MacKenzie Scott's giving, in QALYs

An evidence-weighted, auditable cost-effectiveness model of a $26 billion portfolio

Max Ghenis*

MacKenzie Scott has disclosed $26.39 billion in gifts to more than 2,500 organizations since 2019 — her 2025 giving alone was about a third of the value of that year's U.S. megagifts — yet almost none of it is denominated in outcomes. This paper prices the entire portfolio in quality-adjusted life-years. A Monte Carlo model allocates dollars across 14 intervention archetypes using Scott's own gift-level database, imputes the undisclosed third from recipients' pre-gift IRS 990 revenue, assigns each archetype a cost-per-QALY anchor from published causal estimates, and shrinks every effect toward zero in proportion to the causal credibility of its evidence. The best-guess estimate is a median of roughly 202,000 QALYs (90% interval 94,000–419,000), a blended $150,000 per QALY, and a median benefit–cost ratio of 4.9 under the valuation distribution HHS uses for regulatory analysis. Three findings follow. First, estimated impact is geographically concentrated: the 4.8% of dollars funding health and development delivered outside the United States, priced at global-health anchors, accounts for 69% of estimated health impact. Second, gift size scales with recipient size to only the 0.41 power — a 10× larger organization receives about 2.5× more money. Third, the answer is dominated by evidence philosophy rather than arithmetic: restricting credit almost entirely to randomized evidence collapses the median to ~9,000 QALYs while taking every study at near face value raises it to ~415,000 — a 45× range, larger than the spread from any single distributional parameter tested. All parameters, code, data snapshots, and an AI-audit trail are public, and the model runs interactively in the browser.

Source: Article Notebook

1 Introduction

MacKenzie Scott gave away $7.17 billion in 2025 (Huddleston 2025); Giving USA counts $6.65 billion of it against $19.2 billion in U.S. megagift dollars — about a third of that year's megagift value (Indiana University Lilly Family School of Philanthropy 2026; Childress 2026). (The bases

*max@maxghenis.com. I am CEO of PolicyEngine and president of the Axiom Foundation; this paper is a personal research project and does not represent the views of either organization. Neither I nor the organizations I lead have received funding from MacKenzie Scott, Yield Giving, or GiveWell. Interactive version, code, data, and full audit trail: https://maxghenis.com/mackenzie-scott-qaly and https://github.com/MaxGhenis/mackenzie-scott-qaly.

differ: Yield Giving's own annual total is the larger figure.) Her disclosed lifetime total is $26.39 billion across 2,711 gifts to 2,545 organizations since 2019 (Yield Giving 2026). Almost none of it is denominated in health, and, to my knowledge, no published estimate prices the portfolio in outcomes of any kind.

This paper does so in the unit health economists use to compare a death averted against years lived in better health: the quality-adjusted life-year. The exercise is deliberately in the style of a charity evaluator's cost-effectiveness analysis (GiveWell 2025), scaled from a program to a portfolio, with one addition the setting forces: because most of the portfolio funds interventions whose health effects have never been causally measured, the model makes its epistemics explicit. Every effect is shrunk toward zero in proportion to how well its supporting study identifies causation, the shrinkage schedule is a slider rather than a buried constant, and the headline is reported across the full range of evidence philosophies — from counting only randomized evidence to taking every published estimate at face value.

The paper makes four contributions. It provides what I believe is the first outside, portfolio-level cost-effectiveness estimate of Scott's giving, and I have found no comparable outside estimate for any other individual mega-donor — funder-run portfolio metrics exist, including GiveWell's aggregate impact estimates (GiveWell 2025), Robin Hood's benefit–cost framework (Robin Hood Foundation 2023), the Constellation Fund's portfolio reporting (Constellation Fund 2026), and social-rate-of-return analysis of USAID's Development Innovation Ventures portfolio (Center for Effective Global Action 2022); the difference here is independence from the donor, a subject who publishes no outcome data, and explicit evidence weighting: a best-guess median of 202,000 QALYs from $26.39 billion nominal ($30.3B in 2026 dollars), a blended $150,067 per QALY, and a median benefit–cost ratio of 4.9 under the HHS regulatory value-per-QALY distribution (Kearsley 2026). It estimates extreme geographic concentration of impact under its pricing assumptions: the 4.8% of dollars funding global health and development delivered outside the United States accounts for 69.1% of estimated health impact. It reports a new empirical regularity in the gift data itself: across 1,313 disclosed gift–filing pairs, gift size scales with recipient pre-gift revenue to the power 0.41 (R^2 0.37) — Scott's giving is far flatter across organization size than proportional. And it demonstrates a working method for auditable, AI-assisted cost-effectiveness modeling: the model was built with a frontier language model, reviewed and corrected in public through the revision history summarized in the robustness section, audited by two independent model families against primary filings, and ships with versioned parameters, machine-checked citations, and a browser implementation that reruns the full Monte Carlo under any reader's assumptions.

The unit of account deserves a caveat before any number. QALYs measure health. Most of Scott's portfolio aims at other things — education, economic mobility, equity, civic infrastructure — whose primary benefits are not health benefits. A finding that a dollar of education philanthropy produces few QALYs is a statement about the lens, not a verdict on the gift. The estimates are a partial, health-only accounting under positive-effect priors — not a lower bound on total wellbeing: the lens omits the primary intended benefits of most of the portfolio, and the model places no probability on zero or harmful effects, a structural choice the limitations section returns to.

2 Data

2.1 The ledger

Scott's essays disclose giving in dated tranches totaling $26.39 billion nominal over 2020–2025. The model inflates each tranche — booked by disclosure year — to 2026 dollars (CPI-U, anchored to May 2026), so dollars and cost-effectiveness evidence share one price level: $30.3B total. All derivations below, and the replication repository, start from a verbatim snapshot of the Yield Giving gift database retrieved July 11, 2026 (Yield Giving 2026).

2.2 Allocation across causes

Scott publishes no aggregate dollars-by-cause breakdown, but she publishes something more granular: a gift-level database listing every recipient organization, its self-reported focus areas, and — for 2,035 of 2,711 gifts ($17.46B, about two-thirds of the nominal total) — the dollar amount. Each organization's dollars are split equally across its focus areas; the database's 53 areas map onto the model's 14 intervention archetypes under a documented, sync-tested crosswalk.

The undisclosed third is imputed rather than dropped. Because the essays give each year's total, the undisclosed residual per period is a known dollar amount; it is distributed across that period's undisclosed gifts in proportion to each recipient's pre-gift IRS Form 990 total revenue (ProPublica 2026), raised to an elasticity fit on the disclosed pairs. Pre-gift means the latest filing strictly before the gift year, since a Scott gift mechanically inflates receipt-year revenue. The fit is a finding in its own right: across 1,313 disclosed gift–filing pairs, log gift size rises with log revenue with slope 0.41 (R^2 0.37). A 10× larger organization receives about 2.5× more money, not 10× more. The fit is descriptive and conditional: it is estimated on disclosed gifts with a matched pre-gift filing (1,313 of 2,035 disclosed and 2,711 total gifts), treats repeated gifts to an organization as independent, and pools years; 287 of the 676 undisclosed gifts lack a usable filing and receive their period's median weight. It is a within-sample regularity used for imputation, not a validated prediction rule. Fuzzy name-to-EIN matches were audited by a third, smaller model against the live IRS API, with corrections committed as a reviewable overlay. The imputed third moves each cause share by at most about 1.5 percentage points — under the imputation model, the disclosed two-thirds is representative.

2.3 Delivery geography

Each organization reports service locations. Weighting disclosed dollars by the audited share of each organization's locations outside the United States, an estimated 19% of disclosed dollars ($3.34B of $17.46B) sits with organizations reporting non-U.S. service locations — 20.4% under the raw rule that counts a bare "global" listing as fully non-U.S. — led by environment, health, and funding intermediaries. Locations are an equal-split proxy: a listed location is not an expenditure share, and no organization reports how Scott's own gift was deployed. The model prices geography through a single channel: health-relevant dollars route to a global-health archetype in proportion to the recipient's non-U.S. service share, because global health delivery

is the one category with a published cost-per-QALY evidence base an order of magnitude below U.S. anchors. Education and climate dollars keep their U.S.-study anchors wherever they are delivered; if geography changes their effectiveness, the model does not capture it.

The coarsest geographic field is the bare label "global." An audit of the 50 organizations with the most dollars attached to that label — 87% of its dollars; 20 of the 50 also list other locations — replaced the blanket assumption of fully-abroad delivery with fractions coded independently by two model families from each organization's filings and program reports, then reconciled — the second coder confirmed 33 within tolerance and disputed 17; four disputes were adjudicated to documented filings arithmetic, and final fractions average the two coders everywhere else, with per-organization sources and confidence grades — 16 high, 17 medium, 17 low — committed to the repository. Many "global" organizations operate substantially in the United States: Planned Parenthood and Crisis Text Line, for example, list "global" service but are roughly 95% and 93% U.S. in their most recent organization-wide reporting — later-period proxies for how earlier gifts were deployed, since neither reports gift-level allocation. The audited allocation routes 4.82% of the ledger to the global health and development archetype. That routed bucket is itself mixed: its dollars sit 42% in Sub-Saharan Africa, 20% in South Asia, and 11% in Latin America, but 18% is unattributable multi-region activity and roughly 4% sits in Europe & Central Asia plus Canada — regions that are not classified by income here, so the bucket's true high-income share is unidentified — all priced at the same low-and-middle-income anchors, a coarseness the limitations section flags.

3 Model

Each Monte Carlo draw allocates the giving across archetypes (a Dirichlet centered on the derived shares), assigns each archetype a cost per QALY drawn from a distribution anchored in published estimates, and multiplies the implied QALYs by two independent discounts. Causal credibility asks whether the measured effect is real. Each archetype carries the identification tier of its best supporting evidence, and the tier sets a Beta-distributed shrinkage factor toward the null: a lottery-randomized wealth shock (Cesarini et al. 2016) or the Medicaid mortality quasi-experiments — the ACA expansions (Miller et al. 2021), with cost–benefit arithmetic from the pre-ACA state expansions (Sommers 2017) whose spending component draws on the Oregon Health Insurance Experiment (Finkelstein et al. 2012) — keeps most of its effect; a child-targeted cash quasi-experiment (Aizer et al. 2016), with related designs catalogued in the parameter file, anchors the cash bucket; the community-health-center rollout design (Bailey and Goodman-Bacon 2015) and similar contested quasi-experiments keep less; associational and assumption-only categories are shrunk near zero. Realization asks whether a marginal philanthropic dollar delivers the studied effect at all — transport to different populations, overhead, and funging, net of possible capacity effects: surveyed Scott recipients report stronger capacity and reach (Buteau et al. 2023), self-reports on whose basis the model permits realization above one (an author judgment, not a CEP effect estimate). It is one global triangular draw with mode 0.80 on support 0.55–1.10.

The two multipliers are designed to partition the doubt — transportability concerns in realization, identification concerns in credibility, cost-per-QALY priors stated conditional on the effect being real and delivered — though the separation is imperfect at the prior boundaries: several anchors

already average over mixed null-and-positive evidence before the credibility discount applies (a limitation noted below). Allocation uncertainty is a Dirichlet with concentration 60 around the derived centers — a judgment prior, implying for example a roughly 1.3–10% ninety-percent range on the global-health share — and represents subjective allocation uncertainty rather than sampling uncertainty propagated from the gift data pipeline. The tier ordering is the defensible part; the levels are judgment, proposed by a language model and reviewed by the author, and the results section reports the full consequence of disagreeing with them.

Health effects arrive in different natural units — lives saved, life-years, depression-free days — and are converted to discounted QALYs at a 3% annual rate under source-specific conventions: adult mortality anchors sample roughly 26 remaining life-years at a utility draw centered near 0.78, the frontier's child-mortality benchmark uses a fixed ~65 remaining years at utility 0.87, and direct cost-per-QALY anchors keep their sources' conventions. Every parameter's unit and dollar vintage is machine-validated at load. The global-health frontier benchmark is handicapped within the same framework: GiveWell's reported recent program averages (GiveWell 2025), converted to roughly $175–$241 per discounted QALY-equivalent under this model's own conventions (a model conversion, not a GiveWell QALY figure), modeled as loguniform $150–$260, sharing each draw's realization multiplier, and receiving RCT-grade credibility — the strongest tier — where portfolio archetypes receive their own.

4 Results

Source: Article Notebook

Source: Article Notebook

At the best-guess defaults (100,000 draws), the portfolio produces a median of 201,923 QALYs, with a 90% interval of 94,000 to 419,000 and a mean of 222,060. The interval is a probabilistic-sensitivity range under the stated priors, not a confidence interval; Monte Carlo error on the median is a few hundred QALYs. The blended cost-effectiveness — dollars given over QALYs produced, at the median — is $150,067 per QALY. Monetized under the value-per-QALY distribution HHS recommends for regulatory analysis (central $726,000 at 3%; Kearsley (2026), Table 3), the median draw-level benefit–cost ratio is 4.9 (90% interval 2.1–11.0, reflecting both health-output and valuation uncertainty). Figure 1 shows the full distribution.

4.1 Where it comes from

Table 1 decomposes the mean estimate by cause. The result is the paper's central contrast: the global health and development slice — 4.8% of dollars — contributes 69.1% of estimated QALYs, at a modeled cost of about $9,477 per QALY against $75,673–$305,164 for the causes with quasi-experimental or better U.S. evidence. The 1,278-fold spread in modeled cost-effectiveness within a single donor's portfolio, under stipulated priors, is partly the lens speaking — its low end is dominated by assumption-tier buckets whose near-zero credibility is an input, and whose main benefits QALYs do not count — but the spread among evidenced health buckets alone exceeds an order of magnitude, restating in one column an empirical regularity long emphasized

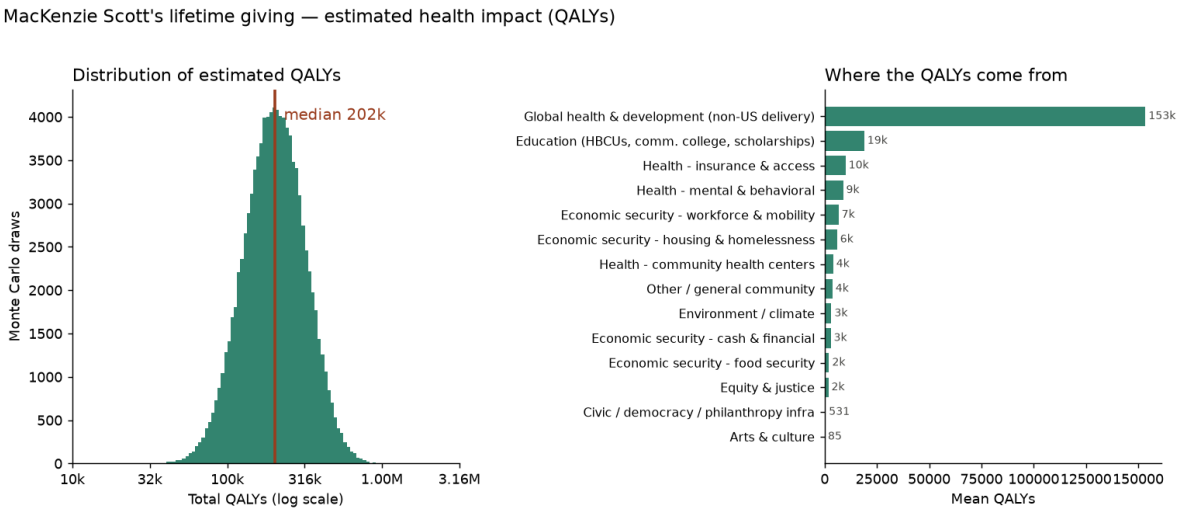

Figure 1: Distribution of total QALYs across 100,000 Monte Carlo draws at the best-guess defaults. The dashed line marks the mean; the solid line the median.

in the effective-altruism literature (Ord 2013): the cheapest health in the world is not bought where most philanthropic dollars are spent.

Source: Article Notebook

Table 1: Dollars and estimated QALYs by cause, best-guess defaults. Dollars apply the allocation centers to the 2026-dollar total; QALYs are mean contributions (only means are additive). $/QALY is the ratio of center dollars to mean QALYs after credibility and realization adjustments — a ratio of expectations, not the raw study anchor.

Cause	Dollars	% of $	QALYs	% of QALYs	$/QALY	Evidence
Global health & development (non-US delivery)	$1.45B	5%	153,469	69%	$9,477	strong_quasi
Education (HBCUs, comm. college, scholarships)	$5.70B	19%	18,939	9%	$300,788	moderate_quasi
Health - insurance & access	$1.12B	4%	9,977	4%	$112,381	strong_quasi
Health - mental & behavioral	$667M	2%	8,810	4%	$75,673	randomized
Economic security - workforce & mobility	$4.06B	13%	6,632	3%	$612,212	observational
Economic security - housing & homelessness	$1.09B	4%	5,918	3%	$184,341	moderate_quasi
Health - community health centers	$1.12B	4%	4,062	2%	$276,005	moderate_quasi
Other / general community	$1.67B	6%	3,708	2%	$449,423	observational
Environment / climate	$2.48B	8%	3,193	1%	$778,250	projection

Table 1: Dollars and estimated QALYs by cause, best-guess defaults. Dollars apply the allocation centers to the 2026-dollar total; QALYs are mean contributions (only means are additive). $/QALY is the ratio of center dollars to mean QALYs after credibility and realization adjustments — a ratio of expectations, not the raw study anchor.

Cause	Dollars	% of $	QALYs	% of QALYs	$/QALY	Evidence
Economic security - cash & financial	$879M	3%	2,880	1%	$305,164	strong_quasi
Economic security - food security	$576M	2%	2,093	1%	$275,096	observational
Equity & justice	$5.21B	17%	1,763	1%	$3M	assumption
Civic / democracy / philanthropy infra	$3.24B	11%	531	<1%	$6M	assumption
Arts & culture	$1.03B	3%	85	<1%	$12M	assumption
Total	$30.30B	100%	222,060	100%	$136,458	

Source: Article Notebook

Source: Article Notebook

Table 2 decomposes the same run by delivery geography. The table is a by-construction decomposition, not evidence of regional differences: within each cause the model prices every delivery location identically — the health routing to global anchors is its only geographic pricing — so a cause's QALYs follow its dollars across regions. Under that construction, Sub-Saharan Africa receives 6% of dollars and 30% of estimated QALYs at $27,031 per QALY; nationally-scoped U.S. giving receives 27% of dollars and 9% of QALYs at $403,000.

Source: Article Notebook

Table 2: A deterministic post-estimation decomposition of mean archetype QALYs by delivery geography — no region-level effect or uncertainty is estimated, and geography affects pricing only through the global-health routing. Asterisked rows are reporting granularity (national or multi-region listings), not places. Geography derives from each organization's reported service locations, equal-split, with audited fractions for the 50 largest "global"-listed organizations.

Region	Dollars	% of $	QALYs	% of QALYs	$/QALY
Sub-Saharan Africa	$1.81B	6%	67,007	30%	$27,031
South Asia	$1.01B	3%	32,006	14%	$31,577
Global (multi-region) *	$1.53B	5%	28,288	13%	$54,164
US — national / unspecified *	$8.24B	27%	20,443	9%	$403,000
Latin America & Caribbean	$1.11B	4%	17,472	8%	$63,620
US — South	$5.96B	20%	16,041	7%	$371,678
US — West	$4.02B	13%	10,966	5%	$366,611

Table 2: A deterministic post-estimation decomposition of mean archetype QALYs by delivery geography — no region-level effect or uncertainty is estimated, and geography affects pricing only through the global-health routing. Asterisked rows are reporting granularity (national or multi-region listings), not places. Geography derives from each organization's reported service locations, equal-split, with audited fractions for the 50 largest "global"-listed organizations.

Region	Dollars	% of $	QALYs	% of QALYs	$/QALY
US — Midwest	$2.75B	9%	7,828	4%	$351,406
US — Northeast	$2.20B	7%	5,631	3%	$390,341
Europe & Central Asia	$299M	1%	5,453	2%	$54,776
Middle East & North Africa	$153M	1%	4,177	2%	$36,670
East Asia & Pacific	$335M	1%	3,339	2%	$100,390
North America (unspecified) *	$698M	2%	1,535	1%	$454,724
Other / unmapped *	$40M	<1%	1,435	1%	$27,621
Canada	$142M	<1%	439	<1%	$323,208

Source: Article Notebook

4.2 Against the global-health frontier

Handicapped within the same framework — sharing each draw's realization multiplier, with RCT-grade credibility — the GiveWell-derived frontier benchmark delivers roughly 521× more health per marginal dollar than the blended portfolio; the same sum directed entirely to the frontier would imply on the order of 105.5 million QALYs. That figure is arithmetic, not a forecast, and not an attainable counterfactual: frontier-priced opportunities are scarce — GiveWell's directed funding runs to hundreds of millions of dollars per year, not tens of billions (GiveWell 2025) — and funding beyond identified room drives toward cash-transfer cost-effectiveness, whose GiveWell estimate rose 3–4× in a 2024 re-evaluation (GiveWell 2024). Converting the full portfolio at cash-transfer cost-effectiveness under this model's own conventions leaves a multiple in the tens — roughly 20–40×, model arithmetic rather than a GiveWell figure. The defensible statement is directional: the frontier multiple is several hundred at today's margin and remains large at any scale the cited benchmarks speak to.

5 Robustness

5.1 Evidence philosophy dominates everything else

Source: Article Notebook

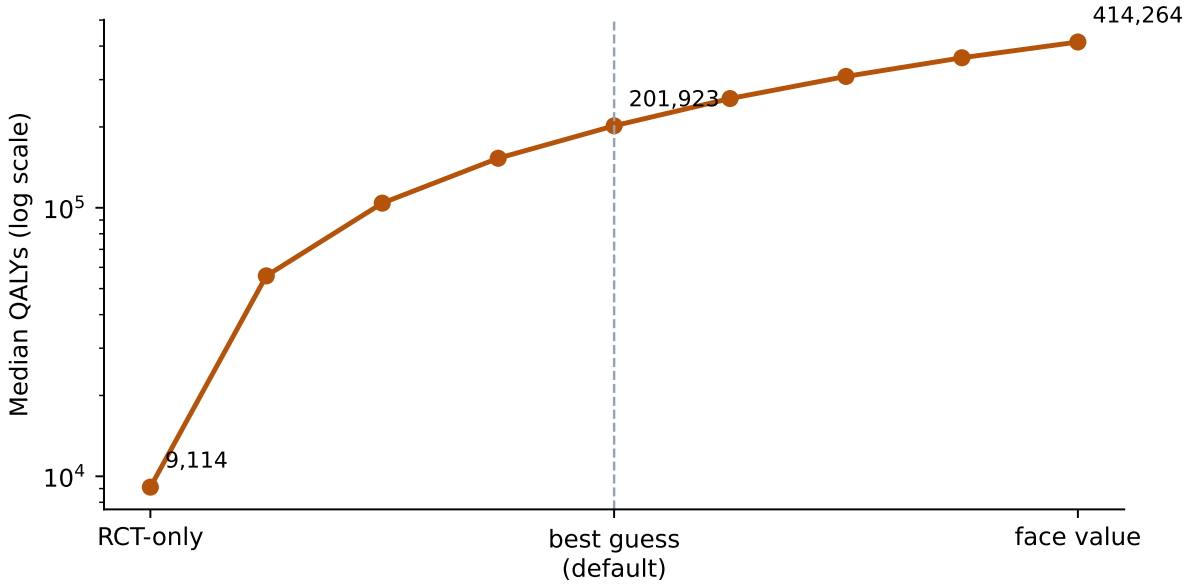

Figure 2: Median total QALYs across the evidence-stance range, 100,000 draws per point, seed 0. The stance interpolates each tier's credibility mean piecewise-linearly from a near-RCT-only floor (non-randomized tiers at mean 0.01; randomized tier at its drafted mean) through the drafted best-guess tiers (0.5) to near face value (all means 0.999). A scenario path, not a bound; tier concentrations are held fixed.

Source: Article Notebook

The dominant sensitivity is not a parameter but a philosophy. At the near-RCT-only floor — non-randomized tiers shrunk to a credibility mean of 0.01, the randomized tier kept at its drafted mean — the median collapses to 9,114 QALYs, dominated by the randomized-tier mental-health bucket; at near face value (all tier means 0.999) it rises to 414,264 (Figure 2). That 45-fold range exceeds the spread from any single distributional parameter — though structural choices can move the headline more: with no geographic repricing at all, every health dollar at U.S. anchors, the median is roughly 70,000 (the pre-correction model in the §5.3 table), so the routing assumption alone spans ~3× — and the gap between the face-value and best-guess estimates is a direct reading of how much of the estimated impact rests on evidence below randomized or strong quasi-experimental identification standards. The Spearman tornado (Figure 3) tells the same story within the default stance: the top three drivers are the global-health allocation share, the realization factor, and global-health credibility.

5.2 The audit moved the headline by less than 2%

The geography audit is the largest data correction the model has absorbed since publication, and its effect is small: recentering the allocation on audited service geographies moved the abroad share from 4.93% to 4.82%, the median from about 205,000 to 201,923, and the blended figure from about $148,000 to $150,067 per QALY. Both allocations ship with the model, the browser

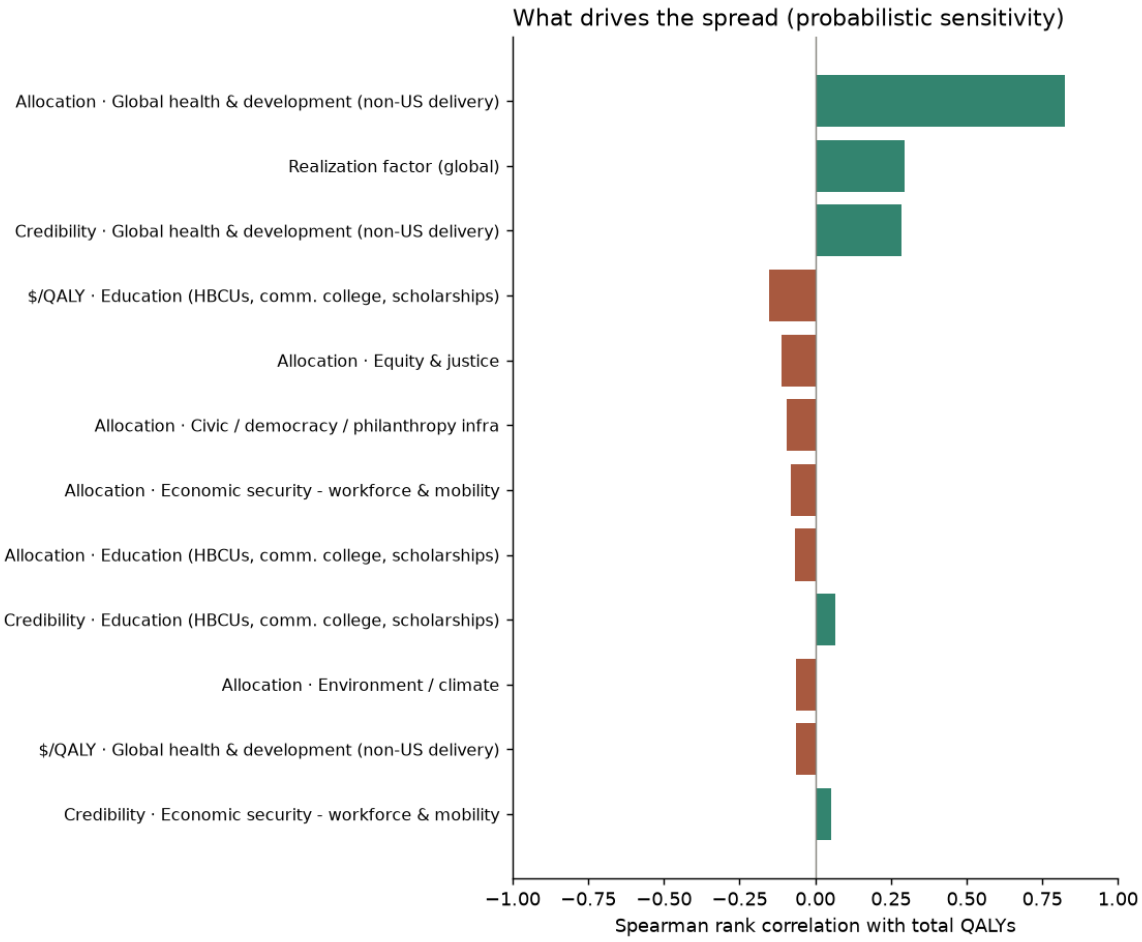

Figure 3: What drives the spread: Spearman rank correlations between each input draw and total QALYs at the best-guess defaults.

version exposes them as a version toggle, and the pre-audit parameters remain pinned at tag `v2026.07.20` (audited: `v2026.07.22`).

5.3 Corrections history

The model reached this state through a series of public revisions — counting a revision as a committed change that moved the headline — and the trajectory is itself evidence about the method. As recorded in the repository history:

Change	Median before → after	Caught by
Cross-model review: dollar vintages normalized, a life-year→QALY conversion made explicit, frontier benchmark re-derived, one citation re-attributed	~98,000 → ~87,000	second model family
Allocation replaced with the derivation from Scott's own gift database	~87,000 → ~70,000	author + model
Health dollars had been priced at U.S. anchors regardless of delivery location	~70,000 → ~205,000	a reader
Two-model geography audit of "global"-listed organizations (tags v2026.07.20 → v2026.07.22)	~205,000 → 202,000	model audit

Every correction is documented in the repository and summarized on the public tool; the detector column is the record, and none of these was caught by the author unaided.

6 Construction and auditability

The model was built with a frontier language model (Anthropic's Claude), with the author reviewing every parameter and assumption; adversarial review by a second model family (OpenAI's), plus reader corrections, produced the revision history above. Three mechanisms make the result checkable rather than merely stated. Parameters live in one typed file in which every value carries its unit, dollar vintage, and inline citation, machine-validated at load; an annotated bibliography maps each quantitative input to its primary source and price-level conversion. Sync tests force the published web parameters, allocation shares, and geography exports to equal fresh derivations from the committed data snapshots, and the body of this manuscript interpolates its model-derived numbers from those same artifacts at render time (the abstract's figures are typed copies of the interpolated values; externally sourced constants carry citations). And the audit trail is data: the 990 match audit ships as its applied per-row overlay, and the 50-organization

geography audit ships with per-organization sources and confidence grades — both applied by the pipeline rather than folded silently into parameters.

The browser implementation is a checked TypeScript port that reruns the full Monte Carlo client-side; every figure in this paper can be reproduced, and every assumption overridden, at maxghenis.com/mackenzie-scott-qaly. The repository, including this manuscript's source, is at github.com/MaxGhenis/mackenzie-scott-qaly.

7 Limitations

The credibility tiers and realization distribution are judgment priors with no literature calibration; the ordering is defensible, the levels are not measured, and the stance sweep is the honest bound on their influence. The QALY lens counts only health, so the estimate is a partial accounting that omits both the non-health benefits where Scott concentrates — education, mobility, equity — and any possible harms; it is not a floor on total wellbeing. Credibility shrinks toward zero but never below it, so the model places no mass on harmful giving; the closest available zero scenario is the near-RCT-only stance floor, which drives unsupported archetypes to ~1% credit. Allocation splits each organization's dollars equally across its focus areas and service locations; both are coarse, and the geography audit covers only the 50 largest "global"-listed organizations. Realization is one global draw, so it cannot vary by cause. Archetype anchors treat heterogeneous organizations as draws from one distribution. Within each cause, delivery regions are priced identically, so the geographic decomposition exposes assumptions rather than measuring regional differences. The undisclosed third of dollars is imputed from a fitted elasticity, not observed. The routed global-health bucket prices its unattributable and Europe-&-Central-Asia-plus-Canada dollars — roughly a fifth of the bucket, whose true income mix is unidentified — at low-and-middle-income anchors, an upward-bias channel. Several cost priors average over mixed null-and-positive evidence before the credibility discount applies, so the two-multiplier separation of doubts is imperfect at those boundaries. The Dirichlet allocation spread is a subjective prior; uncertainty from the elasticity fit, unmatched filings, the focus-area crosswalk, and the geography audit is not propagated. Allocation shares are derived in nominal dollars and applied to the real total; re-deriving in real dollars moves shares by hundredths of a percentage point. The geography audit's scalar U.S. shares and regional decompositions were reconciled separately and can differ by fractions of a point per organization. Finally, several anchors rest on a small number of quasi-experimental papers whose external validity to Scott's grantees is exactly what the realization factor is guessing at.

8 Conclusion

Measured in health, MacKenzie Scott's $26 billion buys an estimated median of about 202,000 quality-adjusted life-years — several times its cost under the government's own regulatory valuation — and most of that estimated health comes from the twentieth of her dollars funding health and development delivered abroad. The number is a model output under stated priors, not a fact, and the model's most important output is not the point estimate but the machinery around it: explicit evidence weights, a 45-fold stance range that any reader can traverse, versioned

corrections, and audits committed as data. Portfolio-scale philanthropy is measurable in outcomes without waiting for donors to publish outcome data — their own gift lists, recipients' tax filings, and the published causal literature suffice — and the same machinery applies to any large giver whose public disclosures carry comparable gift-level allocation and geography fields. The invitation of the interactive version is the paper's real conclusion: disagree with a parameter, move it, and watch what survives.

Acknowledgments

I thank readers on LinkedIn and the EA Forum whose corrections materially improved the model, including the reader who identified the original mispricing of internationally delivered health dollars. The model, audits, and this manuscript were built with Anthropic's Claude, with additional adversarial review by OpenAI models; all judgments and errors are mine. This is a personal research project and does not represent the views of PolicyEngine or the Axiom Foundation. Neither I nor those organizations have received funding from MacKenzie Scott, Yield Giving, or GiveWell.

References

Source: Article Notebook

Aizer, Anna, Shari Eli, Joseph Ferrie, and Adriana Lleras-Muney. 2016. "The Long-Run Impact of Cash Transfers to Poor Families." American Economic Review 106 (4): 935–71. https://doi.org/10.1257/aer.20140529.

Bailey, Martha J., and Andrew Goodman-Bacon. 2015. "The War on Poverty's Experiment in Public Medicine: Community Health Centers and the Mortality of Older Americans." American Economic Review 105 (3): 1067–104. https://doi.org/10.1257/aer.20120070.

Buteau, Ellie, Elisha Smith Arrillaga, and Christina Im. 2023. Emerging Impacts: The Effects of MacKenzie Scott's Large, Unrestricted Gifts: Results from Year Two of a Three-Year Study. Center for Effective Philanthropy. https://cep.org/wp-content/uploads/2023/11/BigGiftsStudy_Report_Y2_FNL.pdf.

Center for Effective Global Action. 2022. Donors Can Spend More Money, Better, with Cost Evidence. Https://cega.berkeley.edu/article/donors-can-spend-more-money-better-with-cost-evidence/.

Cesarini, David, Erik Lindqvist, Robert Östling, and Björn Wallace. 2016. "Wealth, Health, and Child Development: Evidence from Administrative Data on Swedish Lottery Players." Quarterly Journal of Economics 131 (2): 687–738. https://doi.org/10.1093/qje/qjw001.

Childress, Rasheeda. 2026. Donors Gave U.S. Charities $617 Billion in 2025, According to the New Giving USA Report. Https://apnews.com/article/8363b76bc8cf854f6865c31129e8a4b1.

Constellation Fund. 2026. 2025 Investment Report. Https://constellationfund.org/impact-news-insights/investment-report-2025/.

Finkelstein, Amy, Sarah Taubman, Bill Wright, et al. 2012. "The Oregon Health Insurance Experiment: Evidence from the First Year." Quarterly Journal of Economics 127 (3): 1057–106. https://doi.org/10.1093/qje/qjs020.

GiveWell. 2024. Re-Evaluating the Impact of Unconditional Cash Transfers. Https://blog.givewell.org/2024/11/12/re-evaluating-the-impact-of-unconditional-cash-transfers/.

GiveWell. 2025. Impact Estimates. Https://www.givewell.org/impact-estimates.

Huddleston, Jr., Tom. 2025. MacKenzie Scott Announced Another $7.1 Billion in 2025 Charitable Donations—She's Now Given Away $26.3 Billion Since 2019. Https://www.cnbc.com/2025/12/13/mackenzie-scott-revealed-her-total-charitable-donations-for-2025.html.

Indiana University Lilly Family School of Philanthropy. 2026. Giving USA 2026: The Annual Report on Philanthropy for the Year 2025. Giving USA Foundation.

Kearsley, Aaron. 2026. HHS Standard Values for Regulatory Analysis, 2026. ASPE. https://aspe.hhs.gov/reports/standard-ria-values.

Miller, Sarah, Norman Johnson, and Laura R. Wherry. 2021. "Medicaid and Mortality: New Evidence from Linked Survey and Administrative Data." Quarterly Journal of Economics 136 (3): 1783–829. https://doi.org/10.1093/qje/qjab004.

Ord, Toby. 2013. "The Moral Imperative Toward Cost-Effectiveness in Global Health." In Essays. Center for Global Development. https://www.cgdev.org/publication/moral-imperative-toward-cost-effectiveness-global-health.

ProPublica. 2026. Nonprofit Explorer. Https://projects.propublica.org/nonprofits/.

Robin Hood Foundation. 2023. Our Approach: Metrics and Benefit–Cost Analysis. Https://robinhood.org/our-work/our-approach/.

Sommers, Benjamin D. 2017. "State Medicaid Expansions and Mortality, Revisited: A Cost-Benefit Analysis." American Journal of Health Economics 3 (3): 392–421. https://doi.org/10.1162/ajhe_a_00080.

Yield Giving. 2026. Gifts Database. Https://yieldgiving.com/gifts.