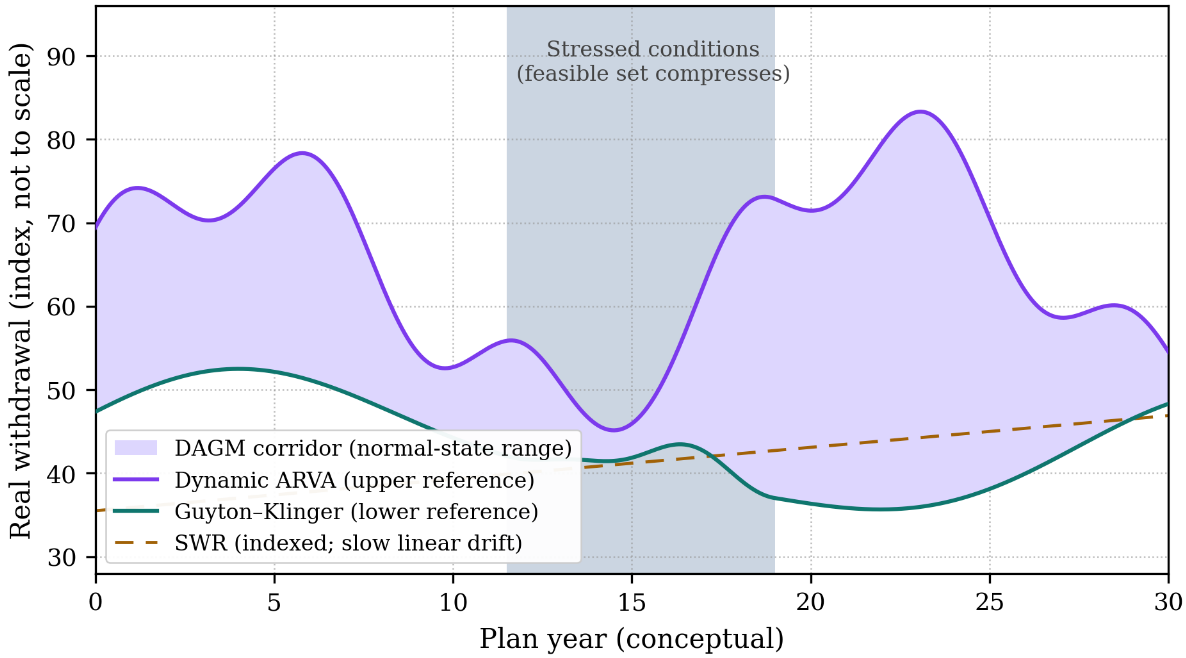

Figure 1: Conceptual DAGM corridor on a realized path: normal-state band between Dynamic ARVA (upper) and Guyton–Klinger (lower); stressed window with compressed feasible set; inflation-indexed SWR with slow real drift. Not to scale. Colours match the Sandbox policy key (DARVA, GK, SWR).

onto governance states subject to $\theta_3 < \theta_2 < \theta_1$; and (iii) the **per-badge model-weighting schedules**, which determine how the four reference models (ARVA, GK, SWR, CPW) are combined within the corridor at each governance state.

Dimension 1 — Expected-capital trajectory $E(t)$. The trajectory determines the denominator of the robustness ratio and therefore what constitutes adequate capital at each review. Three families are considered: a *linear* schedule $E(t) = C_0 \cdot (T - t)/T$ (uniform depletion); an *annuity-consistent* schedule $E(t) = W_0 \cdot A(T - t,\, d)$ where $A(\cdot)$ is the annuity present-value factor at discount d (internally consistent with the ARVA withdrawal formula); and a *real annuity* variant deflated by realized CPI.

Dimension 2 — Badge thresholds $(\theta_1, \theta_2, \theta_3)$. Given $E(t)$, the mapping from ratio to badge state is:

$$\begin{aligned} r(t) \geq \theta_1 &\;\Rightarrow\; \text{GREEN} \\ \theta_2 \leq r(t) < \theta_1 &\;\Rightarrow\; \text{AMBER} \\ \theta_3 \leq r(t) < \theta_2 &\;\Rightarrow\; \text{RED} \\ r(t) < \theta_3 &\;\Rightarrow\; \text{SUPER RED} \end{aligned}$$

where θ_1, θ_2, and θ_3 are proprietary calibration outputs with $\theta_3 < \theta_2 < \theta_1$. Exact threshold values are not disclosed in this version.

Dimension 3 — Per-badge model weights. At each badge state, the governance engine positions withdrawals within the feasible corridor using a weighted combination of the reference models. The exact weighting schedule is calibrated jointly with Dimensions 1 and 2.

Calibration methodology. Parameter selection was conducted via an **iterative multi-agent search** over the three design dimensions. The DAGM engine was instrumented to accept all three design dimensions as runtime inputs, exposing the full parameter space as a callable interface. **Qwen2.5-Coder:32B** and

DeepSeek-R1 (both run locally via Ollama) served as the two active LLM agents in a closed loop: at each iteration an agent proposed a candidate configuration, the engine executed a historical rolling simulation on in-sample data, and the agent evaluated the resulting multi-criteria outcome — mean real withdrawals, worst-window floor, cut-year frequency, and badge health. Configurations that improved on the current best were retained; the search proceeded iteratively over several months until convergence.

The implementation methodology is worth noting: to the authors' knowledge, this is one of the first applications of a closed-loop LLM-driven parametric search to the calibration of a retirement governance framework. Earlier attempts with the generation of locally-deployable models available in 2024 were unsuccessful; the methodology became consistently productive from late 2025 onward, when open-weight reasoning models reached sufficient capability for multi-step parametric search, code execution, and self-correcting evaluation on domain-specific simulation outputs — all running on consumer hardware without external API dependencies.

Parameter freeze. After exploratory calibration converged, the governance parameters — badge thresholds $(\theta_1, \theta_2, \theta_3)$, the expected-capital trajectory $E(t)$, and the per-badge reference-model weighting schedules — were **held fixed** for all empirical work reported in §§5–7. No threshold, trajectory choice, or weighting schedule was revised for individual rolling configurations, regions, horizons, asset menus, lifecycle designs, legacy settings, bootstrap paths, or stress-tape windows.

The exploratory search that produced this configuration was **not archived in fully reproducible form**. The paper therefore reports **historical performance of a fixed governance rule** on a broad test matrix, not a formally documented train–validation split. Exact parameter values remain proprietary and are not disclosed in this version. The resulting configuration does not claim to be the global optimum of the parameter space; it represents a calibration that is sufficiently good to produce a consistent improvement over static reference models across the historical matrix used in this paper. ## Data and Empirical Design

4.4 Datasets

The empirical analysis uses two historical datasets designed to represent distinct macro-financial regimes. **Appendix B** documents their construction (sources, indices, and bridge years). The same material is provided to application users as the Sandbox **Dataset Reference** (PDF). Extended empirical tables, supplementary figures (glide path, lifecycle twins), and editorial source notes are published in a single appendix supplement PDF (see **Appendices** at the end of this document); the main text cites dataset construction in this section and in **Appendix B** of the unified supplement (same text as the Sandbox **Dataset Reference** PDF).

USD (`USD_1928_2025`). Annual calendar-year returns use December-to-December index changes unless noted otherwise. Equities, nominal bonds, gold, and US inflation follow Aswath Damodaran (NYU Stern) through 2024; calendar 2025 is appended from independent market sources. Inflation-linked bonds use a TIPS ETF proxy from 2004 onward and the nominal Treasury series before that (see **Appendix B** §2).

EUR (`EUR_1999_2025`). The euro-area sample starts at 1999 (euro cash introduction and available index history in this build). Series are aligned to EUR and, where applicable, built from monthly index levels aggregated to December-to-December annual returns (R = Index_Dec_Y / Index_Dec_{Y−1} − 1). Equities use MSCI ACWI in EUR (1999–2024 from the Curvo-exported series; 2025 follows the continuous historical EUR row in the repository build for consistent methodology). Nominal bonds use Bloomberg Euro Aggregate where the index history supports clean Dec/Dec years (2010–2025 in the current snapshot); 1999–2009 use the historical EUR dataset values because the imported index strip does not cover those years on the same convention. Inflation-linked bonds follow Bloomberg Euro Government Inflation-Linked for the covered window, with 1999–2005 and 2025 filled from the historical EUR dataset where December levels are missing. Gold uses gold bullion in EUR (Dec/Dec). Inflation for real/nominal conversion uses the euro-area Harmonised Index of Consumer Prices (HICP, all-items, annual rate), sourced from Eurostat. Full detail: **Appendix B** §3.

For each dataset we construct rolling historical windows with a one-year step.

The core empirical matrix (§5.2–§5.3) uses retirement horizons of 10, 15, 20, and 30 years where dataset length allows; **§5.4** adds USD 60/40 and 40/60 sleeves at **40- and 50-year** horizons on the same rolling-window logic. Five equity/bond allocations from 100/0 to 20/80 are tested in the main matrix. **Appendix B** is the authoritative, line-by-line construction note (sources, indices, and bridge years).

5 Empirical Analysis

Data note. §5.1–§5.7 report rolling historical simulations on the §4.5 datasets. §6 reports lifecycle twin experiments; §7 reports the intentional legacy-reserve extension. Extended tables and construction notes are in the appendix supplement.

5.1 Epoch 1 Static Withdrawal Models

Epoch 1 evaluates the behavior of four static withdrawal rules across the common set of 25 historical rolling retirement runs. The purpose is not to identify a universal winner, but to characterize the distinct economic behaviors of the reference models under different horizons, allocations, and market histories.

Across the matrix, the models display clearly differentiated spending profiles. ARVA generally produces the highest withdrawal levels because it recalculates spending directly from remaining capital and remaining horizon. CPW usually ranks next on average withdrawals and upside participation because spending scales proportionally with portfolio value. GK is more compressed and guardrail-driven, while SWR remains the most conservative benchmark in terms of spending. This diversity is precisely what makes the later corridor interpretation meaningful: the policies should be read as complementary reference signals rather than rival candidates for a single universally optimal rule.

Table 1 — First-place counts across the 25 historical runs

Metric	ARVA	GK	SWR	CPW
W_avg_real	25	0	0	0
W_min_real	22	1	2	0
W_max_real	25	0	0	0

The ranking evidence remains strongly ARVA-dominated. ARVA places first in all 25 runs for mean real withdrawal and in all 25 runs for maximum real withdrawal. On the downside metric, ARVA places first in 22 of 25 runs: GK retains one first place in the EUR 20-year 100/0 configuration, and SWR posts a slightly higher average floor in two USD configurations (`usd_2080_rolling_30y`, `usd_4060_rolling_30y`) after correcting USD CPI in the historical dataset. This confirms that the static rules are not interchangeable: ARVA is the dominant income-maximizing reference, but it is not designed to preserve capital, while the more conservative rules give up spending in exchange for other objectives.

Illustrative historical configuration

The contrast between models is visible even in a single representative run. The table below reports the USD dataset, 30-year horizon, 60/40 allocation, which is a useful mid-spectrum case because it shows clear separation across all three dimensions without relying on an extreme allocation.

Metric	ARVA	GK	SWR	CPW
W_avg_real	67,005	53,489	40,221	53,520
W_min_real	41,990	38,410	38,350	33,354
W_max_real	96,924	69,611	43,823	76,900

Interpretation

Taken together, the Epoch 1 results show that static withdrawal rules encode different economic trade-offs rather than variations of a single optimal rule. ARVA is the most aggressive income reference, CPW is the

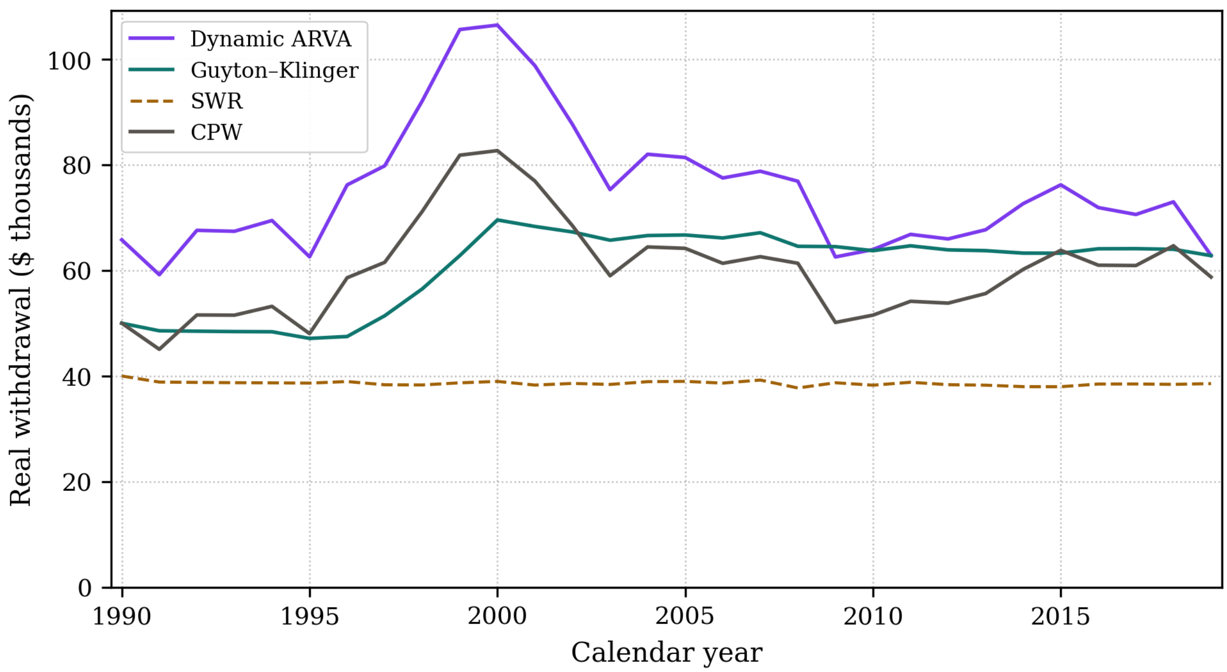

Figure 2: Real annual withdrawals for four static rules on one historical window (USD 60/40, 30-year horizon, retirement starting in 1990; calendar years on the horizontal axis). The table above summarizes metrics aggregated across rolling start years; this path illustrates year-by-year separation on a single realization. Line colours follow the Sandbox policy key (Dynamic ARVA, GK, SWR dashed, CPW). Source: Epoch 1.

most directly capital-adaptive rule, GK emphasizes guardrail discipline, and SWR acts as a conservative safety benchmark.

Plan duration emerges as the dominant fragility driver. The step from 20 to 30 years, especially in the USD dataset, materially reduces sustainable spending and increases exposure to sequence risk. This is also where adaptation speed becomes more consequential: ARVA reacts immediately through capital re-annuitization, while GK adjusts more gradually through guardrails.

The results also challenge simple intuitions about safety. In the USD 30-year 20/80 configuration, the guardrail rule shows visible fragility, suggesting that very bond-heavy portfolios may fail to sustain withdrawals when real growth is insufficient. **Appendix A.4** gives exhaustion timing, calendar clustering of worst cuts, and the worked stagflation vignette on this sleeve. Conversely, conservative rules such as SWR and, at times, GK may accumulate substantial terminal wealth in favorable markets, which is another way of saying that they can underconsume when realized returns are strong.

CPW highlights the opposite trade-off. Because withdrawals move mechanically with capital, the rule shows strong adaptability and very low apparent failure pressure, but this comes at the cost of higher income variability. More broadly, the simulations indicate that model ordering is not fully stable under stress: the rules should be interpreted as dynamic references, not as a fixed hierarchy. This observation provides the conceptual bridge to the corridor governance framework developed in the next epoch.

5.2 Epoch 2 - Governance Outcomes with DAGM

Epoch 2 evaluates the empirical outcomes produced by the DAGM governance framework relative to four static withdrawal rules: Dynamic ARVA, Guyton–Klinger (GK), Safe Withdrawal Rate (SWR), and Constant Percentage Withdrawal (CPW). The purpose of this section is to examine whether introducing a governance layer that combines multiple policy references alters the realized distribution of withdrawals across historical retirement paths.

The analysis uses the same 25 rolling historical runs introduced in Epoch 1 and focuses on three metrics: mean real withdrawal (W_avg_real), mean minimum real withdrawal (W_min_real), and mean maximum real withdrawal (W_max_real). These capture central consumption, downside spending floor, and upside participation, respectively.

Note (level metrics, rolling windows). For each historical start window, the engine already reports W_avg_real, W_min_real, and W_max_real within that window. Unless this paper states otherwise, configuration-level figures for these three quantities are means across rolling start windows (i.e. the mean of window-level averages, of window-level minima, and of window-level maxima). They are not the global minimum of W_min_real across windows nor the global maximum of W_max_real across windows—summaries that would stress the worst start or best upside across the full rolling set rather than the typical window-level profile. The same convention is used in §5.4 and §5.5.

Note (depletion flags). **Early depletion** means capital is exhausted before the planned horizon T is fully executed; **terminal depletion** means exhaustion occurs only after a full T-year run. Table 6 (§5.5, DAGM-only rolling windows) applies the same split on windows that complete T years. **§5.6** and **Appendix A.6** use the same definitions on bootstrap paths and Möbius windows, including truncated paths (essential for SWR on the stress tape).

Because Epoch 2 uses the same 25 historical runs as Epoch 1, the static-model values reported there remain unchanged here. The main text therefore focuses on governance evidence, while detailed compact tables are moved to the appendix.

Win-count evidence

Table 1. Win counts across historical runs

Metric	DAGM vs ARVA	DAGM vs GK	DAGM vs SWR	DAGM vs CPW
W_avg_real	18 / 25	25 / 25	25 / 25	25 / 25
W_min_real	7 / 25	24 / 25	24 / 25	25 / 25
W_max_real	25 / 25	25 / 25	25 / 25	25 / 25

A complementary perspective is obtained by examining how often DAGM produces higher withdrawals than the reference rules across the full set of historical runs.

The figure below plots configuration-level mean W_avg_real for DAGM against each static reference (the metric in the first row of Table 1).

The results of Epoch 2 illustrate the role of governance in retirement spending policy. While the static withdrawal rules examined in the previous section prioritize specific objectives, the DAGM framework combines these policy signals into a structured decision process. The resulting withdrawal paths reflect a dynamic balance between consumption efficiency, spending stability, and sustainability constraints rather than the rigid application of a single rule.

Illustrative comparative case

Table 2 reports one representative run chosen for readability: a USD case with a 30-year horizon and a traditional 60/40 allocation. It combines a long retirement horizon with a balanced portfolio mix and shows the governance pattern clearly without relying on an extreme allocation.

Table 2 — Illustrative case (USD, 30-year horizon, 60/40 allocation)

Policy	W_avg_real	W_min_real	W_max_real	Terminal wealth (real)
DAGM	71,781	42,804	143,661	90,862
ARVA	67,005	41,990	96,924	0

Policy	W_avg_real	W_min_real	W_max_real	Terminal wealth (real)
GK	53,489	38,410	69,611	1,098,110
SWR	40,221	38,350	43,823	1,789,910
CPW	53,520	33,354	76,900	1,070,226

Terminal wealth: mean real capital remaining at end of horizon, across rolling windows. ARVA depletes to zero by actuarial design — full amortization of spendable capital is its explicit objective. DAGM also targets progressive depletion of spendable capital, but does so less abruptly: the governance corridor distributes path variability across years, so residual capital reflects the unplanned remainder of an explicit consumption-first trade-off rather than algorithmic conservatism. GK, SWR, and CPW accumulate large terminal balances as an unintended by-product of their withdrawal algorithms — not as a planned bequest.

In this representative case, DAGM produces the highest average withdrawal, preserves a minimum withdrawal slightly above ARVA, and delivers the highest maximum withdrawal among all policies. The case therefore makes visible the basic effect of the governance layer: it does not simply reduce spending to gain control, but reshapes the spending profile across the realized path.

Interpretation

The win-count evidence reinforces this interpretation. DAGM generates higher average withdrawals than Guyton–Klinger, SWR, and CPW in all simulations, and exceeds the Dynamic ARVA baseline in the majority of runs (18 out of 25). On the minimum withdrawal metric, DAGM improves the withdrawal floor relative to GK in 24 of 25 runs, relative to SWR in 24 of 25, and relative to CPW in all 25, while Dynamic ARVA produces a higher minimum withdrawal in 18 of 25 configurations because of its direct amortization of remaining capital. On the maximum withdrawal metric, DAGM consistently dominates all reference policies across the full simulation set, reflecting the ability of the corridor structure to preserve upside participation during favourable market environments.

Taken together, these results confirm the intended role of the governance framework: DAGM does not aim to dominate a single static rule on every metric, but rather to combine the strengths of multiple reference policies — its structured decision process is precisely what makes that combination possible.

More specifically, relative to GK, SWR, and CPW, DAGM improves the joint profile of average withdrawals, downside floor, and upside participation. Relative to ARVA, the comparison is more nuanced: ARVA sometimes delivers a slightly higher spending floor, but DAGM more often improves the average withdrawal outcome and systematically preserves more upside in favourable historical paths. The governance layer therefore changes the distribution of withdrawals rather than merely shifting their level up or down.

This is the central result of Epoch 2. The contribution of DAGM is not that it replaces the reference rules with a single universally superior withdrawal formula. Its contribution is that it uses the reference rules as complementary policy signals and transforms them into a real-path governance process that yields a more balanced realized spending profile.

5.3 Epoch 3 Additional Metrics

Objective and evaluation lens

Epoch 3 extends the empirical analysis from withdrawal levels to withdrawal stability. Where Epoch 1 and Epoch 2 focus on average, minimum, and maximum real withdrawals, this section examines the quality of the year-to-year withdrawal path generated under realized historical market sequences. The comparison is conducted primarily between DAGM and Dynamic ARVA, since ARVA is the most natural benchmark for assessing whether the governance layer changes the stability properties of a capital-responsive spending rule.

The purpose is not to ask whether DAGM produces a perfectly smooth withdrawal stream. Instead, the goal is to determine whether the governance layer improves the downside profile of spending adjustments while preserving the capacity to adapt along the realized path.

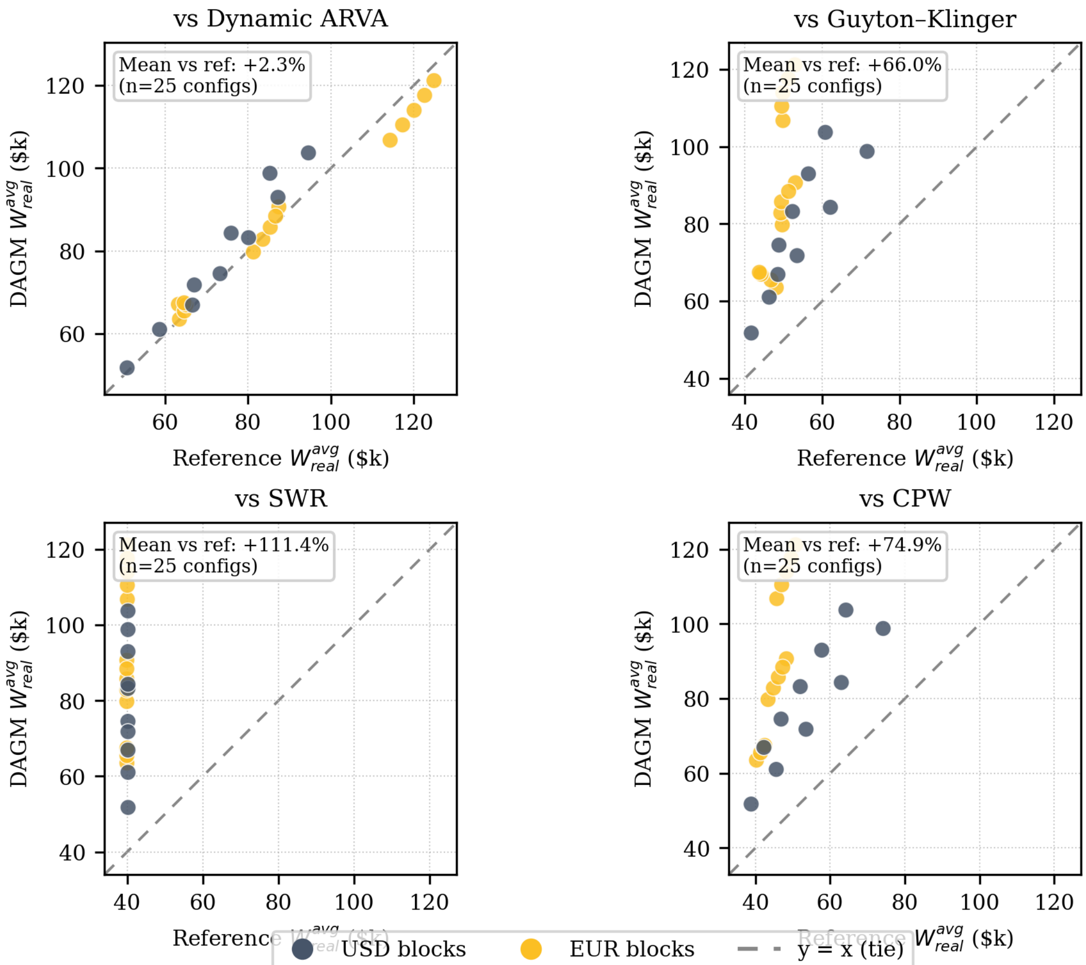

Figure 3: Epoch 2 — Mean real withdrawal (W_{real}^{avg}) at configuration level: DAGM versus each static reference, in $ thousands. Each point is one of the 25 Lab configurations; scatter colours distinguish USD vs EUR blocks (not model identity colours). Points above the $y = x$ line are wins; vertical distance from the line is the margin on this metric. Box: mean percentage gap vs that reference. Source: Epoch 2.

Definition of additional stability metrics

The new metrics introduced in Epoch 3 are designed to separate downside spending pressure from total dynamism.

Table 1 — Definition of additional stability metrics

Metric	Definition and interpretation
Cut-year frequency	Share of plan years in which the real withdrawal falls relative to the previous year. Lower values mean fewer spending cuts.
DWII	Downside Withdrawal Instability Index: a summary measure of the cumulative burden of negative year-to-year withdrawal revisions. Higher values indicate more frequent and/or larger downside adjustments.
avg_abs_change_pct	Average absolute annual percentage change in withdrawals, regardless of sign. This captures total dynamism rather than downside instability specifically.
Mean worst-cut metric	Average across historical paths of each path's largest single-year withdrawal cut. It summarizes tail severity at the run level.
Conditional average cut size	Average size of withdrawal cuts, conditional on a cut occurring. This isolates the magnitude of cuts rather than their frequency.
max_cut_pct (path level)	Largest single negative annual percentage change in withdrawals along one individual historical path. This metric is useful for path-by-path tail inspection.

Summary evidence across the completed matrix

Across the completed historical rolling matrix, the most robust empirical result of Epoch 3 is a systematic improvement in downside withdrawal stability relative to ARVA. DAGM reduces the frequency of withdrawal cuts and improves the Downside Withdrawal Instability Index in every run of the matrix. Mean absolute annual percentage changes in withdrawals are higher for DAGM than for ARVA in 21 of 25 configurations (the exceptions are four USD 30-year high-equity blocks). The evidence therefore indicates that DAGM does not eliminate variability; rather, it reallocates variability away from recurrent downside consumption shocks and, in most configurations, toward broader adaptive adjustments along the realized path.

Table 2 — Summary evidence: DAGM vs ARVA (wins out of 25)

Metric	DAGM vs ARVA	Interpretation
Lower cut-year frequency	25 / 25	Systematic reduction in the incidence of withdrawal cuts.
Lower DWII	25 / 25	Systematic reduction in downside instability burden.
Higher avg_abs_change_pct	21 / 25	DAGM is more dynamic overall on this summary in most runs, not flatter in absolute terms.
Improved mean worst-cut metric	21 / 25	Frequent, but not universal, improvement in cut severity.

Metric	DAGM vs ARVA	Interpretation
Improved conditional average cut size	14 / 25	Mixed evidence on the magnitude of cuts once cuts occur.

The figure below visualizes the same configuration-level comparisons summarized in Table 2 (DAGM minus ARVA means across rolling paths).

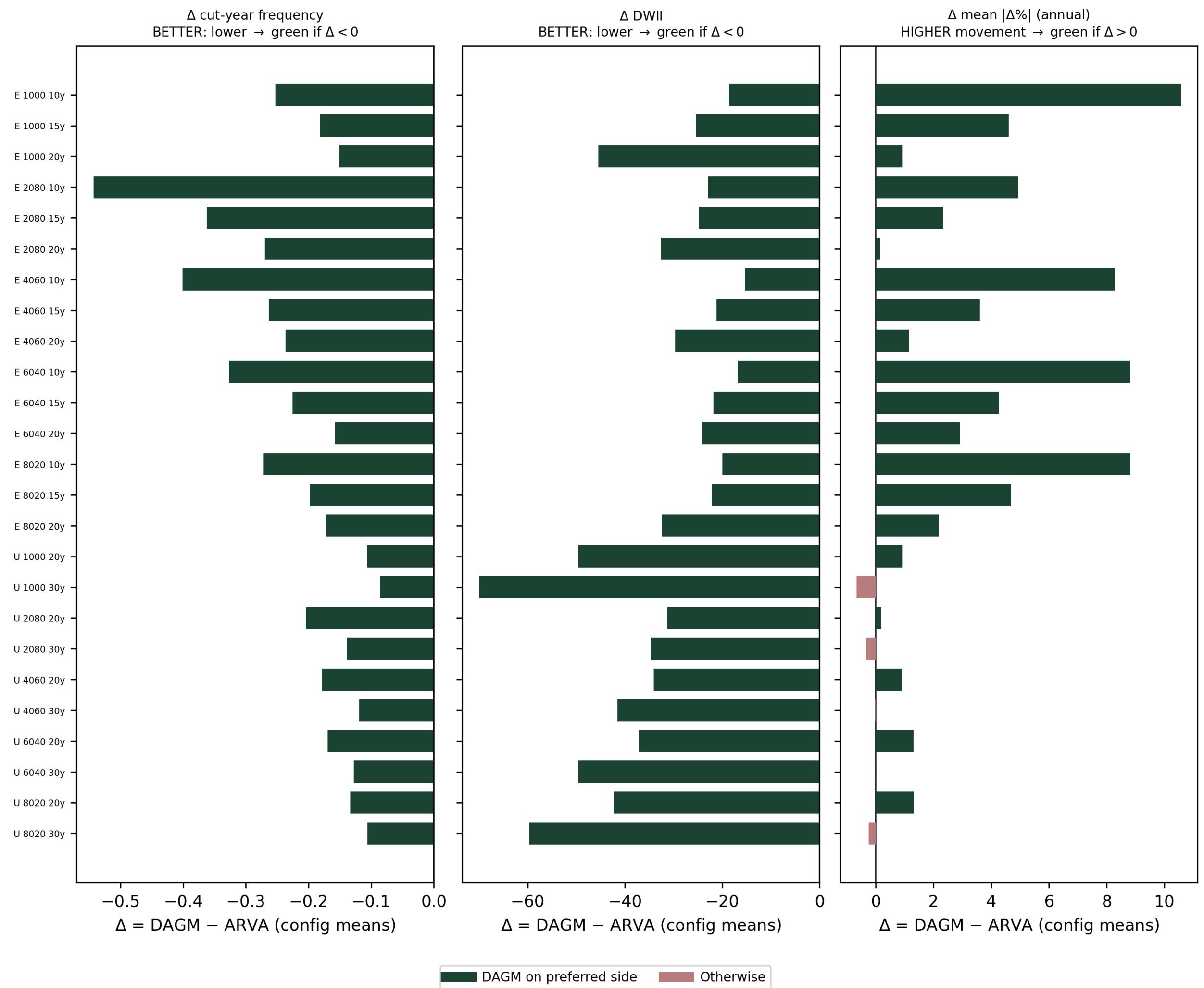

Figure 4: Epoch 3 — For each of the 25 Lab configurations, horizontal bars show Δ = DAGM minus ARVA (means across rolling paths). **Cut-year frequency and DWII:** lower is better; green where $\Delta < 0$. **Mean absolute annual % change:** green where $\Delta > 0$ (more aggregate year-to-year movement; 21 / 25 configs). Source: Epoch 3.

Cut severity results

Evidence regarding the severity of cuts is favorable but not universal. DAGM improves the mean worst-cut metric in 21 of 25 runs and improves the conditional average cut size in 14 of 25 runs: it consistently reduces the incidence and cumulative downside burden of cuts, while improving cut severity in a majority, though not all, of the completed runs.

Dataset asymmetry

The USD dataset shows particularly strong evidence. Across the USD runs, DAGM improves not only cut frequency and DWII in every case, but also improves cut severity metrics in the vast majority of configurations. The EUR dataset confirms the downside-control effect but shows more mixed results on cut severity. In EUR, the most consistent improvement lies in the reduction of downside withdrawal instability rather than in uniform improvements in the magnitude of cuts. This asymmetry strengthens rather than weakens the overall interpretation: the cross-dataset result is downside control, while improvements in cut severity depend more on dataset structure and return regimes.

Pathwise non-dominance

Path-level analysis shows that DAGM does not dominate ARVA on every historical start window. This should not be interpreted as a weakness of the framework. A rule-based governance system designed for real-path implementation should be evaluated by the distributional trade-offs it generates, not by unrealistic expectations of universal pointwise superiority.

Table 3 — Pathwise non-dominance (worst-cut comparison, DAGM vs ARVA)

Configuration	DAGM better	DAGM worse	Comment
EUR 60/40, 20y	2 / 8	6 / 8	Small sample; pathwise split does not overturn matrix-level downside gains.
USD 60/40, 20y	53 / 79	26 / 79	Useful example of non-dominance despite better average downside stability.
USD 60/40, 30y	59 / 69	10 / 69	Governance benefit becomes much clearer at longer horizon.
USD 80/20, 30y	60 / 69	9 / 69	Strongly favorable block for worst-cut outcomes.

Tail-risk case and design boundary

Across the completed matrix, the main empirical exception is the USD 20/80 30-year configuration, where DAGM shows a heavy-tail deterioration in worst-cut outcomes. Pathwise inspection suggests that the typical path remains favorable to DAGM, but a minority of extreme historical paths produce very large late-stage withdrawal collapses. Detailed ledger inspection indicates that the most severe collapse occurs near the end of the plan horizon, after capital has already been nearly depleted. In such cases, the RED governance rule can still select a withdrawal that is large relative to the remaining capital base, producing a sharp decline in the following year. This phenomenon is best understood as a late-stage depletion cliff rather than as a mid-horizon governance failure, and it highlights an important design boundary for automated withdrawal governance near the final years of a retirement plan. A systematic decomposition—exhaustion timing versus floors, calendar clustering of worst YoY cuts, and the worked 1950 start vignette—is in **Appendix A.4**.

This tail case is a terminal-phase boundary (short remaining horizon, depleted capital, bond-heavy paths), not evidence of mid-horizon instability; **Appendix A.4** shows that the worst real cuts can occur even when the badge stays GREEN.

Final interpretation

Taken together, the evidence from Epoch 3 supports a clear and defensible interpretation of the DAGM framework. DAGM does not eliminate spending variability, and it does not dominate ARVA path by path on every historical window. Instead, DAGM systematically reduces downside spending pressure under realized historical paths. The framework achieves this by reallocating variability away from repeated downward spending shocks and toward explicit adaptation of withdrawals as the realized capital trajectory evolves.

The strongest conclusion that emerges from the completed matrix is therefore not universal smoothness or universal dominance, but a robust improvement in the governance of withdrawal decisions under real-world market paths. This result also motivates keeping stability metrics separate from level metrics: a withdrawal policy may be more dynamic overall while still being materially better on the downside margin that retirees are most likely to perceive as painful.

Note: detailed formulas can be moved to the technical appendix; this section keeps only paper-level definitions and empirical interpretation.

5.4 Epoch 4 Horizon Robustness

Objective and evaluation lens

Epoch 4 tests whether the empirical governance benefit identified in the previous sections remains visible when the retirement horizon is extended beyond the 30-year configurations already examined. To isolate the horizon effect cleanly, the analysis is restricted to the USD historical dataset, to equity/bond portfolios only, and to four mandatory rolling configurations: 60/40 and 40/60 allocations at 40-year and 50-year horizons. The evaluation retains the same distinction developed in the earlier epochs between level metrics, downside stability, and pathwise tail behavior. This focus is especially relevant because the earlier results had already suggested that plan duration is a primary fragility driver and that, in the USD 60/40 case, the governance benefit became more visible when moving from 20 to 30 years.

Summary evidence across the longer-horizon block

Across the four mandatory runs, DAGM remains empirically strong on the three core level metrics. Table 1 reports means across rolling start windows for W_avg_real, W_min_real, and W_max_real; Table 2 summarizes pathwise worst-cut dominance for the same four configurations.

Table 1 — USD long-horizon level metrics (mean across rolling windows; real)

Values are means of W_avg_real, W_min_real, and W_max_real across historical start windows, in thousands (rounded). Same definition of level metrics as in the §5.2 note. Rows at 30 years use Epoch 2 summaries; rows at 40 and 50 years use Epoch 4 summaries.

Configuration	Policy	W_avg_real (k)	W_min_real (k)	W_max_real (k)
USD 60/40, 30y	ARVA	67.0	42.0	96.9
USD 60/40, 30y	DAGM	71.8	42.8	143.7
USD 60/40, 40y	ARVA	58.1	32.1	86.7
USD 60/40, 40y	DAGM	62.9	35.7	123.9
USD 60/40, 50y	ARVA	48.0	24.0	72.9
USD 60/40, 50y	DAGM	51.9	28.6	106.3
USD 40/60, 30y	ARVA	58.6	37.4	81.2
USD 40/60, 30y	DAGM	61.0	39.5	112.2
USD 40/60, 40y	ARVA	49.3	27.1	70.9
USD 40/60, 40y	DAGM	51.4	29.8	95.1
USD 40/60, 50y	ARVA	39.0	19.0	58.3
USD 40/60, 50y	DAGM	40.3	21.1	78.6

Table 2 — Pathwise worst-cut dominance (DAGM vs ARVA, four mandatory configurations)

Path-level **max_cut_pct** (largest single-year real withdrawal cut along one rolling window). **DAGM better** counts windows where DAGM's pathwise worst cut is smaller than ARVA's; **DAGM worse** the converse. Source: Epoch 4 rolling-window exports (appendix supplement).

Configuration	DAGM better	DAGM worse	Comment
USD 60/40, 40y	57 / 59	2 / 59	Strongly favorable at 40-year horizon.
USD 60/40, 50y	41 / 49	5 / 49	Still favorable; smaller rolling set at 50y.
USD 40/60, 40y	47 / 59	10 / 59	Favorable on average with more pathwise exceptions.
USD 40/60, 50y	42 / 49	5 / 49	Tail stress visible in a minority of starts.

The figure below shows, for Dynamic ARVA and DAGM separately, the change in the Table 1 rolling-mean level metrics relative to the USD 30-year Epoch 2 baseline (40y—30y and 50y—30y), using Epoch 2 summaries at 30y and Epoch 4 summaries at 40y and 50y.

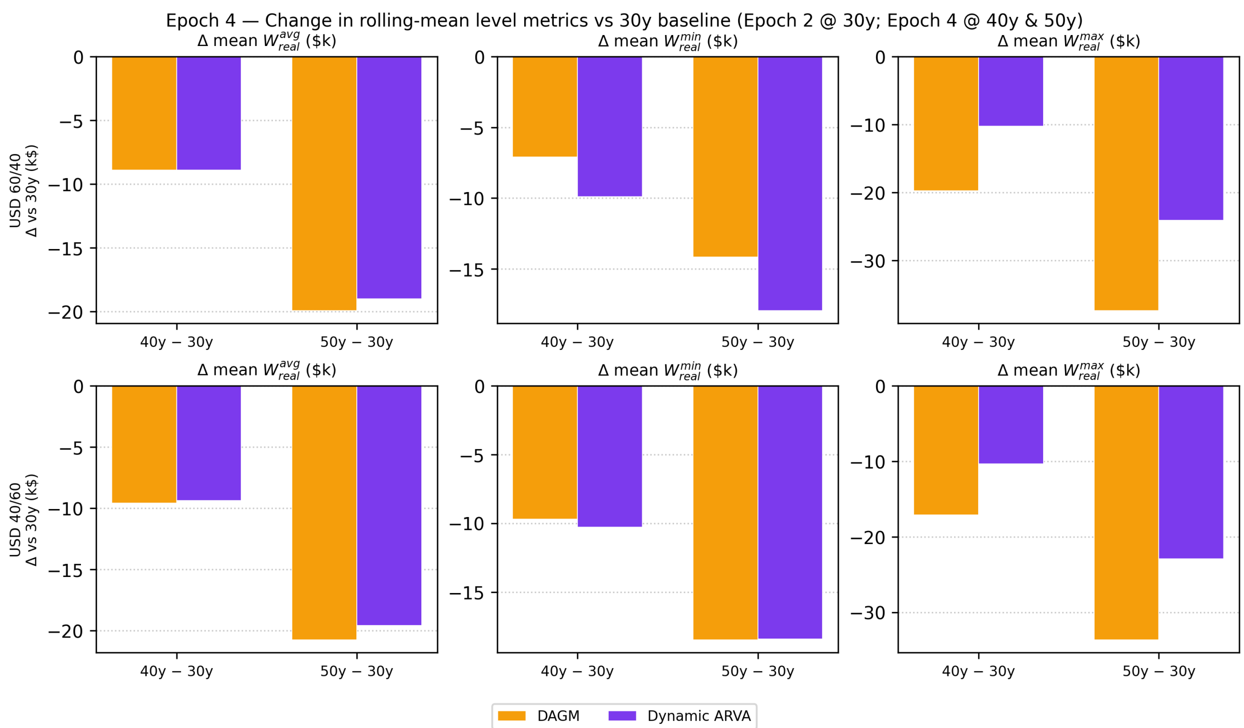

Figure 5: Epoch 4 — Change in rolling-mean level metrics (real k$) vs the 30-year Epoch 2 baseline for USD 60/40 and 40/60. For Dynamic ARVA and DAGM separately, each bar is the mean at 40y or 50y minus the mean at 30y for W_{real}^{avg}, W_{real}^{min}, or W_{real}^{max}.

Downside stability at long horizons

The additional stability evidence remains favorable to DAGM. Across all four mandatory configurations, the governance layer continues to reduce the incidence and cumulative burden of withdrawal cuts relative to Dynamic ARVA, while preserving a stronger overall withdrawal profile.

The 60/40 block is especially strong: the pathwise worst-cut comparison in Table 2 remains overwhelmingly favorable to DAGM at both the 40-year and 50-year horizons, with only a small minority of historical starts showing deterioration.

The 40/60 block also remains favorable on average, but with more visible stress in the tails at the 50-year horizon. This reinforces the interpretive discipline established in Epoch 3: the relevant question is not whether DAGM achieves universal path-by-path dominance, but whether it improves the realized trade-off between consumption level, downside pressure, and adaptive behavior along the realized path.

Tail boundary and qualification

The main qualification in Epoch 4 arises in the USD 40/60 50-year configuration. Even there, the average level metrics remain favorable to DAGM, and the overall downside stability profile remains improved relative to DARVA. However, a small number of extreme historical starts still produce worse pathwise worst-cut outcomes for DAGM. Inspection of the yearly ledgers shows that the most problematic cases are concentrated in Great Depression start windows, particularly 1929 and 1930, and that the deterioration occurs late in the plan horizon under prolonged stress states, including repeated SUPER RED conditions.

This pattern is best interpreted in the same way as the tail exception identified in Epoch 3: not as evidence of a general mid-horizon governance failure, but as a late-stage depletion cliff that remains a design boundary for the framework under very long and adverse realized paths.

Interpretation

Taken together, the evidence from Epoch 4 strengthens the core empirical case for DAGM. Extending the horizon from 30 years to 40 and 50 years does not erase the governance benefit. On the contrary, the mandatory USD block shows that DAGM remains strong on average spending, generally preserves a better downside floor, and continues to deliver markedly stronger upside participation than the static reference. At the same time, the longer-horizon extension sharpens an important qualification: the framework does not dominate ARVA on every path, and in highly adverse long-horizon bond-heavy regimes a minority of late-stage tail cases can still deteriorate sharply. **DAGM improves the governance of withdrawals under longer realized paths by reducing downside spending pressure and preserving a more balanced spending profile, even though very long horizons can still expose tail-risk boundaries in extreme historical windows.**

5.5 Epoch 5 Asset Class Extension

Epoch 5C extends the asset-class robustness test to a four-asset opportunity set including equity, nominal bonds, inflation-linked bonds, and gold. The objective is not to optimize the allocation but to evaluate whether the broader opportunity set alters the empirical ordering of withdrawal policies observed in earlier epochs.

5.5.1 Inflation Linked Extension

Epoch 5A should be interpreted primarily as a robustness test on policy ordering, rather than as a portfolio-optimization exercise. The central question is whether replacing part of the nominal bond sleeve with inflation-linked bonds alters the empirical structure documented in earlier epochs: namely, that corridor-governed DAGM tends to dominate the simpler fixed-reference rules (GK, SWR, and CPW) on average withdrawals and upside participation, while the comparison with ARVA remains more nuanced because ARVA can retain a stronger floor.

Across the inflation-linked runs, the evidence suggests that the answer is no. Introducing an IL sleeve does not materially change the relative ordering among withdrawal policies. The corridor continues to operate between the same ARVA and GK anchors, the internal corridor logic remains unchanged, and DAGM continues to outperform GK, SWR, and CPW along the dimensions that proved most relevant in earlier epochs. ARVA remains the primary comparator on downside protection.

This pattern appears consistently across the tested geographical and horizon blocks. In the EUR 20-year runs, the corridor remains a genuine intermediate governance rule: it consistently sits between the ARVA and GK references and continues to deliver higher withdrawals than the more conservative fixed rules without displacing ARVA as the stronger floor-oriented benchmark. The same structure is observed in the USD 20-year runs, where the corridor again stands above GK, SWR, and CPW across representative cases, while

ARVA continues to dominate the minimum-withdrawal side of the distribution. The USD 30-year runs display the same ordering: the corridor is still governed by the same ARVA–GK relationship and the same governance logic, rather than by a new regime induced by the asset mix.

From a portfolio perspective, the IL extension behaves primarily as a bond-sleeve substitution rather than as a transformation of the opportunity set. In the tested runs it perturbs outcome levels locally but does not generate a materially different environment for the governance layer. Corridor widths, anchor ordering, and policy relationships remain broadly similar to the stock–bond baseline. This is consistent with the behavior expected from a near-neutral diversification extension: it may slightly modify outcomes, but it does not rewrite the decision structure governing withdrawals.

The most defensible interpretation is therefore that the IL extension does not create a new empirical regime. The earlier conclusion remains intact: DAGM should be understood as a governance layer that combines the strengths of simpler rules rather than as a rule that must dominate every reference on every metric. Under the IL extension, as under the stock–bond baseline, DAGM continues to dominate GK, SWR, and CPW in joint distributional terms, while the familiar caveat versus ARVA remains concentrated in floor-type outcomes. In this sense, Epoch 5A functions primarily as a stability check: the asset extension modestly perturbs the portfolio, but it does not alter the substance of the earlier DAGM result.

5.5.2 Gold

Epoch 5B asks the same robustness question under a more economically meaningful asset extension. As in Epoch 5A, the central issue is not whether Gold improves individual metrics in isolation, but whether introducing a Gold sleeve alters the relative empirical position of DAGM compared with the reference rules.

The evidence suggests that it does not. Across the Gold runs, the relative ordering among withdrawal policies remains broadly unchanged. The corridor-governed policy continues to dominate GK, SWR, and CPW on the distributional dimensions that proved most relevant in earlier epochs, while the qualitative comparison with ARVA remains similar. Gold therefore does not overturn the earlier governance result; instead, it preserves it under a somewhat richer portfolio opportunity set.

This continuity is visible across the tested blocks. In the EUR 20-year runs, representative corridor paths continue to sit above GK, SWR, and CPW in terms of average and maximum withdrawals, while ARVA remains the more floor-oriented benchmark. The same structure appears in the USD 20-year runs, where DAGM again dominates the simpler rules in the center and upper tail of the withdrawal distribution. In the USD 30-year runs the same relationship persists: the corridor continues to operate between ARVA and GK, inherits the same governance logic, and leaves ARVA as the primary downside comparator.

Where Gold differs from the IL extension is at the portfolio level. Whereas inflation-linked bonds behave mainly as a substitution inside the bond sleeve, Gold appears to act as a more visible distribution-shaping diversifier. In representative runs the corridor is associated with stronger upside participation, somewhat wider corridors, and higher residual wealth capacity, particularly in medium- and long-horizon cases. These changes do not eliminate the trade-off with ARVA, and they do not imply monotonic improvements in every consumption metric. Rather, they suggest that Gold alters the portfolio environment in a more economically meaningful way than IL, without modifying the governance logic itself.

The most defensible reading is therefore that Gold introduces a stronger portfolio perturbation without altering the governance conclusion. DAGM continues to dominate GK, SWR, and CPW in joint distributional terms, particularly when average spending and upside participation are considered together. Relative to ARVA, the same non-trivial trade-off remains: ARVA retains its role as the more aggressive income-maximizing reference, while DAGM continues to offer a more balanced spending profile, accepting some loss of floor dominance in exchange for stronger participation in favorable paths. Gold therefore strengthens the portfolio environment but does not change the empirical role of the DAGM governance layer.

5.5.3 Four-Asset Extension

Opening: asset-class robustness question

Epoch 5C evaluates the asset-class robustness of the DAGM framework. The objective is not to identify an optimal portfolio allocation, but to test whether the governance logic of the corridor remains effective when the opportunity set of the portfolio expands beyond a traditional equity–bond mix. The previous epochs focused primarily on two-asset portfolios. In this extension the non-equity sleeve is broadened to include inflation-linked bonds and gold, creating a four-asset allocation structure. The purpose of this experiment is therefore not portfolio optimization, but to test whether the empirical ordering of withdrawal policies identified in earlier epochs remains stable when the asset universe becomes more diversified.

To answer this question, the same canonical research protocol is applied across EUR and USD historical datasets with equity allocations ranging from 20% to 80%. For each configuration the analysis compares five withdrawal policies: ARVA, GK, SWR, CPW, and the DAGM corridor-governed strategy. The experiment therefore provides a direct test of whether the corridor framework depends on a specific asset structure or instead represents a portfolio-agnostic governance mechanism.

Core empirical result

The central empirical result of the four-asset extension is straightforward and remarkably stable across datasets and horizons.

First, DAGM remains the strongest model on average real withdrawals (W_avg_real). Across the full set of runs the corridor-governed strategy consistently delivers the highest or near-highest level of average real consumption.

Second, DAGM clearly dominates the maximum withdrawal metric (W_max_real). This reflects the ability of the corridor framework to capture favorable market paths while maintaining governance discipline. When realized returns are strong, the corridor expands and allows withdrawals to increase, producing significantly higher upside consumption levels than the static rules.

Third, ARVA continues to maintain an advantage on the minimum withdrawal metric (W_min_real). ARVA recalculates withdrawals as a dynamic amortization over the remaining plan horizon, which tends to preserve higher minimum withdrawal levels during adverse sequences. However, this floor protection is typically achieved at the cost of lower participation in favorable paths.

Taken together, these results reproduce the same pattern observed in earlier epochs: DAGM optimizes average and upside consumption, while ARVA retains a relative advantage on the minimum-withdrawal dimension. Importantly, the introduction of additional asset classes does not alter this fundamental trade-off.

Table 1 — Aggregate performance across the four-asset runs

Values are the mean across the twelve four-asset configurations of, for each policy, the mean across rolling start windows of W_avg_real, W_min_real, and W_max_real (see §5.2 note). Withdrawals are in real units (same numeraire as elsewhere in the paper).

Model	Mean W_avg_real	Mean W_min_real	Mean W_max_real
DAGM	71,896.2	48,388.9	133,835.5
ARVA	69,232.4	50,705.1	90,793.3
CPW	47,759.8	34,246.8	64,533.2
GK	49,171.6	41,581.3	59,940.3
SWR	40,115.0	38,683.8	42,820.3

Three conclusions emerge clearly from this aggregate comparison (Table 1). First, DAGM delivers the highest average real withdrawals among all policies. Second, DAGM exhibits a very large advantage in maximum withdrawals, reflecting strong participation in favorable paths. Third, ARVA remains the strongest model on the minimum-withdrawal dimension, confirming the structural trade-off already observed in earlier epochs.

Table 2 — Win counts (DAGM vs reference policies), four-asset block

Counts are out of 12 configurations. Within each configuration, policies are compared using the same window-mean construction as in Table 1 (mean of W_avg_real, W_min_real, and W_max_real across rolling windows).

Metric	DAGM vs ARVA	DAGM vs GK	DAGM vs SWR	DAGM vs CPW
W_avg_real	12	12	12	12
W_min_real	2	12	12	12
W_max_real	12	12	12	12

The results reinforce the interpretation already visible in Tables 1 and 2. On the average withdrawal metric, DAGM outperforms ARVA in every configuration and consistently dominates the remaining static policies. On the maximum withdrawal metric, DAGM wins in every run against all competing models, demonstrating the corridor's ability to capture favorable market paths. The only systematic exception concerns minimum withdrawals. ARVA retains an advantage on this dimension, winning most comparisons against DAGM. However, DAGM still consistently outperforms the remaining static rules. This pattern confirms that the corridor framework balances adaptability and stability rather than optimizing a single dimension of the withdrawal distribution.

The figure below reproduces the Table 1 aggregates as bar panels (real thousands, mean across the twelve four-asset configurations).

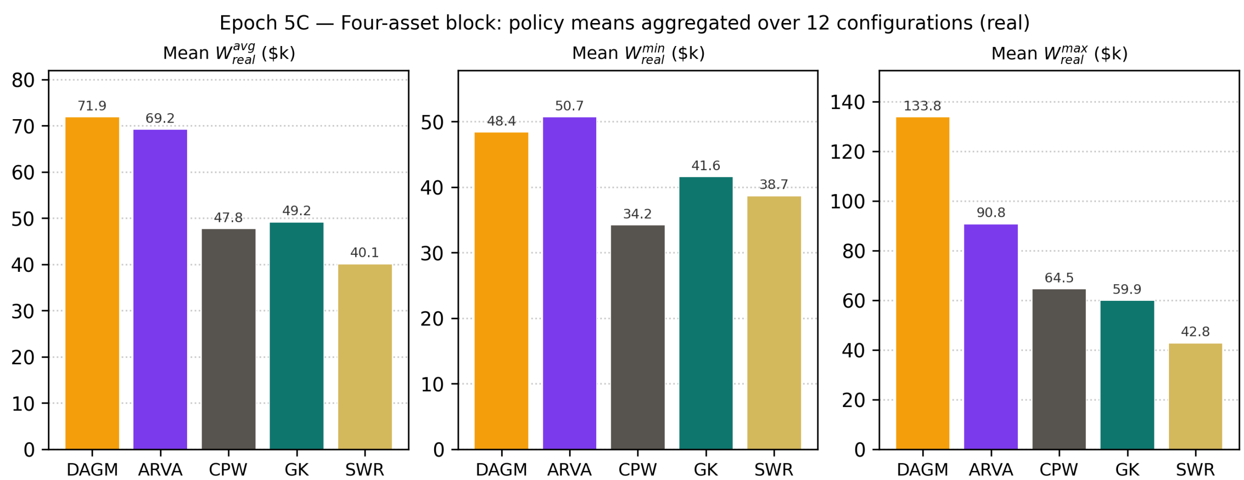

Figure 6: Epoch 5C — Same aggregates as Table 1: mean real withdrawals (thousands) by policy, averaged over the twelve four-asset configurations. Bar colours follow the Sandbox policy key; each panel uses its own vertical scale.

Portfolio-level interpretation

The four-asset extension also reveals several interesting patterns at the portfolio level.

First, there is a visible difference between the EUR and USD historical environments. In the EUR dataset, increasing equity exposure tends to produce relatively moderate changes in average withdrawals while gradually expanding upside potential. In the USD dataset, by contrast, higher equity allocations lead to a more pronounced increase in both average and maximum withdrawals, reflecting the stronger long-term equity premium in the U.S. historical series.

Second, the corridor framework displays a clear sensitivity to portfolio risk. As the equity share increases, DAGM's upside potential expands significantly while the minimum withdrawal level declines only moderately. The corridor therefore allows higher-growth portfolios to translate favorable market outcomes into higher withdrawals without abandoning governance discipline.

Third, the presence of additional defensive assets—specifically inflation-linked bonds and gold—appears to modestly stabilize the lower part of the withdrawal distribution in some configurations, particularly within the EUR dataset. However, this diversification effect does not materially change the relative ordering of withdrawal policies.

Overall, the results indicate that the corridor framework scales naturally with portfolio risk rather than depending on a specific asset allocation structure.

Tables 3–5 report the same twelve four-asset configurations with all five policies side by side, so the DAGM–ARVA trade-off—and gaps versus GK, SWR, and CPW—are visible configuration by configuration. Table 6 retains DAGM-only capital depletion rates (same definitions as §5.2).

Table 3 — Mean W_avg_real by configuration and policy

Configuration-level metrics are means across rolling start windows for each policy (same convention as §5.2). Values are in real units (rounded to integers). Tables 3–5 list all five policies so the cross-model comparison is visible run by run.

Block	EQ	ARVA	GK	SWR	CPW	DAGM
EUR 20y	20%	70,045	48,533	39,924	44,217	70,175
EUR 20y	40%	67,849	46,801	39,924	43,345	68,951
EUR 20y	60%	68,188	44,985	39,924	44,138	70,492
EUR 20y	80%	66,068	44,270	39,924	43,469	69,279
USD 20y	20%	66,993	48,105	40,209	42,410	67,107
USD 20y	40%	73,188	48,884	40,209	46,747	74,292
USD 20y	60%	79,850	51,499	40,209	51,695	82,665
USD 20y	80%	86,940	55,882	40,209	57,454	92,418
USD 30y	20%	50,779	41,012	40,214	38,624	51,415
USD 30y	40%	58,483	46,092	40,212	45,292	60,719
USD 30y	60%	66,738	52,600	40,212	53,094	71,326
USD 30y	80%	75,667	61,396	40,212	62,633	83,914

Table 4 — Mean W_min_real by configuration and policy

Block	EQ	ARVA	GK	SWR	CPW	DAGM
EUR 20y	20%	60,634	46,681	39,192	36,606	56,713
EUR 20y	40%	55,664	44,908	39,192	35,715	50,343
EUR 20y	60%	51,989	42,179	39,192	32,842	46,339
EUR 20y	80%	46,168	40,506	39,192	28,779	42,907
USD 20y	20%	52,809	44,851	38,509	32,831	52,636
USD 20y	40%	55,764	43,655	38,509	36,052	53,544
USD 20y	60%	58,140	42,355	38,509	37,699	53,329
USD 20y	80%	57,396	42,454	38,509	37,364	52,836
USD 30y	20%	37,298	35,024	38,350	27,588	39,078
USD 30y	40%	40,684	37,737	38,350	31,987	42,039
USD 30y	60%	45,538	39,119	38,350	36,178	44,977
USD 30y	80%	46,376	39,507	38,350	37,321	45,925

Table 5 — Mean W_max_real by configuration and policy

Block	EQ	ARVA	GK	SWR	CPW	DAGM
EUR 20y	20%	74,964	51,248	41,484	51,298	111,934

Block	EQ	ARVA	GK	SWR	CPW	DAGM
EUR 20y	40%	79,442	51,022	41,484	52,301	115,737
EUR 20y	60%	86,805	52,349	41,484	56,813	130,728
EUR 20y	80%	93,296	55,841	41,484	61,469	131,674
USD 20y	20%	80,748	52,772	43,154	53,205	104,876
USD 20y	40%	92,036	55,584	43,154	59,473	129,756
USD 20y	60%	104,933	62,866	43,154	68,357	160,735
USD 20y	80%	124,892	73,649	43,154	84,942	192,979
USD 30y	20%	64,947	52,680	43,823	53,232	88,970
USD 30y	40%	77,706	57,150	43,823	61,677	112,226
USD 30y	60%	92,804	67,668	43,823	73,371	146,046
USD 30y	80%	116,946	86,455	43,823	98,261	180,366

Table 6 — DAGM capital depletion (four-asset block)

Fraction of rolling windows on the DAGM track only (see §5.2 notes). **Early depl.** = early plan failure (`dep_prem`; capital depleted before completing T years, or before year T on a full T-year window). **Last-yr depl.** = terminal depletion (`dep_last`; depleted only after a full T-year run).

Block	EQ	Early depl.	Last-yr depl.
EUR 20y	20%	0.00	0.00
EUR 20y	40%	0.00	0.00
EUR 20y	60%	0.00	0.00
EUR 20y	80%	0.00	0.00
USD 20y	20%	0.00	0.05
USD 20y	40%	0.00	0.00
USD 20y	60%	0.00	0.00
USD 20y	80%	0.00	0.00
USD 30y	20%	0.00	0.01
USD 30y	40%	0.00	0.00
USD 30y	60%	0.00	0.00
USD 30y	80%	0.00	0.00

Across Tables 3–5, DAGM scales with portfolio composition in the same way as in the two-asset epochs: in EUR configurations, higher equity exposure mainly lifts the upper tail; in USD configurations, it raises both averages and upside. ARVA's relative strength on W_min_real appears row by row (Table 4), while DAGM's lead on W_avg_real and W_max_real is visible against every static reference in most configurations.

Conclusion

The four-asset extension confirms that the main findings of the paper remain robust under a broader asset universe. Introducing inflation-linked bonds and gold changes the characteristics of the underlying portfolios but does not alter the fundamental hierarchy of withdrawal models. DAGM continues to deliver the strongest performance on average and maximum withdrawals, while ARVA retains a relative advantage on the minimum-withdrawal dimension. This stability is important for the interpretation of the framework. It indicates that the performance of DAGM is driven primarily by its governance structure rather than by a particular asset mix. In other words, the corridor framework functions as a portfolio-agnostic decision architecture capable of combining the adaptive properties of dynamic withdrawal rules with the stability of bounded policy guidance across a wide range of portfolio environments. The four-asset experiment therefore strengthens the central claim of the paper: DAGM operates as a governance layer capable of extracting the strengths of existing withdrawal models under diverse portfolio conditions.

5.6 Depth of Path Analysis: Thousand-Path Resampling and Adversarial Market Tape

5.6.1 Purpose: From Window Count to Path Depth

A conventional rolling design samples on the order of seventy overlapping 30-year cohorts per configuration. That window count is adequate for level comparisons on realized history, but it leaves two questions under-explored:

1. How much depends on the particular ordering of regimes? A few dozen windows cannot exhaust the combinatorics of multi-decade sequence risk.
2. What happens when the path is not sampled from history at all, but forced through a pre-arranged chain of severe bear and stress episodes?

This section increases path depth along two complementary axes. We keep the same engine, policies, initial capital, 30-year horizon, and USD menus as in §5.2–§5.5. We change only how return sequences are supplied:

Axis	Construction	Effective depth
A. Block bootstrap	Resample 5-year blocks from `USD_1928_2025`	1,000 independent paths per menu
B. Möbius stress tape	Roll 30-year windows on a fixed 50-year cyclic tape of concatenated crisis blocks	50 windows on an adversarial sequence

We study two allocation menus: 60/40 equity–bond (literature benchmark) and 50/30/10/10 equity–bond–linker–gold. Summary metrics follow one aggregation rule throughout: real withdrawals deflated from cohort start; across paths or windows, W_{avg} is the mean of unit averages, W_{min} the minimum of unit minima, W_{max} the maximum of unit maxima. **Early depletion** (`dep_prem`): capital is depleted while the simulation executed fewer than the planned horizon $T = 30$ years—truncated paths count (e.g. SWR on the Möbius tape stopping near year 18). **Terminal depletion** (`dep_last`): capital is depleted after a full T-year run. The tables below use the same depletion definitions as §5.2 and Appendix A.6.

5.6.2 Regime A: Thousand-Path Block Bootstrap

Method. Annual equity, bond, gold (and linker where applicable), and CPI series are drawn from `USD_1928_2025`. Paths are generated by block bootstrap with block length five years (seed 42, 1,000 replications per menu). Each path is a full 30-year simulation of all policy tracks. Long-run chronological narrative is destroyed; short-block joint dynamics of returns and inflation are preserved.

Question answered. *If history were replayed with reshuffled macro segments—many more times than rolling windows allow—does the corridor layer keep paths solvent while static rules do not?*

5.6.3 Möbius Stress Tape on an Extreme Bear-Alternating Sequence

Method. We construct the Möbius stress tape dataset `USD_STRESS_TAPE_10x5_v1`: ten blocks of five years (50 tape positions) arranged as a cyclic Möbius strip of macro stress. Six blocks lift realized crisis episodes (e.g. 1929–33, 1973–77, 2008–12); four are synthetic scenarios (hyperinflation trap, Japan-style stagnation, flash crash, correlated stagflation). The tape is not a forecast; it is a deliberate alternation of bear and tail-risk environments in fixed order. We run 50 rolling 30-year windows on tape positions 1–50. Block definitions and sources are documented in the appendix supplement.

Question answered. *When the plan must march through a chained sequence of harsh bears and stress regimes—without the relief of reshuffling—where does rigid spending fail and where does governance compress consumption instead?*

Withdrawal levels on the Möbius stress tape are not comparable to Regime A; only within-tape rankings and depletion counts are meaningful.

5.6.4 Results

5.6.4.1 Table 6.6a — Block bootstrap (1,000 paths, 5-year blocks, seed 42) Withdrawal levels in thousands of real units (rounded). Same numeraire and horizon as §5.2.

Menu	Policy	W_{avg} (k)	Worst-path W_{min} (k)	Early depletion	Terminal depletion
60/40	DAGM	72.1	9.7	0 / 1,000	26 / 1,000
	DARVA	66.9	7.9	0 / 1,000	1,000 / 1,000
	GK	53.1	12.2	2 / 1,000	3 / 1,000
	SWR	35.9	11.1	32 / 1,000	7 / 1,000
	CPW	53.5	7.1	0 / 1,000	0 / 1,000
50/30/10/10	DAGM	66.9	17.3	0 / 1,000	13 / 1,000
	DARVA	63.3	12.9	0 / 1,000	1,000 / 1,000
	GK	49.3	13.5	0 / 1,000	1 / 1,000
	SWR	35.9	11.1	21 / 1,000	3 / 1,000
	CPW	49.6	10.9	0 / 1,000	0 / 1,000

DAGM path-level medians of W_{avg}: 60/40 p50 65.6k (p10 43.4k); 50/30/10/10 p50 63.2k (p10 44.9k). Appendix A.6 gives full policy rows and percentiles.

Increasing path count from order 10^2 (rolling windows) to 10^3 (bootstrap) sharpens inference on tail solvency: SWR accumulates early depletion on a non-negligible share of paths; DAGM records zero early depletion on both menus. Dynamic ARVA again shows terminal depletion on every path—its amortization anchor—not a uniform early-collapse pattern. On the four-asset menu, DAGM's mean W_{avg} is slightly below 60/40, but the worst-path floor rises (17.3k vs 9.7k) and terminal depletion falls (13 vs 26 per thousand), indicating more stable governed outcomes under reshuffled history when diversification is active.

5.6.4.2 Table 6.6b — Möbius stress tape on an extreme bear-alternating sequence (50 windows) Same units as Table 6.6a. **Do not compare levels to Table 6.6a** (Möbius stress tape vs reshuffled history). **Early depletion** = capital exhausted before completing the planned 30-year horizon (same rule as block-bootstrap MC).

Menu	Policy	W_{avg} (k)	Worst-window W_{min} (k)	Early depletion	Terminal depletion
60/40	DAGM	18.2	1.7	0 / 50	0 / 50
	DARVA	18.1	1.7	0 / 50	0 / 50
	GK	19.0	1.6	0 / 50	0 / 50
	SWR	40.3	31.4	50 / 50	0 / 50
	CPW	17.6	1.8	0 / 50	0 / 50
50/30/10/10	DAGM	20.1	3.3	0 / 50	0 / 50
	DARVA	19.8	3.3	0 / 50	0 / 50
	GK	21.2	3.1	0 / 50	0 / 50
	SWR	40.1	34.8	50 / 50	0 / 50
	CPW	20.2	3.4	0 / 50	0 / 50

SWR on the Möbius stress tape: 50 / 50 early depletion — CPI-indexed 4% rule, mean executed span 18 years before capital exhaustion; W_{avg} averages only the executed prefix. Appendix A.6 reports the same flags.

On 60/40, governed policies (DAGM, DARVA, GK, CPW) compress mean real spending to roughly 18–19k—far below bootstrap means on the same menu—while recording no early or terminal depletion on any window. SWR records 50 / 50 early depletion: higher mean W_{avg} on the truncated prefix, but the plan does not complete 30 years on this Möbius stress tape design. Substantially all DAGM years on 60/40 fall below a 12k real comfort threshold: consumption is sacrificed to keep the plan executable, not hidden behind a success label.

On 50/30/10/10, the same pattern holds after correcting cyclic-tape portfolio math for NaN linker rows (weights renormalized when linker is missing on a tape year): zero early depletion for DAGM and peers on all 50 completed windows, with mean W_{avg} near 20k versus 67k under bootstrap on the same menu. The four-asset row is not a diversification "win" on the Möbius stress tape; it documents how linker/gold weights interact with the adversarial sequence once the engine runs full horizons.

5.6.5 Synthesis

This chapter deepens the empirical analysis by moving from window-limited historical coverage to 1,000 bootstrap paths and, separately, to the Möbius stress tape on an extreme bear-alternating sequence. Together they show:

1. **Resampling depth:** With two orders of magnitude more paths than rolling panels typically provide, governed withdrawal remains early-solvent on both menus while static SWR does not; the four-asset menu improves tail floors under reshuffling.
2. **Sequential stress depth:** On the fixed Möbius stress tape, governed and guardrail policies compress spending and preserve zero early depletion on both menus; SWR shows 50 / 50 early depletion (plan truncated 18 years, CPI-indexed), aligned with bootstrap's early-ruin reading.
3. **Regime discipline:** Bootstrap levels and Möbius stress-tape levels must not be merged into one table; each regime answers its own question.

Full numerical exports: Appendix A.6.

5.7 Sequence of return risk (glide path vs. flat allocation)

Many households reduce equity exposure gradually into retirement rather than jumping immediately to a long-run strategic weight. A glide path changes when equity risk is borne and may interact with sequence-of-returns risk and with dynamic withdrawal rules. This section isolates a narrow but policy-relevant comparison: conditional on ending at the same equity target, does a glide from 20% equity to that target change real withdrawal outcomes relative to holding a flat allocation at the target throughout?

Design. For each terminal equity target in {40%, 60%, 80%}, we pair a glide experiment with a matched flat experiment at that target. The two arms share the same dataset, horizon, inflation treatment, initial capital, and policy parameters; only the return path to the terminal mix differs. Policies include dynamic ARVA and corridor-governed DAGM (wealth-dependent rules); static SWR coincides across arms by construction, so we do not emphasize null differences for that benchmark.

Metrics (this section). To describe glide vs. flat on realized paths, we report withdrawals in real terms. Let each rolling start window produce within-window summaries; we then aggregate across windows as follows: W_{avg} is the mean across windows of each window's average real withdrawal; W_{min} is the minimum across windows of each window's minimum real withdrawal (worst-window floor); W_{max} is the maximum across windows of each window's maximum real withdrawal (best-window upside). This "tail-across-windows" summary differs from the convention in the §5.2 note (mean of within-window minima and maxima) used for the level tables in §5.2–5.5. The two conventions answer different questions: level tables emphasize typical within-window profiles; here we emphasize how glide reallocates risk across which historical start years look worst or best for floors and upside.

Implementation. Historical rolling-window simulations use the U.S. sample (30-year horizons) and the EUR sample (20-year horizons). For EUR we vary years-to-target (5, 10, and 20). For the headline aggregate in Table 2, the USD block uses the 5-year glide-to-target specification with three terminal targets. **Appendix A.3** reports the full cell-by-cell percentage differences for every glide speed and terminal target (nine EUR rows; nine USD rows), including 10y and 20y USD glides omitted from the headline mean.

Table 1 — Glide vs. flat experiment: sample coverage

Sample	Dataset	Horizon (years)	Rolling windows
EUR	EUR_1999_2025	20	8
USD	USD_1928_2025	30	69

Table 2 reports the arithmetic mean of $100 \times (\text{glide} - \text{flat})/\text{flat}$ (percent), where glide and flat refer to the paired experiments at the same terminal target. For EUR, the mean is taken over nine cells (three targets $\times$ three glide speeds). For USD, the mean is taken over three cells (5y glide to target, three targets).

Table 2 — Mean percentage difference (glide vs. flat), dynamic ARVA and DAGM

Region	Policy	Mean $\Delta\%$ W_{avg}	Mean $\Delta\%$ W_{min}	Mean $\Delta\%$ W_{max}
EUR	Dynamic ARVA	5.0	27.4	−8.6
EUR	DAGM	2.7	23.3	−3.2
USD	Dynamic ARVA	1.1	5.5	10.6
USD	DAGM	−2.4	−2.4	8.9

Results. On EUR, headline mean glide-minus-flat $\Delta\%$ is strongly positive for worst-window W_{min} under both dynamic ARVA and corridor-governed DAGM (Table 2 reports 27.4% and 23.3%), positive also for mean W_{avg} for both rows (larger under ARVA), and negative on average for best-window upside W_{max} under both policies. On USD, dynamic ARVA shows positive mean $\Delta\%$ across all three aggregates; DAGM shows negative $\Delta\%$ for W_{avg} and W_{min} and positive $\Delta\%$ for W_{max}. Glide therefore heterogeneously reallocates risk across regions, policies, and tail margins rather than supporting a single uniform ranking.

Glide speed (EUR). Stretching the transition to the same terminal target moves the glide experiment. For a 60% terminal equity target, dynamic ARVA shows approximate percentage glide-minus-flat changes in worst-window W_{min} of about +18% (5y glide), +24% (10y), and +36% (20y), while the corresponding changes in best-window W_{max} move from about −6% to −16% in this three-point comparison. The complete EUR grid appears in **Appendix A.3**; qualitatively, longer transitions can deepen left-tail gains while changing upside at the window-max margin.

Interpretation. Matched-terminal-equity glide paths materially reshape withdrawal summaries even when spending is dynamically governed. In these EUR headline cells (post–Epoch 6), glide-minus-flat W_{min} averages are large and positive for both dynamic ARVA and DAGM, while W_{avg} and W_{max} follow the Table 2 pattern above—consistent with allocation paths that temporarily emphasize defensive portfolios improving worst-start floors alongside governance. USD remains heterogeneous, especially under DAGM. Welfare conclusions remain preference-dependent, but narratives that portrayed glide mainly as disposable once withdrawal governance is active—because floor gains clustered on ARVA and far less on DAGM—fit the EUR row of Table 2 less well now: glide and corridor-style rules reinforce each other there on worst-window floors, whereas USD shows little uniform glide benefit at this headline aggregation.

Numerical results in this section are recomputed from fresh Epoch 6 historical rolling simulations aligned with the current engine (including current robustness-badge logic).

6 Lifecycle horizon

In fixed mode, the plan's economic horizon is calendar-anchored (`end_year`). In lifecycle mode, the engine computes a residual horizon T_{run} from official survival tables, the client's age and sex at each review, and a policy threshold on cumulative mortality within T years (`mortality_confidence` / `prob_dead` in product JSON). The corridor end and DARVA inputs use this effective horizon so T can move year-on-year as the client ages, unlike a single fixed calendar span chosen once.

6.1 Robustness baseline under lifecycle

The engine uses a single robustness baseline anchored at plan start and initial capital, with the effective corridor end updated each review; for lifecycle plans, the corridor end follows the demographic T_{run} schedule.

6.2 Twin experiment: lifecycle vs. fixed matched on T

We compare two plans with the same initial capital, same 60/40 allocation, and the same simulated length T (full horizon, not a 20-year prefix): a lifecycle plan and a fixed plan whose calendar horizon is set so that, at the first rolling start, its residual span equals the initial T_{run} from the mortality rule (matched fixed). Withdrawals are evaluated under `corridor_governed_dynamic` (DAGM). Historical returns use **USD** (`USD_1928_2025`). Demographics are aligned with that market: the lifecycle arm uses the U.S. Social Security Administration period life table for 2022 (2025 Trustees Report; Social Security area population), implemented in the engine as `us_ssa_2022_period_tr2025_mf`. The mortality threshold is `mortality_confidence` = 90 (probability-of-death interpretation in engine JSON, consistent with product defaults).

Design. Rolling windows with step 1; anchor review date 1928-12-31 for the first start; male cohorts with completed age in 1928 of approximately 50, 55, 65 years (birth dates 1878-06-15, 1873-06-15, 1863-06-15). Initial T_{run} from the SSA table at that anchor differs from the earlier Italian ISTAT illustration, so T and rolling coverage change accordingly. Simulation length T equals `fixed_horizon_years` for the fixed twin. The last admissible start year is set so that the full T years fit inside 1928–2025 and completed age at the end of the horizon stays within the table's operational range.

Table 7.1 — Twin experiment design (USD historical Lab, SSA 2022 mortality)

Age profile (~ at start 1928)	T (years)	`fixed_horizon_years`	Rolling windows	Start-year range
~50	43	43	18	1928–1945
~55	39	39	18	1928–1945
~65	29	29	18	1928–1945

Table 7.2 — Means across all windows (`corridor_governed_dynamic`)

Chart versions of the two blocks below appear in **Appendix C** of the supplement (`DAGM_v2.1_appendix_supplement.pdf`). Real withdrawals use the Lab's historical deflator. "Mean terminal wealth" is the mean of end-of-horizon real wealth, one value per rolling window (last simulated year). % column: fixed relative to lifecycle for W^{avg}; lifecycle relative to fixed for terminal wealth.

Age ~	T	Mean real W^{avg} — LC	Mean real W^{avg} — FX	FX vs LC %	Mean real TW — LC	Mean real TW — FX	LC vs FX %
~50	43	50 356	54 547	+8.3%	115 930	39 073	+196.7%
~55	39	52 989	58 306	+10.0%	148 384	60 629	+144.7%
~65	29	56 724	66 199	+16.7%	197 586	106 996	+84.7%

On average across windows, fixed shows higher mean real withdrawals than lifecycle (+8–17%); the gap widens as T shortens (older profile at start). Lifecycle shows substantially higher real terminal wealth at T_0 in all three profiles (+85–197%). The correct reading is conditional on survival: terminal wealth is observed at T_0 years with the client alive. The fixed plan anchors to a calendar end date and ARVA converges toward full depletion there by design; whatever residual remains is incidental to the market path. The lifecycle plan, finding the client alive at T_0, holds a positive residual survival time and structurally reserves capital for those remaining years—not as a bequest, but as the buffer the engine must retain to cover post-T_0 withdrawals. In historically adverse paths (e.g., stagflation windows), insufficient real returns prevent lifecycle from accumulating that buffer and the gap narrows or reverses.

Table 7.3 — Start year 1928, last year of plan (after T years), real `terminal_wealth` (corridor policy)

Age ~	T	Real final wealth — LC	Real final wealth — FX	LC / FX ratio
~50	43	371 616	81 550	4.6×
~55	39	426 928	104 445	4.1×
~65	29	430 609	74 495	5.8×

For the 1928 window, lifecycle ends with 4–6× more real wealth than fixed. The mechanism is visible in the final years: at age 90–92, the fixed corridor reaches $T_{\text{remaining}} = 1$–3 and ARVA forces withdrawal rates of 20–60% of remaining capital, converging on full depletion. The lifecycle corridor, by contrast, still shows a residual survival time of 8–9 years for a client alive at that age at $p = 90\%$—the corridor end has not arrived and the engine withdraws at roughly 10–12% of capital. The higher terminal wealth is the capital reserved for those post-T_0 years: not a windfall, but a structural consequence of a plan that has not yet reached its survival boundary.

Source. Rolling-window twin inputs and summary exports are documented in the appendix supplement. Mortality uses SSA 2022 period tables with a pre-computed T_{run} schedule at each review.

6.3 Twin experiment (EUR): older retirees at the anchor year

To complement the USD design (ages ~50–65 at start 1928), we run the same matched-twin protocol on EUR returns (`EUR_1999_2025`) with male cohorts whose completed age in the first calendar year of the rolling window is approximately 75, 80, and 85—implemented by birth dates 1924-06-15, 1919-06-15, and 1914-06-15 when the first start is 1999. The mortality threshold `mortality_confidence` is varied at $p \in \{85\%, 90\%, 95\%\}$ (survival mass within T years). Initial T_{run} from the ISTAT table bundled in the engine at 1999-12-31 depends on p and cohort (see Table 7.4). Simulation length equals T for both twins; the last admissible start year is set so that the simulated horizon fits inside 1999–2025 and completed age at the end of the horizon stays within the table's operational range.

Comparability. Within each $(\text{age profile}, p)$ pair, lifecycle and fixed share the same T and same rolling windows. Across p, T and the number of windows change: absolute withdrawal levels are not read as a causal "effect of p" on identical historical paths; the sweep shows sensitivity of the twin gap and levels to the demographic calibration jointly with the EUR window structure.

Table 7.4 — Twin experiment design (EUR historical Lab, anchor age in 1999)

Age (~ at start 1999)	p	T (years)	Rolling windows	Start-year range
~75	85%	20	8	1999–2006
~75	90%	21	7	1999–2005
~75	95%	23	5	1999–2003
~80	85%	15	13	1999–2011
~80	90%	17	11	1999–2009
~80	95%	19	9	1999–2007
~85	85%	11	16	1999–2014
~85	90%	13	14	1999–2012
~85	95%	14	13	1999–2011

Table 7.5 — Means across all windows (`corridor_governed_dynamic`)

A chart of the fixed − lifecycle gap in mean real W^{avg} (nine cells) is in **Appendix C** of the supplement. Mean real W^{avg} is the mean of window-mean real withdrawals; mean real TW is the mean of real terminal wealth. % column: fixed relative to lifecycle for W^{avg}; lifecycle relative to fixed for TW.

Age ~	p	T	Mean real W^{avg} — LC	Mean real W^{avg} — FX	FX vs LC %	Mean real TW — LC	Mean real TW — FX	LC vs FX %
~75	85%	20	59 656	66 941	+12.2%	182 207	72 016	+153.0%
~75	90%	21	57 434	64 921	+13.0%	214 235	68 830	+211.2%
~75	95%	23	51 492	58 065	+12.8%	213 830	56 674	+277.3%
~80	85%	15	75 643	85 650	+13.2%	189 164	95 681	+97.7%
~80	90%	17	68 732	77 382	+12.6%	174 766	84 343	+107.2%
~80	95%	19	61 660	69 192	+12.2%	175 735	72 569	+142.2%
~85	85%	11	93 693	106 809	+14.0%	212 536	125 954	+68.7%
~85	90%	13	83 776	94 716	+13.1%	185 238	110 489	+67.7%
~85	95%	14	77 191	89 112	+15.4%	224 202	100 243	+123.7%

The same qualitative pattern as in §6.2 appears within each p: fixed shows higher mean real withdrawals than lifecycle (+12–15% across all cells), while lifecycle shows substantially higher real terminal wealth at T_0 (+68–277%). The terminal wealth gap grows at longer T (younger cohort at $p = 95\%$): a longer conservative-spending phase compounds into a larger capital reserve at T_0. The interpretation follows §6.2—terminal wealth is measured conditional on the client's survival; lifecycle's residual is the capital reserved for the remaining survival time, not incidental accumulation. The W^{avg} gap is broadly stable across p within each age profile.

Table 7.6 — Start year 1999, last year of plan, real `terminal_wealth` (corridor policy)

Age ~	p	T	Real final wealth — LC	Real final wealth — FX	LC / FX ratio
~75	85%	20	238 568	53 094	4.5×
~75	90%	21	274 924	55 280	5.0×
~75	95%	23	294 756	56 399	5.2×
~80	85%	15	252 086	51 352	4.9×
~80	90%	17	253 303	53 766	4.7×
~80	95%	19	246 577	48 244	5.1×
~85	85%	11	294 824	69 654	4.2×
~85	90%	13	238 397	57 446	4.1×
~85	95%	14	264 458	53 574	4.9×

For the 1999 start window, lifecycle ends with 4–5× more real wealth than fixed at each p, consistent with §6.2. As in the USD case, the divergence is structural: the fixed corridor reaches $T_{\text{remaining}} = 0$ at T_0 and ARVA targets full depletion; the lifecycle corridor still holds a positive residual survival time for a client alive at T_0 and reserves capital accordingly. **Source.** EUR twin experiment inputs and summary exports are documented in the appendix supplement. Mortality uses Italian ISTAT tables bundled in the engine with the same pre-computed T_{run} schedule convention as §6.2.

6.4 Couple plans: the annual-rebalance architecture

6.4.1 Motivation

Extending a lifecycle plan to cover a couple introduces two distinct challenges. First, the plan must anchor its horizon to the joint second-death distribution — the p90 quantile of $\max(T_{\text{primary}}, T_{\text{spouse}})$ — which is longer than either individual's horizon and permanently depresses current withdrawals: this is the *spousal coverage cost*. For the reference M65+F65 pair (SSA 2022, p90), anchoring to the joint horizon reduces mean real W^{avg} by approximately 3% relative to the single M65 plan; for younger spouses (M65+F55), the reduction reaches 10%. Second, at the moment of first death, the transition to the survivor's individual lifecycle produces a mechanical withdrawal jump whose sign and magnitude are fully determined by the ratio of the old joint horizon to the survivor's new individual horizon — a ratio close to 2 in the standard single-plan architecture, generating a factor-of-two step in withdrawals that is difficult to reconcile with observed household expenditure after bereavement (empirically, surviving households spend roughly 65–80% of their pre-death couple level, not 200%).

The *annual-rebalance* architecture addresses both challenges simultaneously.

6.4.2 Architecture

The key idea is to replace the fixed joint-second-death anchor with a dynamically blended effective horizon $T_{\text{eff}}(t)$, recomputed at each annual review from the two members' current individual mortality residuals:

$$\frac{1}{a_{\text{eff}}(t)} = \frac{1-w}{a(T_1(t))} + \frac{w}{a(T_2(t))},$$

where $T_1(t)$ and $T_2(t)$ are the individual $p = 90\%$ residual survival quantiles of primary and spouse at year t, $a(T) = d/(1-(1+d)^{-T})$ is the annuity factor, and $w \in (0,1)$ is the demographic weight reflecting the relative importance of the survivor's longevity. The inverse $T_{\text{eff}}(t) = a^{-1}(a_{\text{eff}}(t))$ is the fixed-horizon equivalent that produces the same annuity factor as the blend. The full DAGM governance engine runs on a fixed-horizon plan with $T_{\text{remaining}} = T_{\text{eff}}(t)$ updated annually.

Phase structure. Each plan divides into two continuous phases that share the same capital and governance state:

- **Phase A** ($t < T_{\text{death}}$): both alive; withdrawals governed by $T_{\text{eff}}(t)$.
- **Phase B** ($t \geq T_{\text{death}}$): survivor alone; T_{run} switches to the survivor's individual lifecycle. Capital carried over exactly from Phase A; governance badge — including the GK anchor and badge state — carried forward continuously with no re-initialisation.

Advisory parameters. Two inputs are communicated at plan inception: the demographic weight w (default 0.70, reflecting the probability-weighted importance of the survivor's horizon) and the lifestyle step-down $w_{\text{lifestyle}} \in (0,1)$, specifying the advisor's expectation of the survivor's spending relative to the couple's pre-death level. The lifestyle parameter is declarative: it does not alter the capital trajectory or governance state but is recorded in the plan document and communicated to both members ex-ante. The actuarial expected first-death year $E[T_{\text{first death}}]$ serves as the planning horizon for Phase A, computed once from the SSA tables at inception.

6.4.3 Results (USD 1928–2025)

Design. Seven M65+spouse configurations spanning the demographic matrix plus the M65 single baseline. Rolling windows 1928–1993 (couple) and 1928–1996 (single). Parameters: $w = 0.70$, $w_{\text{lifestyle}} = 0.70$, $T_{\text{death}} = E[T_{\text{first death}}]$ (actuarial). Policy: `corridor_governed_dynamic` (full DAGM). SSA 2022 period tables, USD 1928–2025, $1M initial capital, 60/40 allocation.

Table 7.10 — Annual-rebalance couple plan: initial withdrawal, full-plan average, withdrawal jump at first death

W_0^{med}: median initial Phase A withdrawal (both alive). Coverage cost: couple W_0^{med} vs. M65 single (58 391). $W_{\text{full}}^{\text{med}}$: median real W^{avg} over the full plan (Phase A + Phase B combined). $\bar{W}_{\text{jump}}$: mean percentage change in nominal withdrawal at first death (Phase A last year to Phase B first year; GK state carried forward). T_{death}: actuarial expected first-death year.

Configuration	n	W_0^{med}	Coverage cost	$W_{\text{full}}^{\text{med}}$	$\bar{W}_{\text{jump}}$	T_{death}
M65 single	69	58 391	—	62 110	—	—
M65 + F70	66	58 927	+0.9%	66 335	+44.1%	13

Configuration	n	W_0^{med}	Coverage cost	$W_{\text{full}}^{\text{med}}$	$\bar{W}_{\text{jump}}$	T_{death}
M65 + M65	66	58 391	0.0%	64 224	+35.9%	13
M65 + F65	66	57 897	−0.8%	62 291	+31.2%	14
M65 + M60	65	57 018	−2.3%	55 732	+25.1%	14
M65 + F60	63	56 261	−3.7%	53 773	+25.0%	15
M65 + M55	60	55 606	−4.8%	51 469	+22.2%	15
M65 + F55	58	55 312	−5.3%	49 395	+19.2%	16

Supplementary bar charts for coverage cost and mean W_{jump} by configuration appear in **Appendix C.3** of the supplement.

6.4.4 Interpretation

Near-elimination of Phase A coverage cost. For same-age or older spouses (M65+F65, M65+M65, M65+F70), the initial withdrawal is within 1% of the single baseline — essentially no penalty during the go-go years. For younger spouses, the cost scales monotonically with the demographic gap: −2.3% for a 5-year age difference (M65+M60), −5.3% for a 10-year gap with female longevity (M65+F55). The mechanism: at $w = 0.70$ the blended T_{eff} at inception lies substantially below the joint second-death horizon, pulling the initial ARVA annuity factor toward the single-member value.

Positive, moderate W_jump at first death. The annual-rebalance architecture produces a demographic W_jump of +19% to +44% across all configurations — positive and far smaller than the factor-of-two seen in the fixed joint-horizon approach. Three forces drive the sign: Phase A conserves capital under the blended horizon; at transition the survivor inherits the full accumulated pool; and the GK prosperity ratchet carried forward from Phase A provides an elevated governance baseline. The advisor's stated lifestyle expectation ($w_{\text{lifestyle}} = 0.70$, i.e., −30%) frames the behavioral target; the engine's realized withdrawal typically exceeds it in good-market windows, delivering additional income or a capital buffer for the survivor's longevity tail.

Full-plan real average meets or exceeds single. For M65+F65, M65+F70, and M65+M65, the median real W^{avg} over the full plan (both phases combined) equals or exceeds the single baseline of 62 110: the couple does not sacrifice long-run income relative to a single plan. For younger-spouse configurations the full-plan average is lower, reflecting more Phase A conservatism and a shorter actuarial Phase B.

Three independent parameters. The architecture separates three previously conflated effects into orthogonal controls:

1. **Spousal coverage cost** — the intertemporal price of joint longevity coverage, reduced here to 0–5% on initial withdrawals; controlled by w.
2. **Demographic W_jump** — the mechanical capital-release at first death, driven by the T_eff dynamics and GK continuity; controlled by w and the Phase A return path.
3. **Lifestyle step-down** ($w_{\text{lifestyle}}$) — the advisor's declarative expectation of survivor spending; communicated at plan inception, not imposed on the capital.

A plan that uses a single "survivor fraction" to manage all three simultaneously must trade off Phase A spending efficiency against survivor income adequacy. The annual-rebalance design avoids this by separating the demographic and behavioral layers.

6.4.5 Relation to the literature

Joint life annuities and the second-death loading. In actuarial pricing, the premium of a joint-life second-death annuity over a single-life annuity — the *joint-life loading* — compensates the insurer for the longer expected coverage period (Milevsky 2006; Yaari 1965). The spousal coverage cost in the standard joint-horizon approach is the exact historical analogue of this loading, measured on realized return paths rather than in closed form. The annual-rebalance architecture reduces this loading without eliminating the longevity guarantee: the blended horizon is shorter than the joint second-death quantile at inception, but $T_{\text{eff}}(t)$ continues to weight the survivor's individual residual life throughout Phase A, providing implicit joint coverage at a lower current cost.

Widow/widower household economics. Empirical household-finance studies consistently find that surviving households spend 65–80% of their pre-death couple expenditure (Blanchett 2012; Bernheim et al. 2003), not the factor-of-two implied by a full horizon switch at first death. The $w_{\text{lifestyle}}$ parameter is the plan's explicit interface with this behavioral reality: it allows the advisor to record the household's expected step-down at plan inception, creating a client-facing commitment that is auditable and revisable without altering the capital mechanics.

Deterministic-conditional horizon updating. The proposal that the planning horizon should move at each review is well established in probabilistic decumulation models (Pfau 2012; Chen and Milevsky 2003). The single-member framework here extends this idea: T_{run} is the p90 conditional-survival quantile updated annually — a deterministic-conditional rule requiring only a mortality table lookup, without Monte Carlo simulation. The annual-rebalance couple plan generalizes the same rule to a blended joint horizon: $T_{\text{eff}}(t)$ is updated each year from the two members' current ages, capturing the gradual demographic convergence as both spouses age rather than locking in a joint horizon at plan signature.

Source. Annual-rebalance couple runs on USD 1928–2025; $E[T_{\text{first death}}]$ from SSA 2022 period tables. Full exports are documented in the appendix supplement.

---

7 DAGM with Intentional Legacy Reserve

The sections above treat retirement income as the sole inter-temporal objective. This section introduces a **ring-fenced legacy reserve**: a fraction λ of total wealth set aside as an explicit bequest target, managed separately from the withdrawal engine.

7.1 Mechanism

At each review, total portfolio wealth W_t is partitioned into spendable capital and legacy reserve:

$$W_t = S_t + L_t, \quad S_t = \max(0,\, W_t - L_t)$$

The legacy reserve is **CPI-indexed from plan creation**, so its real value is constant at $\lambda \cdot C_0$:

$$L_t = \lambda \cdot C_0 \cdot \frac{\text{CPI}_t}{\text{CPI}_0}, \quad L_t^{\text{real}} = \lambda \cdot C_0$$

All withdrawal references — ARVA, GK, SWR, CPW — are computed on S_t, the spendable base. The withdrawal corridor, governance badges, and robustness diagnostics operate exclusively on S_t; L_t accumulates separately unless portfolio losses force $W_t < L_t$, partially impairing the reserve.

Terminal wealth at end of horizon includes both components:

$$W_T^{\text{real}} = \underbrace{\max(0,\, W_T^{\text{real}} - \lambda C_0)}_{\text{residual spendable}} + \underbrace{\min(W_T^{\text{real}},\, \lambda C_0)}_{\text{legacy realized}}$$

7.2 Empirical evidence

Setup. Rolling-window simulations on USD historical data (1928–2025, $n = 69$ windows of 30 years, $C_0 = 1{,}000{,}000$). λ is varied from 0 to 50% in 10 pp steps. Benchmark policies (ARVA, GK, SWR) are reported at $\lambda = 0$ on the full capital — their natural domain, since none carries a bequest mechanism. All figures are real annual withdrawals deflated to plan-start; terminal wealth figures are real. Level metrics $W_{\text{avg}}^{\text{real}}$, $W_{\text{min}}^{\text{real}}$, and $W_{\text{max}}^{\text{real}}$ follow the §5.2 convention: **means across rolling windows** of each window's within-path summary (not global extrema across all windows).

Table 8.1 — DAGM with legacy vs. benchmark policies (USD 60/40, $T = 30$, $n = 69$)

Policy / scenario	$W_{\text{avg}}^{\text{real}}$	$W_{\text{min}}^{\text{real}}$	$W_{\text{max}}^{\text{real}}$	Residual spendable	Total final
DAGM	**71,781**	42,804	143,661	90,862	**90,862**
DAGM $\lambda = 10\%$	**73,320**	40,038	162,102	48,028	**148,028**
DAGM $\lambda = 20\%$	**74,892**	37,093	183,421	7,201	**207,201**
DAGM $\lambda = 30\%$	**76,526**	33,471	204,853	(34,675)	**265,325**
DAGM $\lambda = 40\%$	**78,143**	29,166	226,643	(76,219)	**323,781**
DAGM $\lambda = 50\%$	**79,860**	24,469	249,043	(117,457)	**382,543**
ARVA	67,005	41,990	96,924	—	—
GK	53,489	38,410	69,611	1,098,110	**1,098,110**
SWR	40,221	38,350	43,823	1,789,910	**1,789,910**

Residual spendable = Total final $-\lambda C_0$. Values in parentheses: mean terminal wealth falls short of the real legacy target λC_0. ARVA terminal wealth = 0 by actuarial design. GK/SWR terminal wealth is a by-product of algorithmic conservatism, not a planned bequest. Appendix A.5b reports full tables for alternative asset allocations.

7.3 Three categories of terminal capital

The results in Table 8.1 reveal a structural distinction that is absent from standard withdrawal-rule comparisons.

DAGM $\lambda > 0$ — *intentional and quantified.* The bequest target λC_0 is declared at plan inception, CPI-indexed throughout, and ring-fenced from withdrawals. The trade-off is visible year by year: the governance corridor adjusts to $S_t = W_t - L_t$, so the client observes at each review exactly how much capital is reserved for heirs and how much governs spending. Notably, W_{avg} rises with λ in the long USD sample: the corridor applied to a smaller S_t extracts retirement income more efficiently per unit of spendable capital, so the income stream is not penalised by the bequest commitment.

GK / SWR — *accidental and unplanned.* Both policies leave substantial terminal capital—means of 1.1M and 1.8M respectively for USD 60/40—but as a structural consequence of conservative withdrawal formulas, not as a bequest plan. No portion is designated, communicated, or guaranteed. The realized amount depends entirely on the market path.

ARVA — *zero by design.* ARVA's actuarial amortization formula exhausts spendable capital over the horizon by construction. No bequest is possible, even in strongly favorable market environments.

7.4 Operating range

At $\lambda \leq 20\%$ the mean residual spendable is positive, indicating the legacy target is on average preserved with margin. At $\lambda \geq 30\%$ mean terminal wealth falls short of λC_0: in historically adverse 30-year windows, returns are insufficient to sustain full real legacy integrity alongside DAGM withdrawals. This is not a plan failure—W_{avg} continues to rise and terminal wealth at $\lambda = 30\%$ (265k) remains in the order of magnitude of the target (300k)—but it reflects the genuine actuarial tension between consumption and bequest at high ring-fence levels.

At $\lambda > 40\%$, mean within-window $W_{\text{min}}^{\text{real}}$ falls materially (29k at $\lambda = 40\%$, 24k at $\lambda = 50\%$ in Table 8.1) and some windows record capital depletion. **Values of λ above 40% are not operationally indicated for the main retirement plan.** Clients with strong bequest intent beyond that threshold are better served by a **dedicated inheritance portfolio** — a separate account with allocation and time horizon aligned to the heirs' objectives — leaving the DAGM plan to operate without reserve constraints.

---

8 Discussion and limitations

8.1 Conclusions

Retirement spending rules are evaluated here on a criterion that matches implementation reality: they must be executable year after year on a single realized path of wealth and horizon, without foresight of future returns. That lens distinguishes this work from exercises that summarize rules only through simulated distributions of outcomes. Within that frame, the paper develops Dynamic ARVA with a portfolio-consistent conversion anchor (DARVA), embeds it in a corridor together with other reference policies, and adds a transparent governance layer—DAGM—that maps capital position and remaining horizon into spending states (GREEN through SUPER RED) using a simple path-based robustness diagnostic.

The static reference models examined in §5.1 are not interchangeable: they encode different trade-offs among average consumption, downside floors, upside participation, and income variability. ARVA emerges as the strongest income-oriented amortization benchmark; GK and SWR lean toward discipline and conservatism; CPW embodies full capital responsiveness. That heterogeneity motivates not choosing a single "winner" rule, but combining signals—precisely what the corridor structure formalizes.

The empirical core (§5.2–§5.7) shows that DAGM reshapes realized withdrawal profiles relative to static rules. Against GK, SWR, and CPW, DAGM improves joint outcomes on average and maximum real withdrawals across the historical matrix while preserving competitive floors in most configurations; against ARVA, the comparison is nuanced: ARVA can retain a slight edge on minimum withdrawals in some runs, whereas DAGM more often raises average withdrawals and systematically preserves stronger upside in favorable paths. Epoch 3 adds that DAGM does not promise universal smoothness: it reallocates variability—reducing the incidence and cumulative burden of downward spending shocks (cut-year frequency, DWII) in every run versus ARVA, while total year-to-year dynamism is higher than ARVA in most configurations (21 / 25). That separation between downside stability and overall variability is central to interpreting the framework as governance rather than as a flat spending target.

Robustness exercises reinforce the same message. Longer horizons (Epoch 4) and richer asset sets including inflation-linked bonds and gold (Epoch 5) do not overturn the qualitative ordering: DAGM remains strong on level and upside metrics; ARVA remains the primary floor comparator; tail exceptions—late-stage depletion cliffs under extreme histories—appear as design boundaries to monitor rather than as refutations of average-case governance gains. The glide-path comparison (§5.7), after the Epoch 6 refresh, underscores that EUR glide can amplify worst-window W_{min} under both ARVA-style and corridor-governed withdrawals; allocation-based glides are not best read as interchangeable substitutes that governance fully replaces on that EUR margin. Conversely, headline USD cells mix signs across policies and tail summaries—so glide need not mechanically improve every scalar once the corridor is active on that sample.

Section 6 connects the spending machinery to horizon. When the residual horizon is demographic rather than calendar-fixed, the residual survival time can move at each review; the matched-twin experiments (§6.2–6.3) isolate the contrast between lifecycle and fixed plans that share the same initial T and simulation length. On average, the fixed twin shows higher mean real corridor withdrawals than the lifecycle twin, but at the cost of much lower terminal wealth. The correct reading is conditional on survival: at T_0 years the client is alive; the fixed plan, having depleted toward its calendar anchor by design, leaves little residual; the lifecycle plan, finding the client alive at T_0, holds a positive residual survival time and reserves capital structurally for those remaining years — not as a bequest but as a buffer for post-T_0 withdrawals. In adverse historical paths (stagflation windows), insufficient real returns prevent lifecycle from sustaining that buffer and the advantage narrows or reverses. The EUR extension (§6.3) varies mortality confidence (85%, 90%, 95%): within each (age, p) cell the qualitative twin pattern holds; across p, T and rolling coverage change, so levels must be read as sensitivity to calibration and window structure. §6.2 pairs SSA 2022 U.S. period mortality with USD returns; §6.3 pairs Italian ISTAT mortality with EUR returns — the two blocks are not interchangeable as a single population design. The structural discount from the longer demographic anchor remains by design and should not be removed.

§6.4 extends the framework to couple households via the annual-rebalance architecture (contribution (g)). The key result, across seven M65+spouse demographic pairs (USD 1928–2025): Phase A coverage cost — the reduction in initial withdrawals when a spouse is added — falls to 0–5%, compared to 1–10% under a fixed joint-horizon approach. The withdrawal jump at first death is positive and moderate (+19–44%), driven by the capital accumulated under the conservative blended horizon being reallocated to the survivor's shorter individual lifecycle, with governance badge state carried forward continuously. For same-age or older spouses (M65+F65, M65+M65, M65+F70), the full-plan real average meets or exceeds the single baseline — the couple plan does not sacrifice long-run income. The three phenomena that traditional approaches conflate — coverage cost, demographic W_jump, and lifestyle step-down — are made independent, explicitly parameterized plan inputs: the advisor communicates each one at plan inception, and all three are auditable on the realized path.

Legacy and the explicit life-style vs. bequest trade-off. Contribution **(d)** is not only definitional: at each annual review the framework recomputes spendable capital net of a CPI-indexed legacy reserve, then maps that state into corridor bounds, governance colour, and the withdrawal the household may select. That structure makes the tension between sustaining lifestyle spending along the realized path and honoring a bequest intention visible year by year, rather than burying it in a single initial withdrawal rate. **Appendix A.5** reports a historical sensitivity in which the initial legacy share   is raised from 0% to 70% on three rolling sleeves while the DAGM track is held fixed. Across configurations, mean within-window minimum real withdrawals fall as   rises, while mean terminal wealth (real) rises sharply—consistent with weaker cash-flow floors alongside larger residual balances when more capital is ring-fenced at plan open. Mean average withdrawals can move in opposite directions on the short EUR slice versus longer USD menus, so the tables should be read as illustrating the trade-off, not as recommending a   level.

Practical takeaway. DAGM is best understood as a transparent governance architecture: it combines amortization and guardrail ideas into a rule that households and advisers can audit on the realized path, while making explicit the trade-off between consumption efficiency, stability on the downside margin, and participation in good outcomes. It is not a proof of optimality under a social welfare criterion, nor a substitute for client-specific constraints, taxes, and non-portfolio income.

8.2 Limitations

- **Evidence type.** Results are historical rolling-window simulations on two regional samples. They describe how rules would have behaved on past paths; they do not deliver out-of-sample predictive probabilities of retirement "success."
- **Welfare.** The paper reports withdrawal levels and stability metrics, not a utility-based ranking of consumption paths. Gains on W^{avg} or cut-frequency can conflict with preferences over volatility or

legacy.
- **Implementation gap.** Lab/engine defaults simplify taxes, fees, liquidity, and discrete life events; live advice may require buffers and overrides not modeled here.
- **Governance calibration.** Structural badge parameters were fixed before the reported Lab matrix was executed; the exploratory search that set them was not independently reproducible from archived logs (see §4.4).
- **Tails.** A minority of extreme paths—especially very long horizons, bond-heavy allocations, or late-stage depletion—can produce large single-year cuts or cliff-like outcomes; DAGM improves average downside metrics but does not dominate ARVA on every window. **Appendix A.4** disaggregates exhaustion timing versus real-consumption shocks for USD 20/80 and gives a worked stagflation vignette. The reported Lab matrix does **not** include an automatic terminal-phase actuarial safeguard; bond-heavy sleeves near plan end should be monitored on remaining horizon and spendable capital, with adviser override where appropriate (§5.3).
- **Lifecycle twins.** §6.2 uses U.S. SSA period mortality with USD returns; §6.3 uses Italian ISTAT mortality with EUR returns. The two blocks are not interchangeable as a single population design.
- **Couple plans: EUR not usable.** The annual-rebalance couple plan (§6.4) is validated on the USD 1928–2025 dataset only. The EUR dataset covers 1999–2025 (27 years), which is shorter than the minimum horizon required for a two-phase M65+F65 plan (( )31 years); no rolling window of full length fits within the available history. By analogy with the structural consistency between USD and EUR observed in §6.2–6.3, the coverage-cost reduction mechanism is expected to generalise qualitatively, but empirical validation on EUR data requires either a longer return series or profiles with shorter joint horizons (e.g. M80+F80). In addition, the wider male–female longevity gap in ISTAT tables relative to SSA 2022 may produce larger coverage costs for female-spouse configurations; this differential has not been quantified.
- **Glide vs. flat (§5.7).** Aggregation conventions for min/max across windows differ from the §5.2 level tables; headline glide–flat numbers are sensitive to that choice and to glide speed (Appendix A.3).
- **Path depth (§5.6).** Block-bootstrap and **Möbius stress-tape** regimes in **Appendix A.6** (`A6_epoch12_validation_matrix.md`) are not probabilistic forecasts. Möbius stress-tape withdrawal levels are not comparable to bootstrap levels; interpret each regime on its own depletion counts and within-regime rankings.

8.3 Future work

- Extend the §6.4 couple framework to **stochastic first-death** timing and validate on **EUR/ISTAT** mortality to assess the effect of the wider gender gap on both W_jump and the spousal coverage cost (see §8.2 for the current data constraint).
- Extend $w_{\text{lifestyle}}$ to a **hard governance constraint**, allowing advisors to cap the Phase B withdrawal as an explicit fraction of the final Phase A level rather than a declarative expectation.
- Specify and test a **terminal-phase actuarial safeguard** (remaining-horizon and spendable-capital triggers) on bond-heavy configurations identified in Epoch 3 and **Appendix A.4**.

References

The list below is illustrative, not exhaustive: it connects the policy families in §§2–3 to widely cited practitioner and research sources on withdrawal rates, guardrails, valuations, and dynamic spending. Dynamic ARVA, DARVA, and DAGM are constructs defined for this framework and its software.

Bengen, W. P. (1994). Determining withdrawal rates using historical data. *Journal of Financial Planning*, 7(4), 14–24.

Cooley, P. L., Hubbard, C. M., & Walz, D. T. (1998). Retirement savings: Choosing a withdrawal rate that is sustainable. *AAII Journal* (February 1998).

Daga, A., Clarke, A. S., Pakula, D., & Bupp, J. (2021). Sustainable spending rates in turbulent markets. *Vanguard Research Note* (March 2021).

Finke, M., Pfau, W. D., & Blanchett, D. M. (2013). The 4% rule is not safe in a low-yield world. *Journal of Financial Planning*, 26(6), 46–55.

Guyton, J. T. (2004). Decision rules and portfolio management for retirees: Is the "safe" initial withdrawal rate too safe? *Journal of Financial Planning*, 17(10), 52–61.

Guyton, J. T., & Klinger, W. J. (2006). Decision rules and maximum initial withdrawal rates. *Journal of Financial Planning*, 19(11), 50–58.

Kitces, M. (2009). Dynamic asset allocation and safe withdrawal rates. *The Kitces Report* (April 2009). Kitces.com.

Kitces, M. (2014). The ratcheting safe withdrawal rate: A more complete approach to navigating the risk of retirement failure. *Nerd's Eye View* (blog). Kitces.com.

Milevsky, M. A., & Young, V. R. (2007). Annuitization and asset allocation. *Journal of Economic Dynamics and Control*, 31(9), 3138–3157.

Pfau, W. D. (2011). Safe withdrawal rates from a stock and bond portfolio: Assessing the classic 4% rule. *Journal of Financial Planning*, 24(2), 46–55.

Pfau, W. D., Blanchett, D. M., & Finke, M. (2014). Asset valuations and safe portfolio withdrawal rates. *Retirement Management Journal*, 4(1), 21–34.

Scott, J. S., Sharpe, W. F., & Watson, J. G. (2009). The 4% rule—at what price? *Journal of Investment Management*, 7(3), 29–39.

Sharkansky, S. (2025). The only other spending rule article you will ever need. *Financial Analysts Journal*. https://doi.org/10.1080/0015198X.2025.2541567

Waring, M. B., & Siegel, L. B. (2015). The only spending rule article you will ever need. *Financial Analysts Journal*, 71(4), 91–107.

Data availability

Simulation results for the configurations reported in this paper are available from the author upon reasonable request. Researchers interested in running the DAGM framework on additional configurations may contact the author to discuss feasibility.

Appendices (unified supplement PDF)

Extended empirical tables (**Appendix A**), the full dataset construction note (**Appendix B**), supplementary figures—glide path, lifecycle twins, and couple-plan charts (**Appendix C**, including **C.3** cited from §6.4), optional additional charts (**Appendix D**), and editorial source notes are published in a single file:

`DAGM_v2.1_appendix_supplement.pdf` (bundled with this release). Tables A.2–A.6 reproduce the extended empirical grids underlying §5.

Contents (section titles):

1. **Appendix A — Empirical supplements** — `A_empirical.md`: A.1 Validation and reproducibility; A.2 win-count tables (Epochs 2, 3, 5C); A.3 Glide vs. flat (full Epoch 6 grid, Tables A.5–A.6).
2. **Appendix A.4 — Bond-heavy USD pathology** — `A4_bond_heavy_pathology.md`; cited from §5 / §8.
3. **Appendix A.5 — Legacy share sensitivity (Epoch 11)** — `A5_legacy_lambda_sensitivity.md`; cited from **§8.1**.
4. **Appendix A.5b — DAGM with legacy vs. benchmark policies** — `A5b_legacy_policy_compare.md`; cited from **§7**.

5. **Appendix A.6 — Epoch 12bis path-depth (60/40 and 50/30/10/10; block MC + Möbius stress tape)** — `A6_epoch12_validation_matrix.md`; cited from **§5.6**. *(Not Table A.6 in §A.3 of `A_empirical.md`, which is the §5.7 glide grid.)*
6. **Appendix B — Historical market datasets (full construction)** — `B_datasets.md` (Sandbox Dataset Reference PDF).
7. **Appendix C — Supplementary figures** — `C_figures.md`: glide vs. flat (C.1); USD and EUR lifecycle twin charts (C.2); couple-plan coverage cost and first-death withdrawal jump (C.3, §6.4).
8. **Appendix D — Additional charts** — `D_additional_charts.md`: optional Lab figures.
9. **Source notes and change log** — `Z_source_changelog.md`.

The main paper PDFs do not duplicate these materials; cite the supplement when referencing extended grids or supplementary figures.