Shortfall Risk and Retirement Spending: An Efficient Frontier Approach

April 2026

Martin-D. Lacasse

Corresponding author
[8905 Ashford Gables Ct]
[Tampa, FL 33626]
[martin.d.lacasse@gmail.com]

Martin-D. Lacasse, Ph.D., is a former Associate Director in the Department of Mathematics at Johns Hopkins University and a retired Distinguished Research Scientist from ExxonMobil Research & Engineering Company.

John D. Leonard II

[2121 Stuart Ave]
[Richmond, VA 23220]
Professor Emeritus, Virginia Commonwealth University

John D. Leonard, Ph.D., is Professor Emeritus in the Department of Computer Science in the College of Engineering at Virginia Commonwealth University. He previously served as Executive Dean of the College. Prior to VCU, Dr. Leonard served as Associate Dean for the College of Engineering and held several tenured faculty positions at the Georgia Institute of Technology.

Shortfall Risk and Retirement Spending: An Efficient Frontier Approach

April 2026

Abstract

Probability of success (PoS) is the dominant risk metric in retirement income planning, yet it is fundamentally incomplete: it records only whether a plan fails, not by how much. A plan with a PoS of 98% could involve either catastrophic portfolio exhaustion or merely a modest spending reduction in a few scenarios — two very different risk profiles that PoS cannot distinguish. This article introduces a complementary framework that makes the magnitude of potential shortfalls directly visible. Drawing on an analogy with the Markowitz mean-variance portfolio theory, we construct a spending/shortfall efficient frontier that traces the trade-off between committed first-year real spending and the mean shortfall when that commitment is too ambitious for some scenarios. The framework is implemented in Owl (Optimal Wealth Lab), an open-source mixed-integer linear programming (MILP) engine that models US federal tax law, Social Security benefits, required minimum distributions (RMDs), and Medicare costs. We illustrate with representative historical and Monte Carlo case studies and show that the efficient frontier enables more nuanced and actionable spending decisions than PoS alone.

Key Takeaways

- Probability of success is an incomplete retirement risk measure: it records only whether a plan fails, not by how much. Two plans with identical PoS values can have dramatically different risk profiles depending on shortfall magnitude.
- We introduce a spending/shortfall efficient frontier, formally analogous to the Markowitz mean-variance frontier, that identifies the maximum committed spending for any target success rate and makes shortfall magnitude and tail risk directly visible to retirees and their advisors.
- The safe withdrawal rate is the frontier's most conservative special case; the full framework makes the cost of any target success rate explicit and actionable, enabling retirees and their advisors to evaluate a range of defensible spending options rather than a single binary pass/fail threshold.

Keywords: retirement income planning, efficient frontier, probability of success, shortfall risk, mixed-integer linear programming

JEL Codes: C61 (Optimization Techniques; Programming Models), D91 (Intertemporal Household Choice; Life Cycle Models and Saving), G11 (Portfolio Choice; Investment Decisions)

How much should a retiree spend each year? This question sits at the heart of retirement income planning, and yet no single answer satisfies everyone. Future investment returns, inflation, tax laws, and how long the retiree will live are all unknown at the moment of retirement. Any spending decision involves a bet on an uncertain future.

The most widely used tool for navigating this uncertainty is Monte Carlo simulation, which generates thousands of hypothetical future scenarios and reports a probability of success (PoS) — the fraction of scenarios in which the portfolio is not exhausted over the planning horizon. A PoS of 90% is a common planning target, interpreted loosely as "there is only a 10% chance this plan fails."

This framing has a subtle but important flaw. PoS is a binary measure: each scenario is either a success or a failure, with no account taken of the severity of failure. A scenario in which the portfolio is depleted in the very last month of a 30-year retirement counts identically to one in which it runs out 15 years early. A scenario that ends with a 3% spending reduction in the final two years counts as a failure alongside one that forces a 40% spending cut at age 75. PoS cannot distinguish these cases.

In practice, this matters. Advisors routinely encounter clients who are reluctant to accept a PoS of 80% even when the "failing" scenarios are mild, or who feel false confidence from a high PoS when the tail scenarios are severe. What is missing is a measure of *how much* a plan falls short in bad scenarios, not just *whether* it falls short.

This article proposes such a measure. Drawing on the logic of Modern Portfolio Theory (MPT), we construct a spending/shortfall efficient frontier that makes the trade-off between committed spending and shortfall risk explicit. Just as Markowitz (1952) showed that portfolio selection is a two-dimensional problem — balancing expected return against variance — we show that retirement spending commitment is also a two-dimensional problem: balancing committed spending against the mean shortfall when that commitment is too ambitious for some scenarios.

The frontier is computed using Owl (Lacasse, 2026b), which solves a mixed-integer linear program (MILP) for each scenario in a historical or Monte Carlo ensemble. Crucially, the framework does not replace PoS — it enriches it. Any target success rate maps directly to a committed spending level and its associated shortfall distribution, giving retirees and their advisors a more complete picture without discarding a metric they already understand.

The remainder of this article is organized as follows. We first review existing approaches to retirement spending commitment. We then introduce the Owl platform and describe how it computes the efficient frontier. Representative case studies follow. We conclude with practical implications for retirees and their advisors.

EXISTING APPROACHES TO RETIREMENT SPENDING COMMITMENT

The Safe Withdrawal Rate

The safe withdrawal rate (SWR) concept originates with Bengen (1994), who examined all 30-year historical windows in US return data since 1926 and found that an initial annual real withdrawal of 4.15% of a 50/50 stock-bond portfolio never exhausted funds in any historical window. The Trinity Study (Cooley et al., 1998) extended this analysis across five allocation mixes and multiple time horizons.

The SWR has the appeal of simplicity and a worst-case guarantee. Its limitation is that same guarantee: by ensuring survival in the single most adverse historical sequence (the 1966 US start year), it may cause the typical retiree to spend far less than their plan can comfortably sustain (Kitces, 2008). The SWR says nothing about how outcomes are distributed between the worst case and the best case.

Monte Carlo Simulation and Probability of Success

Monte Carlo simulation addresses this limitation by generating a distribution of outcomes. The most common summary statistic is PoS: the fraction of simulated scenarios in which the portfolio does not run out. The approach is widely available in planning software and is the current standard of practice.

As noted above, PoS discards information about the magnitude of failure. Blanchett (2007) observed that PoS should be supplemented with a measure of shortfall magnitude, a theme further developed by Blanchett et al. (2012), who used expected shortfall to rank the efficiency of alternative withdrawal strategies, and by Gardner and Pittman (2013), who formalized expected shortfall as a mortality-adjusted metric — the average dollar gap in desired spending weighted by the probability of being alive to experience it. The present article takes a different approach: rather than evaluating predefined withdrawal rules or adjusting for longevity uncertainty within the metric, it finds the optimal committed spending level for any target PoS, making shortfall magnitude directly visible within the framework practitioners already use — and giving retirees a richer basis for spending decisions without requiring them to abandon a familiar metric.

Prior Efficient Frontier Work

Pfau (2013) proposed a broader framework for retirement income in which spending power, portfolio sustainability, and expected bequests form a three-dimensional efficient surface, with the retiree choosing a point based on preferences. That framework uses simulation to map the surface and relies on product allocation (systematic withdrawal, annuities, guaranteed living benefits) as the decision variable. The present article differs in two ways: the decision variable is a committed spending level rather than a product mix, and shortfall magnitude replaces sustainability as the risk measure, enabling a precise two-dimensional frontier with a formal LP solution.

Dynamic Spending Rules

Scott et al. (2009) demonstrated that financing a constant spending plan with a volatile investment strategy is fundamentally inefficient: it generates unnecessary shortfalls when markets underperform and wasteful surpluses when they outperform, and a retiree following such a rule could typically achieve the same spending distribution at lower cost. A family of dynamic spending strategies — Guyton-Klinger guardrails (Guyton and Klinger, 2006), the required minimum distribution (RMD) method (Larson, 2022), spending-flexibility rules (Pfau, 2012), and others — address this by adjusting spending annually in response to portfolio performance.

The efficient frontier framework proposed here is compatible with dynamic spending rules. Because spending in each scenario is computed by a full MILP retirement optimizer — which accepts any normalized spending profile, including dynamic ones — the framework accommodates flexible spending shapes. The committed first-year basis is the decision variable; the shape of spending over time is a separate input. The flat spending profile used in the case studies below is adopted for simplicity of exposition; in practice, the framework is equally applicable to any normalized spending shape.

OWL: OPTIMAL WEALTH LAB

The case studies in this article use Owl (Lacasse, 2026b), an open-source retirement financial planning tool available at `https://github.com/mdlacasse/Owl`. Owl is freely available under the GNU General Public License and can be used as a Python library, a command-line tool, or through a browser-based interface.

Owl builds on a line of LP-based retirement optimization research. Ragsdale et al. (1994) were among the first to formulate tax-efficient withdrawal scheduling as a linear program; Coopersmith and Sumutka (2011) extended this approach, demonstrating gains over conventional account-ordering rules. Welch developed the Optimal Retirement Planner (ORP), a freely accessible LP-based retirement planning tool that popularized the approach among practitioners (Welch, 2015, 2016, 2017). ORP's source code was never publicly released, and the tool went offline following the author's passing. DiLellio and Ostrov (2019) developed withdrawal algorithms across traditional, Roth, and taxable accounts that account for dividends, taxable lots, and RMDs, with explicit bequest optimization. Owl extends this line of work by incorporating a full federal tax model — including Medicare IRMAA, net investment income tax, ACA premium credits, and SS taxability — and by enabling the stochastic spending optimization described in this article. A parallel line of research has addressed decumulation through stochastic programming and optimal control (Konicz and Mulvey, 2013; Forsyth, 2020; Forsyth et al., 2021); the MILP approach adopted here is complementary, enabling exact tax and benefit modeling in place of stylized assumptions.

What Owl Computes

At its core, Owl solves a mixed-integer linear program that determines the optimal allocation of annual withdrawals across a retiree's accounts — taxable, traditional tax-deferred, Roth, and health savings accounts — to maximize first-year real spending (or, alternatively, estate value at death) (Lacasse, 2026a). The model incorporates:

- **Federal income taxes:** Ordinary income brackets (including projected 2026 law changes), long-term capital gains brackets, and the 3.8% net investment income tax.
- **Social Security:** Benefit computation, taxability rules (0%, 50%, or 85% of benefits subject to income tax), spousal and survivor benefits, and optional optimization of claiming ages.
- **Required minimum distributions:** Enforced as hard lower bounds on traditional account withdrawals.
- **Medicare IRMAA:** Income-related surcharges for Parts B and D, computed from modified adjusted gross income with a two-year lag.
- **Roth conversions:** Optimized annually within tax bracket constraints.
- **ACA premium tax credits:** Modeled for pre-Medicare retirees.

The result is a plan-year-by-year schedule of optimal withdrawals and a maximum first-year real spending basis. For a typical 25-year plan, the computation takes a few seconds on a laptop.

Scenario Analysis and the Efficient Frontier

Owl supports three modes of scenario analysis. In *historical mode*, the plan is solved for each start year in a user-specified range (1928 to the present), using the actual US equity, bond, and inflation returns for the corresponding sequence (Damodaran, 2026). In *Monte Carlo mode*, the plan is solved for a user-specified number of synthetic scenarios drawn from a stochastic return model. In *spending optimization mode*, Owl collects the optimal spending basis from each scenario and then solves a second, much faster LP to find the committed spending level corresponding to the retiree's chosen risk tolerance — producing the efficient frontier.

THE SPENDING/SHORTFALL EFFICIENT FRONTIER

The Core Idea

Suppose Owl is run for each of N historical scenarios. For each scenario, Owl finds the maximum *spending basis* — the highest real first-year spending level the plan can sustain under that scenario's return and inflation sequence. Call this value g_s for scenario s. The collection of these N values summarizes the retiree's spending potential across the historical record.

Now suppose the retiree must commit to a single first-year spending level g^* before knowing which scenario will actually occur. In any scenario where g^* exceeds what that scenario can support — where $g^* > g_s$ — the plan faces a *shortfall*: a gap between what was committed and the optimal spending the MILP computes for that scenario. This shortfall is:

$$\sigma_s = \max(0,\ g^* - g_s).$$

A scenario with no shortfall ($g^* \le g_s$) is a "success" in the PoS sense. A scenario with shortfall ($g^* > g_s$) is a "failure." But unlike PoS, the shortfall σ_s also carries information about *how much* the plan falls short.

The Trade-Off and the Efficient Frontier

The retiree faces a clear trade-off: a higher committed spending g^* is desirable, but it implies larger shortfalls in poor scenarios. A lower g^* is safer but leaves spending on the table in typical scenarios. This is structurally identical to the trade-off between expected return and variance in Markowitz's portfolio theory.

Owl sweeps a risk-aversion parameter λ from zero (maximize spending, ignore shortfalls) to large values (minimize any shortfall, even at the cost of much lower spending). For each value of λ, a fast LP selects the optimal g^*. Plotting the resulting pairs of committed spending and mean shortfall traces the *spending/shortfall efficient frontier*. The role of λ mirrors the risk-aversion coefficient in Milevsky and Huang (2011), who showed using a utility-based framework that a retiree's optimal withdrawal rate declines as aversion to longevity risk increases — a continuous-time analogue of the discrete-time trade-off made explicit by the frontier here.

Exhibit 1 summarizes the analogy with Modern Portfolio Theory.

Concept	**Modern Portfolio Theory**	**This Article**
Decision	Portfolio weights	Committed spending g^*
Maximize	Expected return	Committed spending g^*
Risk measure	Variance (standard deviation)	Mean shortfall $\bar{\sigma}$
Risk aversion	λ (penalty on variance)	λ (penalty on shortfall)
Efficient frontier	Return vs. standard deviation	Spending vs. mean shortfall
Conservative extreme	Minimum-variance portfolio	Safe withdrawal rate
Aggressive extreme	Maximum-return portfolio	Best-case scenario spending

Exhibit 1: Analogy between Modern Portfolio Theory and the spending/shortfall efficient frontier.

The two extreme points of the frontier have intuitive interpretations. At the most aggressive extreme, g^* equals the best-case scenario basis, and the plan accepts large shortfalls in all scenarios worse than the best. At the most conservative extreme, g^* equals the worst-case scenario basis — every scenario has zero shortfall — and this is precisely the safe withdrawal rate. Between these extremes, every point on the frontier represents the maximum committed spending achievable at that mean shortfall level.

Connecting the N per-scenario bases $\{g_s\}$ produced by Owl to the efficient frontier requires solving a small linear program for each value of λ, selecting the committed spending level that best balances spending and mean shortfall across scenarios. Sweeping λ from zero to infinity traces the full frontier from the best-case spending level down to the safe withdrawal rate. The complete formulation, including the treatment of infeasible scenarios and the connection to Conditional Value-at-Risk, is given in the Appendix.

Selecting a Point on the Frontier

Rather than asking a retiree to interpret a risk-aversion parameter, Owl translates the frontier into a familiar choice: a target PoS. For any target success rate ρ (say, 85%), Owl identifies the highest committed spending that achieves that PoS. The retiree then sees not just the committed spending and PoS, but also:

- The *mean shortfall* $\bar{\sigma} = \frac{1}{N}\sum_s \sigma_s$ — the average dollar gap across all N scenarios, assigning zero to scenarios with no shortfall. Dividing by $(1-\rho)$ gives the conditional shortfall per failing scenario, which is formally equivalent to the Conditional Value-at-Risk (CVaR) of spending losses at confidence level ρ (Artzner et al., 1999; Rockafellar and Uryasev, 2000, 2002) — a coherent risk measure widely used in financial risk management.
- The *worst-case scenario spending* (historical mode) or *5th-percentile scenario spending* (Monte Carlo mode) — a tail-risk indicator. For historical scenarios, the sample is fixed, so the minimum is stable and identifies the single worst sequence on record. For Monte Carlo simulations, the sample minimum converges toward zero as the number of draws grows — not because spending can go negative, but because independent random sampling will eventually produce sequences of many consecutive poor years that real market dynamics would suppress through mean reversion and serial correlation. The 5th percentile converges to a stable quantile and provides a more meaningful tail indicator for large Monte Carlo ensembles.[1]
- The *median scenario spending* — showing how conservative g^* is relative to a typical outcome.

These three additional statistics are entirely invisible in a standard PoS analysis. They are the information that allows a retiree to make an informed spending commitment.

CASE STUDIES

We illustrate the framework using a representative retired couple, Chris and Pat. Chris, an engineer, works full-time through the end of 2027, retiring at 64 just as Medicare eligibility begins; Pat, a marketing professional, continues through 2028, retiring at 62. Both file jointly and plan to age 89 and 92, respectively. Their nest egg is concentrated in Chris's tax-deferred accounts, reflecting decades of 401(k) contributions, with smaller taxable and Roth accounts. They own their home outright and plan to leave it to their heirs. Pat claims Social Security upon retirement

[1] The 5th percentile characterizes shortfall severity only when the target success rate $\rho < 95\%$, so that more than 5% of scenarios fall below committed spending. When $\rho \geq 95\%$, the 5th percentile lies in the success zone and does not reflect failure severity; in that case the mean shortfall divided by $(1-\rho)$ — the average shortfall per failing scenario — is the more informative tail statistic.

to receive early income while Chris's larger benefit continues to grow. Chris is inclined to claim at age 70 to maximize the household survivor benefit; Owl's SS age optimizer confirms this choice in a preliminary run. Both claiming ages are then fixed for the frontier analysis. Their financial profile is shown in Exhibit 2. The asset allocation is held constant at 60% equity / 40% bonds throughout the planning horizon; while Owl supports user-defined gliding paths, a fixed allocation was chosen here to isolate the effect of sequence-of-returns risk from allocation drift.[2]

Parameter	**Value**
Ages at plan start (2026)	63 (Chris), 60 (Pat); retire end of 2027 and 2028 respectively
Planning horizon	Age 89 (Chris), 92 (Pat)
Filing status	Married filing jointly
Taxable accounts	\$100,000 (Chris), \$50,000 (Pat)
Tax-deferred (Traditional IRA / 401(k))	\$700,000 (Chris), \$150,000 (Pat)
Roth IRA	\$70,000 (Chris), \$40,000 (Pat)
Primary residence	\$450,000 (held to death; left as bequest)
401(k) contributions	Employee max + 4% employer match (Chris through 2027, Pat through 2028)
Social Security PIA	\$2,000/mo (Chris, claiming age 70)†, \$1,400/mo (Pat, claiming age 62)
Spending profile	Flat (constant real spending)
Asset allocation	60% equity / 40% bonds
Heirs' tax rate on IRA estate	30%
Bequest target	\$400,000 (excluding primary residence)

Exhibit 2: Chris and Pat's financial profile. †Chris's claiming age confirmed by Owl's SS age optimizer in a preliminary run and fixed for the frontier analysis; Pat's claiming age chosen by the user. PIA = Primary Insurance Amount (monthly benefit at full retirement age); actual benefits are computed by Owl from the PIA and the chosen claiming age, applying the standard Social Security actuarial adjustments for early or delayed claiming.

Historical Scenarios

Setup. Using the financial profile in Exhibit 2, Owl solves the full MILP for each start year from 1928 to 1993, yielding 66 historical scenarios spanning the full range of available US market data (Damodaran, 2026). The range ends in 1993 because later start years do not provide a complete return history sufficient to cover the couple's 33-year planning horizon. Each per-scenario solve jointly optimizes Roth conversions and savings withdrawals, subject to US tax rules and Medicare costs, producing the maximum sustainable spending basis g_s for that scenario. These 66 bases then feed the commitment LP described in the previous section, sweeping λ across a logarithmically spaced grid to trace the efficient frontier with uniform visual density across the full curve.

Results. Exhibits 3 and 4 show the three-panel output from Owl at target success rates of 85% and 70%, respectively. Each exhibit shows the PoS curve (left), the efficient frontier (center), and a bar chart of committed spending and shortfall by historical start year (right). The left

[2] The complete case file (`Case_chris+pat.toml`) is included in the Owl repository and can be used to reproduce all results in this article. Results were generated with Owl version 2026.04.08.

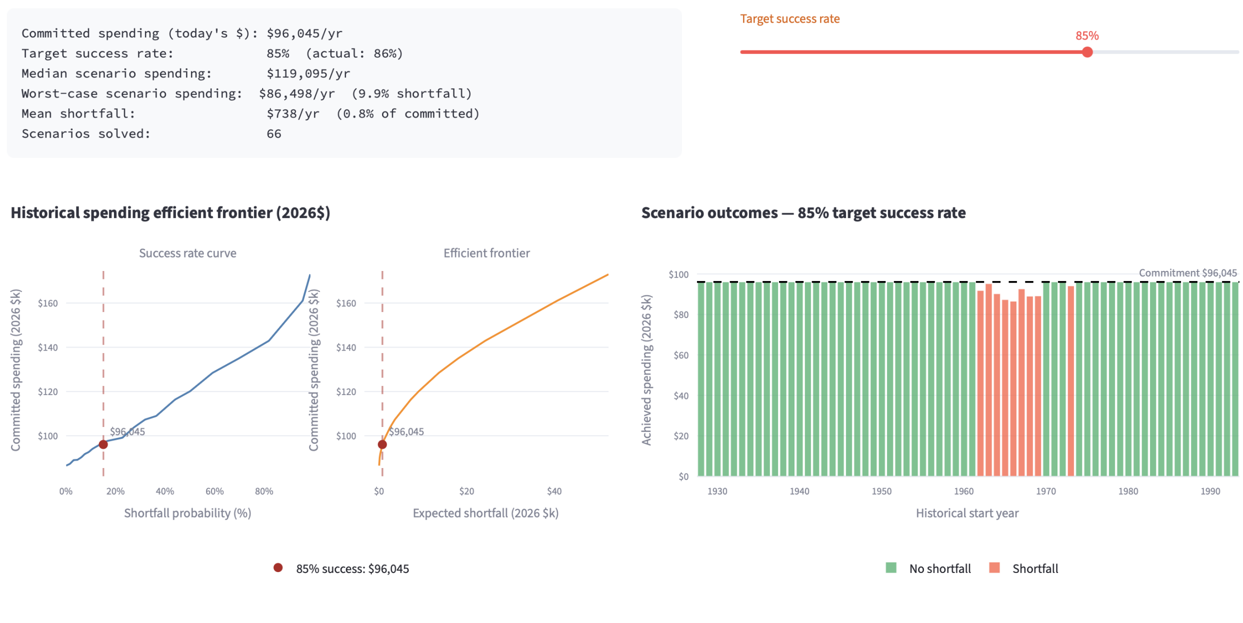


Exhibit 3: Forward historical scenarios (1928–1993), 85% target success rate as shown in Owl. *Left:* Committed spending versus shortfall probability. *Center:* Committed spending versus mean shortfall. *Right:* Committed spending and shortfall by historical start year.

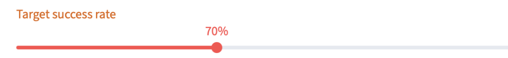
and center panels are the same curve with the selected point shifted; the bar chart on the right changes dramatically between the two targets, making the cost of accepting more risk concrete and scenario-specific.

At an 85% target, Chris and Pat can commit to $96,045/yr in today's dollars, with a mean shortfall of only $738/yr (0.8% of committed spending) and a median scenario spending of $119,095/yr — well above the commitment. Accepting a 70% target raises committed spending to $103,712/yr (+8%), but the mean shortfall triples to $2,485/yr (2.4%) and the worst-case shortfall grows from 9.9% to 16.6% of committed spending. The median and the worst-case scenario — the 1966 start year — are the same under both targets, since they reflects properties of the scenario set rather than the chosen commitment level.

Time-Reversed Historical Scenarios

Setup. To illustrate Owl's ability to augment the historical scenario set, we repeat the analysis using the same 66 return sequences played in reverse chronological order. Reversing a return sequence preserves the marginal distribution of annual returns — the same mean, variance, and cross-asset correlations — and, for a stationary process, the autocorrelation structure. What changes is the *ordering*: a sequence that historically presented strong early returns followed by a weak decade now presents that weak decade first, and vice versa. For a retiree drawing down a portfolio, this directly reverses the sequence-of-returns risk profile, yielding 66 additional stress scenarios that are structurally distinct from the forward set without fabricating any return values. The forward and reversed sets are not statistically independent — they share the same annual returns — so the combined 132 scenarios modestly overstate the effective sample size; we treat them as a richer

```
Committed spending (today's $): $103,716/yr
Target success rate:            70%  (actual: 73%)
Median scenario spending:       $119,095/yr
Worst-case scenario spending:  $86,498/yr  (16.6% shortfall)
Mean shortfall:                 $2,485/yr  (2.4% of committed)
Scenarios solved:               66
```



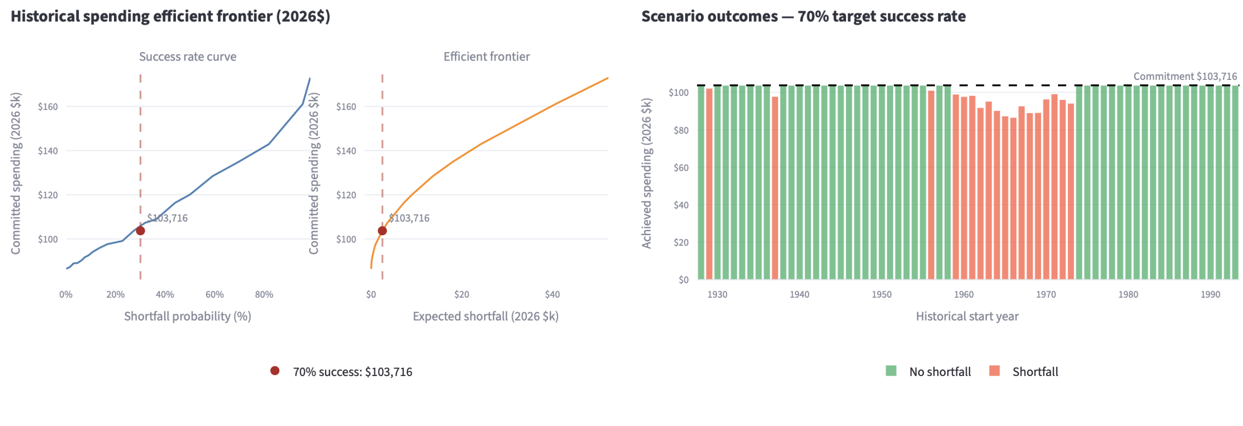


Exhibit 4: Forward historical scenarios (1928–1993), 70% target success rate. *Left:* Committed spending versus shortfall probability. *Center:* Committed spending versus mean shortfall. *Right:* Committed spending and shortfall by historical start year. Compared to Exhibit 3, the bar chart reveals a substantially broader and deeper shortfall distribution, illustrating the concrete cost of targeting a lower success rate.

description of sequence risk rather than as independent draws. Other resampling approaches, such as the stationary block bootstrap of Politis and Romano (1994), are also available in Owl but not illustrated here.

Results. Exhibits 5 and 7 show the corresponding three-panel output for the time-reversed scenarios at 85% and 70% targets. Exhibit 6 summarizes all four cases.

The reversed sequences tell a strikingly different story. At 85%, the committed spending drops slightly to $93,017/yr (vs. $96,045 forward), yet the median scenario spending *rises* to $125,793/yr (vs. $119,095 forward). The reversed set has a wider spread: more scenarios are very favorable — early-retirement years that were historically weak become strong when reversed, allowing the portfolio to grow before drawdown — but the worst case is more severe ($79,450/yr, a 14.6% shortfall, vs. 9.9% forward).

The contrast is sharpest at the 70% target. The reversed sequences allow a committed spending of $111,071/yr, nearly $7,400 more than the forward $103,712/yr — a counterintuitive result explained by the higher median: there are more strong scenarios in the reversed set, which pulls the 70th-percentile threshold upward. But the downside is commensurately more severe: the worst-case shortfall reaches 28.5% of committed spending, compared to 16.6% for the forward set. This highlights a general principle — a more favorable median does not imply a safer tail.

These differences merit further investigation. The sensitivity of committed spending and short-

```
Committed spending (today's $): $93,017/yr
Target success rate:           85%  (actual: 86%)
Median scenario spending:      $125,793/yr
Worst-case scenario spending:  $79,450/yr  (14.6% shortfall)
Mean shortfall:                $1,060/yr  (1.1% of committed)
Scenarios solved:              66
```

Exhibit 5: Time-reversed historical scenarios (1928–1993), 85% target success rate. *Left:* Committed spending versus shortfall probability. *Center:* Committed spending versus mean shortfall. *Right:* Committed spending and shortfall by historical start year.

fall severity to sequence ordering — with the same marginal return distribution — underscores how strongly decumulation outcomes depend on the timing of returns rather than their average level. Whether the reversed set is a useful stress test or an artifact of the reversal depends on whether real return sequences exhibit the kind of autocorrelation structure that reversal preserves or distorts, a question that goes beyond the scope of this article (Zumbach, 2009). The goal here was to demonstrate the effect of sequence by using two sets that only differ by their ordering.

Case	Target PoS	Committed	Mean shortfall	Shortfall (%)	Worst case	Median
Forward	85%	$96,045	$738	0.8%	$86,498 (9.9%)	$119,095
Forward	70%	$103,712	$2,485	2.4%	$86,498 (16.6%)	$119,095
Reversed	85%	$93,017	$1,060	1.1%	$79,450 (14.6%)	$125,793
Reversed	70%	$111,071	$4,494	4.0%	$79,450 (28.5%)	$125,793

Exhibit 6: Summary of historical frontier results for forward and time-reversed scenarios. Worst-case percentages are shortfalls as a fraction of committed spending. All dollar amounts are in today's dollars per year. (Monte Carlo results are in Exhibit 9.)

The 1966 sequence revisited. A key advantage of the efficient frontier is its ability to contextualize the worst historical scenario. The 1966 US start year is widely cited as the most dangerous sequence for retirees — the sequence that defines the safe withdrawal rate. Under the forward scenarios at an 85% target, the 1966 sequence produces a worst-case spending of $86,498/yr against a commitment of $96,045/yr — a shortfall of $9,547/yr, or 9.9% of committed spending. This is a meaningful but manageable reduction, not a catastrophic failure. Knowing that the worst case in a century of history amounts to a 10% spending reduction is a fundamentally different framing than being told "there is a 15% chance of failure."

```
Committed spending (today's $): $111,071/yr
Target success rate:            70%  (actual: 73%)
Median scenario spending:       $125,793/yr
Worst-case scenario spending:  $79,450/yr  (28.5% shortfall)
Mean shortfall:                 $4,494/yr  (4.0% of committed)
Scenarios solved:               66
```

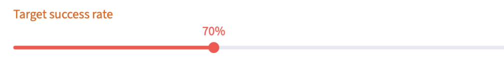


Exhibit 7: Time-reversed historical scenarios (1928–1993), 70% target success rate. *Left:* Committed spending versus shortfall probability. *Center:* Committed spending versus mean shortfall. *Right:* Committed spending and shortfall by historical start year.

Monte Carlo Scenarios

Setup. The Monte Carlo scenarios use Owl's *histolognormal* model: log-returns for equities, bonds, and inflation are assumed jointly normally distributed, with the mean vector and covariance matrix estimated by maximum likelihood from the full 1928–2025 historical record or a contiguous subset of it. Each scenario is an independent draw of a synthetic return series of the same length as the planning horizon. This is the standard parametric baseline in the retirement literature: Finke et al. (2013) use a similar multivariate normal calibration (with arithmetic means adjusted to prevailing bond yields); Collins et al. (2015) survey the sensitivity of retirement outcomes to the choice of return model and recommend the lognormal as more appropriate than the normal for asset returns, since portfolio values are products of gross-return factors and cannot turn negative. More sophisticated models — including vector autoregressions that capture mean reversion in equity returns (Campbell and Viceira, 2002) and regime-switching models that better replicate volatility clustering — exist but are not required for the illustrative purposes of this section.

Owl ran 500 simulations, completing in approximately 2 minutes on a standard laptop; each simulation requires a full MILP solve of the retirement plan under a synthetic return sequence.

Results. Exhibit 8 shows the Monte Carlo efficient frontier for Chris and Pat at an 85% target success rate. A second figure at 70% is omitted, as the frontier shape is nearly identical; numerical results for both targets appear in Exhibit 9.

Comparison with historical scenarios. At an 85% target, the Monte Carlo frontier yields a committed spending of $94,833/yr, nearly identical to the $96,045/yr from the forward historical

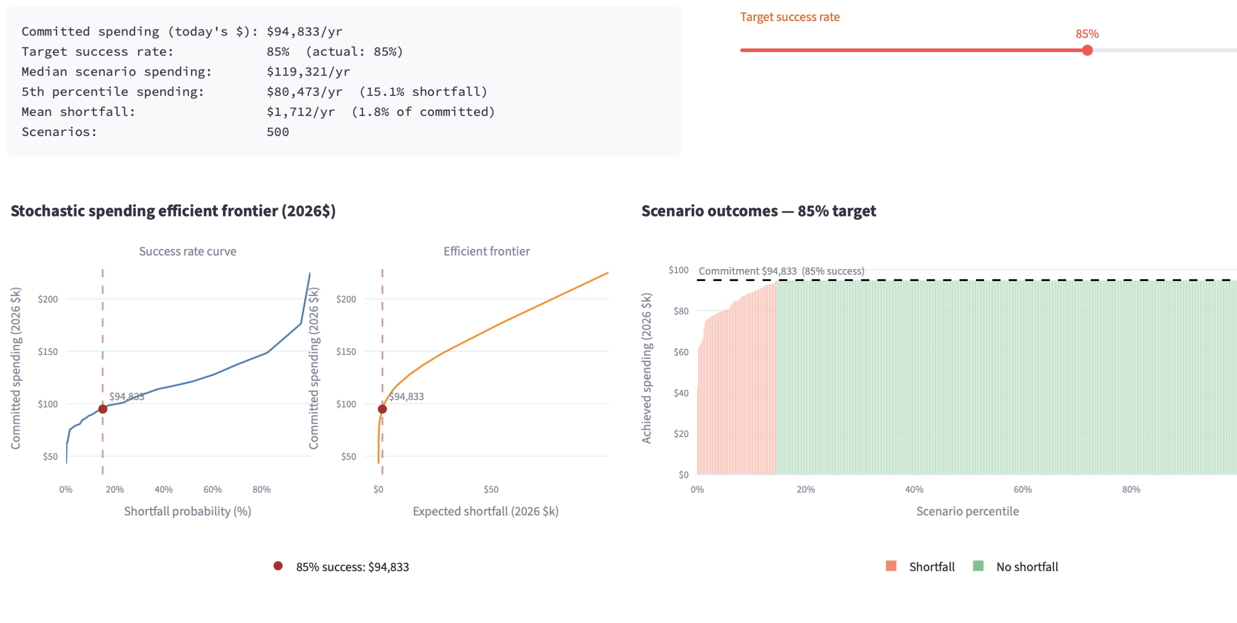


Exhibit 8: Monte Carlo scenarios (500 simulations, *histolognormal* model), 85% target success rate as shown in Owl. *Left:* Committed spending versus shortfall probability. *Center:* Committed spending versus mean shortfall. *Right:* Achieved spending by scenario percentile, sorted from worst (0th) to best (100th); red bars indicate shortfall scenarios.

Case	Target PoS	Committed	Mean shortfall	Shortfall (%)	5th percentile	Median
Monte Carlo	85%	$94,833	$1,712	1.8%	$80,473 (15.1%)	$119,321
Monte Carlo	70%	$105,526	$4,000	3.8%	$80,473 (23.7%)	$119,321

Exhibit 9: Summary of Monte Carlo frontier results (500 simulations, multivariate lognormal model (termed histolognormal in Owl). The 5th-percentile column replaces the worst-case column used for historical scenarios (Exhibit 6); percentages are shortfalls as a fraction of committed spending. All dollar amounts are in today's dollars per year.

ensemble. The mean shortfall is $1,712/yr (1.8% of committed spending), somewhat higher than the historical $738/yr, reflecting the broader spread of synthetic scenarios. The 5th percentile of scenario spending lands at $80,473/yr — a 15.1% shortfall from committed — a more stable tail indicator than the sample minimum for large simulations. At 70%, committed spending rises to $105,526/yr (+11.3% over the 85% target), with a mean shortfall of $4,000/yr (3.8%). The 5th-percentile level is unchanged — it is a property of the scenario ensemble, not of the commitment level — but it now represents a 23.7% shortfall, illustrating how the same tail outcome cuts deeper as the spending commitment rises.

A qualitative difference between the two frontiers is visible in the shape of the PoS–shortfall curve. The historical frontier is approximately linear across the full range of success rates: each percentage-point reduction in the target buys a roughly proportional increase in committed spending. The Monte Carlo frontier shares this linear behavior in its middle range but exhibits sharp nonlinearity near the extremes — a steep drop in committed spending as PoS approaches 100%, and a rapid acceleration in mean shortfall as PoS approaches 0%. This shape is a direct signature of the independent random sampling assumption. Achieving near-perfect success requires guarding against increasingly extreme tail draws — sequences with many consecutive poor years that independent sampling will eventually generate, but that real market dynamics suppress through mean reversion and autocorrelation. The historical frontier does not display this curvature because its 66 scenarios are bounded: the worst is the 1966 sequence, a real draw subject to the same autocorrelation structure as all other historical periods. The nonlinearity of the synthetic frontier near its extremes is therefore a property of the return model rather than a feature of the retiree's actual risk landscape. This reinforces the recommendation to use correlated return models — block bootstrap, stationary bootstrap, or GARCH — when the behavior of the frontier at extreme success-rate targets is of primary concern (all three model families are available in Owl).

On simulation size and tail reliability. A common instinct in Monte Carlo practice is that more scenarios improve accuracy. This holds for central statistics: the standard error of an estimator such as mean shortfall scales as $1/\sqrt{N}$, so a few hundred scenarios already yields tight estimates of the quantities that matter most (Fitzpatrick and Tharp, 2022). Tail statistics, however, behave differently. The worst-case spending in a sample of N independent draws is estimated from roughly one observation — and the sample minimum decreases without bound as N grows, generating increasingly extreme scenarios that real market dynamics, through mean reversion and autocorrelation, would never produce (Tharp, 2017; Collins et al., 2015; Campbell and Viceira, 2002). In conventional Monte Carlo practice, this means that a planner who runs 10,000 simulations instead of 500 obtains a more alarming portfolio-depletion chart, not a more accurate one.

The efficient frontier sidesteps this problem by design. Mean shortfall $\bar{\sigma}$ is computed across *all* N scenarios — assigning a dollar value to every outcome, success and failure alike — and its $1/\sqrt{N}$ convergence is fast. The 5th-percentile statistic is similarly robust: estimated from $0.05N$ scenarios, it retains meaningful sample depth even for modest ensemble sizes. Shifting the conversation from worst-case spending (unreliable, alarming) to mean shortfall and a stable quantile is therefore not merely a presentational choice — it is the statistically appropriate way to summarize tail risk from independent Monte Carlo draws. And while MILP solves are slower than conventional simulation — a few hundred scenarios in minutes rather than thousands in seconds — this constraint aligns naturally with the scale at which tail statistics are actually trustworthy.

PRACTICAL IMPLICATIONS

Reframing the Spending Conversation

The efficient frontier changes the retirement spending conversation in a meaningful way. Under the conventional PoS paradigm, the retiree negotiates a success rate threshold without any sense of what happens in the scenarios below that threshold. The efficient frontier makes the cost of those scenarios visible: a mean shortfall of $X per year, with spending in the bottom 5% of scenarios reaching $Y per year. Both statistics converge rapidly with sample size, unlike the worst-case outcome in a large Monte Carlo ensemble of independent draws — which grows more extreme with every additional simulation run.

For many retirees, the efficient frontier will reveal that a lower PoS target is entirely reasonable once the magnitude of potential shortfalls is understood. A 70% success rate may sound alarming in isolation; it sounds very different when the frontier shows that the mean shortfall across all scenarios amounts to only 2–4% of committed spending, with even the most extreme historical sequence falling short by at most 17%, while the median scenario sustains roughly 15% more than the committed level — equivalently, committing to the median would correspond to a 50% target success rate. This contextual information allows retirees to make genuinely informed decisions rather than reacting to a single percentage.

Annual Re-Evaluation

Annual re-evaluation transforms a static spending commitment into a dynamic spending strategy. Each year, as new information arrives — updated portfolio values, revised return forecasts, tax-law changes, or unexpected expenses — the frontier is recomputed on current conditions and a new committed spending level is selected for the coming year. The result is a plan that adapts continuously to the retiree's evolving situation, without imposing a mechanical adjustment rule.

The computational structure of the framework supports this cycle directly. Constructing the frontier requires solving one MILP per scenario — a few minutes for a historical ensemble on a laptop. Once those solves are complete, however, exploring the full range of PoS targets is instantaneous: the commitment step is a simple LP that recombines the existing g_s values. The entire frontier can be explored interactively — adjusting the target success rate and watching the committed spending and mean shortfall update in real time — without any additional computation.

The Safe Withdrawal Rate as a Special Case

The efficient frontier nests the safe withdrawal rate as its most conservative point. At sufficiently high risk aversion, the committed spending converges to the worst-case scenario basis — the spending level that survives every historical scenario. This is precisely Bengen's 4% rule (adjusted for the specific plan's tax situation, account structure, and horizon).

The efficient frontier makes visible how much spending is given up by insisting on zero shortfall in all scenarios. The analogy to Modern Portfolio Theory is instructive: an investor who insists on holding all assets in cash to eliminate any possibility of loss achieves zero variance — but at the cost of dramatically lower expected wealth. Markowitz showed that rational investors need not accept this extreme; the efficient frontier offers higher expected returns for only a modest increase in risk. The SWR occupies the same position on the spending/shortfall frontier: it eliminates all shortfall risk, but at the cost of committed spending that may be far below what a typical scenario could comfortably sustain. For many retirees, accepting a small mean shortfall in the worst-case scenarios in exchange for meaningfully higher spending is the rational choice — and the frontier makes that trade-off visible.

Compatibility with Existing Planning Tools and Future Extensions

The spending/shortfall efficient frontier is not inherently limited to MILP-based planners, though

MILP offers the most precise implementation. Because each scenario solve maximizes spending subject to all tax, RMD, and Social Security constraints, it returns g_s — the true maximum the plan can sustain — making the shortfall measure exact and the frontier well-defined. Traditional forward-projection tools simulate a fixed spending level and record only whether and when the portfolio is exhausted; they do not expose the dollar gap.

That said, advisors using other planning software can in principle extract per-scenario spending outcomes and use them as inputs to the commitment step — a simple LP that could be implemented in a spreadsheet for small scenario sets. Owl provides this full capability as an open-source tool (Lacasse, 2026b). A natural extension would go further: for tools that report only the year of portfolio exhaustion, the gap between that year and the planning horizon — discounted at an inflation-adjusted rate — could serve as a proxy for shortfall magnitude. While cruder than the MILP-based measure, such an approximation could bring shortfall-aware frontier analysis to the much wider universe of simulation tools already in use by advisors, meaningfully broadening access to the framework introduced here.

A second extension is to embed longevity risk directly into the scenario set. The current framework fixes the planning horizon and varies only the return and inflation sequence across scenarios. By also drawing a random lifespan for each scenario from standard mortality tables — so that each market sequence is paired with its own planning horizon — the optimizer naturally produces a g_s that reflects both sequence-of-returns risk and the risk of outliving the plan. The commitment step is unchanged; the resulting frontier penalizes shortfalls arising from either source of uncertainty in a unified way, giving retirees a single, comprehensive picture of retirement risk (Simsek et al., 2018). Should the larger scenario set slow the computation, run time scales linearly with the number of scenarios and remains practical on a standard laptop for typical planning horizons.

CONCLUSION

Probability of success has long served as the primary risk metric in retirement income planning. It is a useful measure, but an incomplete one. It records only whether a plan fails, not by how much. Two retirees with identical PoS values — one whose “failures” are mild and one whose “failures” are severe — face very different retirement risks, but PoS cannot distinguish them.

The spending/shortfall efficient frontier fills this gap. By making the expected magnitude of potential shortfalls visible alongside committed spending, it gives retirees and their advisors the information needed to make genuinely informed spending decisions. The framework is formally analogous to the Markowitz mean-variance frontier: committed spending plays the role of expected return, and mean shortfall plays the role of variance. The efficient frontier is the set of spending commitments that are Pareto-optimal in this two-dimensional space.

Owl implements this framework as free, open-source software, incorporating a comprehensive model of US tax law, Social Security, RMDs, and Medicare. As an open-source platform, it also provides a common substrate where researchers and practitioners can share, reproduce, and test new retirement modeling ideas. It is available at `https://github.com/mdlacasse/Owl`.

DECLARATION OF SUBMISSION

The authors confirm that this manuscript is original work that has not been published elsewhere. It is not currently under review at any other publication in the Portfolio Management Research family or any other journal. The authors have read the PMR publication agreement and agree to its terms. The authors' institutions do not require that this paper be posted on any public-facing website, including SSRN, ResearchGate, or SciHub, prior to publication.

References

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance 9*(3), 203–228.

Bengen, W. P. (1994, October). Determining withdrawal rates using historical data. *Journal of Financial Planning*, 171–180.

Blanchett, D. M. (2007). Dynamic allocation strategies for distribution portfolios: Determining the optimal distribution glide path. *Journal of Financial Planning 20*(12), 68–81.

Blanchett, D. M., M. Kowara, and P. Chen (2012). Optimal withdrawal strategy for retirement portfolios. *Retirement Management Journal 2*(3), 7–20.

Campbell, J. Y. and L. M. Viceira (2002). *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*. Oxford: Oxford University Press.

Collins, P. J., H. D. Lam, and J. Stampfli (2015). How risky is your retirement income risk model? *Financial Services Review 24*, 193–216.

Cooley, P. L., C. M. Hubbard, and D. T. Walz (1998, February). Retirement savings: Choosing a sustainable withdrawal rate. *Journal of the American Association of Individual Investors*, 16–21. Also known as the Trinity Study.

Coopersmith, L. W. and A. R. Sumutka (2011, September). Tax-efficient retirement withdrawal planning using a linear programming model. *Journal of Financial Planning*.

Damodaran, A. (2026). Annual returns on stock, t.bills, t.bonds and inflation—united states (1928–present). NYU Stern School of Business, available at `https://pages.stern.nyu.edu/~adamodar/`. Historical annual returns dataset, updated annually.

DiLellio, J. A. and D. N. Ostrov (2019). Constructing tax-efficient withdrawal strategies for retirees with traditional 401(k)/iras, roth 401(k)/iras, and taxable accounts. *Financial Services Review 28*(2), 67–95.

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

Fitzpatrick, J. and D. Tharp (2022, May). Evaluating retirement spending risk: Monte carlo vs historical simulations. Kitces.com. Available at https://www.kitces.com/blog/monte-carlo-simulation-historical-returns-sequence-risk-calculate-sustainable-spending-levels/.

Forsyth, P. A. (2020). A stochastic control approach to defined contribution plan decumulation: "the nastiest, hardest problem in finance". *ASTIN Bulletin: The Journal of the IAA*. arXiv:2008.06598.

Forsyth, P. A., K. R. Vetzal, and G. Westmacott (2021). Optimal control of the decumulation of a retirement portfolio with variable spending and dynamic asset allocation. *ASTIN Bulletin: The Journal of the IAA 51*(3), 905–938.

Gardner, G. and S. Pittman (2013). Measuring the risk of running out of money in retirement. *Journal of Financial Planning 26*(12), 38–44.

Guyton, J. T. and W. J. Klinger (2006, March). Decision rules and maximum initial withdrawal rates. *Journal of Financial Planning*.

Kitces, M. E. (2008, May). Resolving the paradox – is the safe withdrawal rate sometimes too safe? *The Kitces Report*.

Konicz, A. K. and J. M. Mulvey (2013). Applying a stochastic financial planning system for an individual: Immediate or deferred life annuities? *Journal of Retirement 1*(1), 90–103.

Lacasse, M.-D. (2026a). Optimal wealth lab (Owl): Mathematical foundations of a retirement planning optimizer. Technical report, GitHub. Available at https://github.com/mdlacasse/Owl/blob/main/papers/owl.pdf.

Lacasse, M.-D. (2026b). Owl: Optimal wealth lab – a mixed-integer linear programming retirement financial planner. https://github.com/mdlacasse/Owl. Version 2026.04.08. Additional contributors are listed in the AUTHORS file at https://github.com/mdlacasse/Owl/blob/main/AUTHORS. The case file Case_chris+pat.toml is included in the repository for reproducibility.

Larson, S. J. (2022, January). Required minimum distributions as a retirement strategy: The tradeoff between RMD volatility and the expected number of dollars paid out. *Journal of Financial Planning 35*(1), 60–67.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance 7*(1), 77–91.

Milevsky, M. A. and H. Huang (2011). Spending retirement on planet Vulcan: The impact of longevity risk aversion on optimal withdrawal rates. *Financial Analysts Journal 67*(2), 45–58.

Pfau, W. D. (2012, March). Spending flexibility and safe withdrawal rates. *Journal of Financial Planning*.

Pfau, W. D. (2013, February). A broader framework for determining an efficient frontier for retirement income. *Journal of Financial Planning*. SSRN 2151259.

Politis, D. N. and J. P. Romano (1994). The stationary bootstrap. *Journal of the American Statistical Association 89*(428), 1303–1313.

Ragsdale, C. T., A. F. Seila, and P. L. Little (1994). An optimization model for scheduling withdrawals from tax-deferred retirement accounts. *Financial Services Review 3*(2), 93–109.

Rockafellar, R. T. and S. Uryasev (2000). Optimization of conditional value-at-risk. *Journal of Risk 2*(3), 21–41.

Rockafellar, R. T. and S. Uryasev (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance 26*(7), 1443–1471.

Scott, J. S., W. F. Sharpe, and J. G. Watson (2009). The 4% rule—at what price? *Journal of Investment Management 7*(3), 31–48.

Simsek, K. D., M. J. Kim, W. C. Kim, and J. M. Mulvey (2018). Optimal longevity risk management in the retirement stage of the life cycle. *Journal of Retirement 5*(3), 73–92.

Tharp, D. (2017, July). Fat tails in Monte Carlo analysis vs. safe withdrawal rates. Kitces.com. Available at `https://www.kitces.com/blog/monte-carlo-analysis-risk-fat-tails-vs-safe-withdrawal-rates-rolling-historical-returns/`.

Welch, J. S. (2015). Mitigating the impact of personal income taxes on retirement savings distributions. *Journal of Personal Finance*.

Welch, J. S. (2016). Measuring the financial consequences of IRA to Roth IRA conversions. *Journal of Personal Finance*.

Welch, J. S. (2017, August). A 3-step procedure for computing sustainable retirement savings withdrawals. *Journal of Financial Planning*.

Zumbach, G. (2009). Time reversal invariance in finance. *Quantitative Finance 9*(5), 505–515.

APPENDIX: COMMITMENT LP FORMULATION

Given N per-scenario bases $g_s \geq 0$ (the optimal spending level Owl finds for scenario s) and a risk-aversion parameter $\lambda \geq 0$, the commitment LP is:

$$
\begin{aligned}
\text{maximize} \quad & g^* - \frac{\lambda}{N}\sum_{s=1}^{N} \sigma_s & & \\
\text{subject to} \quad & \sigma_s \geq g^* - g_s, & s = 1,\ldots,N & \qquad (1)\\
& \sigma_s \geq 0, & s = 1,\ldots,N & \qquad (2)\\
& 0 \leq g^* \leq \max_s g_s. & & \qquad (3)
\end{aligned}
$$

The decision variables are the scalar committed spending level g^* and the N shortfall variables σ_s, giving $N+1$ variables and $2N+1$ constraints in total. Constraints (1)–(2) force $\sigma_s = \max(0, g^* - g_s)$ at optimality: σ_s equals the spending shortfall in scenario s if that scenario falls below the commitment, and zero otherwise. Constraint (3) prevents g^* from exceeding the best achievable outcome.

Sweeping the frontier. At $\lambda = 0$ the penalty vanishes and the LP simply maximizes g^*, yielding the most aggressive point on the frontier (best-case scenario spending, full shortfall accepted in all worse scenarios). As $\lambda \to \infty$ the penalty dominates and g^* is driven down to $\min_s g_s$, the safe withdrawal rate, where every scenario has zero shortfall. Solving the LP for a grid of λ values traces the full Pareto curve of committed spending versus mean shortfall $\bar{\sigma} = \frac{1}{N}\sum_s \sigma_s$.

Infeasible scenarios. If a scenario is infeasible — no plan satisfying all financial constraints exists under that rate sequence — its basis is set to $g_s = 0$. This contributes a full shortfall $\sigma_s = g^*$ for any positive commitment and counts as a failure in the shortfall probability $P(\sigma_s > 0)$. The normalization constant N always equals the total number of scenarios requested, not the number that solved successfully.

Connection to Conditional Value-at-Risk. The mean shortfall $\bar{\sigma}$ divided by $(1-\rho)$, where ρ is the target success rate, equals the Conditional Value-at-Risk (CVaR) of spending losses at confidence level ρ (Rockafellar and Uryasev, 2000, 2002) — a coherent risk measure widely used in financial risk management. Minimizing $\bar{\sigma}$ for fixed g^* is therefore equivalent to minimizing CVaR, and the commitment LP is structurally identical to the CVaR minimization LP of Rockafellar and Uryasev (2000), with g^* playing the role of the Value-at-Risk threshold.