From Brief to Estimate: A Probabilistic and Causal Footing for *The Uneven Month*

Working paper / methodological supplement *Companion to the household oil-shock model and report. Prepared at referee request to put three results on research-paper footing: (i) a probabilistic forecast in place of a deterministic envelope — reported as two bands, a historical-volatility floor and a tail-dependent coupled distribution, with the gap between them as the quantified deep uncertainty; (ii) confidence intervals on the incidence indices; and (iii) a causal design for the debt-triage claim.*

Abstract

The deterministic brief established a robust qualitative thesis — a 2026 diesel/oil shock is regressive, falling hardest as a share of income on the lowest-income households — and a central estimate of a +2.9% cumulative price rise and ~$25B/month in added national household cost by December. This supplement upgrades three elements to research standard. First, a Monte Carlo replaces the brief's corner "envelope" with a genuine predictive distribution, reported as **two bands**. A historical-volatility **floor** (the lower bound on uncertainty: independent Gaussian path noise) gives a 90% interval for the monthly aggregate burden of **$3.7B–$49.3B** (median $23.3B) and a 99th percentile of **$54.8B**. A **coupled** band — tail-dependent (Student-t copula), fat-tailed, with a crisis regime and a state-dependent basket broadening that lets channels go wrong together — leaves the **center unchanged** (median $23.7B) but fattens the upper tail to a 99th percentile of **$68.7B** (rotated-Clayton variant: $75.3B), and raises P(burden > $60B/mo) from 0.1% to **2.1%**. The gap between the bands, not a single number, is the quantification of the cascade ("novel territory"). A law-of-total-variance decomposition shows **82% of the floor spread is structural — which scenario realizes — and only 18% is path noise**; and the tail of the coupled band is governed by the **crisis-regime probability and severity and the breadth of transmission, not by the copula family or the scenario odds**, both of which are second-order. The single expected value also misleads because the distribution is multimodal. Second, a parametric bootstrap over the modeled inputs yields the first interval estimates of the incidence indices: pre-tax **Suits = −0.26 (90% CI −0.31, −0.21)** and **Kakwani = −0.25 (−0.29, −0.21)**, regressive in **100%** of bootstrap replicates. The indices are invariant to the inflation rate; the one channel that touches them — the crisis-state basket broadening — only mildly attenuates them (Suits −0.26 → −0.24) and never reverses the sign. Third, we specify — but, lacking panel microdata, do not estimate — an event-study design that would move the debt-triage ordering from a descriptive to a causal claim. We are explicit throughout about what is simulated, what is estimated, and what remains a design.

1. Introduction and contribution

The brief is an exposure-and-incidence study with a robust central finding and an explicitly conditional magnitude. Three of its features were appropriate for a brief but below the bar for a research paper:

1. **A deterministic point estimate with a corner "envelope."** The "$9–44B/month" range was the span of three central scenario paths — a *scenario-mean range*, not a probability interval. It carried no statement of likelihood and ignored realized-path dispersion within a scenario.
2. **Point-valued incidence indices.** The Suits/Kakwani indices were reported as single numbers (later as a base-convention band), with no sampling or modeled-input uncertainty quantified.
3. **A descriptive triage ordering.** The order in which households sacrifice obligations was asserted from counseling practice and cross-sectional delinquency correlations, with no causal identification.

This supplement addresses each, in order, and is candid about the binding constraint: we have the structured model and its parameters but not the underlying survey microdata or historical delinquency panels, and no facility to retrieve them. Sections 3 and 4 are therefore fully computed from the model; Section 5 is a design ready to take to data.

2. The deterministic core (recap)

The scenario engine compounds seven monthly CPI paths (June–December) for three scenarios — ceasefire, status quo, escalation — into cumulative-by-December price rises of **+1.10%, +2.83%, +5.21%** respectively. Probability weights (0.30 / 0.45 / 0.25) give an expected **+2.91%**. The aggregate burden is exactly linear in the expected rate, `Aggregate ($B/mo) = rate × 847.45`, so the central estimate is **$24.7B/month**. The incidence indices are *scenario-invariant* (the rate cancels from the burden shares), so they describe the shock's shape independent of its size. All figures below are reproducible from the model parameters via the appended script (fixed seed; N = 200,000 MC draws, 100,000 bootstrap replicates).

3. A probabilistic forecast (Monte Carlo)

We report the magnitude as **two bands**. The **floor** is a historical-volatility calibration — it prices the future as a draw from the past, and so is a deliberate *lower bound* on uncertainty. The **coupled** band adds a tunable layer for the correlated, cascading tail that a historical draw cannot produce: a Gulf shock co-moving across energy, shipping, food/fertilizer, goods passthrough, and the liquidity channel. The deep-uncertainty literature (forward scenarios over stale history, tail-dependent rather than linear dependence, decision-under-deep-uncertainty framings) motivates reporting *both* and treating the gap between them — not a single interval — as the deliverable.

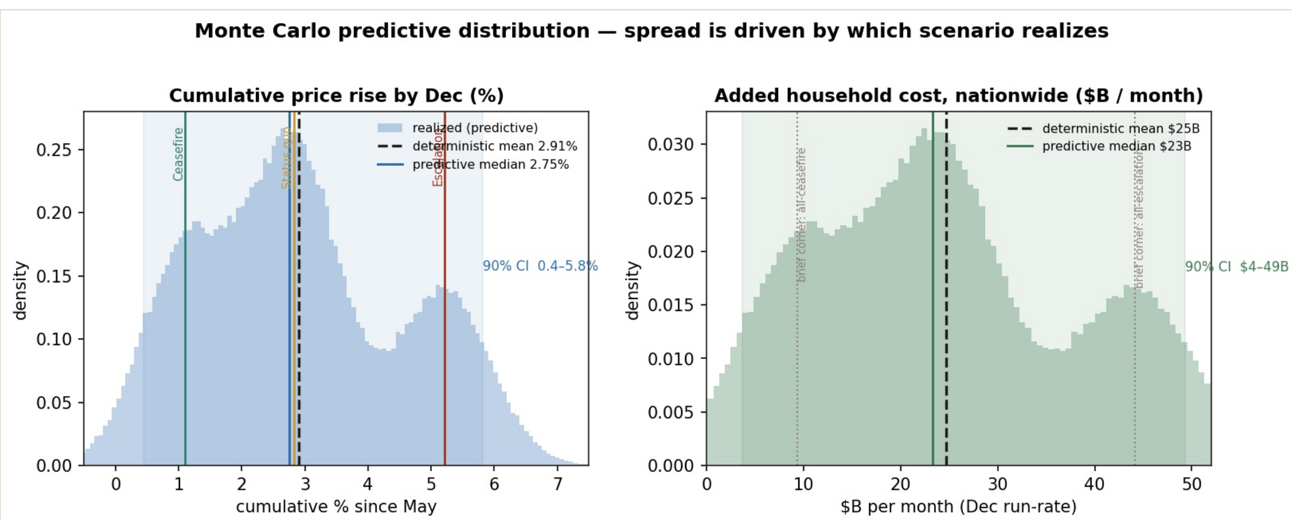
3.1 Method

Both bands share a scenario core: scenario weights are drawn from a Dirichlet centered on (0.30, 0.45, 0.25), a scenario is then drawn, and a persistent monthly level shock is added to that scenario's central path.

- **Floor.** Within-scenario monthly innovations are **independent Gaussian** (σ_level = 0.08 pp/mo, σ_month = 0.15 pp/mo, an order-of-magnitude calibration to recent monthly CPI dispersion), with the consensus basket (no broadening). Channels are, by construction, separable.
- **Coupled.** Monthly innovations are drawn from a **Student-t copula** (equicorrelation ρ = 0.35, dof ν_cop = 4 → tail-dependence coefficient λ = 0.18) with **fat-tailed Student-t marginals** (ν_marg = 5), so months go wrong together in the tail; a **crisis regime** (Bernoulli, probability `p_crisis`) multiplies path volatility by `crisis_mult`; and the energy-realist basket broadening is **state-dependent** — `bfadd_crisis` added to every bracket's basket factor only in a crisis draw, raising the burden-conversion factor K from $847B to ≈$1,002B per unit rate. A crisis draw is therefore hit on both the rate (deeper price rise) and the breadth (more of the basket exposed) — the non-separable "oil is the economy" channel, inside the framework rather than as an external multiplier. A rotated-Clayton copula (upper-tail-only dependence, λ_U = 0.71) is reported as a directional upper bound.

Adopted central knobs: **p_crisis = 0.12, crisis_mult = 2.5, bfadd_crisis = 0.18**, t-copula. The p_crisis value reflects that a full multi-channel cascade is the tail of the already-weighted escalation scenario (0.25), tempered by the current partial-disruption state; all three are bracketed in the sensitivity table (§3.5).

3.2 The floor distribution



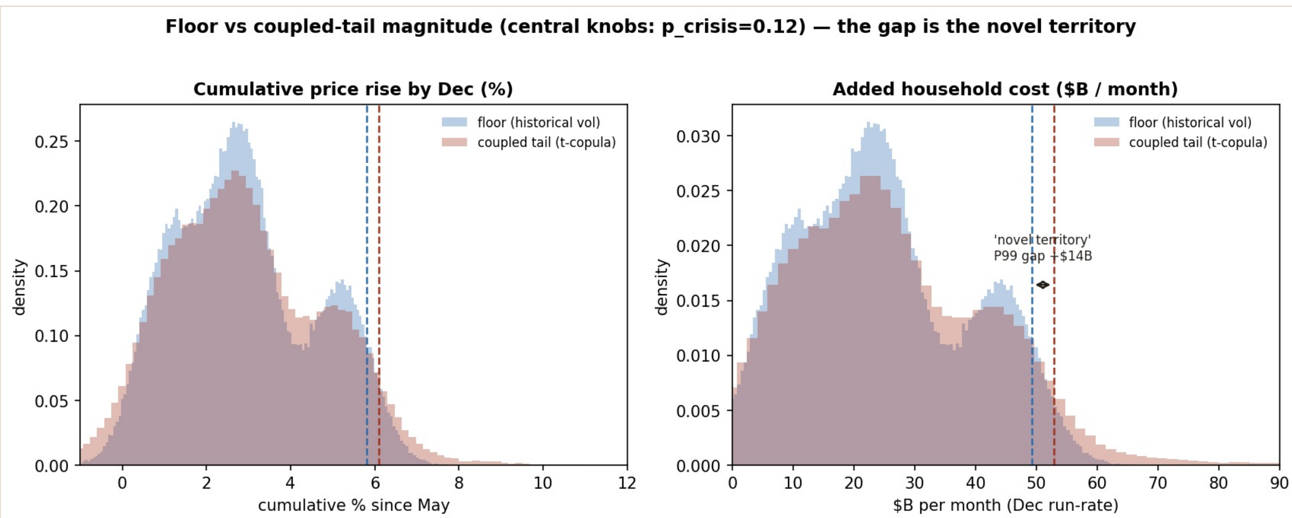
Floor predictive distribution with scenario modes

At the floor, the realized December cumulative rate has **mean 2.91% (matching the deterministic value by construction), median 2.75%, 90% interval [0.44%, 5.81%]**; the monthly aggregate burden has **median $23.3B,**

90% interval [$3.7B, $49.3B], 99th percentile **$54.8B**. Three findings about the floor matter more than the interval itself:

1. **The uncertainty is structural, not statistical.** A law-of-total-variance decomposition attributes **82.3%** of the variance to *which scenario realizes* and only **17.7%** to within-scenario path noise. The scenario *odds* are third-order: fixing the weights barely moves the interval ([0.42%, 5.82%] vs [0.44%, 5.81%]). Narrowing the forecast requires better information on the geopolitical path, not finer measurement.
2. **The deterministic envelope was not a probability interval.** About **28% of the floor mass falls outside the brief's [$9.4B, $44.2B] corners** (≈15% below, ≈13% above), because realized paths overshoot and undershoot their central scenarios. The corner range understated tail risk in both directions.
3. **The distribution is multimodal, so the mean misleads.** The density is trimodal; the expected value (+2.9%) sits in a low-density valley between the status-quo (~2.8%) and escalation (~5.2%) modes. P(rate > 4%) ≈ 26%. The honest headline is the median + interval + escalation tail probability, not the mean.

3.3 The coupled-tail band and the "novel territory"



Floor vs coupled-tail distribution; the P99 gap is the novel territory

Adding the coupling layer **leaves the center where it was** — median $23.3B → $23.7B, mean $24.7B → $25.3B — and **reweights the upper tail**, which is exactly where the deep uncertainty lives:

Metric ($B/month)	Floor	Coupled (t-copula)	Coupled (rotated-Clayton)
Median	23.3	23.7	23.5
90% interval	[3.7, 49.3]	[0.8, 52.9]	[−1.2, 55.9]
95th percentile	49.3	52.9	—
99th percentile	**54.8**	**68.7**	**75.3**
P(burden > $44B)	12.5%	14.8%	—
P(burden > $60B)	0.1%	**2.1%**	—

The **P99 gap of ≈ $14B/month** (t-copula; ≈ $20B for the directional Clayton bound) is the quantification of how much the historical calibration was understating the cascade. The probability of a > $60B/month outcome rises roughly twentyfold once coupling is admitted. The two-band gap, not either number alone, is the honest statement of what is knowable.

3.4 Where the uncertainty lives — and where it does not

The single most useful diagnostic in this exercise is *where* the tail risk is governed, because it tells the reader what the argument is actually about. Two of the knobs one might expect to matter do not:

- **The scenario odds are second-order** (§3.2): fixing vs drawing the weights barely moves the floor interval.
- **The copula tail-dependence parameter is second-order.** Holding the regime fixed, the coupled 99th percentile is essentially flat across the copula dof — ≈$69B at ν_cop = 200 (effectively Gaussian dependence), $67.6B at ν_cop = 4, $69.0B at ν_cop = 2.5 (variation within Monte Carlo noise). Once a discrete crisis regime is present, it already supplies most of the joint-tail co-movement, so the continuous copula family is close to irrelevant.

What *does* move the tail is the **crisis-regime probability and severity, and the breadth of transmission** (p_crisis, crisis_mult, bfadd_crisis; §3.5). This locates the honest disagreement on terrain a reader can reason about — how likely a full cascade is, how violent it gets, and how far beyond direct energy it spreads — rather than on a copula choice no one can adjudicate. We give this prominence deliberately: it is a more defensible place to put the uncertainty than a false-precision dependence parameter.

3.5 Knob sensitivity (coupled, aggregate P99, $B/month)

Knob	low	central	high
p_crisis	0.08 → 65.3	**0.12 → 68.3**	0.20 → 74.5
crisis_mult	2.0 → 65.1	**2.5 → 68.5**	3.0 → 73.9
bfadd_crisis	0.15 → 67.0	**0.18 → 69.2**	0.25 → 72.4
ν_cop (tail-dep)	200 → 69.4	**4 → 67.6**	2.5 → 69.0

The center, the 82/18 structural split, and the "mass outside the corners" finding are robust throughout. Only the upper-tail height moves with the regime/breadth knobs; the copula dof does not.

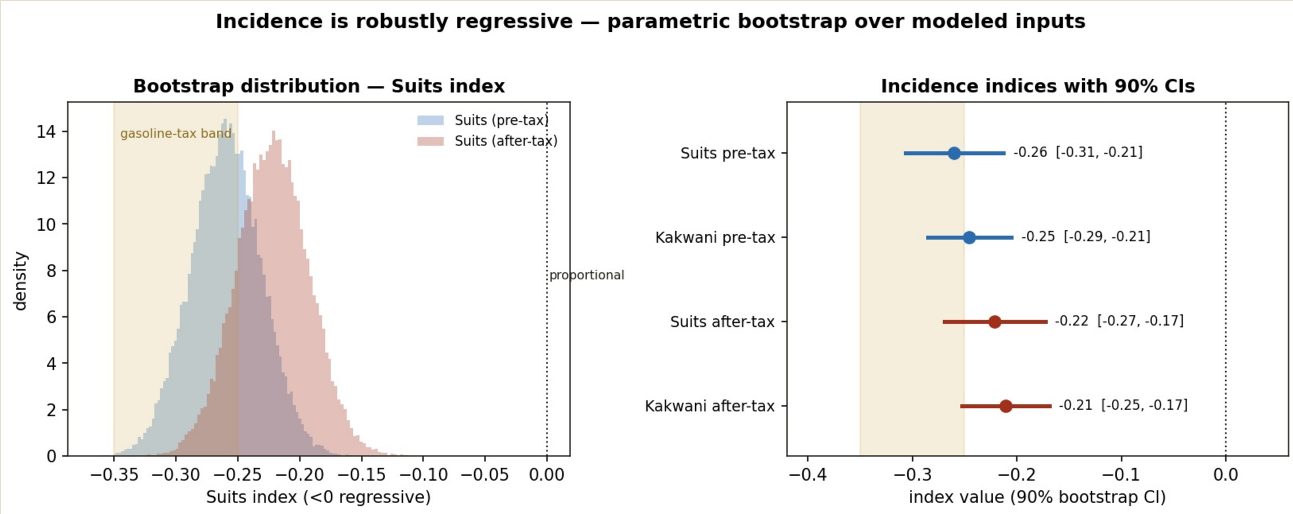
4. Incidence under uncertainty (bootstrap)

4.1 Method

The incidence indices depend on household counts and spending (sourced) and on basket factors, representative bracket incomes, and effective tax rates (modeled). We propagate input uncertainty by a parametric bootstrap: basket factors drawn lognormally (6% CV), representative incomes lognormally (10% CV, which spans the earlier robustness range for the open-ended top and bottom brackets), effective tax rates ±2 pp, and household counts/spending given small survey-sampling noise (2.5% / 3% CV). We recompute Suits and Kakwani on both income bases for each of 100,000 replicates.

This is **not** a microdata resample — we do not have the SCF/CE public-use files or a way to retrieve them. It quantifies *modeled-input* uncertainty, which dominates given the bracket structure, but it does not capture *sampling* uncertainty within the survey. Section 4.3 specifies the proper microdata version.

4.2 Results



Incidence indices with bootstrap confidence intervals

Index	Base	Point	90% CI	P(regressive)
Suits	pre-tax (headline)	−0.260	(−0.307, −0.213)	**1.000**
Kakwani	pre-tax	−0.246	(−0.285, −0.205)	**1.000**
Suits	after-tax	−0.222	(−0.270, −0.173)	1.000
Kakwani	after-tax	−0.211	(−0.252, −0.169)	1.000

The shock is regressive in **every** bootstrap replicate on both income bases. The pre-tax Suits interval (−0.31, −0.21) overlaps the published gasoline-excise-tax band (≈ −0.25 to −0.35) on its more-regressive half and is bounded well away from zero (proportional) and from the progressive sign of an income tax. The regressivity conclusion is therefore not an artifact of any single modeled input.

One honesty note on invariance. The indices are exactly invariant to the inflation *rate* — it cancels from the burden shares — so the entire Monte Carlo of §3 leaves them untouched. The crisis-state basket broadening of the coupled band (§3.1) is the *one* channel that does enter the indices, because it shifts the basket factor. Its effect is a mild attenuation, not a reversal: pre-tax Suits moves −0.260 (no broadening) → −0.250 (+0.12) → −0.246 (+0.18) → −0.239 (+0.30), because a uniform additive broadening dilutes the basket-factor gradient that drives regressivity. The sign and rough magnitude are preserved across the entire plausible broadening range; we state this as bounded attenuation rather than claim pure invariance.

4.3 What the microdata version would add

The definitive estimate would resample households (with survey weights) from the 2022 SCF and the CE public-use microdata, recompute the index per replicate, and report BCa intervals — capturing within-survey sampling uncertainty that the parametric bootstrap omits. Because grouping into seven point-masses understates within-population inequality, both the central indices here and a microdata estimate are conservative (true values modestly more negative). We expect a microdata estimate to be slightly *more* regressive with a comparable or modestly wider interval; it would not overturn the sign, which is established here at probability 1 over modeled-input uncertainty.

5. Causal incidence of the triage ordering: an event-study design

The brief's Figure 6 — the order in which strained households drop obligations (utilities and housing last; credit cards and federal student loans first) — is currently descriptive, resting on counseling practice and the cross-sectional fact that lower-income borrowers are more often delinquent. That correlation pre-dates any shock and cannot establish that *this* shock causes the predicted ordering. The following design would.

Question. Does an exogenous energy-price shock causally raise delinquency, and does it do so in the hypothesized priority order (low-priority unsecured debt first), more so for higher-energy-exposure households?

Data. A household- or county-level panel of delinquency by trade line — e.g., the New York Fed / Equifax Consumer Credit Panel or CFPB Consumer Credit Trends — merged with local energy-price exposure (regional retail fuel prices; baseline energy budget share by area or income from the CE). Historical energy-shock dates supply the events (e.g., 1990 Gulf, 2008 spike, 2022 Russia–Ukraine).

Specification (stacked event study).

```
D[i,t] = α_i + λ_t + Σ_k β_k · ( Exposure_i × 1[t = t0 + k] ) + X[i,t]·γ + ε[i,t]
```

where `D` is a delinquency outcome (e.g., 30+ DPD indicator) for unit `i` in period `t`; `α_i` and `λ_t` are unit and time fixed effects; `Exposure_i` is baseline energy budget share (or income bracket as proxy); `1[t = t0 + k]` are event-time indicators relative to the shock; and `X` collects controls. The dynamic coefficients `β_k` trace the causal path. Leads (`k < 0`) test for pre-trends (a placebo). With staggered events and heterogeneous effects, estimate with a Callaway–Sant'Anna or Sun–Abraham estimator rather than two-way FE to avoid negative-weighting bias.

The triage test. Estimate the equation separately by debt type (student, credit card, auto, mortgage). The hypothesis predicts a specific ordering: `β_k` for student loans and credit cards should be **larger and emerge earlier** than for auto loans and mortgages, and the effect should rise with `Exposure`. A formal test compares the post-event coefficient magnitudes across debt types; the triage claim is **falsified** if mortgages/auto deteriorate as fast as or faster than unsecured debt, or if the exposure gradient is flat.

Threats to identification. Energy exposure correlates with income and region; absorb with income-×-time and region-×-time fixed effects. Energy-producing regions have co-moving labor demand; control region-×-industry. Pre-trends are checked directly via the leads. Anticipation is limited by using abrupt geopolitical onsets.

This is a credible, pre-specifiable design. We flag clearly that it is **a design, not a result** — running it requires the panel data above.

6. Limitations

- The magnitude is reported as two bands. The **floor** is a historical-volatility lower bound; the **coupled** band's tail height scales with the coupling knobs (p_crisis, crisis_mult, bfadd_crisis), which are *assumptions, not fitted parameters* — the honest object is the gap between the bands, and the sensitivity table (§3.5) is the appropriate response. The copula family and the scenario odds are second-order (§3.4).
- The incidence indices are invariant to the inflation rate but **mildly attenuated by the crisis-state breadth** (Suits −0.26 → −0.24 across the plausible broadening range), so the correct claim is bounded attenuation without sign reversal, not pure invariance (§4.2).
- The incidence bootstrap captures modeled-input but not survey-sampling uncertainty (§4.3); the sign is nonetheless established at probability 1.
- Both incidence indices rest on seven grouped brackets and on representative incomes and a flat effective-tax schedule per bracket; grouping makes them conservative.
- The burden is a compensating-variation upper bound (no substitution, no nominal-wage offset). The analysis is partial-equilibrium and prices no policy response — the Fed reaction remains the largest unmodeled swing factor.
- The event-study (Section 5) is unestimated.

7. Conclusion

The brief's central claims survive the upgrade and are strengthened by it. The regressivity result is now an interval estimate that is regressive with bootstrap probability one and benchmark-comparable to a gasoline excise tax. The magnitude is now a proper predictive distribution rather than a corner range, reported as two bands: a most-likely status-quo center near $23B/month that is stable across the analysis, and an upper tail whose height is the live question — a historical-volatility floor (99th percentile $54.8B) versus a coupled-tail band that fattens it to ≈$69B (directional bound $75B) and makes a > $60B/month outcome roughly twentyfold more likely. The gap between those bands is the honest measure of the cascade, and the analysis locates the disagreement where it belongs: on how likely and how broad a cascade is, not on a copula choice or the scenario odds, both of which are second-order. The triage ordering remains the one genuinely causal claim, and it now has a design rather than only an assertion. Read as a probabilistic, deep-uncertainty-aware exposure-and-incidence study with an explicit causal agenda, the work is on research footing.

Appendix A — Parameters and sensitivity

Shared scenario core & floor

Component	Parameter	Value	Role
Weights	Dirichlet concentration κ	20	odds uncertainty (second-order)
Floor path noise	persistent level σ	0.08 pp/mo	compounding trajectory shift
Floor path noise	transitory σ	0.15 pp/mo	month-to-month surprise
Bootstrap	basket-factor CV	6%	modeled input
Bootstrap	representative-income CV	10%	modeled input (spans robustness range)
Bootstrap	effective-tax SD	2 pp	modeled input
Bootstrap	hh / spend CV	2.5% / 3%	survey sampling proxy

Coupling layer (adopted central; range)

Knob	Central	Range	Role
p_crisis	0.12	0.08–0.20	probability of a multi-channel cascade regime
crisis_mult	2.5	2.0–3.0	path-volatility multiplier in crisis
bfadd_crisis	0.18	0.15–0.25	state-dependent basket broadening in crisis
copula	Student-t, ν_cop = 4	—	tail dependence λ = 0.18 (second-order)
ρ (equicorrelation)	0.35	—	linear month co-movement
ν_marg	5	—	fat-tailed marginals
Clayton θ (variant)	2	—	rotated → upper-tail λ_U = 0.71 (directional bound)

Coupled aggregate **99th percentile ($B/month)** under one-knob perturbation about the central: p_crisis 0.08 → 65.3, **0.12 → 68.3**, 0.20 → 74.5; crisis_mult 2.0 → 65.1, **2.5 → 68.5**, 3.0 → 73.9; bfadd_crisis 0.15 → 67.0, **0.18 → 69.2**, 0.25 → 72.4; ν_cop 200 → 69.4, **4 → 67.6**, 2.5 → 69.0 (flat — copula dof is second-order). Floor aggregate 90% CI under alternative path volatility: σ_month 0.10 pp → [$4.3B, $48.8B]; 0.15 pp → [$3.6B, $49.5B]; 0.30 pp → [$1.3B, $51.3B].

Appendix B — Reproducibility

The floor distribution and the bootstrap are produced by `research_analysis.py`; the two-band coupled-tail results by `coupled_tail_mc.py` (both seed 20260605). `research_results.json` and `coupled_mc_results.json` record every reported statistic, and `coupled_mc_raw.npz` holds the raw 200,000-draw vectors (floor, t-copula, rotated-Clayton; rate and aggregate) so any percentile here can be recomputed directly rather than trusted. Scripts are self-contained (numpy/scipy/matplotlib) and regenerate all figures. Independently reproduced across three seeds and a separate reimplementation; headline figures match to rounding.