Who Gets Credit When Carbon Scores Fall?

Pouya Behmaram* Evan Jo†

September 2026

Abstract

A fund can look greener without its manager moving toward lower-emissions firms. Independent measures agree that U.S. mutual fund carbon intensity fell by more than a third, yet portfolio reweighting explains only 1.2 percent. Most of the decline comes from changing issuer emissions, sales denominators, estimates, and coverage. These inputs also reverse the institutional sector trend and reorder manager rankings. In a hypothetical selection, attribution replaces two in five managers and improves subsequent reweighting toward firms with lower emissions conditional on industry and firm characteristics. Carbon scores describe measured exposure; evaluating implementation requires attribution and comparable coverage.

Keywords: delegated portfolio management, performance attribution, ESG ratings, carbon footprints, sustainability ratings, institutional investors
JEL Codes: G11, G14, G23, Q54

*Université du Québec à Montréal, École des sciences de la gestion (behmaram.pouya@uqam.ca).
†Smith School of Business, Queen's University (evan.jo@queensu.ca).

1 Introduction

When a fund's carbon score falls, who gets the credit? A client evaluating a decarbonization mandate needs to distinguish a manager's portfolio decisions from changes in the firms it holds and the information used to measure them. Selecting a low-carbon portfolio and selecting a manager who repeatedly reweights toward lower-emissions firms are different decisions. A footprint alone need not supply the information required for both.

A footprint, however, records more than what the manager did. It combines a portfolio chosen by the manager with issuer-level measurement inputs, including emissions values, sales denominators, coverage, and estimates. We call the assignment of these inputs to securities the emissions map. Both holdings and measurement inputs move every year. Firms report new emissions, vendors re-estimate missing ones, coverage expands, denominators grow, and managers trade. The reported score sums these movements into one number, so on its own it cannot say whether the manager changed the holdings or issuer developments and measurement changes lowered the reported exposure.

This joint production could be harmless. If map movement were small, or common to every portfolio, or if carbon systems disagreed so deeply that no shared verdict existed, decomposing the score would refine familiar conclusions without changing them. The important case is the opposite one. Evaluators agree that carbon intensity fell, the decline is read as delegated performance, and yet attribution changes which portfolios and managers appear to have improved.

The data show why the distinction matters. Two separately produced systems agree that fund carbon intensity declined, yet portfolio reweighting explains little of that decline. Attribution changes measured trends and manager rankings, and improves a hypothetical selection rule for subsequent implementation. A separate comparison of dated emissions archives shows that coverage changes can alter the assessment of the same completed portfolios. Making coverage comparable eliminates those assessment changes. These results connect the composition of a carbon record to two decisions, selecting a manager for implementation and assessing progress against a stated target.

Two separately produced systems deliver the fund verdict. We compare Morningstar's published Scope 1 and 2 carbon intensity with an independent Refinitiv-based construction for the same fund on the date of its disclosed portfolio. Across 12,113 common portfolio-years, the median annual rank correlation between the two systems is 0.922, and among funds observed at both endpoints, median intensity falls 37.7 percent under Morningstar and 37.8 percent under the construction from 2019 to 2024. The verdict is credible, and it is shared. For the 1,065 funds

passing the matched-date screen, however, portfolio reweighting accounts for 1.2 percent of the constructed decline. Growth in issuers' nominal sales accounts for 71.4 percent, mostly through broad-industry output prices, and movement in issuer emissions values accounts for 29.5 percent. The average fund's carbon intensity fell mainly because the denominator beneath its holdings grew.

Where the decline came from then determines who appears responsible for it. Holding the evaluator fixed, the Refinitiv-based total and its portfolio-originated component rank funds with correlations of 0.804 to 0.827 across attribution orders, and 28.6 to 29.8 percent of funds receive opposite signs. Evaluator disagreement alone, measured before any attribution, produces opposite signs for 14.3 percent. Comparing Morningstar's published score directly with the portfolio-originated record combines both forces, and rank correlations fall to 0.598 to 0.625 with opposite signs for 33.1 to 33.5 percent of funds. A fund can improve under the score investors see while its portfolio-originated record deteriorates, and the reverse.

The institutional record shows the same anatomy at the scale of a sector. We combine quarterly Form 13F holdings with Refinitiv emissions from 2007Q1 through 2023Q4, a panel of reporting managers with $43.9 trillion of assets at the endpoint. The observed raw footprint becomes browner by 0.993 log-exposure units. Holding finite-emissions coverage fixed, while keeping observed portfolios and the emissions of continuously covered firms, instead yields an improvement of 2.292 units. Coverage supplies 90.7 percent of raw map movement, vendor estimates supply 79.0 percent of the coverage anatomy, and map movement accounts for 71.5 percent of the aggregate continuing-manager anatomy under our primary symmetric convention, with directional-order bounds of 60.6 to 71.5 percent. The map determines the sector's sign as well as its scale.

Attribution also changes individual manager rankings. Among 861 reporting entities observed over the full 68 quarters, removing cumulative map movement changes the measured trend sign for 59.7 percent of entities and leaves measured and portfolio-originated rankings correlated at 0.406 under the symmetric convention, with directional-order bounds of 36.9 to 69.8 percent and 0.269 to 0.494. Under the lagged-map order the measured record explains 27.4 percent of the variance in the portfolio-originated record. The best linear projection puts positive weight on the total, so it changes scale while preserving ranks. Recovering the ordering that excludes realized map exposure requires information beyond an increasing recalibration of the total. Attribution supplies that information directly; statistical models with additional inputs could also change ranks.

The Morningstar Sustainability Rating, the fund-level globe rating of relative ESG risk, provides a separate setting for studying investor response. It is distinct from carbon intensity and from the Star rating of past performance. The flow measure records 36.5 basis points of inflow after Star upgrades and 48.1 basis points of outflow after Star downgrades. A published globe downgrade classified as driven by changing peer cutoffs produces a response of −1.9 basis points in the first fully public month, with a flat pre-event path and a one-sided bound that rules out monthly outflows worse than about 27 basis points. This estimate identifies bounded nonresponse to evaluator-caused changes in the linked population and period. Broader investor responses to ESG ratings lie outside the design. Nonresponse can reflect limited attention or a decision to disregard changes in relative labels; this comparison does not distinguish those explanations.

We separately examine the information supplied by attribution for a hypothetical manager-selection rule. We rank 2,154 reporting entities on observations dated through 2013Q4 in the current data vintage, omit 2014, and evaluate their portfolio reweighting from 2015Q2 through 2023Q4. The outcome values actual-minus-passive weights at midpoint issuer scores, where higher scores mean lower emissions conditional on industry and firm characteristics. The conventional top quintile subsequently trails its bottom quintile by 0.193 standard deviation, while the spread based on attributed action is positive at 0.194. Removing map movement accounts for about 77 percent of this reversal. Attribution replaces 171 of 431 managers, and the entrants subsequently outperform the displaced managers by 0.552 standard deviation. Their 39.7 percent replacement share implies a 0.219-standard-deviation gain in future action for the attributed top quintile. The gain remains positive when mandate capacity ranges from 5 to 50 percent of managers. The conventional record therefore selects a weaker tail for this particular implementation objective.

A performance measure is useful for delegation when it is informative about the action the client values (Holmström, 1979). Rewarding outcomes also raises the familiar problem of rewarding luck (Bertrand and Mullainathan, 2001). In a carbon record, the remainder includes issuer developments, valuation denominators and measurement updates. Managers can choose their exposure to these components even though the components are not themselves purchases or sales. The selection test establishes the value of separating them for the stated implementation objective, without requiring that all map movement be exogenous luck.

The outcome matters for interpreting the gain. Attribution-selected managers also show less subsequent browning when evaluated on a fixed issuer score vector, but the physical-emissions comparison is statistically unresolved. The attributed top group also has higher

average covered-portfolio carbon intensity over the evaluation window. The evidence supports selection for implementation within the conditional score, leaving a reduction in physical exposure unestablished. Conventional-only managers subsequently grow reported 13F equity AUM 4.02 percentage points per year faster than attribution-only managers. That descriptive difference may reflect manager type or secular composition; it identifies neither a response to carbon scores nor an observed mandate allocation.

The data vintage can change an assessment even when the completed investment record is held fixed. We evaluate the same 861 managers' actual 2019Q4 and 2023Q4 portfolios using emissions archives collected on July 18 and September 22, 2026. Under an illustrative seven percent annual decline rule, 17 managers lose a passing assessment and eight gain one. Coverage changes reproduce all 25 switches; revisions to values on common coverage produce none. Restating both vintages on common coverage eliminates the switches. This is a concrete comparability problem in the observed equity record, whose emissions coverage is incomplete, rather than evidence of an actual mandate decision.

The core decompositions are accounting identities. For institutional portfolios, we carry lagged positions forward at security returns, introduce current holdings, and attach the current map, so each transition splits into valuation, trading, and map movement. For fund carbon intensity, a product decomposition separates portfolio weights, issuer emissions, nominal sales, coverage, source composition, and their interaction. Each identity closes at the portfolio transition before aggregation. The exercise is the carbon analogue of classical performance attribution, and it is an index-number problem, because portfolio quantities and the per-security measurement map change together. The portfolio term bundles trading, valuation, and investor flows, so we read it as the path of the portfolio, a broader object than manager trading.

The identities leave the attribution order open. A lagged-map order assigns the portfolio-by-map interaction to the map, a current-map order assigns it to the portfolio, and a symmetric order splits it. The fund conclusions are stable across the three orders. The manager conclusions are severe under every order but differ materially in magnitude across them. We use the symmetric midpoint, the two-factor Shapley allocation, as the primary convention. Both passive valuation and active reweighting use midpoint issuer scores. The two directional orders form bounds. The resulting range measures economically relevant ambiguity in responsibility. A report must name its convention to resolve that ambiguity.

The nominal-sales result is mechanical in the accounting sense and consequential in the economic sense. Weighted average carbon intensity divides issuer emissions by nominal sales, so

revenue growth lowers the ratio with constant emissions. This arithmetic measures no physical abatement, and the price component uses a broad-industry benchmark rather than a firm-specific inflation estimate. The economic fact is that a ratio used to evaluate delegated portfolios can improve because an issuer-side denominator grew. Arithmetic is the mechanism through which the reported record assigns credit.

A small framework organizes what an evaluator needs. Write the measured record as purchases and sales, plus valuation, plus map movement. Removing the map recovers a portfolio-originated record; removing valuation as well isolates purchases and sales. No strictly increasing correction of the total can undo a rank reversal. For any fixed mandate size, the difference in future performance between conventional and attributed selection equals the share of managers replaced times the future-performance gap between entrants and displaced managers. The framework also shows why a reward based on the total loads on predictable map exposure; initial uncovered portfolio share predicts subsequent map movement. The evidence establishes exposure to that incentive, while deliberate gaming would require contract terms and point-in-time information.

Three features of the data define the scope. Morningstar and Refinitiv deliver current historical series. Our two 2026 archives identify changes between those downloads, but leave the map observed during the 2007–2024 investment history unknown. The historical decompositions and selection exercise use their existing input vintages; the September archive enters the separate completed-record comparison. The two fund systems diverge in 2022, so the bridge supports the five-year endpoint verdict; it cannot attribute every annual movement. The ranking samples contain funds and entities observed continuously over their windows. The incidence estimates apply to these stable populations, with broader populations left for future work.

The evidence speaks to four literatures, on carbon-footprint attribution, mutable financial data, climate finance, and delegated performance evaluation. Carbon-footprint attribution already has a practical foundation. Wang and Nagy (2025) separate investment decisions, financing changes, and real-world decarbonization in a total-portfolio reporting framework, which establishes the institutional demand for attribution. Our increment is empirical and delegated. We carry changing issuer maps through actual U.S. fund and Form 13F portfolios, separate coverage, source, and denominator movement, and measure what the separation does to sector conclusions and manager classifications.

A second strand documents that financial and ESG data are mutable. Analyst databases have been rewritten (Ljungqvist et al., 2009), separately dated Refinitiv downloads deliver different ESG histories (Berg et al., 2021), and ESG ratings disagree across providers (Berg et al., 2022).

Reported and estimated emissions behave differently in return tests (Aswani et al., 2024), and firms revise self-reported emissions (Cohen et al., 2026). We extend this literature by carrying a mutable map through portfolios and measuring how the resulting record allocates responsibility.

Climate-finance research studies how environmental preferences and risks shape holdings, prices, and returns (Hong and Kacperczyk, 2009; Bolton and Kacperczyk, 2021; Ilhan et al., 2021; Pástor et al., 2021, 2022), and recent work distinguishes portfolio reallocation from real-economy decarbonization (Pástor et al., 2025; Atta-Darkua et al., 2025). Our distinction comes one step earlier. Before asking whether investors changed firms or firms changed emissions, an evaluator must know whether the score moved because of weights, issuer quantities, denominators, coverage, or the map that combines them.

The finance home is delegated performance evaluation. Performance signals direct capital toward managers perceived to have skill (Berk and Green, 2004; Berk and van Binsbergen, 2015), sustainability labels and ratings can redirect fund flows (Hartzmark and Sussman, 2019; Ceccarelli et al., 2024; Del Guercio and Tkac, 2008), and the pooled globe-flow relation weakens once sustainability trades reveal a performance cost (Gantchev et al., 2024). Classical attribution separates the sources of portfolio performance (Brinson et al., 1986, 1991), and benchmark design shapes manager compensation and portfolio choice (Admati and Pfleiderer, 1997; Cuoco and Kaniel, 2011; Basak and Pavlova, 2013). That literature generally takes the characteristic map used in evaluation as given. We show that when the map moves, the realized record is a joint product of agent choice, issuer developments and evaluator-side data production, and we measure how much of the record each side produced. Relative to pooled rating tests, the flow design isolates published changes produced by the evaluator's ruler with the fund's own inputs held fixed in the classification counterfactual.

Two recent studies use mechanical rating changes to identify investor responses. Choi et al. (2026) find that upgrades make ESG funds more likely to add a stock, while Rzeźnik et al. (2026) trace the price response to a rating-methodology change to retail trading. Lowry et al. (2026) instead distinguish ESG funds by their incentives to engage and show that committed funds acquire more information, engage more, and have greater firm-level effects. We ask whether the fund-level record assigns responsibility correctly in the first place, and whether it selects managers who subsequently implement greener portfolios.

The results carry one practical prescription. A portfolio footprint should be accompanied by its map contribution, coverage path, denominator, source composition, and the convention that allocates their interaction. This reporting discipline leaves the familiar total intact while making

responsibility visible.

2 Data and Evaluation Objects

2.1 Institutional holdings and the reporting unit

The institutional sample combines quarterly U.S. equity holdings with stock characteristics from 2007Q1 through 2023Q4. The unit is an actual Form 13F reporting entity identified by FactSet. We neither consolidate entities to an asset-management parent nor aggregate small filers into artificial portfolios. We retain positive common-stock positions with valid CRSP links. Synthetic residual accounts, short positions, and the source panel's household placeholder are excluded. Our population statements therefore concern the retained reporting entities.

The reporting entity is the unit at which holdings are filed and at which the portfolio decomposition closes. This choice matters when one asset-management organization contains several filing entities or when a filer changes its security set. Consolidating to an ultimate parent would require an additional ownership map that can change over time and would mix the footprint attribution with parent reconstruction. We instead preserve the entity delivered by the FactSet panel and require the same identifier through the manager comparisons. The resulting object is a filing-level performance record, which is the level directly observable from the holdings.

The construction preserves the accounting unit before forming weights. If an investor-stock-quarter appears more than once, position values are summed and one manager AUM value is retained. Stock-level joins use CRSP `permno` and calendar quarter. Position values and manager AUM enter in the same dollar units. Lagged and current holdings are joined by reporting entity, `permno`, and adjacent quarter, preserving complete purchases and sales. CRSP ex-distribution returns carry lagged positions forward and match 99.73 percent of lagged position value.

The union of lagged and current holdings is essential for the trading term. An inner join would discard complete sales and new purchases precisely when the portfolio changes most. In the union, a position present only at the lagged date is carried forward and then reduced to zero, while a position present only at the current date enters from zero. Stock returns are used only to construct the passive counterfactual; current reported position value determines the actual endpoint. AUM is attached at the reporting-entity quarter, not summed across position rows. These conventions allow valuation, trading, and reclassification to be expressed on the same scale.

The retained panel covers entities with consistent identities, holdings, AUM, and stock links,

a subset of all 13F filings. We use the full retained panel for sector accounting and a balanced subset of 861 actual entities for long-horizon manager comparisons. Market capitalization less retained positions is a residual containing institutions outside the panel and other ownership. We make no household-ownership or marginal-buyer inference from that residual.

The sector and manager samples serve different estimands. Sector accounting allows the set and AUM weight of eligible entities to evolve and then separates within-entity change from weight change and entry or exit. The manager exercise holds the population fixed for all 68 quarters so that reordering cannot be caused by reporters appearing or disappearing. This balanced sample is demanding by design: it asks whether the same filing entities receive different long-run classifications under moving and fixed maps.

2.2 The emissions map and its lineage

Firm emissions come from the current Refinitiv ESG database. For each firm-year, reported total CO_2-equivalent emissions take priority when available; the vendor estimate is used otherwise. If neither is finite, the observation remains uncovered. We assign fiscal-year $Y-1$ information to calendar year Y and use $\log(1+\mathrm{CO}_2)$ for the raw map. The annual lag limits mechanical look-ahead; historical entry dates in the vendor system remain unavailable. We retain the reported-versus-estimated distinction because the two sources can behave differently (Aswani et al., 2024). Here the source flag attributes map movement and carries no judgment about estimate quality.

Each security-quarter observation therefore has both a numerical state and a source state. The numerical state is missing or finite. Conditional on being finite, the source is firm reported or vendor estimated. Between adjacent quarters, the map can change because a continuously covered value changes, because a security crosses the missing-to-finite boundary, or because it loses finite coverage. This lineage rule is applied before manager aggregation. It lets the source components add exactly to the parent map term and avoids classifying a newly covered estimate as a revision to a previously observed number.

For the standardized-score accounting, quarterly log emissions are residualized on log market equity, log book equity, operating profitability, investment, passive ownership, missing-control indicators, and two-digit SIC effects, then standardized. A higher residual denotes browner exposure; its negative is the emissions-green orientation used in several exhibits. Unlike fund carbon intensity, which divides issuer emissions by issuer sales, the institutional score is built on emissions levels alone, so no sales denominator enters the institutional decomposition. The raw

and residualized measures answer complementary questions. Raw emissions retain a transparent source and physical interpretation. Residualization makes manager comparisons less mechanical with respect to observable scale and industry characteristics. Because controls, coverage, and cross-sectional standardization vary over time, both remain evaluator-constructed characteristics rather than observed contractual targets.

The cross-sectional transform belongs to the map. A firm's standardized score can change with constant raw emissions because controls, the covered comparison set, or the cross-sectional mean and variance change. The residualized map decomposition therefore displays a standardization component in addition to reported, estimated, entry, and exit components. The raw decomposition omits that component by construction. We use the raw map to interpret economic sources and physical scale, and the residualized map to compare managers after common firm characteristics are removed.

2.3 Mutual funds and published carbon intensity

The primary fund analysis uses Morningstar's historical Scope 1 and 2 weighted average carbon intensity in U.S. dollar terms from 2019 through 2024. Each Morningstar fund is linked to one CRSP mutual-fund portfolio under a fixed identifier crosswalk. We use the latest CRSP holdings report that matches Morningstar's disclosed portfolio date; nearest-date matches are excluded. Positive holdings must sum to between 80 and 120 percent of TNA, at least 80 percent of TNA must link to CRSP `permno`, and at least 50 percent must carry the constructed Refinitiv metric.

For fund i at date t, the observed carbon score has the form

$$S_{it} = \sum_{n} w_{int} I_{nt}, \tag{1}$$

where w_{int} is security n's positive position value divided by fund TNA and I_{nt} is issuer emissions divided by nominal sales. Refinitiv supplies the reported-or-estimated emissions numerator under the same source priority used in the institutional analysis. Compustat supplies same-fiscal-year sales. The construction is scaled to tonnes of Scope 1 and 2 emissions per million U.S. dollars of sales, matching the orientation and unit of the published Morningstar measure.

The cross-evaluator sample contains 12,113 fund-date observations for 2,765 CRSP portfolios. The responsibility decomposition uses the stricter balanced population of 1,065 funds observed at every December endpoint from 2019 through 2024; 996 of those funds also pass every screen in 2025. A chained midpoint identity separates changes in Equation (1) into portfolio reweighting,

issuer emissions, nominal sales, coverage, reported-versus-estimated source composition, and their interaction. Because fund TNA is the weight denominator, the portfolio term combines trading, valuation, investor flows, and changes in non-equity assets. We call it portfolio reweighting to distinguish it from manager trading.

Morningstar and Refinitiv both supply current historical series. The fund comparison aligns portfolio dates and carbon objects; the vendor state visible to investors at each historical date remains unobserved. The same current-vintage limitation applies to the institutional sample.

2.4 Three evaluation objects and their denominators

We keep three denominators distinct. The *covered-equity style* is the position-weighted residualized emissions score among covered positive equity holdings. It asks whether the measured portion of a manager's equity portfolio tilts toward high- or low-emissions firms. Manager-style exercises require at least 25 percent position-value coverage, at least ten covered stocks in a usable quarter, and at least eight usable quarters in each estimation period. This object is conditional on vendor coverage and does not represent the footprint of total AUM.

Covered-equity style is used only when the economic object is a portfolio tendency among securities with observable scores. Dividing by covered position value prevents changes in the amount of missing data from being labeled a change in style. It also means that style cannot be interpreted as total financed emissions. The screens require enough covered securities and time-series observations for a reporting entity to have an estimable type.

The *AUM-scaled residualized exposure* divides the covered-position score numerator by total reported manager AUM and assigns zero to positions without a finite score. It asks how much measured exposure the portfolio carries per dollar managed. Zero marks absence from the measurement system and makes no assumption about an uncovered firm's emissions. This denominator makes coverage movement visible and allows valuation, trading, and map terms to add exactly. It is the primary object for the standardized accounting and manager-monitoring results.

Here the zero contribution records what the measurement system includes in the reported footprint at that date. When a previously uncovered position receives a finite score, the measured numerator changes with an unchanged holding. The fixed-coverage path removes that coverage change while preserving observed holdings and changes in emissions intensity among the securities that remain covered. It is an attribution counterfactual inside the measured object and leaves the missing firm's emissions unidentified.

The *raw financed footprint* attributes issuer emissions by the institution's equity ownership share, using CRSP equity market capitalization as the attribution denominator, and scales by manager AUM when expressed as intensity. It translates the measured record into tonnes of CO_2 equivalent. The market-capitalization denominator, AUM scaling, missing-coverage convention, and data vintage travel with every physical estimate. The three objects answer different economic questions: raw scores trace sector attribution, residualized scores compare managers, and the financed footprint expresses physical scale.

2.5 Historical inputs, dated archives and data availability

The current vintage permits a decomposition of the footprint it generates. Recovering the historical information set would require archived vintages. Map movement can reflect either a revision to a value already in the contemporaneous record or backfill of a previously absent value. The database used here identifies their sum, not the two terms. We therefore call the first finite observation the first covered fiscal year in the current vintage. Archived vendor vintages would be required to classify it as a live disclosure or a later backfill. This scope follows the database-history problem identified by Ljungqvist et al. (2009) and Berg et al. (2021).

We additionally archive the same six Refinitiv emissions fields on July 18 and September 22, 2026, 66 days apart. The source-data queries are identical. These downloads identify revisions, deletions and additions between collection dates, without reconstructing the information investors possessed during the earlier investment history. The original historical inputs remain in the fund, sector and selection analyses. The new archives enter only the separate comparison of completed portfolio assessments.

That comparison fixes the 861 balanced managers, endpoint holdings and denominators, and a date-valid organization–CUSIP–PERMNO link based on July metadata. Fiscal year-end must fall within the CRSP name interval; fiscal-year overlap and normalized company-name similarity resolve competing links. Distinct organizations are never averaged. Within each organization-year-field, conflicting finite values are unavailable rather than averaged, and reported emissions take priority over estimates. Common coverage retains securities with usable emissions under both archives at each portfolio endpoint, leaving their original AUM weights unchanged. Coverage shares distinguish represented security-position value from the full reported equity-asset denominator.

Raw Refinitiv, FactSet, CRSP, and Compustat records cannot be redistributed. The replication package therefore supplies construction code, licensed-data acquisition instructions, dis-

tributable table aggregates, and schema-compatible inputs for the figure code. Researchers with the relevant entitlements can regenerate the empirical exhibits without placing licensed rows or credentials in the public release.

3 Evaluating managers when measurement inputs change

3.1 A jointly produced record

A portfolio footprint can evaluate a manager only after the reported change is assigned to the portfolio and the measurement system. Let V_{int} denote reporting entity i's dollar position in stock n at quarter t, A_{it} its reported AUM, and M_{nt} the emissions score attached to that stock. The measured AUM-scaled exposure is

$$B_{it} = \frac{\sum_n V_{int} M_{nt}}{A_{it}}. \tag{2}$$

The emissions map M_{nt} assigns each security its date-specific score, including coverage, reported and estimated emissions, firm controls, and the cross-sectional transform. We use "map" as shorthand for this assignment. It includes issuer developments as well as evaluator decisions; a change in the map need not be a revision or an error. Sales denominators enter the fund carbon-intensity measure. Portfolio AUM instead enters the weights in Equation (2).

For a continuing entity, define B_{it}^{p} by carrying lagged positions forward with security returns while holding the lagged map fixed. Define $B_{it}^{1|0}$ by replacing that passive portfolio with current positions and current AUM, again at the lagged map. Adding and subtracting these intermediate states gives

$$\begin{aligned} B_{it} - B_{i,t-1} = \underbrace{(B_{it}^{p} - B_{i,t-1})}_{\text{valuation}} + \underbrace{(B_{it}^{1|0} - B_{it}^{p})}_{\text{purchases and sales}} \\ + \underbrace{(B_{it} - B_{it}^{1|0})}_{\text{map movement}}. \end{aligned} \tag{3}$$

Purchases into and sales out of securities remain in the portfolio term because the lagged and current security universes are joined. All components use compatible dollar units and the same AUM convention. The decomposition closes for every manager before aggregation.

Orient each component so that a larger value denotes environmental improvement and suppress time subscripts. Equation (3) becomes

$$y_i = a_i + v_i + m_i, \tag{4}$$

where y_i is the measured record, a_i is the purchases-and-sales contribution, v_i is valuation, and m_i is map movement. The *portfolio-originated record* is

$$p_i \equiv a_i + v_i = y_i - m_i. \tag{5}$$

A mandate concerned with the contribution of changing portfolio weights targets p_i. A mandate concerned with manager purchases and sales targets a_i. The evaluator can report either target without discarding the total.

3.2 Attribution as an index-number problem

The footprint aggregates a quantity vector, the portfolio, against a per-unit value vector, the emissions map. A change in B_{it} therefore compares two aggregates built from different quantities and different per-unit values. This is the problem index numbers are designed to resolve. Valuation and purchases and sales vary the portfolio at the lagged map, while map movement varies the map on the current portfolio. Equation (3) is a chained Laspeyres–Paasche factorization (Diewert, 1976).

The index-number interpretation distinguishes measurement levels from performance changes. Today's map may measure a portfolio's emissions more accurately than yesterday's map. That improvement belongs in the current exposure level. A statement that the portfolio became greener between the two dates, however, requires a common ruler. Otherwise, portfolio change and measurement change enter the same performance record.

The portfolio-by-map interaction leaves an attribution convention to choose. A lagged-map order assigns the interaction to map movement, a current-map order assigns it to the portfolio, and their midpoint splits it equally. The midpoint is the two-factor Shapley allocation. We use it as the primary convention and report the two directional orders as bounds.

To separate valuation from reweighting under this convention, let w_n^0 denote the initial AUM weight, w_n^p the weight after carrying initial holdings at security returns and dividing by passive AUM, and w_n^1 the observed final weight. With $\bar{M}_n = (M_{n,t-1} + M_{nt})/2$, the primary

contributions, before orienting them toward greening, are

$$\begin{aligned} v_{it}^{\text{mid}} &= \sum_n (w_n^p - w_n^0)\bar{M}_n, \\ a_{it}^{\text{mid}} &= \sum_n (w_n^1 - w_n^p)\bar{M}_n, \\ m_{it}^{\text{mid}} &= \sum_n \frac{w_n^0 + w_n^1}{2}(M_{nt} - M_{n,t-1}). \end{aligned} \tag{6}$$

Both portfolio components use the same score vector, and their sum with map movement equals the observed exposure change. If actual weights equal passive weights, attributed reweighting is zero even when returns and issuer scores change together. The action term measures implementation relative to passive holdings. Changes in portfolio scale or client capital can affect those weights; the term does not identify trading intent or engagement.

Changing coverage creates the same attribution problem. When a security first receives a finite emissions value, the measured footprint changes even if the portfolio does not. The new value may improve the exposure estimate, but the associated historical movement did not originate with the manager. Component reporting preserves both facts.

3.3 Why the total cannot recover the intended ranking

Let $x_i \in \{p_i, a_i\}$ denote the performance target and define the remainder $u_i(x) \equiv y_i - x_i$. The remainder is m_i when the target is the contribution of changing portfolio weights and $v_i + m_i$ when the target is purchases and sales.

Proposition 1 (Rank reversal and nonrecoverability)**.** *For two managers with $x_i > x_j$, the measured record reverses their ordering if and only if*

$$u_j(x) - u_i(x) > x_i - x_j. \tag{7}$$

For any strictly increasing transformation h, $h(y_i)$ and y_i produce the same manager ordering. Once the measured record reverses a target ranking, no strictly increasing recalibration of the total can recover that ranking.

The first statement follows by substituting $y_k = x_k + u_k(x)$ for managers i and j. The second follows from monotonicity. Against a zero benchmark, the measured and intended trend signs differ exactly when $x_i y_i < 0$.

Attribution recovers the two natural targets directly:

$$p_i = y_i - m_i, \qquad a_i = y_i - v_i - m_i. \tag{8}$$

A linear projection of x_i on y_i illustrates the distinction. A positive projection coefficient can improve the scale of a forecast, but it preserves the ordering of y_i. Statistical calibration and responsibility attribution are different operations. A positive recalibration can change passage of a fixed numerical threshold even while it preserves ranks, and a forecast using additional covariates can reorder managers. The nonrecoverability result concerns increasing transformations of the total alone.

3.4 The consequence of manager selection

Ranking disagreement becomes economically consequential when it changes who is selected and the replacement managers perform differently. Suppose a mandate selects K managers. Let $S_y(K)$ be the managers selected using the measured record and $S_x(K)$ those selected using the intended target. Let g_i denote a subsequent mandate outcome and $\overline{g}(S)$ its mean among managers in set S.

Define the entrants $E = S_x(K) \setminus S_y(K)$ and the displaced managers $D = S_y(K) \setminus S_x(K)$. The two selected sets have the same size, so $|E| = |D| = J$.

Proposition 2 (Selected-set consequence)**.** *If $J > 0$, the difference in subsequent mandate performance between the two selection rules is*

$$\overline{g}(S_x(K)) - \overline{g}(S_y(K)) = \frac{J}{K}\left[\overline{g}(E) - \overline{g}(D)\right]. \tag{9}$$

If $J = 0$, the difference is zero.

Managers common to both sets cancel when their selected-set means are subtracted, which proves Equation (9). The expression separates an extensive margin, the share of managers replaced, from an intensive margin, the future-performance gap between entrants and displaced managers. A large ranking disagreement has no selection cost when the two replacement groups perform alike. Conversely, a modest replacement rate can matter when their performance differs sharply.

Reporting (x_i, u_i) retains the total because $y_i = x_i + u_i$ and also makes the intended performance target available on its own. Whether selection on that target improves subsequent

outcomes is empirical. Our held-out tests estimate both terms in Equation (9), first for a top-quintile mandate and then across manager capacities. The nested records y_i, p_i, and a_i also show whether the selection gain enters when map movement or valuation is removed.

3.5 Incentive implication

The selection result does not require managers to anticipate map movement. A separate incentive implication arises when compensation depends on the total and expected map exposure varies with portfolio choice. Let w_i denote the portfolio and let a score-based reward be $R_i^S = \beta y_i(w_i)$, with $\beta \geq 0$. Its marginal evaluation payoff is

$$\nabla_{w_i} E[R_i^S] = \beta \nabla_{w_i} E[x_i(w_i)] + \beta \nabla_{w_i} E[u_i(w_i)]. \tag{10}$$

A reward written on x_i removes the second term. The two rules generate the same marginal incentive when expected movement in the remainder is locally unrelated to the portfolio.

Initial uncovered portfolio share predicts subsequent map movement in the data, so managers enter the evaluation period with heterogeneous exposure to the second term. The current-vintage histories do not establish what managers knew at the time or whether they pursued that exposure. The incentive equation therefore motivates separating the components; the empirical consequence we measure is the performance of a hypothetical manager-selection rule.

4 Agreement on the Verdict, Attribution of Responsibility

The attribution problem matters for delegated finance only if it appears in a score that investors and asset managers actually observe. We therefore compare Morningstar's published Scope 1 and 2 carbon intensity with an independent construction for the same mutual fund on its disclosed portfolio date. The construction applies Refinitiv issuer emissions and Compustat sales to CRSP mutual-fund holdings.

Carbon-footprint attribution already has a useful institutional foundation. Wang and Nagy (2025) separate investment decisions, financing changes, and real-world decarbonization in a total-portfolio reporting framework. We take the complementary empirical step: using disclosed fund holdings and issuer-year records from one fixed database vintage, we separate portfolio reweighting from changes in the data attached to the portfolio and ask whether the distinction changes fund rankings and trend signs.

The two systems agree closely. Table 1 contains 12,113 common portfolio-years for 2,765

portfolios. The median annual rank correlation is 0.922, and the median absolute rank gap is 4.4 percentile points. Among the 1,467 funds observed at both endpoints, the median Morningstar decline from 2019 to 2024 is 37.7 percent; the Refinitiv-based decline is 37.8 percent. Agreement makes the fund-level verdict credible while leaving the source of the decline unresolved.

Table 1. Agreement between two fund carbon-intensity systems

Year	Portfolios	Rank correlation	Median rank gap	Refinitiv coverage
2019	1,596	0.933	4.7	87.5%
2020	1,652	0.937	3.7	87.2%
2021	2,355	0.899	5.4	86.8%
2022	1,984	0.911	5.5	86.2%
2023	2,040	0.945	3.2	85.2%
2024	2,486	0.903	3.9	84.4%
Balanced 2019–2024: 1,467 portfolios; Morningstar change -37.7%; constructed change -37.8%.				

Notes: The sample matches the same CRSP mutual-fund portfolio on Morningstar's disclosed portfolio date. Rank correlation is the annual Spearman correlation between Morningstar's published Scope 1+2 USD weighted-average carbon intensity and a construction using CRSP holdings, the current Refinitiv emissions map, and same-fiscal-year Compustat sales. The rank gap is in percentile points. Refinitiv coverage is the median share of portfolio TNA with a valid constructed metric. Across all 12,113 portfolio-years, the median annual rank correlation is 0.922 and the median absolute rank gap is 4.4 percentile points. Both systems provide current-vintage histories; archived point-in-time records are unavailable.

4.1 Who produced the decline?

Figure 1 answers that question for 1,065 funds with matched holdings and valid measures in every December from 2019 through 2024. Panel A shows the published Morningstar path beside the construction that can be decomposed. Mean published intensity falls from 193.1 to 128.3, while constructed intensity falls from 164.5 to 104.2. The endpoint changes differ by 4.5 intensity points, an absolute residual equal to 7.0 percent of the published decline. The two paths diverge in 2022. Attribution therefore pertains to the five-year decline, not each intermediate annual movement.

Panel B separates the sources of the constructed decline. Growth in issuers' nominal sales denominators contributes 43.1 intensity points, or 71.4 percent of the decline. Movement in issuer emissions values contributes 17.8 points, or 29.5 percent. Portfolio reweighting contributes 0.8 points, or 1.2 percent. Coverage, source composition, and their interaction account for the small remaining net difference. The reported-source share of fund weight rises from 61.5 to 77.9 percent over the window, while the estimated share falls from 23.6 to 5.6 percent.

The sales result can be separated without changing the parent identity. A single public BEA gross-output benchmark maps 98.63 percent of absolute position-weighted sales contributions.

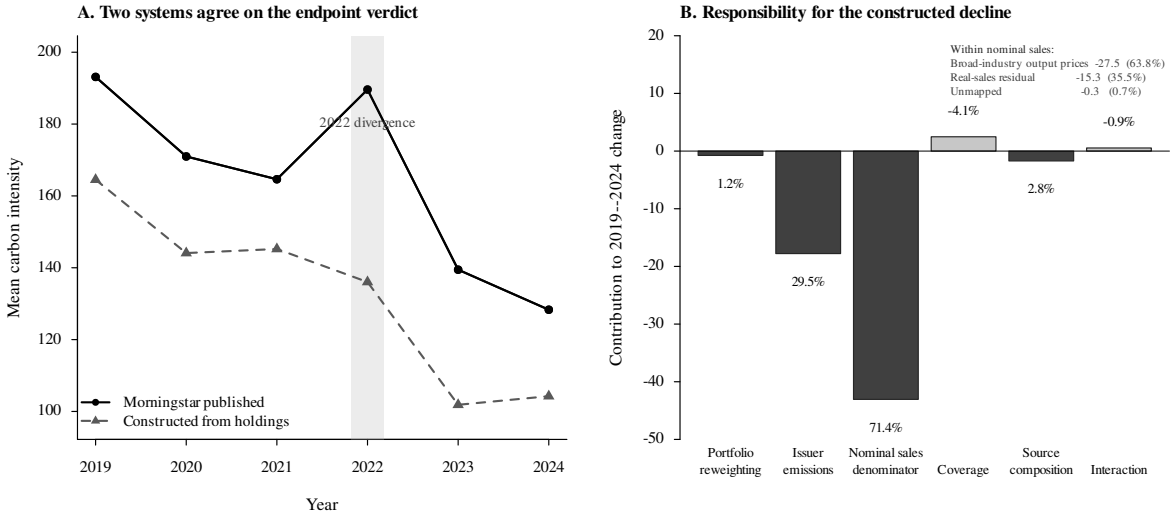

Figure 1. Agreement on the fund carbon verdict and responsibility for its decline. Panel A reports equal-fund mean Morningstar Scope 1 and 2 weighted average carbon intensity and the Refinitiv-based construction at each matched December portfolio date. Panel B decomposes the constructed 2019–2024 change into six contributions; labels are signed shares of the total constructed decline. The inset splits the nominal-sales contribution into broad-industry output-price, real-sales residual, and unmapped subrows. The balanced sample contains 1,065 funds. Portfolio reweighting combines trading, valuation, flows, and changes in the non-equity TNA denominator. The shaded 2022 observation marks the interim divergence, so the bridge applies to the five-year decline.

Broad-industry output prices account for −27.465 intensity points, or 63.8 percent of the nominal-sales term. The real-sales residual accounts for −15.273 points, or 35.5 percent, and the unmapped remainder is −0.319 points, or 0.7 percent. The first subrow uses a broad-industry price benchmark, not firm-specific transaction prices. The residual uses deflated sales and carries no physical-output interpretation.

The small equal-fund portfolio contribution masks active reweighting. Its annual contributions alternate sharply: −13.7, 1.2, 19.5, −11.8, and 4.0 intensity points. Across funds, the normalized portfolio term has an interquartile range from −22.7 to 32.5 percent, while the sales term has a median of −29.3 percent. Reweighting is active and dispersed. It contributes almost nothing to the industry's five-year equal-fund mean because positive and negative fund-level movements offset.

The offsetting mean conceals a sharp evaluation consequence. Three records separate the sources of that consequence. First, holding the evaluator fixed, the Refinitiv-based total and its portfolio-originated component have a Spearman correlation of 0.804, and 28.6 percent of funds receive opposite signs. Removing the nonportfolio components therefore changes responsibility even within one measurement system. Second, Morningstar's published total

and the Refinitiv-based total have a correlation of 0.711 and opposite signs for 14.3 percent, which measures evaluator disagreement before attribution. Finally, the published score and the portfolio-originated benchmark have a correlation of 0.598 and opposite signs for 33.5 percent. This last comparison is the decision-facing responsibility gap; it combines evaluator choice with attribution.

Appendix Table 22 reports the full published-to-portfolio transition. The portfolio-originated contribution combines trading, valuation, flows, and changes in the non-equity TNA denominator, so it is broader than pure trading or manager action. These are classifications of the same 1,065-fund population; no sampling inference is attached.

4.2 A moving ruler

Morningstar's Sustainability Rating, the fund-level globe rating of relative ESG risk, supplies a separate evaluation system from carbon intensity. Between 2020 and 2025, the median corporate ESG risk score of continuously rated funds falls from 24.50 to 21.00. The median peer threshold that converts the score into a published rating falls by almost the same amount. The published share of funds in the top two categories falls 4.5 percentage points. Applying each fund's own 2020 thresholds to the same later scores instead would raise that share by 46.3 to 72.4 percentage points, depending on the choice of 2020 anchor.

Table 2 reports the paths and their incidence. Morningstar assigns fixed percentile buckets within a global category. The published label therefore orders funds within a peer set rebuilt around them while leaving the level of that peer set unreported. A fund whose own score is unchanged over the six years is downgraded 94.3 percent of the time and loses 1.64 rating categories on average. A fund whose score improves by at least two points is still downgraded one third of the time.

Table 2. A rebased ruler conceals a change in the underlying score

Panel A. Scores, thresholds, and rating shares

Year	Median score	Median threshold	Top two published	Top two 2020 ruler
2020	24.50	25.03	29.1	45.0
2021	23.20	23.79	27.0	76.0
2022	22.46	23.53	25.6	86.0
2023	22.21	23.02	24.7	89.1
2024	21.60	22.18	24.2	93.3
2025	21.00	21.51	24.6	94.7

Panel B. Rating change by own-score change, 2020 to 2025

Own-score change	Funds	Downgraded	Mean rating change
Near stationary, $\lvert\Delta\rvert \leq 0.5$	88	94.3	-1.64
$\lvert\Delta\rvert \leq 1.0$	171	90.1	-1.43
$\lvert\Delta\rvert \leq 2.0$	559	87.7	-1.17
Improved by 2.0 or more	4,407	33.2	0.11

Balanced panel of 4,972 funds carrying a Morningstar sustainability rating, a portfolio corporate ESG risk score, and all four global-category thresholds in every year from 2020 through 2025. Lower scores denote lower assessed risk. The published rating is assigned on fixed percentile buckets of the global category, so its distribution is close to constant by construction. The final column of Panel A bands each observed score against that fund's category thresholds at its first 2020 observation. Across four monthly 2020 anchors, the 2020 to 2025 change in the counterfactual top-two share ranges from 46.3 to 72.4 percentage points, compared with a published change of -4.5. The final column is a counterfactual classification; all other labels are published. The score is a holdings-weighted aggregate of vendor company scores and can fall because funds trade or because held-company scores change. The data leave those channels combined.

The score can fall because a fund reweights its portfolio or because company scores attached to existing holdings change. The available fund-level rating history cannot separate those channels. The conclusion holds under either one. If funds improve, the relative label conceals that improvement. If company scores and peer thresholds drift together, the label nets two map movements against each other and displays neither.

The fund layer separates two questions that a total normally bundles. Distinct carbon systems agree that intensity fell. Attribution shows that the average decline came mainly from nominal sales growth and movement in issuer emissions values; common portfolio reweighting contributes almost nothing. Relative fund labels then show why agreement on a rank or category is different from identifying progress in levels. The verdict can be stable while responsibility remains hidden.

5 Who Moved the Institutional Footprint?

The fund evidence shows that a published carbon verdict can survive while its responsibility assignment changes. The longer institutional record reveals why. Figure 2 separates the component that moves the sector average from the components that move manager rankings. Panel A decomposes the average record into valuation, portfolio movement, and map reclassification under the lagged-map directional order used for the nested source anatomy. Panel B asks which components preserve the initial ordering of managers. Map reclassification moves the center of the distribution, while trading accounts for most of the compression in manager ordering.

The sector's average residualized emissions-green exposure falls from 0.0107 to −0.1168, a change of −0.1275. A manager-composition decomposition assigns −0.1316 to changes within continuing entities. Changes in continuing managers' AUM shares contribute 0.0035, and entity entry and exit contribute 0.0007. Composition changes therefore offset the decline slightly; the browning occurs inside the portfolios of continuing entities.

The continuing-manager accounting attributes a change of −0.1311 to valuation, portfolio movement, and map reclassification. Under the primary symmetric allocation, valuation contributes 0.0587, attributed reweighting contributes 0.0283, and reclassification contributes −0.2182. The map accounts for 71.5 percent of the sum of absolute components, with a range of 60.6 to 71.5 percent across attribution orders. Panel A uses the lagged-map endpoint of that range to retain the nested source anatomy: its portfolio and map terms are −0.0458 and −0.1137. Under that directional order, the map share is 89.8 percent during the concentrated 2016Q4–2019Q3 rise window. From 2019Q3 to 2023Q4, reclassification remains the largest component, at 51.7 percent of absolute anatomy (−0.0364 versus −0.0242 for portfolio movement). The

Figure 2. Anatomy of a changing institutional carbon footprint. Panel A decomposes the 2007Q1–2023Q4 change in average residualized emissions-green exposure among continuing eligible reporting entities, weighted by lagged assets under management (AUM). It uses the lagged-map directional order so that map reclassification can be connected to the nested source chain; the symmetric manager-layer allocation is primary in Section 6. Panel B decomposes the 2023Q4-on-2007Q1 cross-manager exposure slope for balanced entities. Bars separate valuation, trading or portfolio movement, and emissions-map reclassification. Shaded reclassification segments distinguish firm-reported emissions, vendor-estimated emissions, cross-sectional standardization, coverage entry, and coverage exit. The components are accounting contributions and assign no trading motive.

finding concerns a relative measured emissions score. Corporate-emissions quantities and investor trading are separate objects.

Within-industry movement drives the result. Applying the same accounting to negative standardized raw log emissions yields an aggregate decline of −0.5421. A Fama–French 12-industry midpoint split assigns 0.0582 to industry allocation and −0.6003 to movement within industries. Industry allocation offsets measured browning. This companion decomposition matters because emissions levels differ sharply across sectors. The aggregate direction appears within the industry cells used for the split, beyond a rotation toward high-emissions industries.

Table 3 decomposes the model's map term m_i into its empirical inputs. The table aligns the standardized-score anatomy and the raw-log source chain while keeping their units distinct. On the raw scale, the parent map term is −2.949. Changes in continuously covered firm-reported values contribute −0.125 and explain only 3.4 percent of absolute reclassification. Continuously covered vendor-estimated values contribute −0.135. First coverage in the current vintage contributes −3.076, while loss of coverage offsets 0.388. Coverage entry therefore dominates changes in reported emissions among firms already represented.

Table 3. What moves the institutional emissions footprint

Component	Contribution or share	90% interval
Panel A. Continuing-manager standardized-score anatomy		
Valuation	0.0283	
Trading	−0.0458	
Emissions-map reclassification	−0.1137	
Total change	−0.1311	
Reclassification share of absolute anatomy	60.6%	
Panel B. Raw emissions-map source chain		
Firm-reported share of reclassification	3.4%	[3.0%, 3.7%]
Coverage share of reclassification	90.7%	[90.1%, 91.3%]
Vendor-estimated share of coverage anatomy	79.0%	[78.0%, 80.1%]

Notes: The sample contains continuing eligible Form 13F reporting entities from 2007Q1 through 2023Q4. Panel A uses residualized emissions-green exposure per dollar of manager assets under management (AUM); missing scores contribute zero. Valuation carries lagged positions forward with returns, trading replaces passive positions with current holdings at lagged scores, and reclassification attaches current scores to current holdings. Panel B uses negative $\log(1 + CO_2)$, where CO_2 denotes carbon dioxide. Coverage and source labels describe the current Refinitiv vintage, not historical availability. Shares use the sum of absolute components. Intervals resample managers. Vendor estimates are classified inputs, not presumed errors.

A midpoint factorization makes the coverage result sharper. The two coverage figures sit at different levels of one nested decomposition. The source chain above classifies movement by origin among firms represented in the map, while the midpoint product split separates the measurement

perimeter from intensity, so each is exact within its own level rather than a restatement of the other. Coverage entry contributes −3.707 raw log-emissions units and coverage exit offsets 0.422, summing to the −3.285 coverage contribution. Exits thus undo 11.4 percent of entry-driven movement without any portfolio trade in either cell. The perimeter is selected: entrants carry 175.0 tonnes of CO_2e per $1 million of issuer equity value, compared with 138.4 for covered incumbents, while exiters carry 152.7 compared with 140.9 for their incumbents. Brown entry mechanically worsens the measured path; brown exit mechanically improves it.

Changes in the amount of covered position value account for 90.7 percent of absolute raw reclassification, with a 90 percent manager-bootstrap interval of 90.1 to 91.3 percent. Vendor estimates make up 75.5 percent of entering position value. This value share differs from the 79.0 percent contribution share obtained after estimated entries and exits are grouped in the same coverage anatomy. Estimated coverage entry contributes −2.802 and reported entry contributes −0.904; estimated and reported exits offset 0.207 and 0.215. Appendix Table 16 reports the reconciliation and physical boundary comparisons. The cumulative raw coverage path moves another −0.431 units after 2019Q3, while covered-position intensity offsets 0.055; both firm-reported and vendor-estimated entries continue through 2023Q4.

Expanding estimation can make a footprint more complete. It can also make an unchanged holding appear to brown when a relatively high-emissions issuer first receives a finite value. The accounting leaves the choice between broad estimated coverage and narrower reported coverage to the evaluator. It shows that a monitoring record should disclose the coverage path and source composition before assigning the resulting movement to the manager.

Figure 3 isolates the coverage component inside m_i and shows how it changes the measured record y_i. The solid path is cumulative map reclassification. The figure separates coverage from intensity, divides coverage by source, and adds a full footprint path that prevents changes in finite coverage from entering. The observed continuing-manager raw footprint becomes browner by 0.993 log-exposure units over the common endpoints. Holding coverage fixed while retaining observed portfolios, valuation, and intensity among covered positions produces an improvement of 2.292 units. Both manager-bootstrap intervals exclude zero. Coverage reverses the sign.

The accounting also closes on the residualized, within-quarter standardized scale. Firm-reported change, vendor-estimated change, standardization, coverage entry, and coverage exit contribute −0.039, −0.037, −0.026, −0.008, and −0.003 and reconcile to the −0.114 reclassification term. Their proportions differ from the raw source split because residualization and standardization distribute cross-sectional movement across continuously covered firms. We use

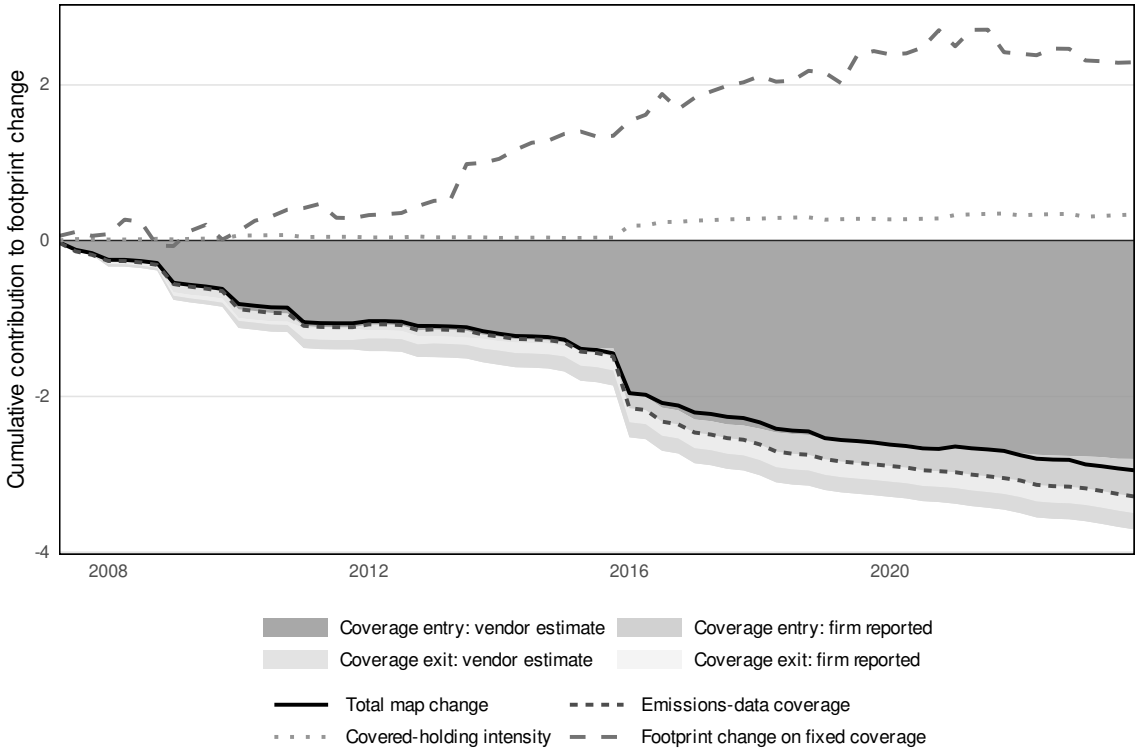

Figure 3. Institutional decarbonization under moving and fixed emissions coverage. The solid line is cumulative map reclassification for continuing eligible reporting entities; the dashed and dotted lines split it into finite-emissions coverage per dollar of manager AUM and emissions intensity among covered positions. Shaded bands divide coverage entry and exit by firm-reported versus vendor-estimated values. The long-dashed gray line is the full footprint change with coverage held fixed. It preserves the observed portfolio, valuation, and covered-position intensity while leaving uncovered firms outside the footprint. All paths use negative $\log(1 + \mathrm{CO}_2)$, where CO_2 denotes carbon dioxide, so more negative values denote browning. All displayed decompositions close at each transition.

the raw split to classify the footprint inputs and the residualized split to connect those inputs to the manager score.

The source chain also explains why the endpoint share of estimated values is insufficient. Changes in estimates for firms covered at both dates are small relative to coverage entry. The main time-series contribution comes when a security moves from no finite value in the current-vintage panel to a finite value, most often an estimate. Attribution therefore requires the coverage state as well as the source label.

Panel B of Figure 2 establishes a separate portfolio fact. The 2023Q4-on-2007Q1 cross-manager exposure slope is 0.159, with a 90 percent interval of $[0.072, 0.251]$. Trading contributes -0.867 to that slope, while valuation and reclassification contribute 0.107 and -0.081. Manager ordering therefore converges substantially through portfolio trading even though map movement dominates the sector mean.

The contrast prevents two common interpretations from being collapsed. A component can move the sector's average without determining which managers converge, and a component can reorder managers without explaining the average trend. Here the map governs the aggregate footprint conclusion, while trading erases much of the initial cross-sectional ordering. The accounting identity applies to both questions, but the estimands are different.

6 The Consequence for Manager Accountability

The sector result matters for delegated finance only if map movement loads differently across managers. A common additive map shift would change the aggregate footprint but preserve every relative evaluation. The empirical question is therefore whether cross-manager differences in m_i are large enough to overturn differences in the portfolio-originated record p_i. Proposition 1 implies that manager i is ranked below manager j despite $p_i > p_j$ exactly when $m_j - m_i > p_i - p_j$. The decomposition observes each term in this condition.

We evaluate the balanced sample of 861 actual FactSet reporting entities present in all 68 quarters from 2007Q1 through 2023Q4. For each entity, we cumulate the measured record y_i and compare it with the portfolio-originated record $p_i = y_i - m_i$. The comparison holds fixed the manager population, endpoints, denominator, and current data vintage. It changes only whether cumulative map reclassification is included in the performance record.

6.1 Trend signs and manager assignments

Table 4 reports the classification consequences. Our primary convention splits the portfolio-by-map interaction equally, the two-factor Shapley allocation averaging the two directional orders. Map movement then accounts for 71.5 percent of aggregate absolute movement, changes the long-run trend sign for 59.7 percent of managers, and leaves a measured-to-portfolio rank correlation of 0.406. The lagged- and current-map orders bound these quantities: the map share is 60.6 to 71.5 percent, sign reversals are 36.9 to 69.8 percent, and rank correlations are 0.269 to 0.494.

We use the symmetric point estimate because it is the unique equal allocation of the interaction. The two directional orders show how far responsibility can move when one side receives the full cross term. A manager can appear to green under the measured record and brown under the portfolio record, or the reverse, even though the observed portfolio path is the same. Both records use the same vendor and emissions definition; only the contribution produced by movement in the map is shown separately.

Table 4. Manager responsibility under the symmetric convention and directional bounds

Convention	Map share	Spearman	Flip EW	Flip AUM	Top exit EW	Top exit AUM	Bottom exit EW	Bottom exit AUM
Symmetric midpoint (primary)	71.5%	0.406	59.7%	80.9%	60.5%	45.9%	71.3%	78.8%
Lagged map (bound)	60.6%	0.494	36.9%	32.5%	53.5%	37.5%	66.7%	76.5%
Current map (bound)	62.7%	0.269	69.8%	90.1%	70.9%	56.5%	73.6%	84.5%

Notes: The sample contains the same 861 long-lived actual reporting entities. The symmetric midpoint splits the portfolio-by-map interaction equally and is the primary convention; the lagged-map and current-map rows provide directional bounds. AUM-weighted incidence uses 2023Q4 manager AUM. Deciles remain equal-count assignments before AUM weights are applied to exit incidence.

Figure 4 displays the complete transition matrix. The diagonal is visible, but it is not dominant. Managers initially placed in the same measured decile disperse across several symmetric portfolio-record deciles, including movements across the center of the distribution. Appendix Table 17 reports every cell. The figure summarizes a monitoring consequence: the identity of the managers singled out by the measured record changes after attribution. Its estimand is the change in monitoring classifications.

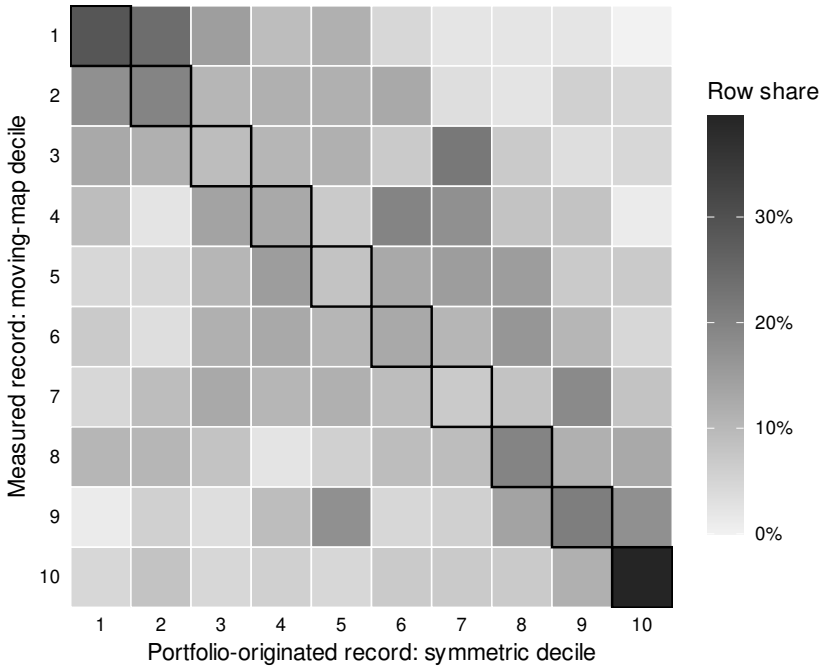

Figure 4. Manager assignments after emissions-map movement is removed. Rows rank managers by the measured record y and columns rank the same managers by the portfolio-originated record $p = y - m$ under the symmetric convention. Each cell reports the share of managers in its moving-map row assigned to the corresponding portfolio-originated decile. Darker cells contain larger shares; outlined cells preserve the original decile. Decile 1 is the most-browning trend and decile 10 the most-greening trend. Only 17.9 percent of managers remain on the diagonal. The sample contains 861 actual reporting entities present in all 68 quarters.

6.2 Why the reordering is quantitatively large

Table 5 connects these classifications to the cross-manager moments in the framework. Panel A reports the six-term expansion of

$$\mathrm{Var}(y) = \mathrm{Var}(a) + \mathrm{Var}(v) + \mathrm{Var}(m) + 2\,\mathrm{Cov}(a, v) + 2\,\mathrm{Cov}(a, m) + 2\,\mathrm{Cov}(v, m).$$

The positive variance terms are individually large relative to $\mathrm{Var}(y)$ because the covariance terms offset them. In particular, $2\,\mathrm{Cov}(a, m)$ equals -0.0376, or -1.61 times the variance of the total measured record. Aggregating the identity into one footprint therefore hides economically large and opposing manager, valuation, and map components.

Map dispersion is 44.9 percent of $\mathrm{Var}(m_i) + \mathrm{Var}(p_i)$, while the portfolio share is 55.1 percent. The standard deviation of map movement is 0.903 times the standard deviation of portfolio-originated movement. These are the moments that give Proposition 1 quantitative force. Cross-manager map differences are nearly as dispersed as the record the evaluator intends to compare.

Table 5. Signal content of the measured manager record

Object	Contribution or estimate	Share of Var(y)
Panel A. Variance anatomy of the measured record		
Var(a)	0.043316	1.858
Var(v)	0.011828	0.507
Var(m)	0.028117	1.206
$2\,\mathrm{Cov}(a, v)$	−0.020682	−0.887
$2\,\mathrm{Cov}(a, m)$	−0.037633	−1.614
$2\,\mathrm{Cov}(v, m)$	−0.001630	−0.070
Var(y)	0.023316	1.000
Panel B. Dispersion and signal extraction		
Map share of map-plus-portfolio dispersion	44.9%	
Portfolio share of map-plus-portfolio dispersion	55.1%	
Map-to-portfolio standard-deviation ratio	0.903	
Corr(p, y)	0.523	
Variance of p explained by y	27.4%	
Optimal projection weight on centered y	0.636	

Notes: The sample contains the same 861 balanced reporting entities as Table 4. The measured record is $y = a + v + m = p + m$, where a is trading, v is valuation, m is map reclassification, and $p = a + v$. Panel A expands Var(y) into the six variance and covariance terms; shares use the common denominator Var(y). In Panel B, map and portfolio dispersion shares divide Var(m) and Var(p) by their sum. The projection weight is $\mathrm{Cov}(p, y)/\,\mathrm{Var}(y)$.

Under that same lagged-map order, the moments measure the information in the standard record. Corr(p_i, y_i) is 0.523, so a linear projection on the measured record explains 27.4 percent of the cross-manager variance in the portfolio-originated record. The best linear predictor places weight 0.636 on centered measured changes. Because this weight is positive, it shrinks the scale without changing any rank. The contrast isolates the value of attribution. A statistical correction based only on y_i can improve a point forecast, but it cannot undo the manager reordering produced by m_i.

6.3 Which parts of the map drive the ordering?

Portfolio-originated and map changes have a correlation of −0.631. Table 6 applies the nested decomposition to that correlation. Cross-sectional standardization contributes −0.336, vendor-estimated emissions changes contribute −0.155, coverage entry contributes −0.146, and coverage exit contributes −0.003. Changes in firm-reported emissions offset these negative contributions by only 0.009. The component rows add to the parent correlation up to rounding.

Table 6. Which map components change manager ordering?

Map component	Contribution to $\mathrm{Corr}(p, m)$
Firm-reported emissions changes	0.009
Vendor-estimated emissions changes	−0.155
Cross-sectional standardization	−0.336
Coverage entry	−0.146
Coverage exit	−0.003
Total, $\mathrm{Corr}(p, m)$	−0.631

Notes: The sample contains the same 861 balanced reporting entities as Table 4. Each row divides the covariance between the portfolio-originated record p and the indicated map component by $\mathrm{SD}(p)\,\mathrm{SD}(m)$. The five rows sum to the parent correlation up to rounding. Reported and estimated emissions changes refer to continuously covered observations in the current Refinitiv vintage; coverage entry and exit record changes between missing and finite scores.

The reordering spans several parts of the moving map. These components load most negatively on managers whose portfolios have more favorable portfolio-originated changes. Standardization is the largest contributor, but estimated values and coverage entry are separately material. The aggregate source result and the manager covariance result answer different questions: coverage dominates the level movement in the raw sector record, while several map components jointly determine which managers move in the ranking.

Some of this heterogeneous loading is visible before the subsequent map movement occurs. Table 7 uses one predetermined exposure: a manager's average uncovered position-dollar share in 2007Q1–2007Q4. Its Spearman correlation with subsequent cumulative map movement is 0.233. The Pearson correlation is 0.158, and the univariate slope is 0.168 with an HC2 standard error of 0.046. Initial coverage therefore has moderate predictive content for subsequent map exposure.

Table 7. Predetermined coverage exposure and subsequent map movement

Statistic	Estimate	HC2 standard error
Spearman correlation	0.233	
Pearson correlation	0.158	
Univariate slope	0.168	0.046

Notes: Predetermined exposure is a manager's average uncovered position-dollar share in 2007Q1–2007Q4. The outcome is cumulative map reclassification through 2023Q4. The slope comes from a univariate regression with an intercept and an HC2 standard error. The association measures whether initial portfolio composition predicts later map movement and carries no causal interpretation.

Two managers with different initial security sets therefore need not receive a common map

shift. The association establishes predictable exposure to map movement, which is the condition in Equation (10). Deliberate pursuit of that exposure remains untested.

6.4 What an attributed manager report reveals

Removing m_i recovers $p_i = a_i + v_i$. Isolating purchases and sales also requires removing valuation. A mandate that rewards the path of the held portfolio can compare y_i with p_i. A mandate that rewards trading must also show valuation and use $a_i = y_i - v_i - m_i$. The evaluator chooses the target; the decomposition supplies the information needed to implement that choice. A total footprint describes the measured portfolio; its components assign responsibility.

6.5 Economic scope of the manager comparison

The reordering extends to the largest reporting entities. Under the symmetric convention, 80.9 percent of 2023Q4 AUM belongs to managers whose measured and portfolio-originated trend signs differ. The two directional orders bound this incidence at 32.5 to 90.1 percent.

The convention-invariant core clarifies what survives the order choice. In all, 502 of 861 managers, or 58.3 percent, keep the same portfolio-originated trend sign under all three orders. They represent 41.2 percent of the balanced sample's $25.5 trillion in endpoint AUM. The symmetric point estimate and the directional bounds therefore report different economic objects: the former provides a principled responsibility allocation, while the latter shows how much classification depends on assigning the interaction to one side.

The reporting entities also exhibit persistent portfolio heterogeneity. Appendix Table 20 evaluates emissions styles on observations not used to estimate the type. Odd- and even-quarter ranks correlate 0.982, while nonoverlapping formation and evaluation periods correlate 0.436. The early top-minus-bottom quintile remains 1.266 standard deviations apart in the later period. The decline across horizons is informative. Reporting-entity styles persist without implying permanent climate preferences.

A total footprint rank leaves the contribution of the evaluator's intended target hidden. In this sample, attribution reverses trend signs for 59.7 percent of managers under the symmetric convention, with 36.9 to 69.8 percent across the directional orders. The total remains intact, but responsibility no longer belongs implicitly to the manager.

7 Economic Consequences

An allocator needs different information to select a low-carbon portfolio, select a manager for repeated implementation, and assess progress against a target. The total footprint and its components answer these questions differently. The selection exercise measures the information supplied by attribution for implementation. Removing map change raises the top-minus-bottom spread in future attributed reweighting by 0.299 standard deviation, and the attributed top quintile improves on the conventional top quintile by 0.219 standard deviation. The dated-archive comparison below tests a second decision, whether the same completed portfolio record passes a fixed illustrative target. Both exercises evaluate stated rules; the flow analysis separately measures investor response to published globe changes.

7.1 Selection under total and attributed records

For a mandate that rewards portfolio reweighting toward lower-emissions firms, attribution supplies a record of implementation relative to passive holdings. The relevant score here is negative residualized issuer emissions, conditional on industry and firm characteristics. It differs from both absolute financed emissions and the Morningstar globe rating. We compare hypothetical selection on the total record with selection on attributed reweighting. Formation uses observations dated through 2013Q4, evaluation uses 2015Q2 through 2023Q4, and 2014 and the bridge transitions are omitted. The same 2,154 reporting entities enter every ranking. Because issuer histories come from the current vintage, the exercise measures retrospective signal content, rather than an investment strategy implementable with information available at each historical date.

The nested records locate the source of the selection difference. The conventional top quintile subsequently trails its bottom quintile by 0.193 standard deviation of future attributed reweighting. Removing map change increases that spread by 0.299 standard deviation, with a 90 percent interval of $[0.184, 0.425]$. Removing valuation adds another 0.088 standard deviation $[0.015, 0.150]$, leaving a positive attributed spread of 0.194. Map movement accounts for about 77 percent of the full 0.387-standard-deviation improvement. The first step replaces 139 of 431 conventionally selected managers; full attribution replaces 171. The pattern shows that the total obscures information useful for this implementation target.

The replacement margin measures the consequence for the selected set. Managers admitted only by attribution subsequently outperform those selected only by the conventional record by 0.552 standard deviation of attributed reweighting, with a 90 percent interval of $[0.311, 0.745]$.

Equation (9) combines this gap with the 39.7 percent replacement rate. The attributed top quintile improves on the conventional top quintile by 0.219 standard deviation, with a 90 percent interval of $[0.123, 0.298]$ and a 95 percent interval of $[0.109, 0.312]$. These are manager-bootstrap intervals conditional on the observed calendar history.

The gain concerns selection in the tails. The attributed formation record has a full-rank Spearman correlation of -0.011 with subsequent action, with a 90 percent interval of $[-0.050, 0.029]$. It therefore does not provide a monotone ranking of future implementation across managers. Its top quintile nevertheless has positive mean subsequent action of 0.00299, with an interval of $[0.00175, 0.00421]$.

An alternative outcome removes the attribution decomposition from the evaluation period. Holding the formation rankings fixed, we measure the change in portfolio-weighted green score between 2015Q1 and 2023Q4 using the same 2015Q1 score vector from the current vintage at both endpoints. Among managers observed at those endpoints, attribution-only managers exceed conventional-only managers by 0.046 score units, and the attributed top exceeds the conventional top by 0.018 units, with a 90 percent interval of $[0.008, 0.029]$. All four selection groups decline on this fixed score vector, so the contrast is less browning among attribution-selected managers. This outcome removes changing scores and the interaction convention but retains the same issuer characteristic and includes valuation effects. A coarser outcome, the reduction in weight in the dirtiest covered quintile, has the same sign but is statistically unresolved.

The physical outcome sets a material limit on the interpretation. On the existing lagged-score physical trading measure, the attributed top minus the conventional top is -0.18 tonnes of CO_2e per million dollars of beginning AUM per quarter, with a 90 percent interval of $[-0.76, 0.32]$. Higher values denote lower exposure. Thus the selection advantage in conditional issuer scores does not establish an advantage in physical portfolio emissions. The appendix reports the group means and the separate endpoint samples.

The attributed record ranks managers on conditional improvement rather than on the level of exposure, and the two orderings differ. In a retrospective comparison over the same window, equal-manager means among eligible covered sleeves place the attributed top quintile 19.8 tonnes of CO_2e per million dollars of issuer sales above the conventional top quintile, with a 95 percent interval of $[5.2, 33.3]$. Screening for exposure and attributing action are separate objectives and call for separate statistics.

The selection gain extends across mandate capacities. Varying the selected share from 5 to 50 percent, while retaining the sample, dates, outcome, and equal-weight allocation, gives positive

gains at all ten capacities. Figure 5 reports pointwise 90 percent intervals; each lower bound exceeds zero. The gain is 0.372 standard deviation at 5 percent capacity and 0.096 at 50 percent. Removing map movement accounts for most of the gain throughout the curve.

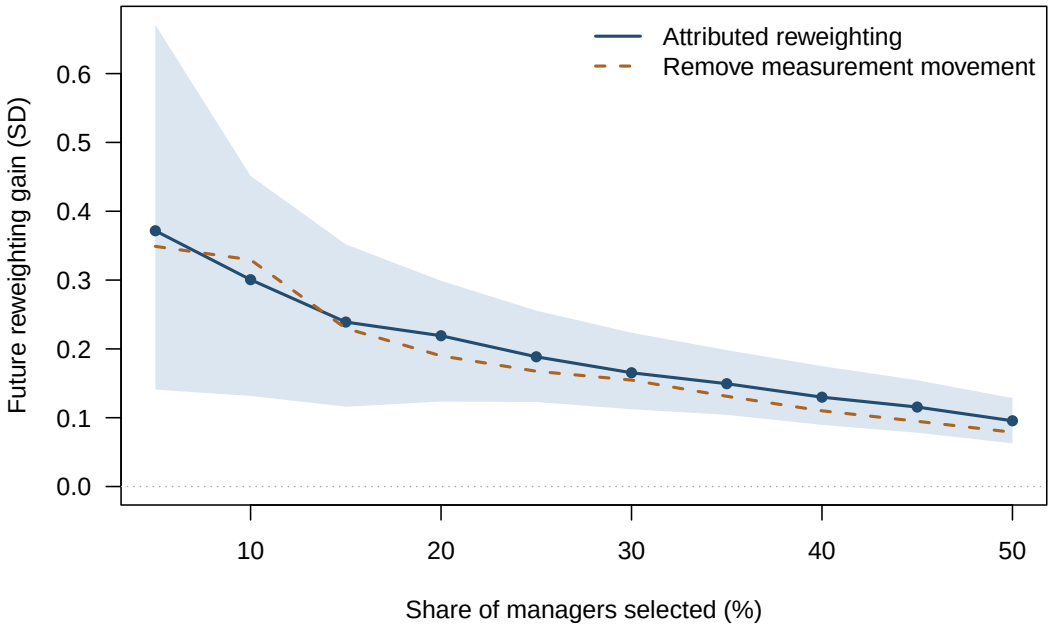

Figure 5. Manager selection as mandate capacity changes. Formation uses observations dated through 2013Q4 and evaluation uses 2015Q2–2023Q4, omitting 2014 and bridge transitions. The solid line compares mean future attributed reweighting under selection on attributed and total formation records. The dashed line removes map movement while retaining valuation. The shaded region is a pointwise 90 percent manager-bootstrap interval for the full gain. Values use the cross-manager standard deviation of future action. The sample contains 2,154 managers, selection is equally weighted, and issuer histories are from the current data vintage.

Table 8. Manager selection and subsequent reported equity asset growth

Panel A. Selection for subsequent attributed reweighting				
Formation record	Selected	Spread (SD)	Increment (SD)	Replaced
Total measured, y	431	-0.193	–	–
Portfolio-originated, p	431	0.106	0.299 [0.184, 0.425]	139
Attributed reweighting, a	431	0.194	0.088 [0.015, 0.150]	171
Full attribution versus total	–	–	0.387 [0.261, 0.502]	171
Panel B. Managers selected by only one record				
Formation selection	Managers	Action (SD)	Asset growth	Net asset growth
Total only	171	-0.070	5.39%	-4.01%
Attribution only	171	0.482	1.37%	-7.59%
Total minus attribution	–	-0.552 [-0.745, -0.311]	4.02 [0.60, 6.57]	3.58 [0.10, 6.86]

Notes: The sample contains 2,154 reporting entities. Formation uses 2007Q2–2013Q4 and evaluation uses 2015Q2–2023Q4, omitting 2014 and bridge transitions. Both passive valuation and actual-minus-passive portfolio reweighting are evaluated at midpoint issuer scores. Panel A reports top-minus-bottom future-action spreads; replacements are relative to the total-record top quintile. The attributed top improves on the total-record top by 0.219 SD [0.123, 0.298]. Asset growth compares mean log AUM in 2015Q2–2016Q1 and 2023Q1–2023Q4; endpoint coverage is 76.6% for total-only and 70.8% for attribution-only managers. Net growth removes holdings-implied price changes and is available for all managers. Growth differences are annualized percentage points. Brackets are 90% manager-bootstrap intervals, conditional on the observed calendar history. Selection is hypothetical and asset growth is descriptive.

Reported assets grew faster for the managers the total record favors. Conventional-only managers, whom the total record selects and attribution removes, grow reported 13F equity AUM 4.02 percentage points per year faster than the attribution-only managers admitted in their place, with a 90 percent interval of $[0.60, 6.57]$. Mean annual growth is 5.39 percent and 1.37 percent, respectively. Endpoint retention is 76.6 percent and 70.8 percent. Growth net of holdings-implied price changes differs by 3.58 percentage points on the full manager sample, with a 90 percent interval of $[0.10, 6.86]$; its 95 percent interval includes zero. Both net-growth group means are negative, so this second comparison describes less contraction among conventional-only managers. Table 8 reports both margins alongside the selection results that define the two groups.

The asset-growth contrast leaves its cause unresolved. Manager type, passive investing, business combinations, and reporting changes can affect 13F AUM. The conventional-only group grows faster, but the comparison identifies neither client demand for its carbon record nor dollars allocated because of the ranking. The evidence establishes useful information for the stated implementation objective. An observed allocation decision or an independently valued mandate outcome would be needed to establish a realized delegation cost.

7.2 Reassessing a completed investment record

A historical portfolio can receive a different target assessment when its measurement inputs change. We hold the same 861 managers' actual 2019Q4 and 2023Q4 holdings, reported AUM, issuer equity market caps and security links fixed across two Refinitiv archives collected on July 18 and September 22, 2026. Only the emissions archive changes. The score uses FY2018 and FY2022 emissions, respectively, consistent with the one-year fiscal lag.

Let $F_{it}^{(v)}$ denote equity-attributed emissions per million dollars of reported AUM at endpoint t, evaluated with archive v. The illustrative target passes when $F_{i,2023Q4}^{(v)}/F_{i,2019Q4}^{(v)} \leq 0.93^4$. All 861 managers have positive starting footprints and positive covered position value at both endpoints under both archives. Missing emissions contribute zero to the measured footprint, so coverage travels with the assessment. The calculation concerns the observed equity record, not actual certification under a regulation or a client contract.

Table 9. Reassessing the same completed portfolio record

Coverage basis	July pass	September pass	Lose pass	Gain pass
Available in each archive	563	554	17	8
Common across archives	554	554	0	0
Emissions coverage as a share of reported AUM (%)				
Endpoint	July	September		
2019Q4	68.5	67.8		
2023Q4	73.3	72.7		

Notes: The same 861 reporting entities' actual 2019Q4 and 2023Q4 holdings are evaluated with July 18 and September 22, 2026 emissions archives. Holdings, reported AUM, issuer equity market caps and organization-to-security links are fixed across archives. Reported emissions take priority over estimates; fiscal years are 2018 and 2022. A pass requires the ending equity-attributed intensity divided by its positive starting value to be at most 0.93^4. Missing emissions contribute zero. Common coverage retains securities measured in both archives at each endpoint, with original AUM weights. Counts describe an illustrative assessment of the observed equity record, whose emissions coverage is incomplete.

Table 9 shows 25 changed assessments, or 2.9 percent of managers. Seventeen lose a pass and eight gain one. Coverage loss can work in either direction. Removing a measured issuer from the final footprint lowers that footprint, while removing one from the baseline can make the required decline harder to achieve.

Changing the coverage mask alone reproduces all 25 switches. Updating values among continuously covered records while retaining the earlier coverage produces none. In this diagnostic, July values are retained for records that disappear in September. A second comparison evaluates both archives on their common covered securities at each endpoint, preserving the original AUM weights. Both then produce 554 passes, with no individual switches. The coverage adjustment

therefore resolves the classification instability in this example, even though some continuous scores still change.

The scope of coverage limits the decision these counts can support. Emissions cover 68.5 percent of reported AUM at the starting endpoint and 73.3 percent at the ending endpoint under July data; the September shares are 67.8 and 72.7 percent. Coverage among represented security positions is much higher, but those positions do not exhaust reported AUM. Thus the switches establish instability in a constructed assessment, not whole-portfolio mandate failure. The independent archive diagnostics and fixed-date rank comparison appear in Appendix A.12. The rank changes over this short interval are small, and common-coverage value revisions do not change target passage.

7.3 Physical scale

Standardized scores establish direction and attribution; physical units reveal the economic scale of the coverage component in m_i. Table 10 translates the measured record y_i and the institutional-sector fixed-coverage adjustment into raw and physical units. The translation uses the ownership approach common in financed-emissions measurement: each portfolio is assigned its equity share of an issuer's emissions, with CRSP equity market capitalization as the ownership denominator. Portfolio intensity divides that quantity by manager AUM, and positions without finite emissions contribute zero. The calculation changes units, not the sample or the underlying coverage counterfactual.

Measured equity-attributed intensity falls from 100.96 to 35.54 tonnes of CO_2 equivalent per million dollars of AUM, a decline of 64.8 percent. At the same time, eligible-manager AUM rises from $13.8 trillion to $43.9 trillion. The corresponding measured financed-emissions quantity therefore increases from 1,397.4 to 1,558.7 $MtCO_2e$. The opposite directions arise because intensity divides by an expanding asset base, while the level assigns emissions to a much larger set of institutional dollars.

The fixed-coverage comparison isolates the map contribution inside this physical object. Expanding coverage adds 56.69 tonnes per million dollars of AUM relative to the path generated by holdings, valuation, and emissions among already covered positions. Multiplying that additive contribution by 2023Q4 endpoint AUM gives 2.49 $GtCO_2e$. This is the scale of the coverage component under the stated attribution rule. Sector endpoint emissions, emissions caused by institutions, and the effect of divestment on production lie outside this calculation.

The raw-log and physical results make the same monitoring point from different directions.

Table 10. The physical scale of the moving-coverage footprint

Object	2007Q1	2023Q4
Raw footprint change, moving coverage	−0.993	
Raw footprint change, fixed coverage	2.292	
Equity-attributed intensity (tCO_2e per \$1m AUM)	100.96	35.54
Eligible-manager AUM (trillion dollars)	13.8	43.9
Equity-attributed financed emissions ($MtCO_2e$)	1,397.4	1,558.7
Coverage contribution (tCO_2e per \$1m AUM)	56.69	
Coverage contribution at endpoint AUM ($GtCO_2e$)	2.49	

Notes: Raw footprint changes compare the common endpoints and use negative $\log(1 + CO_2)$. The fixed-coverage path preserves observed portfolios, valuation, and emissions intensity among covered positions while removing changes in finite-emissions coverage. Physical quantities attribute issuer emissions by the portfolio's equity ownership share, use CRSP equity market capitalization as the attribution denominator, and scale by manager AUM; missing emissions contribute zero. The 2.49 $GtCO_2e$ entry multiplies the coverage-path contribution by 2023Q4 endpoint AUM. It measures the coverage component under this rule, separate from endpoint footprint levels and emissions caused by institutions.

On the raw-log scale, moving coverage turns an improving path into measured browning. In physical intensity, observed coverage expansion adds a large positive emissions quantity relative to fixed coverage even as covered-position intensity improves. The measurement map can therefore move both a relative score and an ownership-scaled footprint against the record implied by positions that were already measurable.

Useful estimates and clean attribution address separate questions. Better coverage can make a database more informative at the endpoint while changing the apparent historical performance of portfolios that held the newly measured issuers. An evaluator may prefer the more complete endpoint. Attribution should show the increase in measured scope as coverage, leaving manager action and firm abatement separate. Completeness and accountability are distinct properties of the record.

Issuer equity market capitalization is another source of measured progress. A rise in that denominator lowers the assigned intensity with unchanged emissions. Appendix A.13 separates this effect from emissions-value changes and coverage in the balanced manager sample. Market-cap movements lower measured intensity substantially, while coverage entry raises it. Neither component can be read as vendor revisions or manager trading. This denominator is distinct from issuer sales in fund WACI and from manager AUM in portfolio weights.

7.4 Flow response to peer-cutoff-driven globe downgrades

We ask whether capital penalizes a published downgrade in the Morningstar Sustainability Rating, the fund-level globe rating of relative ESG risk, classified as driven by changing peer cutoffs. The globe rating is distinct from Morningstar's Star rating, which scores past performance and enters below only as a detection benchmark. We reconstruct each globe from the fund score and its peer breakpoints. A cutoff-driven event requires an actual published downgrade that also occurs when the new cutoffs are applied to lagged own-score inputs, whereas current own-score inputs under lagged cutoffs leave the rating unchanged. The observed score itself need not be constant. Within the CRSP-linked population from 2020 through 2025, each treated fund is matched to nonchangers in the same peer-map episode on its lagged score level and score path. Every treated observation therefore carries an actual published change; the estimand requires no first-stage translation from map exposure to a label.

Table 11 first verifies that the flow measure detects ordinary demand around Morningstar's performance verdict. Star upgrades are followed by 36.5 basis points of net inflow and Star downgrades by 48.1 basis points of net outflow, both with $p < 0.001$. This detection benchmark reproduces the familiar sensitivity of fund flows to performance ratings (Del Guercio and Tkac, 2008) before the same measure is applied to the evaluator-side sustainability event.

An evaluator-caused sustainability downgrade produces a response of -1.92 basis points in the first fully public month, with a standard error of 15.22 basis points and a joint pretrend p-value of 0.950. The expected-negative one-sided 95 percent bound is about -27 basis points of monthly organic growth. Cumulating the first three full public months gives -33.02 basis points, with a standard error of 31.81. At the salient one- and five-globe extremes, the first fully public month estimate is 8.5 basis points, with a standard error of 30.0.

The estimate bounds the average flow response to cutoff-driven globe changes within this population and period. An investor concerned with absolute portfolio risk may disregard a change in a relative label; limited attention can generate the same response. The design distinguishes neither explanation and does not estimate demand for carbon intensity or attributed portfolio implementation. The Star benchmark establishes that the flow measure detects performance-related movements, while the cause-classified globe design differs from pooled rating-flow tests, whose relation weakens after 2018 (Gantchev et al., 2024).

Expense records provide a secondary test on a slower margin. Holding the pre-event share classes and their TNA weights fixed, the first full post-publication fiscal record changes the expense ratio by -0.218 basis points, with a 95 percent interval of $[-0.734, 0.298]$. This estimate

is imprecise but rules out changes larger than about three quarters of a basis point in either direction. The data lack an announced fee change and effective date, so we treat this as supporting evidence, separate from the event-study result.

Table 11. Fund flows after Star changes and cutoff-driven globe downgrades

Signal	Horizon	Flow $t+1$ or $t+3$ (bp)	SE	p	Cum. flow $t+3{:}t+5$ (bp)	SE	p	One-sided 95% lower bound (bp)
Star upgrade	$t+1$	36.5	3.8	< 0.001	–	–	–	–
Star downgrade	$t+1$	-48.1	3.9	< 0.001	–	–	–	–
Cutoff-driven globe downgrade	$t+3$	-1.9	15.2	0.900	-33.0	31.8	0.300	-26.9

Notes: Organic flow is asset growth net of the return on lagged assets. The Star rows are a detection benchmark for the flow measure and carry no causal interpretation of the Star label. The globe row compares an actual cutoff-driven published downgrade with matched nonchangers in the same peer-map episode. The first full public month is $t+3$; $t+1$ is the omitted reference month. Matching weights are used, and standard errors are clustered by treated peer-map episode and fund. The lower bound is the expected-negative one-sided 95% bound for the per-label $t+3$ response.

8 Implications for Delegated Evaluation

8.1 The evaluator's problem

A portfolio footprint can describe current exposure without measuring what a manager did. The measured record y_i combines purchases and sales a_i, valuation v_i, and movement in the evaluator's map m_i. These components belong to different economic actors. Treating their sum as manager performance assigns issuer and evaluator changes to the manager by construction.

The appropriate record depends on the mandate. An investor concerned with the contribution of changing portfolio weights can use $p_i = y_i - m_i$. An investor concerned with purchases and sales can use $a_i = y_i - v_i - m_i$. The first target includes the effect of valuation on portfolio weights; the second excludes it. Neither target requires discarding the familiar footprint. It requires reporting the total alongside the components needed to assign responsibility.

The distinction matters because map movement is neither small nor common across portfolios. Holding coverage fixed reverses the institutional sector's trend. Removing map reclassification changes the trend sign for 59.7 percent of long-lived managers under the symmetric convention, with directional-order bounds of 36.9 to 69.8 percent. A statistical correction based only on the total alone cannot recover a reversed ranking when it is strictly increasing. A positive linear projection rescales the score while preserving ranks; a model with additional information could reorder managers.

Map changes can nevertheless contain useful information. If a previously uncovered brown

issuer enters the database, the new footprint may describe the fund's exposure more accurately. The evaluator should update the exposure measure, while keeping the update separate from manager credit or blame. Coverage exit creates the opposite risk: a footprint can improve when a brown holding disappears from the measurement perimeter. Better measurement and manager attribution are therefore complementary objectives.

Four kinds of movement sit behind a footprint change, and they do not carry the same claim on an investor's attention. Issuer disclosure reports something new about the firm. A vendor estimate reports something new about a model. A standardization rule changes the yardstick, and coverage entry or exit changes which firms are measured at all. The first two can inform a view of what the portfolio holds, while the others move the ruler or the perimeter rather than the firm. None of the four is a manager action, so an evaluator who assigns credit from the total rewards managers for all of them.

8.2 Selection and incentives

Component reporting changes the information available for delegation. A report containing (x_i, u_i) retains the total $y_i = x_i + u_i$, so an allocator can reproduce any decision based on the total. It also permits decisions that condition separately on manager performance and evaluator movement. The empirical question is whether that added information changes which managers a mandate would select.

The held-out exercise answers that question for a simple selection rule. The conventional and attributed records replace 171 of 431 top-quintile managers. Those admitted only by attribution subsequently outperform those selected only by the conventional record by 0.552 standard deviation of attributed reweighting. Equation (9) combines this gap with the 39.7 percent replacement share, yielding a 0.219-standard-deviation gain for the attributed top quintile. The gain remains positive when the mandate holds between 5 and 50 percent of managers, and most of it comes from removing map movement. The exercise measures signal content rather than realized mandate allocation. It shows that attribution improves selection for reweighting toward firms with lower emissions conditional on industry and firm characteristics. The physical-emissions comparison remains unresolved, and the selected group has higher covered-portfolio carbon intensity. The result therefore depends on the mandate valuing this implementation objective.

The incentive implication is narrower. Equation (10) shows that a reward tied to the total loads on both the intended performance target and any expected map movement associated with portfolio choice. Initial uncovered portfolio share predicts subsequent map movement,

establishing predictable exposure to evaluator movement. A direct test of deliberate gaming would require the contracts and point-in-time information available to managers.

The fund-flow result asks a separate behavioral question. Flows respond strongly around Morningstar performance-rating changes, while their average response to globe downgrades classified as driven by peer cutoffs is indistinguishable from zero. This comparison concerns relative ESG-risk labels. It does not establish whether investors value attribution for portfolio implementation or use it when selecting managers.

8.3 A reporting rule

A responsibility report should disclose five objects: the portfolio date, the emissions-data vintage, the denominator, the coverage path, and the split between portfolio and map movement. Within the map term, issuer-reported values, vendor estimates, cross-sectional transformations, coverage entry, and coverage exit should remain distinguishable. For ratio measures such as carbon intensity, the report should also separate portfolio weights, emissions, nominal sales, and their interaction.

The report must state its attribution convention. Under a lagged-map order, the portfolio-by-map interaction is assigned to map movement; under a current-map order, it is assigned to the portfolio; under the symmetric midpoint, it is split equally. We use the symmetric convention as the primary allocation and the two directional orders as bounds. The empirical conclusion is severe under all three, although its magnitude is convention dependent.

A dated vintage is equally important. Our two 2026 archives change 25 illustrative target assessments of unchanged completed portfolios. Making coverage common across vintages eliminates all 25 switches in this sample. Preserving the original assessment and reporting a restated comparison with its coverage share makes the source of the change visible. Common coverage improves comparability but leaves uncovered emissions unknown, so an evaluator must also disclose how much of the portfolio is measured.

For physical intensity, the report should distinguish issuer market-cap movements from changes in emissions. Chaining attributed contributions can produce a nonpositive intermediate factor even when the final index is positive. Such a chain cannot describe a positive emissions trajectory. Actual endpoint portfolios, a named denominator and comparable coverage provide the direct assessment; source components explain its movement.

The principle extends beyond carbon. Any delegated portfolio evaluated with a date-specific characteristic map, including ESG scores, governance ratings, credit classifications, or risk labels,

can inherit the same attribution problem. The magnitudes need not match those for emissions. The general lesson is narrower: when both a portfolio and the ruler used to evaluate it change, the total measures exposure but is insufficient for evaluating manager performance.

9 Conclusion

A declining portfolio carbon score does not identify the contribution of the manager. Morningstar and a Refinitiv-based construction agree that U.S. mutual fund carbon intensity fell by more than a third from 2019 to 2024, but portfolio reweighting accounts for only 1.2 percent of the constructed decline. Issuer emissions and nominal sales denominators explain most of the movement. In institutional portfolios, coverage changes reverse the sector's measured trend, while attribution changes the trend sign for 59.7 percent of continuously observed managers under the symmetric convention.

Separating the components improves a hypothetical selection rule for implementation. Attribution replaces 171 of 431 managers in the conventional top quintile, and the entrants subsequently outperform the displaced managers by 0.552 standard deviation of attributed reweighting. The selected top quintile gains 0.219 standard deviation, with positive gains at capacities from 5 to 50 percent. The objective is reweighting toward firms with lower emissions conditional on industry and firm characteristics. Covered-portfolio carbon intensity is higher for the attributed group, and the physical-action benefit is unresolved. Selection for this implementation objective therefore requires a different statistic from a screen for low portfolio exposure.

A separate dated-archive comparison shows a practical consequence for target assessment. Updating emissions data changes 25 illustrative assessments of the same completed portfolios. Coverage changes reproduce all the switches; common-coverage value revisions produce none. Evaluating both archives on comparable coverage eliminates the switches in this sample. The remedy improves comparability of the observed record while leaving the uncovered portfolio unidentified. These decision exercises establish useful information for specified evaluation rules; realized allocation effects remain unmeasured.

A responsibility report should preserve the current exposure estimate and separately report portfolio implementation, issuer inputs, valuation denominators and coverage. It should name the data vintage and the convention used to allocate interactions, retain the original assessment, and show a restated comparison when historical coverage changes. That distinction allows better information about today's portfolio to coexist with a comparable record of the manager's implementation.

References

A. R. Admati and P. Pfleiderer. Does it all add up? benchmarks and the compensation of active portfolio managers. *Journal of Business*, 70(3):323–350, 1997. doi: 10.1086/209721.

J. Aswani, A. Raghunandan, and S. Rajgopal. Are carbon emissions associated with stock returns? *Review of Finance*, 28(1):75–106, 2024. doi: 10.1093/rof/rfad013.

V. Atta-Darkua, S. Glossner, P. Krueger, and P. Matos. Decarbonizing institutional investor portfolios: Helping to green the planet or just greening your portfolio? Swiss Finance Institute Research Paper 25-42, 2025.

S. Basak and A. Pavlova. Asset prices and institutional investors. *American Economic Review*, 103(5):1728–1758, 2013. doi: 10.1257/aer.103.5.1728.

F. Berg, K. Fabisik, and Z. Sautner. Is history repeating itself? the (un)predictable past of ESG ratings. European Corporate Governance Institute Finance Working Paper 708/2020, 2021.

F. Berg, J. F. Kölbel, and R. Rigobon. Aggregate confusion: The divergence of ESG ratings. *Review of Finance*, 26(6):1315–1344, 2022. doi: 10.1093/rof/rfac033.

J. B. Berk and R. C. Green. Mutual fund flows and performance in rational markets. *Journal of Political Economy*, 112(6):1269–1295, 2004. doi: 10.1086/424739.

J. B. Berk and J. H. van Binsbergen. Measuring skill in the mutual fund industry. *Journal of Financial Economics*, 118(1):1–20, 2015. doi: 10.1016/j.jfineco.2015.05.002.

M. Bertrand and S. Mullainathan. Are ceos rewarded for luck? the ones without principals are. *Quarterly Journal of Economics*, 116(3):901–932, 2001. doi: 10.1162/00335530152466269.

P. Bolton and M. Kacperczyk. Do investors care about carbon risk? *Journal of Financial Economics*, 142(2):517–549, 2021. doi: 10.1016/j.jfineco.2021.05.008.

G. P. Brinson, L. R. Hood, and G. L. Beebower. Determinants of portfolio performance. *Financial Analysts Journal*, 42(4):39–44, 1986. doi: 10.2469/faj.v42.n4.39.

G. P. Brinson, B. D. Singer, and G. L. Beebower. Determinants of portfolio performance II: An update. *Financial Analysts Journal*, 47(3):40–48, 1991. doi: 10.2469/faj.v47.n3.40.

M. Ceccarelli, S. Ramelli, and A. F. Wagner. Low carbon mutual funds. *Review of Finance*, 28 (1):45–74, 2024. doi: 10.1093/rof/rfad015.

S. Choi, F. Ferri, and D. Macciocchi. Do mutual funds respond to mechanical changes in ESG ratings? *Management Science*, 2026. doi: 10.1287/mnsc.2024.06184. Articles in Advance, published January 13, 2026.

L. Cohen, E. Rouen, and A. Sachdeva. Widespread revisions of self-reported emissions by major US corporations. *Nature Climate Change*, 16:33–36, 2026. doi: 10.1038/s41558-025-02494-9.

D. Cuoco and R. Kaniel. Equilibrium prices in the presence of delegated portfolio management. *Journal of Financial Economics*, 101(2):264–296, 2011. doi: 10.1016/j.jfineco.2011.02.012.

D. Del Guercio and P. A. Tkac. Star power: The effect of Morningstar ratings on mutual fund flow. *Journal of Financial and Quantitative Analysis*, 43(4):907–936, 2008. doi: 10.1017/S0022109000014393.

W. E. Diewert. Exact and superlative index numbers. *Journal of Econometrics*, 4(2):115–145, 1976. doi: 10.1016/0304-4076(76)90009-9.

N. Gantchev, M. Giannetti, and R. Li. Sustainability or performance? ratings and fund managers' incentives. *Journal of Financial Economics*, 155:103831, 2024. doi: 10.1016/j.jfineco.2024.103831.

S. M. Hartzmark and A. B. Sussman. Do investors value sustainability? a natural experiment examining ranking and fund flows. *Journal of Finance*, 74(6):2789–2837, 2019. doi: 10.1111/jofi.12841.

B. Holmström. Moral hazard and observability. *Bell Journal of Economics*, 10(1):74–91, 1979. doi: 10.2307/3003320.

H. Hong and M. Kacperczyk. The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, 93(1):15–36, 2009. doi: 10.1016/j.jfineco.2008.09.001.

E. Ilhan, Z. Sautner, and G. Vilkov. Carbon tail risk. *Review of Financial Studies*, 34(3): 1540–1571, 2021. doi: 10.1093/rfs/hhaa071.

A. Ljungqvist, C. Malloy, and F. Marston. Rewriting history. *Journal of Finance*, 64(4):1935–1960, 2009. doi: 10.1111/j.1540-6261.2009.01484.x.

M. Lowry, P. Wang, and K. D. Wei. Are all ESG funds created equal? only some funds are committed. *Review of Financial Studies*, 39(1):79–113, 2026. doi: 10.1093/rfs/hhaf084.

Morningstar Research. Morningstar sustainability rating methodology, October 2019. Version 1.2, October 31, 2019.

L. Pástor, R. F. Stambaugh, and L. A. Taylor. Sustainable investing in equilibrium. *Journal of Financial Economics*, 142(2):550–571, 2021. doi: 10.1016/j.jfineco.2020.12.011.

L. Pástor, R. F. Stambaugh, and L. A. Taylor. Dissecting green returns. *Journal of Financial Economics*, 146(2):403–424, 2022. doi: 10.1016/j.jfineco.2022.07.007.

L. Pástor, R. F. Stambaugh, and L. A. Taylor. Green tilts. *Journal of Financial Economics*, 174: 104173, 2025. doi: 10.1016/j.jfineco.2025.104173.

A. Rzeźnik, K. W. Hanley, and L. Pelizzon. Investor reliance on ESG ratings and stock price performance. *Management Science*, 2026. doi: 10.1287/mnsc.2024.08288. Articles in Advance, published May 29, 2026.

X. Wang and Z. Nagy. Carbon-footprint attribution for total portfolio. Technical report, MSCI Research, February 2025. URL `https://www.msci.com/research-and-insights/paper/carbon-footprint-attribution-for-total-portfolio`.

A Construction and Supporting Evidence

A.1 Reporting-entity and trade construction

The source holdings panel identifies Form 13F reporting entities using FactSet ownership identifiers. Constructed residual portfolios and source placeholders are excluded before estimation. Within an entity-stock-quarter, duplicate position rows are summed before any portfolio weight is formed. The corresponding entity-quarter AUM is required to be unique after this collapse. Security joins use CRSP `permno` and calendar quarter. Row-order and nearest-observation matches are excluded.

For the trade identity, the union of lagged and current security positions preserves purchases into and sales out of securities. Lagged positions are carried forward with CRSP ex-distribution returns. Positions with no current holding enter as sales, and new positions enter as purchases. Return-adjusted passive holdings, current holdings, and scores use the same dollar scale before division by manager AUM. The broad object matches 99.73 percent of lagged position value to a return. Closure is evaluated at the entity-quarter level before weighted means, cumulative paths, or covariance slopes are calculated.

Appendix Table 12 reports the detailed full-sample and subperiod accounting. Panel A separates within-manager change from shifts in continuing-manager AUM shares and entity entry and exit. Panel B reports the valuation, trading, and reclassification identity. Panels C and D provide the raw-industry companion and descriptive timing line. The main text uses only the rows needed for the attribution argument.

Table 12. The anatomy of institutional emissions-green drift

	Full sample 2007Q1–2023Q4	Rise window 2016Q4–2019Q3	Late window 2019Q3–2023Q4
Panel A. Aggregate tilt decomposition			
Aggregate tilt change	-0.1275	-0.1242	-0.0468
Within-manager contribution	-0.1316	-0.1264	-0.0504
Between-manager contribution	0.0035	0.0020	0.0049
Entry-exit contribution	0.0007	0.0001	-0.0013
Panel B. Anatomy of continuing-manager drift, lagged AUM weights			
Trading	-0.0458	-0.0080	-0.0242
Valuation	0.0283	0.0061	0.0098
Score reclassification	-0.1137	-0.1250	-0.0364
Continuing-manager change	-0.1311	-0.1268	-0.0508
Trading share of absolute anatomy	24.4%	5.8%	34.4%
Valuation share of absolute anatomy	15.1%	4.4%	13.9%
Reclassification share of absolute anatomy	60.6%	89.8%	51.7%
Panel C. Raw-emissions and industry companion			
Raw emissions-green tilt change	-0.5421	-0.1695	-0.1304
Within-manager contribution	-0.5554	-0.1670	-0.1469
Industry-allocation component	0.0582	0.0452	-0.0326
Within-industry component	-0.6003	-0.2147	-0.0979
Panel D. Descriptive timing of trading			
Lagged four-quarter relative green return	-0.0083		
Newey–West standard error	(0.0065)		
Observations	64		

The sample contains eligible retained Form 13F reporting managers from 2007Q1 through 2023Q4. Panel A decomposes aggregate emissions-green tilt. Panel B reconstructs continuing-manager changes using lagged assets-under-management (AUM) weights. Valuation carries lagged holdings forward with Center for Research in Security Prices (CRSP) ex-distribution returns at fixed lagged scores. Trading moves that passive portfolio to current reported holdings at lagged scores. Reclassification updates scores on current holdings. Missing scores enter the numerator as zero because total manager AUM is the denominator. Panel C uses negative standardized log carbon dioxide without regression residualization and a Fama–French 12-industry midpoint decomposition. Panel D is descriptive, includes only an intercept, and uses four Newey–West lags. The components are accounting contributions and carry no causal interpretation.

A.2 Sample representativeness

The continuously observed samples differ from their eligible populations in predictable ways. Appendix Table 13 shows that the 1,065 continuously observed funds are larger than the 2019 eligible sample, with median TNA of $900 million versus $623 million. Their equity shares, initial carbon intensities, and endpoint declines are similar. Appendix Table 14 shows stronger selection at the reporting-entity layer. The 861 long-lived actual reporting entities are larger and hold more positions than retained reporting entities overall. Manager rank results therefore describe stable, larger institutions; the sector accounting continues to use the retained evolving population.

Table 13. Representativeness of the continuously observed fund sample

Sample	Funds	TNA, $m	Equity share, %	Morningstar WACI	2019–2024 change, %	Survive to 2024, %
Eligible 2019 annual sample	1,596	623.1 [127.7, 2497.4]	97.9 [95.1, 99.2]	140.6 [65.9, 247.0]	-37.7	91.9
Eligible 2019–2024 endpoints	1,467	696.6 [135.1, 2747.4]	97.9 [95.2, 99.2]	140.9 [65.9, 245.9]	-37.7	100.0
Continuously observed matched-date sample	1,065	899.6 [184.4, 3360.8]	97.3 [94.8, 98.9]	133.1 [61.6, 230.4]	-36.9	100.0

Notes: TNA is aggregate CRSP portfolio TNA across positive share classes on December 31, 2019. Bracketed values are the 25th and 75th percentiles around the median. Equity share is linked common-stock holdings as a percentage of TNA. The carbon change is the percent equivalent of the median within-fund log change among funds observed at both endpoints. The 1,065 continuously observed funds pass all matched-date screens in every year from 2019 through 2024. The table is descriptive and applies no reweighting or inference.

Table 14. Representativeness of long-lived reporting entities

Sample	Entities	Entity-quarters	AUM, $bn	Positions	Observed years
Retained reporting entities	10,228	252,686	0.26 [0.14, 0.75]	56 [24, 128]	4.2 [1.8, 9.5]
861 long-lived actual reporting entities	861	58,548	1.28 [0.45, 6.75]	123 [64, 342]	17.0 [17.0, 17.0]

Notes: The retained reporting entities are the evolving eligible population used in the sector accounting. Entity statistics are medians across each entity's eligible quarters; table entries give the cross-entity median with the 25th and 75th percentiles in brackets. Position count is the number of positive deduplicated PERMNO positions in an eligible entity-quarter. The long-lived sample contains the 861 actual reporting entities observed in all 68 quarters from 2007Q1 through 2023Q4. The table is descriptive; it applies no reweighting or inference.

A.3 Map construction and current-vintage coverage

The raw source anatomy uses negative $\log(1 + CO_2)$ so that higher values are greener. For firms with finite emissions at both adjacent endpoints, the reported component requires reported values at both dates. If either endpoint uses a vendor estimate, the change is classified as estimated. Coverage entry sets the lagged value to zero and the current value to its finite source-specific observation; coverage exit reverses that construction. The raw parent equals reported change, estimated change, coverage entry, and coverage exit at every transition.

The residualized anatomy applies the same source classification to quarterly residualized and standardized scores. A standardization term captures movement caused by the changing cross-sectional transformation and included controls. This term is zero by construction on the raw scale. The coverage-versus-intensity split uses a midpoint product decomposition, so its two components sum to the raw parent. The subsequent reported-versus-estimated coverage split sums to the coverage component. All displayed source and coverage results are therefore nested decompositions of one object.

Appendix Table 15 places the raw source anatomy, the residualized mean anatomy, and the cross-manager covariance-slope anatomy on one surface. The columns use different units and answer different economic questions. The table reconciles the scales; cross-column component rankings would be invalid. The raw column identifies what moves the physical emissions input; the residualized columns show how those inputs enter the manager-evaluation score.

Table 15. The anatomy of emissions-score reclassification

	Raw log mean	Residualized mean	Covariance slope
Parent reclassification	-2.949	-0.114	-0.081
Firm reported	-0.125	-0.039	-0.197
Firm estimated	-0.135	-0.037	-0.116
Standardization	0.000	-0.026	0.059
Coverage entry	-3.076	-0.008	0.176
Coverage exit	0.388	-0.003	-0.003
Firm-reported absolute share	3.4% [90% CI: 3.0%, 3.7%]		
2023 eligible-manager AUM	\$43.9 trillion		
Cumulative score-weighted trading	-\$1.73 trillion score-dollars		

Notes: The sample contains continuing eligible Form 13F reporting managers from 2007Q1 through 2023Q4. The first column uses negative unstandardized log(1+carbon dioxide); the second and third use the residualized emissions-green score. Reclassification holds current positions fixed and moves the lagged score to the current score. Firm reported requires reported Refinitiv totals at both endpoints; firm estimated includes an estimate at either endpoint. Coverage entry and exit use the paper's missing-to-zero convention. Standardization is zero on the raw-log scale. The 90% interval resamples managers. Reported emissions are corporate disclosures, not verified physical abatement. AUM denotes assets under management. The dollar trading line sums lagged standardized emissions scores times net dollar trades; it is an exposure-equivalent accounting quantity, not sale proceeds or price impact.

Appendix Table 16 then separates entry from exit at the measurement boundary and reports the physical-intensity comparison used to interpret those cells. The contribution remains on the raw negative-log emissions scale; the intensity ratios use tonnes of CO_2e per \$1 million of issuer equity value.

Table 16. Coverage entry, exit, and the measured sector path

Object	Estimate	90% interval
Entry contribution	-3.707	–
Exit contribution	0.422	–
Coverage contribution	-3.285	[-3.430, -3.124]
Vendor-estimated share of entering value	75.5%	–
Entrant physical intensity relative to incumbents	1.26	–
Exiter physical intensity relative to incumbents	1.08	–
Actual raw sector path	-0.993	[-1.255, -0.733]
Fixed-coverage raw sector path	2.292	[2.067, 2.532]
Coverage share of absolute map movement	90.7%	[90.1%, 91.3%]

Notes: The sample contains actual Form 13F reporting entities. Entry and exit hold current positions fixed. Physical intensity is tonnes CO_2e per $1 million of issuer equity value. The raw sector path uses negative-log emissions. The estimates are accounting contributions and carry no behavioral interpretation.

A.4 Manager-classification transition

Appendix Table 17 reports the complete mapping from moving-map to fixed-map manager deciles. Cells are conditional on the moving-map row, so the diagonal gives assignment persistence and off-diagonal cells show where the same managers move after cumulative map reclassification is removed. The construction uses the same balanced 861-entity sample and ranks as the main manager exhibit.

Table 17. Manager decile assignments under the symmetric convention

	Symmetric portfolio-originated decile										
Measured decile	1	2	3	4	5	6	7	8	9	10	*N*
1	**28.7**	24.1	14.9	9.2	11.5	4.6	2.3	2.3	2.3	0.0	87
2	17.4	**19.8**	10.5	11.6	11.6	12.8	3.5	2.3	5.8	4.7	86
3	12.8	11.6	**9.3**	10.5	11.6	7.0	22.1	7.0	3.5	4.7	86
4	9.3	2.3	14.0	**12.8**	7.0	19.8	17.4	8.1	8.1	1.2	86
5	4.7	4.7	10.5	15.1	**8.1**	12.8	15.1	15.1	7.0	7.0	86
6	7.0	3.5	11.6	12.8	10.5	**12.8**	10.5	16.3	10.5	4.7	86
7	4.7	9.3	12.8	10.5	11.6	9.3	**7.0**	8.1	18.6	8.1	86
8	10.5	10.5	8.1	2.3	5.8	9.3	9.3	**19.8**	11.6	12.8	86
9	1.2	5.8	3.5	9.3	17.4	4.7	5.8	14.0	**20.9**	17.4	86
10	4.7	8.1	4.7	5.8	4.7	7.0	7.0	7.0	11.6	**39.5**	86

Notes: Cells report the percentage of managers in each measured-record row assigned to each portfolio-originated column after the portfolio-by-map interaction is split equally. Decile 1 is the most-browning trend and decile 10 the most-greening trend. Bold cells preserve the original decile. The sample contains 861 actual reporting entities present in every quarter from 2007Q1 through 2023Q4. Only 17.9 percent remain on the diagonal. The table compares monitoring records; manager action is the narrower trading component.

A.5 Attribution-order sensitivity

The portfolio-by-map cross term has no convention-free owner. Appendix Tables 18 and 19 report the three standard orders. Fund ranks and sign assignments change little. The institutional

conclusion remains severe under every order, but its magnitude is convention dependent.

Table 18. Fund responsibility under three attribution orders

Convention	Comparison	Spearman	Opposite signs	Same decile
Lagged map (Laspeyres)	Constructed total vs portfolio-originated	0.804	28.6%	40.1%
Lagged map (Laspeyres)	Published total vs portfolio-originated	0.598	33.5%	24.8%
Symmetric midpoint	Constructed total vs portfolio-originated	0.827	29.8%	41.0%
Symmetric midpoint	Published total vs portfolio-originated	0.625	33.5%	25.9%
Current map (Paasche)	Constructed total vs portfolio-originated	0.810	28.8%	41.2%
Current map (Paasche)	Published total vs portfolio-originated	0.617	33.1%	24.4%

Notes: The sample contains the same 1,065 continuously observed funds. The three rows allocate none, one half, or all of the portfolio-by-map interaction to the portfolio-originated component. Positive values mean greening. No sampling inference is attached.

Table 19. Institutional responsibility under three attribution orders

Convention	Map share	Spearman	Opposite signs	Same decile	Top exit	Bottom exit
Current map (Paasche)	62.7%	0.269	69.8%	14.8%	70.9%	73.6%
Lagged map (Laspeyres)	60.6%	0.494	36.9%	21.3%	53.5%	66.7%
Symmetric midpoint	71.5%	0.406	59.7%	17.9%	60.5%	71.3%

Notes: Manager classifications use the same 861 long-lived actual reporting entities. The map share uses the retained evolving universe. Both valuation and attributed reweighting use lagged, midpoint, or current scores according to the row's convention. The three orders allocate the portfolio-by-map interaction differently; the portfolio-originated record is their combined valuation and reweighting contribution.

A.6 Persistence of manager climate styles

Manager style is the value-weighted residualized emissions score among covered positive equity positions, divided by covered position value. The interlaced exercise estimates empirical-Bayes manager types on odd quarters and evaluates them on even quarters, then reverses the assignment. The forward exercise uses 2007Q1–2013Q4 for formation, omits 2014, and evaluates holdings in 2015Q1–2023Q4. Both apply the coverage, stock-count, and usable-quarter screens described in Section 2.

Appendix Table 20 reports strong local repeatability and economically meaningful, but lower, portability across nonoverlapping eras. These results establish persistent entity differences without assigning them to permanent preferences or a particular mandate.

Table 20. Persistence of manager climate styles across horizons

Exercise	Managers	Rank persistence	Held-out spread
Interlaced odd/even quarters	5,279	0.982	2.586 SD
Pre-2014 formation to 2015–2023	2,231	0.436	1.266 SD
90% manager-bootstrap interval		[0.403, 0.468]	[1.176, 1.372]

Notes: Manager style is the value-weighted mean residualized emissions score among covered positive equity positions. A higher value is browner. The first row estimates empirical-Bayes types on odd quarters and evaluates them on even quarters, then reverses the exercise. The second row forms the emissions type on 2007Q1–2013Q4, omits 2014, and evaluates future style on 2015Q1–2023Q4. The exercises require at least 25 percent scored-position coverage, ten scored stocks per usable quarter, and eight usable quarters in each estimation period. The spread is the formation-type top-minus-bottom quintile difference scaled by the cross-sectional standard deviation (SD) of held-out style.

A.7 Held-out manager selection

The selection population contains reporting entities with at least eight transition observations in both eras. Formation uses 2007Q2–2013Q4, evaluation uses 2015Q2–2023Q4, and 2014 and both bridge transitions are omitted. The nested formation records are manager means of total measured change, portfolio-originated change after removing the symmetric map component, and attributed reweighting under Equation (6). Future action uses the same actual-minus-passive weights and midpoint scores. Managers are sorted into quintiles separately on each formation record; rankings and outcome standard deviations are reestimated in each of 2,000 manager-bootstrap draws. The resulting 2,154-manager sample and the observed calendar history are common across records.

The four formation groups identify the managers affected by the selection rule. Both records select 260 managers. The conventional record alone selects 171, with $0.39 trillion of representative formation AUM; attribution alone selects another 171, with $0.45 trillion. Attribution-only managers subsequently exceed conventional-only managers by 0.552 standard deviation of attributed reweighting. Formation AUM sums managers' mean beginning AUM across formation transitions. It measures the reporting entities' asset scale, not dollars assigned under an observed mandate.

Removing map movement increases the future top-minus-bottom spread by 0.299 standard deviation and replaces 139 conventionally selected managers. Removing valuation adds 0.088 standard deviation. Full attribution raises the spread by 0.387 standard deviation and replaces 171 managers from the conventional top group. Map movement accounts for about 77 percent of this increase.

The fixed-score endpoint comparison retains the candidate population and formation rankings but requires observed holdings at 2015Q1 and 2023Q4. Its 1,567 observations include 121 of 171 attribution-only managers and 129 of 171 conventional-only managers. The fixed-score means for these groups are -0.0145 and -0.0604 score units, respectively. The issuer score vector is dated 2015Q1 within the current data vintage; it is not an archived 2015 information set. Endpoint attrition and valuation-driven weight changes remain part of this comparison.

The asset-growth outcome compares mean log reported 13F equity AUM in 2015Q2–2016Q1 and 2023Q1–2023Q4, requiring two observations in each block, and divides the difference by 7.75 years. It retains 121 attribution-only managers and 131 conventional-only managers. A second outcome subtracts the holdings-implied price component from each quarterly AUM

change, sums net asset dollars across evaluation transitions, divides by the sum of beginning AUM, and multiplies by four. This net-growth measure is available for all 2,154 managers. The conventional-only group exceeds the attribution-only group by 4.02 percentage points in endpoint growth and 3.58 percentage points in net growth. The latter contrast has a 90 percent interval of $[0.10, 6.86]$ and a 95 percent interval of $[-0.44, 7.45]$. Verified client flows are unavailable, and these comparisons leave a response to sustainability measurement unidentified.

Table 21 translates the same selection groups into physical units. Physical action is the negative lagged-score trading contribution in the physical-emissions decomposition, in tonnes of CO_2e per $1 million of beginning AUM per quarter. The attributed top minus the conventional top is -0.18, with a 90 percent interval of $[-0.76, 0.32]$. The interval spans zero and the point estimate has the opposite sign from the standardized implementation gain. Thus the physical outcome provides no evidence of a portfolio-exposure benefit.

Table 21. Physical exposure under the two manager-selection rules

Selection group	Managers	AUM ($tn)	Attributed action	Physical action
Selected by both	260	0.38	0.00210	0.78
Total only	171	0.39	-0.00063	0.38
Attribution only	171	0.45	0.00435	-0.08
Selected by neither	1552	11.01	0.00100	1.07
Top-set difference	–	–	0.00198 [0.00108, 0.00279]	-0.18 [-0.76, 0.32]
Replacement difference	–	–	0.552 [0.311, 0.745] SD	-0.46 [-1.92, 0.82]

Notes: The 2,154-manager sample, formation and evaluation dates, and hypothetical equal-dollar selection rules match the main selection table. Attributed action uses midpoint issuer scores for both passive and active portfolio changes. Differences subtract total selection from attributed selection. Group means and the top-set action contrast use score units; the replacement action contrast uses SD. Physical action retains the lagged-score physical trading measure, in tonnes CO_2e per $1 million AUM per quarter; higher values indicate lower exposure. Formation AUM sums each manager's mean beginning AUM over formation transitions. Brackets are 90% manager-bootstrap intervals. The physical contrast is statistically unresolved and does not establish issuer abatement.

A.8 Fund responsibility transitions

Appendix Table 22 first separates the three performance records: Morningstar's published total, the matched Refinitiv-based total, and the Refinitiv portfolio-originated component. It then reports the complete mapping from published fund greening deciles to portfolio-originated deciles. The table uses the balanced 2019–2024 population and the main horizon and normalization.

Table 22. Fund rankings across published, constructed, and portfolio-originated records

Panel A: Three-record comparison

First record	Second record	Rank correlation	Opposite signs	Same decile
Morningstar published total	Refinitiv constructed total	0.711	14.3%	28.6%
Refinitiv constructed total	Refinitiv portfolio component	0.804	28.6%	40.1%
Morningstar published total	Refinitiv portfolio component	0.598	33.5%	24.8%

Panel B: Published-to-portfolio decile transition

	Portfolio-originated greening decile										
Published greening decile	1	2	3	4	5	6	7	8	9	10	*N*
1	**56.1**	15.9	12.1	2.8	0.9	0.9	1.9	2.8	2.8	3.7	107
2	21.7	**33.0**	10.4	14.2	5.7	3.8	1.9	5.7	2.8	0.9	106
3	8.4	15.9	**27.1**	23.4	9.3	7.5	4.7	2.8	0.9	0.0	107
4	4.7	10.4	17.9	**9.4**	25.5	11.3	6.6	5.7	3.8	4.7	106
5	4.7	4.7	5.6	16.8	**13.1**	16.8	14.0	13.1	5.6	5.6	107
6	0.9	2.8	3.8	7.5	5.7	**12.3**	40.6	16.0	4.7	5.7	106
7	0.9	3.7	5.6	6.5	17.8	7.5	**15.9**	20.6	12.1	9.3	107
8	0.0	6.6	7.5	9.4	9.4	18.9	5.7	**14.2**	16.0	12.3	106
9	0.9	1.9	7.5	6.5	9.3	9.3	4.7	11.2	**29.0**	19.6	107
10	1.9	4.7	2.8	2.8	3.8	11.3	4.7	7.5	22.6	**37.7**	106

Notes: Cells report row percentages for the same 1,065 funds observed in every annual transition from 2019 through 2024. Decile 10 is the greatest greening and decile 1 the greatest browning; bold cells preserve the published decile. Published greening is the proportional decline in Morningstar carbon intensity. Portfolio-originated greening is the negative cumulative portfolio-reweighting contribution divided by initial constructed intensity. This component combines trading, valuation, flows, and changes in the non-equity TNA denominator; manager trading is a narrower object. Panel A separates evaluator disagreement before attribution (row 1), attribution holding the evaluator fixed (row 2), and their combined responsibility gap (row 3). Panel B expands the third comparison. Of funds in the published top and bottom deciles, 62.3% and 43.9% leave those tails, respectively. These classifications describe the balanced observed population and carry no sampling inference.

A.9 A second delegated evaluation system

Morningstar's fund sustainability ratings provide a separate setting in which delegated portfolios are evaluated through a changing map. The late-2019 methodology replaced the underlying company ESG rating with an ESG Risk Rating and introduced transition buffers and absolute-risk rules (Morningstar Research, 2019). Appendix Table 23 compares the observed reset with adjacent months, reports the complete old-to-new rating transition, and follows the event movers for one year. It also measures the pre-event assets exposed to reclassification and summarizes subsequent rating recurrence and peer breakpoint drift.

The exhibit uses published labels and breakpoints directly, leaving the causes of later rating changes unclassified. The table contains no fund-flow or return outcome. Its role is to provide a second setting in which rebuilding an evaluation system produces large, persistent reclassification.

Table 23. A methodology reset in a second delegated evaluation system

Panel A. Published rating changes around the September 2019 reset			
Transition ending	Fund transitions	Changed globe	Median risk score
June 2019	5,769	5.5%	45.9
July 2019	5,804	5.0%	46.0
August 2019	5,837	4.1%	46.1
September 2019	**5,739**	**58.1%**	**24.9**
October 2019	5,812	2.6%	24.8
November 2019	5,837	3.9%	24.9

Panel B. August-to-September rating transition					
August globe \ September globe	1	2	3	4	5
1	222	240	100	51	13
2	181	576	561	136	33
3	95	428	981	383	79
4	26	142	427	429	150
5	2	14	96	179	195

Panel C. Persistence among the 3,336 event movers			
Horizon after reset	Observed movers	Same new label	Post-reset side
1 month	3,307	96.8%	98.5%
3 months	3,275	88.9%	94.8%
6 months	3,163	79.3%	90.0%
12 months	3,057	63.7%	82.2%

CRSP-linked pre-event assets	$14.00 trillion
Assets assigned a different globe	$6.58 trillion (47.0%)
Old/new globe-rank correlation	0.513
Old globe 1 / globe 5 exit	64.5% / 59.9%
Published globe changes, full history	29,339 / 553,026 (5.3%)
Median middle category breakpoint, 2020 to 2025	25.0 to 21.5

Notes: The exhibit uses directly observed Morningstar Sustainability Rating globes, with 1 denoting Low and 5 High. Morningstar replaced the underlying company ESG rating with the ESG Risk Rating and added rating buffers and absolute-risk rules in late 2019 (Morningstar Research, 2019). Panel A counts consecutive-month transitions. Panel B reports the 5,739 funds rated in both August and September. In Panel C, a mover remains on the post-reset side when its future globe remains strictly above its August globe after an upgrade or strictly below it after a downgrade. Pre-event assets sum August 30, 2019 CRSP TNA across share classes within each uniquely linked portfolio; 4,651 linked portfolios have finite TNA. The observed rating history runs from August 2018 through December 2025. The breakpoint is the fund-weighted median third corporate-risk breakpoint. All panels use published labels and contain no fund-flow or return outcomes.

A.10 Flow response to a moving rating map

We first assign treatment at a mechanical peer-map rebanding episode. Table 24 shows a 5.2-basis-point estimate in the first fully public month and a one-sided lower bound of -7.3 basis points per episode. Its first stage into an actual label change is 0.172, so neither number is a per-label effect.

Table 24. Episode-level flow bound in the exposure design

Design	Estimate (bp)	SE (bp)	p	Lower 95%	Episodes
Rebanding episode	5.2	7.6	0.491	-7.3	651

Notes: The exposure design is an intention-to-treat comparison around mechanical peer-map rebanding episodes. Its first stage into an actual published label change is 0.172, so the bound applies at the episode level. Table 11 conditions on treated observations with an actual published change and reports the per-label design.

The main design conditions on an actual map-driven published change and matches non-changers within the same peer-map episode. Of 29,339 observed rating changes, we classify 13,649 by cause. Classification covers 24.5 percent of changes in the corporate-only regime and 70.2 percent after the sovereign component enters; 56.8 percent of map-driven downgrades link to the CRSP data. These rates describe selection into the estimable population; treatment remains an actual map-driven label change. Table 25 reports the complete label-conditioned downgrade path used in the main text.

Table 25. Event-time flow path after evaluator-caused globe downgrades

Event month	Estimate (bp)	SE (bp)	p
-3	0.9	16.3	0.958
-2	-4.8	15.6	0.758
-1	3.4	15.4	0.823
$+0$	-8.6	15.4	0.578
$+2$	-8.2	14.6	0.575
$+3$	-1.9	15.2	0.900
$+4$	14.6	14.6	0.318
$+5$	9.8	16.0	0.541

Notes: This is the dynamic path for the pooled label-conditioned downgrade design in Table 11. Month $t+1$ is the omitted reference; $t+3$ is the first full month after the rating is public. The joint pretrend test over $t-3$ through t has $p = 0.950$.

A.11 Expense ratios after evaluator-driven downgrades

The expense-ratio test fixes the share classes and TNA weights in the last prepublication month. For each class it selects the latest fiscal record completed by that date and the first complete fiscal period beginning after the downgrade is fully public. At least half of baseline class TNA must pair; median coverage is 100 percent and the post fiscal period begins six months later at the median. The matched-set model contains 1,388 treated events and 3,718 control rows. This design measures the recorded total expense ratio, which combines management fees, distribution fees, and waivers. The data lack an announced effective date.

Table 26. Fee outcomes around evaluator-caused downgrades

Outcome	Horizon	Estimate	SE	95% interval
Expense ratio change	First post fiscal period	−0.218 bp	0.263	[−0.734, 0.298]
Probability of a fee cut	First post fiscal period	−0.355 pp	1.784	[−3.852, 3.142]
Prior expense ratio change	Prior fiscal record	−0.161 bp	0.275	[−0.701, 0.378]

Notes: The design fixes the pre-event share-class set and TNA weights. It compares the latest fiscal record completed before publication with the first full fiscal period beginning after the downgrade is public. The sample contains 1,388 treated events and 3,718 matched control rows. The outcome is the recorded total expense ratio. The data lack an announced fee change and effective date.

A.12 Dated emissions archives and completed-record assessment

The July 18 and September 22, 2026 archives contain the same six Refinitiv emissions fields under identical source-data queries. The later archive has 113,379 rows. Matching on organization, fiscal year, CUSIP and field yields 1,131 revised values and 646 deletions across the six fields. Some additions carry record dates at or before the first download, while revisions and deletions lack a complete version history. The record-date field therefore cannot reconstruct first visibility or the sequence of historical revisions. These observations identify changes between the two downloads; they do not recover the database investors saw during the historical investment sample.

The original fixed-2023Q4 comparison affects 1.04 percent of represented position dollars and gives a manager rank correlation of 0.99732 across archives. Exits from the original brownest and greenest deciles are 8.14 and 2.30 percent. These results use the historical-CUSIP aggregation preserved in the original archive. They are a limited fixed-date diagnostic; the completed-record assessment in Section 7.2 instead fixes an organization-to-security link and never averages distinct organizations. The two constructions should not be treated as interchangeable.

For the completed-record assessment, the fixed link contains 6,302 security-year pairs and no unresolved ties between top-priority organizations. Fiscal-year overlap and company-name similarity choose among date-valid candidates. There are 557 links with normalized name similarity below 0.5; the rule is a reproducible linkage procedure rather than external verification of every issuer identity. Common coverage is the intersection of finite, nonnegative emissions under the two archives at each endpoint. All 861 managers have positive baseline intensity and positive covered position value at both endpoints in both archives.

The coverage mechanism is checked by crossing the July and September masks with the two sets of values. The old-coverage, new-values cell retains July emissions for deleted records. Changing coverage alone gives the same 17 losses and eight gains in target passage as the full September update. Updating common-record values alone changes no classification. This diagnostic was motivated by the available-versus-common-coverage result and isolates its mechanism; it is not an additional independent test of the primary outcome.

Coverage disappears for six held security-years in FY2018, three previously reported and three estimated, and thirteen in FY2022, one reported and twelve estimated. There are no entries in the held endpoint positions under the fixed link and no changed organization/CUSIP sets. The median absolute change in the end/start intensity ratio is 0.00927, with a 95th percentile of 0.15652. On common coverage these statistics are zero and 0.00285. The largest full-record ratio change is 158.69, reflecting sensitivity to a small baseline; the mean therefore gives a misleading account of a typical manager's revision.

Coverage among represented security positions is 96.0 and 95.1 percent at the two endpoints under July data, but these positions do not exhaust reported AUM. The median minimum covered-AUM share across endpoints and archives is 59.9 percent among managers losing a pass and 51.7 percent among gainers. Their associated 2023Q4 reported equity assets sum to $172.1 billion, or 0.68 percent of the sample's summed reported equity-asset denominator. This quantity identifies the scale of the reporting entities attached to changed assessments, without measuring assets reallocated or actually certified.

A.13 Physical sources and target attribution conventions

For the 861 balanced managers, we reconstruct all 57,687 quarterly transitions of the equity-attributed intensity used in the physical analysis. Let E_n^j and K_n^j denote emissions and issuer equity market capitalization at endpoint j, and let $s_n^j = 10^6 E_n^j / K_n^j$ when both inputs are valid, with zero contribution otherwise. The lag-convention input change is $\sum_n w_n^1 (s_n^1 - s_n^0)$. On continuously covered positions, its emissions and market-cap components are

$$C_E = 10^6 \sum_n w_n^1 (E_n^1 - E_n^0) \frac{1/K_n^0 + 1/K_n^1}{2}, \tag{11}$$

$$C_K = 10^6 \sum_n w_n^1 \frac{E_n^0 + E_n^1}{2} (1/K_n^1 - 1/K_n^0). \tag{12}$$

Coverage entry contributes $w_n^1 s_n^1$ and exit contributes $-w_n^1 s_n^0$. These four sources sum to the original input change at every manager transition. The emissions component includes reported changes, estimates and source switches in the historical input vintage. It is not a separate estimate of firm abatement or dated vendor revisions.

Table 27. Physical sources and attribution conventions in target indices

A. Sources of the issuer-input contribution					
Window	Total	Emissions	Market cap	Entry	Exit
2007Q1–2023Q4	-0.570	-0.126	-1.279	0.880	-0.045
2019Q4–2023Q4	-0.888	-0.398	-0.942	0.464	-0.012
B. Target passage on the same eligible managers					
Window	Managers	Lag rule	Midpoint	Lag only	Midpoint only
2007Q1–2023Q4	790	203	300	18	115
2019Q4–2023Q4	825	271	442	17	188
C. Managers with a nonpositive or invalid chain factor					
Window	Midpoint portfolio		Action		
	Invalid	Positive end	Invalid	Positive end	
2007Q1–2023Q4	71	7	28	1	
2019Q4–2023Q4	36	1	12	1	

Notes: The sample contains 861 reporting entities. Panel A reports pooled lag-AUM-weighted mean quarterly contributions in tonnes of CO_2e per million dollars of AUM. At current holdings weights, emissions and reciprocal market-cap changes split their interaction equally; coverage entry and exit are separate. Negative values reduce measured intensity. Panel B applies a seven percent annual decline rule to chained portfolio attribution indices. Each factor is one plus the attributed quarterly change divided by the actual starting footprint. Comparisons retain the same managers with positive factors under both conventions. Panel C counts exclusions from the full sample; "Positive end" counts invalid paths whose final product is positive. These indices attribute the measured record and need not describe feasible alternative portfolios or actual mandate decisions.

Panel A of Table 27 shows that market-cap movements make a large negative contribution, while coverage entry makes a positive contribution. A rising equity market value lowers this intensity with constant emissions. The large input-change contribution in this metric therefore cannot be assigned to coverage expansion or vendor behavior alone.

To examine target sensitivity, define an attribution index whose quarterly factor is $1+c_{it}/F_{i,t-1}$, where c_{it} is a component of the measured change and $F_{i,t-1}$ is the actual positive starting footprint. The product of these factors is compared with 0.93^T, for $T = 16.75$ or four years. Using the full measured change reproduces the actual endpoint ratio. Using only a portfolio component yields an attribution index, which need not be the footprint of a feasible alternative portfolio.

The lag portfolio component is $\sum_n (w_n^1 - w_n^0)s_n^0$; its midpoint counterpart is $\sum_n (w_n^1 - w_n^0)(s_n^0 + s_n^1)/2$. Their difference is half the portfolio-by-score interaction. Panel B compares target passage on the same managers with positive chain factors under both conventions. Passes change from 203 to 300 in the long window and from 271 to 442 in the recent window. Moving valuation between portfolio subcomponents leaves their sum unchanged; changing the score convention changes the portfolio/input split. Thus target classifications based on these indices are convention dependent.

Every chain factor must be positive to describe a positive trajectory. Panel C shows why checking only the endpoint is insufficient. One action path in each window has a positive endpoint despite an invalid intermediate factor. The complete-path screen leaves 273 action passes among 833 eligible managers in the long window and 340 among 849 in the recent window. Midpoint portfolio indices exclude 71 and 36 managers, respectively. These exclusions do not arise in the direct actual-endpoint vintage assessment.

Source-removal indices provide a separate sensitivity within the lag convention. Replacing the measured quarterly change by that change minus C_K reduces passes from 358 to 76 in the long window and from 548 to 287 in the recent window. Removing C_E instead gives 338 and 463. These counts illustrate the contribution of denominator and emissions movements under a specified index; binary threshold effects overlap and cannot be added across sources. They do not count vendor-caused mandate failures.