Monograph

Investment Committees: Governance and Design Choices

Bernhard (Bernd) Scherer

CFA Institute Research Foundation

Investment Committees: Governance and Design Choices

Bernhard (Bernd) Scherer
Head of Portfolio Implementation, ADIA

CFA Institute
Research
Foundation

Statement of Purpose

The CFA Institute Research Foundation is a not-for-profit organization established to promote the development and dissemination of relevant research for investment practitioners worldwide.

Photo credit: Loomis Creative

ISBN: 978-1-952927-78-2

CONTENTS

CFA Institute

PROFESSIONAL LEARNING QUALIFIED ACTIVITY
This publication qualifies for 2 PL credits under the guidelines of the CFA Institute Professional Learning Program.

ABOUT THE AUTHOR

Dr. Bernhard Scherer combines three decades of senior leadership across asset classes and institutions with a sustained record of academic output.

Senior Management Experience

Since September 2023, Dr. Scherer has served as head of portfolio implementation at ADIA in Abu Dhabi, overseeing the core liquid markets portfolio. His responsibilities span internal equity and fixed-income beta teams, quantitative portfolio management, cross-asset implementation, and external manager selection. Previously, he served as Geschäftsführer and chief investment officer (CIO) of LBBW Asset Management in Stuttgart, overseeing EUR10 billion in assets under management and 300 employees across fixed income, equities, multiasset investments, and alternatives. Before that, he was CIO of Lampe Asset Management, head of research at Deutsche Asset Management in Frankfurt and New York City, CIO at FTC Capital in Vienna, and head of quantitative GTAA (Global Tactical Asset Allocation) at Morgan Stanley London. His career began at Morgan Stanley, Sal. Oppenheim, and Schroders, where he built quantitative strategies from the ground up.

Academic Credentials

Dr. Scherer holds a doctorate (summa cum laude) from the University of Giessen, supervised by Professor J. P. Krahnen; an MSc in economics from Queen Mary College London; and a Diplom-Ökonom from the University of Augsburg. He has held full professorships and visiting positions at EDHEC Business School, WU Vienna, and Birkbeck College.

Publication Record

Dr. Scherer has published six books on risk budgeting, portfolio optimisation, and quantitative asset management with Springer, Risk Books, and Oxford University Press. His journal output includes more than 80 papers, which have been published in the *Journal of Banking and Finance*, *Journal of Financial Markets*, *Journal of Empirical Finance*, *Quantitative Finance*, *Financial Analysts Journal*, *Journal of Derivatives*, *Journal of Portfolio Management*, and many others. He sits on the editorial boards of the *Financial Analysts Journal* and *Journal of Asset Management*, and has received several best paper awards.

Dr. Scherer works for Abu Dhabi Investment Authority in the United Arab Emirates. His private email is drberndscherer@gmx.net. All views expressed in this monograph are solely those of the author and do not necessarily reflect those of his employer.

1. INTRODUCTION

Large investors love committees. These committees signal seriousness, distribute blame, and produce copious minutes. Whether they actually improve portfolios is less obvious; however, investment committees (ICs) are ubiquitous in institutional asset management. Asset owners rely on committees to set strategic policy, approve tactical tilts, and demonstrate fiduciary process. Asset managers often maintain ICs to deliver an in-house view, coordinate risk budgets, and reduce key-person dependence. Committees, therefore, sit at the intersection of portfolio choice and organisational design. Despite their prevalence, the economic case for how committees should be structured (and whether they even should centralise investment decisions) remains surprisingly underdeveloped. In practice, ICs are frequently treated as a governance ritual: Meetings, minutes, and consensus language substitute for an explicit mapping from dispersed information to portfolio weights. In this monograph, I take a more austere view and treat an IC as a *technology used to aggregate information into a central portfolio*. I also ask how that technology performs under realistic frictions, including heterogeneous beliefs, career incentives, status hierarchies, and noisy deliberation.

Two ideas motivate my analysis. First, centralisation and decentralisation are not merely stylistic choices but, rather, are choices over *the distribution of outcomes*. A central IC, in principle, can improve the information ratio by pooling independent signals, just as diversifying uncorrelated alpha sources can improve a fund of funds. At the same time, centralisation collapses cross-sectional dispersion across mandates: All mandates win or lose together. This is economically relevant for asset managers who have many clients because profits, flows, and franchise survival depend nonlinearly on the fraction of mandates that outperform in the same period. With many clients, this stops being an abstract "hit-rate" discussion—if decisions are centralized. A single bad year doesn't show up in one corner of the book. It hits a large share of mandates at the same time, and the outflows arrive in a pack.

Second, committees do not aggregate information in a frictionless way. Committee discussions are path dependent and status weighted. A committee can drift toward a position that is more extreme than the arithmetic average of its members' starting position: This phenomenon is known in social psychology as *group shift* or *group polarisation*. If group shift is present, the theoretical benefits of pooling signals can be partially or fully undone by the aggregation mechanism. Three questions frame this analysis:

- **When is it economically justifiable to centralise investment decisions in an IC?** In particular, how does the gain from aggregating informative signals trade off against the higher variance and left-tail risk that come from synchronising mandate outcomes?
- **How should an IC aggregate views into a portfolio?** Can a simple mechanical rule—anonymous portfolio-vector averaging—dominate discussion-based consensus by reducing noise, chief investment officer (CIO) dominance, and group shift?
- **Does group-shift bias occur in discussion-based ICs, and when is it the largest?** What features of the decision environment (narrative salience, speaking order, reputational incentives) generate the strongest drift away from the arithmetic mean of member views?

To answer these questions, I combine a parsimonious governance model with a controlled simulation protocol. In a simple active management setting, I compare decentralised portfolio managers (multi-PM firms) to centralised ICs, highlighting the trade-off between diversification of *opinions* and dispersion of *outcomes*. The governance model compares two stylised organisational structures for an asset manager with N mandates: decentralised management, in which each mandate is run independently, and centralised management, in which an IC generates one model portfolio that is then implemented across mandates (subject to client constraints). Using a standard mapping from information ratios to hit rates, $p(\mathrm{IR}) = \Phi(\mathrm{IR})$, the model shows that pooling N uncorrelated informative signals can raise the hit probability; however, centralisation also increases the variance of the number of outperforming mandates from order N to order N^2. For a weak organisation with negligible skill ($\mathrm{IR} \approx 0$), centralisation does not improve its information ratio. Instead, it dramatically increases the probability of joint underperformance, making decentralisation strictly preferable for any risk-averse owner. For a sufficiently skilful organisation, centralisation can dominate but only above a skill threshold that rises when bankruptcy risk and fixed costs are introduced. The model therefore reframes IC design as an economic decision about risk-bearing capacity and franchise survival, not as an aspiration to managerial "alignment."

The second central contribution of this monograph addresses the aggregation mechanism. The standard industry IC—a monthly meeting with a macro overview, sequential commentary, and a CIO summary that becomes the implemented portfolio—relies on conversation as the aggregation technology. Conversation is nonlinear and opaque: Small differences in sequencing, tone, or status can produce large differences in outcomes, and after the fact, it is difficult to reconstruct how individual information translated into weights. Drawing on the literature on forecast aggregation, decision hygiene, and organisational noise, I propose a deliberately simple alternative: *anonymous portfolio-vector averaging*. Each IC member submits a portfolio weight vector. All vectors are scaled to a common *ex ante* risk level and are averaged; the resulting aggregate is then rescaled to a target tracking error and implemented. This design is narrower than generic "structured decision" prescriptions because portfolio weights are the mediating assessment. It is also more radical in one respect: It removes the final discretionary override that allows rhetoric and hierarchy to dominate outcomes. The goal is not to discover the best forecaster *ex post* and overweight them (a statistically fragile and politically biased procedure in small samples) but rather to impose a transparent and incentive-compatible mapping from dispersed views to a central portfolio.

The third central contribution investigates *group-shift bias* empirically within the constraints of this setting. Direct observational data are rarely available: One seldom observes both the *ex ante* portfolios that committee members would have chosen independently and the counterfactual committee portfolio that would have resulted under an alternative decision rule. Therefore, I use a large language model (LLM) as a controlled narrative simulator. The LLM is given a fixed cast of stylised committee personalities and asked to produce, for each macro scenario, (1) individual starting portfolios, (2) a traditional IC discussion under a specified speaking order and CIO role, and (3) a final traditional IC portfolio. The anonymous IC outcome is constructed mechanically as the equal-weight average of the same starting portfolios. This setup is not offered as a forecasting device. Its purpose is to isolate, in a reproducible environment, the mapping from beliefs to portfolios under different governance rules. Group shift is measured as the distance between the traditional IC outcome and the arithmetic mean of member portfolios; thus, the empirical object is governance-induced drift, not returns.

This analysis yields three main results. First, in the governance model, centralisation is economically unattractive for weak firms: Without skill, centralisation synchronises errors, increases profit volatility, and raises bankruptcy risk. Only sufficiently skilful firms with adequate risk-bearing capacity should rationally bet their franchise on a single IC portfolio. Second, the anonymous portfolio-vector IC provides a robust aggregation mechanism that preserves the diversification benefits of multiple views while limiting the damage from status, sequencing, and free riding. It does not create alpha, but it does reduce the chance that a meeting manufactures unnecessary concentration. Third, the simulations show consistent evidence of group-shift bias in the traditional IC regime: Final discussion-based portfolios deviate systematically from the arithmetic mean of member portfolios, and the magnitude of this drift is largest when a salient one-dimensional narrative (e.g., "AI boom" or "equity bubble") combines with first-mover anchoring and explicit reputational language. In other words, committees do not merely add noise; they add *directional* noise aligned with the narrative of the meeting.

The academic and practitioner body of literature relevant for IC design is sparse. The research comes from three directions that rarely talk to one another: (1) empirical work on team versus individual decision making in asset management; (2) behavioural and social psychology research on group dynamics, noise, and the disposition effect; and (3) portfolio and governance papers on centralisation versus decentralisation. Empirical mutual fund studies have long compared team-managed products to individually managed products. Bliss, Potter, and Schwarz (2008) found that teams do not reliably deliver higher alphas once standard factor controls are applied, but they do hold more diversified portfolios and attract larger flows, which is consistent with a governance and marketing role. Bär, Kempf, and Ruenzi (2011) found that teams exhibit less extreme style tilts and compress performance dispersion, which is consistent with a reduction in idiosyncratic risk taking as well as with a potential loss of upside. These studies are suggestive for IC design but are limited by observability: One rarely observes the individual portfolios that would be needed to assess whether the committee outcome is the average of members' information or a distortion caused by group dynamics.

A second strand of research studies centralisation versus decentralisation formally. Sharpe (1981) shows that full centralisation is trivially optimal in a frictionless mean–variance world but that decentralisation can be rational when information and constraints are heterogeneous. Textbook active-management algebra (e.g., Grinold and Kahn 2000; Lee 2000) clarifies how aggregating uncorrelated informative signals raises information ratios and hit rates. Scherer (2023) emphasised that this gain comes with a distributional cost: Centralisation collapses dispersion across mandates and can increase joint failure risk. Practical accounts of centralised portfolio management (diBartolomeo 2018; Scherer 2022) describe how a single model portfolio can be implemented across many accounts, implying that the missing ingredient is often not technology but governance.

A third strand comes from behavioural and organisational research. Social psychology documents that groups are not frictionless averaging machines: Anchoring, information cascades, and group polarisation can generate decisions that are more extreme than the average of individual views. The organisational "noise" literature (e.g., Kahneman, Sibony, and Sunstein 2021) argues that unstructured deliberation can increase variability in judgement and recommends decision hygiene: This requires independent assessments, structured inputs, and algorithmic aggregation of inputs. Only then can a group meeting be focused on commonly understanding market mechanics instead of individual speculation. Related work in forecasting

(Tetlock and Gardner 2015) shows that independent forecasts that have been mechanically aggregated often dominate deliberative consensus. The proposed anonymous portfolio-vector IC is a direct application of these principles to portfolio choice, with the additional feature of explicit risk normalisation to guarantee equal influence in the portfolio space.

The remainder of this monograph proceeds in three steps: (1) a governance model, (2) an institutional design proposal, and (3) an empirical (simulation-based) mechanism check. In Chapter 2, I develop a parsimonious model of centralised versus decentralised decision making in asset management and make explicit the trade-off between diversification of *opinions* and the dispersion (or synchronisation) of *outcomes* across mandates. In Chapter 3, I treat the traditional IC as an aggregation technology and explain why discussion need not behave like mechanical averaging after anchoring, sequencing effects, status-weighted influence, weak accountability, and organisational noise are introduced.

Building on this diagnosis, in Chapter 4, I propose the algorithmic alternative of anonymous portfolio-vector averaging with risk normalisation and clarify how member-level scaling and committee-level target tracking-error scaling correspond to distinct governance choices. In Chapter 5, I evaluate incremental procedural fixes commonly advocated in practice (speaker rotation, formal dissent, and input standardisation) and argue that although potentially beneficial, these remedies do not change the mapping from dispersed views to implemented weights and therefore operate mainly at the margin. In Chapter 6, I address the organisational and political reasons why anonymous averaging is rarely adopted despite its transparency and simplicity, emphasising the implied loss of CIO discretion, the reintroduction of individual measurability, and the reduced availability of a single *ex post* narrative.

The empirical component begins in Chapter 7, in which I describe the simulation protocol used to generate counterfactual committee outcomes under controlled conditions, including the asset universe, the fixed cast of committee roles, and the catalogue of macro scenarios. In Chapter 8, I report evidence on governance-induced drift (group-shift bias); I first compare discussion-based outcomes with the mechanical equal-weight average of the same *ex ante* member portfolios and then relate the direction and magnitude of drift to scenario salience and meeting structure. In Chapter 9, I discuss internal and external validity of using an LLM as a controlled narrative simulator for IC behaviour, with an emphasis on interpretation as mechanism evidence rather than calibration of magnitudes. In Chapter 10, I conclude with the implications for institutional design and explain when centralisation is economically justified.

A set of appendices provides technical detail and supporting material. Appendix A collects derivations for the governance model. Appendix B refines the profit specification to allow for base fees and demonstrates that the qualitative centralisation trade-offs are unchanged under a more realistic revenue structure. Appendix C derives the anonymous portfolio-vector averaging rule with risk normalisation and target tracking-error scaling. Appendix D illustrates historical episodes consistent with the failure modes emphasised in the main text.

2. DIVERSIFICATION OF OPINION VERSUS DISPERSION OF OUTCOMES

When considering the trade-off between diversification of opinions and the dispersion of outcomes across mandates, the normative argument for ICs is straightforward. If individual portfolio managers (PMs) have heterogeneous but partly informative views on expected returns, then aggregating those views should improve the information ratio of the centralised product, much as diversifying uncorrelated alpha streams improves a fund of funds. My argument is undermined by a key friction. Centralised decision making will underperform or outperform *all* mandates at once, whereas a decentralised structure generates a dispersion of outcomes across clients.[1] I model the impact of this friction as a choice of the economically optimal governance framework. Under which conditions is it preferable to install a centralised IC? I keep the algebra to a minimum and focus on the economic logic. A full technical derivation is provided in Appendix A.

For the purpose of this monograph, an IC does not merely provide yearly outlooks and monthly minutes for marketing purposes. Its core activity is to decide on a set of common active positions across all firm mandates. Under centralised decision making, individual PMs have no discretion and portfolios must differ only because of different investment guidelines. The exact opposite is true in a decentralised firm. Every PM can choose positions on the basis of their personal views rather than implementing an IC portfolio.[2]

At the intuitive level, my analysis can be summarized as follows. Centralising the investment process by pooling all information into a single portfolio replicated across mandates can improve performance if the firm (individual PMs employed by the firm) truly has investment skill. Although combining signals strengthens the probability of beating the benchmark, it also aligns outcomes. As a result, centralisation is attractive only for firms with sufficiently strong skill; in these cases, the higher expected number of winning mandates should outweigh the increased volatility of profits. Firms with little or no skill gain nothing from aggregation but still bear the risk of synchronized failures, so they rationally prefer decentralised structures. In short, centralisation amplifies both skill and mistakes, meaning only genuinely strong firms benefit from this approach. Weaker firms will be better off diversifying outcomes through decentralisation.

Basic Setup: Skill and Hit Ratios

I start with an asset management firm that runs N mandates (funds or accounts), each of size A, charging a fixed fee f per unit of assets. Excess returns on mandate i over one year are written as follows:

$$r_i = \alpha + \varepsilon_i,$$

[1]The portfolio management literature has advanced steadily over the past five decades; academic work on asset management as a *business* has not kept pace. Chapter 2 of this monograph, along with Appendices A and B, narrows that gap.

[2]Whether it is attractive for clients to participate in a portfolio manager lottery is not part of this monograph. In practice, almost all firms will pretend to have an IC even though it is toothless in defining common views let alone enforcing their implementation across all client portfolios.

where α is the firm's common skill and ε_i is idiosyncratic noise. Without loss of generality, I further assume that all PMs have independent views.

For a given information ratio IR, following standard normal distribution arguments, we map this to a *hit ratio*—that is, the probability that a mandate outperforms its benchmark:

$$p(\mathrm{IR}) = \Pr(r_i > 0) = \Phi(\mathrm{IR}), \tag{2.1}$$

where Φ is the standard normal cumulative distribution function. When IR = 0, we get a coin flip ($p = 1/2$); as skill rises, the hit ratio increases monotonically.

We compare two stylised governance structures:

- Decentralised (D): Each mandate is run by its own PM. Conditional on α, mandate outcomes are independent. Let $p_D = \Phi(\mathrm{IR}_D)$ be the hit ratio per mandate.
- Centralised (C): The firm runs a single IC portfolio off a pooled signal set. Every mandate is a scaled clone of this portfolio. Its hit ratio is $p_C = \Phi(\mathrm{IR}_C)$, with $\mathrm{IR}_C > \mathrm{IR}_D$ if real skill exists to aggregate. In this case, we get $\mathrm{IR}_C = \sqrt{N}\mathrm{IR}_D > \mathrm{IR}_D$.

Diversification Across Inputs (opinions) Versus Outcomes (fees)

Let n_g denote the number of mandates that outperform under governance regime $g \in \{D, C\}$.

Under decentralisation, each mandate has hit probability p_D, which is independent of the others. The number of winners, n_D, is then binomially distributed with mean and variance, as follows:

$$\mathbb{E}[n_D] = Np_D;\ \mathrm{var}(n_D) = Np_D(1 - p_D). \tag{2.2}$$

Intuitively, some mandates win and some lose; most of the time, you end up with a middle number of winners.

Under centralisation, all mandates hold the same IC portfolio. Either they all beat the benchmark, or they all fail together:

$$n_C = \begin{cases} N\ \textit{with probability}\ p_C, \\ 0\ \textit{with probability}\ 1 - p_C. \end{cases}$$

The mean and variance are therefore

$$\mathbb{E}[n_C] = Np_C,\ \mathrm{var}(n_C) = N^2 p_C(1 - p_C). \tag{2.3}$$

We obtain the following simple trade-off:

- When $p_C > p_D$ (there is skill, and the IC aggregates it), centralisation raises the *expected* number of winning mandates.
- At the same time, variance explodes because mandates now live and die together. You either have a year in which everybody is a hero or a year in which they are all dogs.

From Hit Ratios to Profits

I start with the assumption that fee income is proportional to the number of winning mandates. In Appendix B, I relax this assumption. Let $\pi_0 = fA$ be the yearly fee income per outperforming mandate, including flows and performance-related economics. Under governance regime g, fee income is as follows:

$$\Pi_g = \pi_0 n_g.$$

Using Equations 2.2 and 2.3, we obtain the following:

$$\mathbb{E}[\Pi_D] = \pi_0 N p_D,\ \mathrm{var}(\Pi_D) = \pi_0^2 N p_D (1 - p_D),\ \text{and}$$

$$\mathbb{E}[\Pi_C] = \pi_0 N p_C,\ \mathrm{var}(\Pi_C) = \pi_0^2 N^2 p_C (1 - p_C).$$

Asset management owners dislike volatility in profits, either because of preferences or because banks and regulators do. A convenient approximation is a simple mean–variance objective:

$$V_g \approx \mathbb{E}[\Pi_g] - \frac{\gamma}{2}\mathrm{var}(\Pi_g), \tag{2.4}$$

where $\gamma > 0$ is a coefficient of risk aversion. The firm prefers centralisation if $V_C > V_D$—that is, if the gain in expected profits from a higher hit ratio is large enough to offset the higher variance of profits.

Weak Firms

A weak firm has no skill: $\mathrm{IR}_D \approx 0$, and aggregation does not magically create skill, so $\mathrm{IR}_C \approx 0$. Then, $p_D \approx p_C \approx 1/2$, and both structures have the *same* expected number of winners and the *same* expected fee income.

The variance term, however, is different: Centralisation replaces a benign world in which roughly half of the mandates win every year with a brutal lottery in which either *all* mandates win or *none* do. For any $\gamma > 0$, a weak firm strictly prefers decentralisation. Centralising without skill just synchronises losses.

Strong Firms

A strong firm has positive skill: $IR_D > 0$, and aggregation pushes IR_C higher, so $p_C > p_D > 1/2$. Centralisation improves the expected number of winners and thus the expected profit. It still raises variance dramatically. Therefore, a *skill threshold* IR* shows that centralisation is optimal only if the firm's baseline IR exceeds IR*. Strong firms centralise; marginal shops rationally stick with decentralisation.

Numerical Illustration: Skilled Versus Weak Firms

Exhibits 1 and **2** show the distribution of the number of outperforming mandates for two extreme cases, following Scherer (2023): (1) a skilled firm with $\mathrm{IR}_D = 0.2$ per PM and $\mathrm{IR}_C = 0.63$ for the IC and (2) a weak firm with $\mathrm{IR}_D = \mathrm{IR}_C = 0$.

Exhibit 1. Distribution of Outperforming Mandates: Skilled Firm ($N = 10$, $IR_D = 0.2$, $IR_C = 0.63$)

No. of Outperforming Mandates	Decentralised	Centralised
10 (all)	0.43%	73.6%
9	3.09%	0.0%
8	10.10%	0.0%
7	19.56%	0.0%
6	24.86%	0.0%
5	21.67%	0.0%
4	13.12%	0.0%
3	5.44%	0.0%
2	1.48%	0.0%
1	0.24%	0.0%
0 (none)	0.02%	26.4%
Expected no. of outperformers	5.79	7.36
Variance	2.44	19.45

Source: Authors' computations.

Under decentralisation, the skilled firm nearly always has some winners and some losers; the mass is concentrated between three and seven outperformers. Under centralisation, either all mandates outperform or none do. The expected number of winners is higher (7.36 versus 5.79), but the variance is roughly eight times larger.

For the weak firm, the mean number of winners is exactly the same under both governance structures (five), but the variance under centralisation is 10 times higher. In half the years, all mandates outperform; in the other half, none do. Expectation offers no benefit—just more ways to blow up.

Exhibit 3 translates these distributions into the mean–variance value, V_g, from Equation 2.4, normalising $\pi_0 = 1$ and setting risk aversion to $\gamma = 0.05$.

For the strong firm, centralisation delivers both higher expected profits and a higher mean–variance value V. For the weak firm, expected profits are the same, but V is strictly lower under centralisation because of the variance penalty. The result is obvious: If you have no alpha, putting everything through one IC does not fix that; it just synchronises your mistakes.

Exhibit 2. Distribution of Outperforming Mandates: Weak Firm ($N = 10$, $IR_D = IR_C = 0$)

No. of Outperforming Mandates	Decentralised	Centralised
10 (all)	0.10%	50.0%
9	0.98%	0.0%
8	4.39%	0.0%
7	11.72%	0.0%
6	20.51%	0.0%
5	24.61%	0.0%
4	20.51%	0.0%
3	11.72%	0.0%
2	4.39%	0.0%
1	0.98%	0.0%
0 (none)	0.10%	50.0%
Expected no. of outperformers	5.00	5.00
Variance	2.50	25.00

Exhibit 3. Mean–Variance Value of Decentralisation versus Centralisation ($N = 10$, $\pi_0 = 1$, $\gamma = 0.05$)

	Decentralised			Centralised		
Firm Type	$\mathbb{E}(\Pi)$	var(Π)	V	$\mathbb{E}(\Pi)$	var(Π)	V
Strong ($IR_D = 0.2$, $IR_C = 0.63$)	5.79	2.44	5.73	7.36	19.45	6.87
Weak ($IR_D = IR_C = 0$)	5.00	2.50	4.94	5.00	25.00	4.38

Bankruptcy Risk and Fixed Costs

The mean–variance objective is still too kind to centralisation. In real life, asset managers have fixed overhead F (people, systems, compliance, the usual bureaucracy). If too many clients fire you in the same bad year, fee income cannot cover F, and the firm has to shrink, sell itself, or shut down.

The extended model adds

- a breakeven number of winners n^* needed to cover fixed costs,
- a bankruptcy probability $q_g = \text{prob}(n_g < n^*)$ under regime g, and
- a bankruptcy cost B capturing the value destroyed if the franchise is lost.

The owner's objective becomes

$$V_g^B \approx \mathbb{E}[\Pi_g] - \frac{\gamma}{2}\text{var}(\Pi_g) - Bq_g.$$

For weak firms with $p_D \approx p_C \approx ½$, centralisation explodes both the variance of profits and the probability that everyone underperforms at once. The mean term does not improve, but the bankruptcy grows larger: q_C is much higher than q_D. Centralisation, in that case, is a fairly direct route to closing down.

For moderately strong firms, the conclusion is more nuanced but no less sobering. The mean–variance gain from centralisation has to compensate not only for higher variance but also for the sharply higher chance of falling below the breakeven point in a single bad IC year. The skill threshold, IR*, at which centralisation becomes optimal is therefore higher once bankruptcy risk is priced in. Only genuinely skilful, scalable firms with strong balance sheets should even consider betting their franchise on a single IC portfolio.

Why All Asset Owners but Not All Asset Managers Have ICs

This governance model explains an otherwise puzzling fact of institutional life: Almost all large *asset owners* (pension plans, endowments, sovereign funds) have ICs, but many *asset managers* have nothing of the sort.[3] The model draws a clean line between situations in which centralisation in an IC is economically justified and situations in which it is mostly theatre.

In the model, an asset management firm with N mandates chooses between decentralised portfolio management (each mandate run by a separate PM) and a centralised IC portfolio. Centralisation not only raises the expected number of outperforming mandates when the firm has skill but also multiplies the variance of outcomes and the probability that every mandate underperforms in the same year. For weak firms with essentially no skill, centralisation offers no gain in expected profit and a large increase in both volatility and bankruptcy risk. These weak firms rationally prefer decentralisation.

The $N > 1$ Case: Asset Managers

For an asset manager with many mandates ($N > 1$), decentralisation versus centralisation is a real economic choice. Under decentralisation, the number of winners, n_D, is binomial with mean Np_D and variance $Np_D(1 - p_D)$; some clients win and some lose each year. Under a central IC, the number of winners, n_C, is either 0 or N with probabilities $1 - p_C$ and p_C; everyone wins

[3] I consider ICs designed for marketing purposes but without the power to enforce a firmwide position to be fake.

or loses together. Centralisation, therefore, trades a higher mean number of winners (if $p_C > p_D$) against much higher variance and a fatter disaster tail.

A profit-maximising, risk-averse asset manager with fixed costs and real bankruptcy risk centralises only if two conditions hold:

- It has enough skill that p_C is meaningfully above p_D.
- Its balance sheet and owners can tolerate highly correlated bad years.

Weak shops with marginal or zero skill have no business running a single IC portfolio for all mandates: They merely synchronise losses and raise bankruptcy risk. Strong, scalable firms can rationally centralise; everyone else should treat ICs with suspicion.

The $N = 1$ Case: Asset Owners

Asset owners effectively manage *one* mandate: a single pension plan, endowment, or sovereign fund balance sheet. From the perspective of this governance model, they are in the $N = 1$ case. That changes the economics completely.

When $N = 1$, there is *no* cross-sectional dispersion of mandates to diversify. Under decentralisation and centralisation, the number of winners is either $n_D \in \{0,1\}$ or $n_C \in \{0,1\}$; the owner either beats the benchmark that year or does not. In this case, no "half the clients win, half lose" scenario has to be traded off. In particular:

- Centralising decisions inside an IC cannot increase the correlation of mandate outcomes across clients, because it has only one mandate.
- At the same time, *any* sensible aggregation of diverse views that raises the information ratio of the central portfolio (moving from IR_D to $\text{IR}_C > \text{IR}_D$) improves the hit ratio for *every* level of skill, however small.

In other words, asset owners face no "diversification across mandates versus diversification across opinions" trade-off. By construction, they cannot diversify across mandates; the only diversification available is across *opinions*. Provided the IC actually aggregates information rather than destroying it, some form of committee is always attractive for any $\text{IR}_D > 0$.

Different Objective Functions and Agency Problems

Different incentives sit on top of this mechanical $N = 1$ argument. Asset managers care about the level and volatility of fee income and the risk of going out of business. Asset owners care about long-term surplus over liabilities and, more immediately, governance risk: the risk that trustees, CIOs, or politicians will be blamed for "going off process" after a bad year.

For an asset manager, a CIO, in principle, can run a mandate alone and sell that concentration of decision rights to clients who can redeem if they dislike it. For an asset owner, a lone CIO structure is politically fragile: Beneficiaries cannot redeem, and boards, auditors, and the press will ask "who approved this?" A formal IC, however imperfect economically, diversifies *blame* and demonstrates process.

When considering why all asset managers do not have ICs but all asset owners do, we can put the following pieces together:

- For asset managers with many products ($N > 1$), ICs are an optional technology. They must earn their keep by raising expected profits enough to compensate for higher profit variance and higher bankruptcy risk. Only firms with strong, scalable skill and robust capital or those forced by demanding institutional clients will rationally centralise.
- Asset owners with one balance sheet ($N = 1$) do not face a downside from "losing diversification across mandates"—it does not exist in the first place. An upside from diversifying opinions exists whenever the IC's aggregation raises the information ratio and political demand is strong for a visible process. Thus, ICs are a near-universal equilibrium institution, regardless of whether they help returns.

The anonymous portfolio-vector design advocated in this monograph simply accepts this asymmetry and tries to make the best of it. If asset owners are going to have ICs anyway, they should at least use them to diversify opinions in a disciplined, algorithmic way, rather than to amplify noise and provide post hoc narratives.

3. CHALLENGES TO INVESTMENT COMMITTEES

This chapter is grounded in my experience across senior investment roles and culminates in the perspective I gained after stepping into the CIO seat for various investment organisations. Seeing committees from both sides of the table clarifies that their performance depends less on who is in the room than on how the room is run. In this discussion, I blend formal arguments with concrete observations about cases in which committees add discipline—and cases in which they quietly manufacture risk. Because failure modes are often more instructive than success stories, I draw examples from both effective and dysfunctional processes. The goal is to distill these lessons into practical design principles that improve decision quality without turning governance into theatre.

Defining Traditional Investment Committees

A traditional IC meeting is a monthly ritual. Members arrive with thick slide decks, most of which are lightly rebranded broker reports, copied charts from last week's strategy notes, and valuation tables whose axis labels are either missing or incompatible. One chart shows price-to-earnings ratios on a log scale, another uses levels; one assumes spot inflation; another assumes long-run targets. All these charts fail to make assumptions explicit.

The meeting begins with a macro overview. Someone speaks first—usually the CIO, sometimes the most confident strategist—and sets the tone. Growth is slowing, inflation is sticky, policy is restrictive, geopolitical risks are rising. Heads nod. A second speaker adds a chart showing that equities are "not cheap." A third shows a different chart suggesting that equities are "not expensive either." The IC agrees that "valuation is mixed." Discussion follows. Opinions are expressed verbally, often in qualitative language: "I feel uncomfortable with duration." "This trade is crowded." "We should keep some dry powder." Dissent is polite but rarely decisive. Nobody writes down an explicit portfolio. Nobody specifies how much risk they would actually like to take if they were accountable for the outcome. At the end, the CIO summarizes. A few tilts are agreed on: Slightly underweight equities, slightly overweight defensives, keep some inflation protection. The final portfolio is not the average of individual views—no such average was ever computed—but a negotiated compromise shaped by hierarchy, rhetoric, and time pressure. The minutes record "robust discussion." The weights are implemented. This process looks like aggregation. It is not.

From an economic perspective, the problem is simple. Aggregation requires a well-defined mapping from individual information sets to a collective action. Traditional ICs do not have such a mapping. They rely on conversation. Conversation has three structural weaknesses.

First, it is *nonlinear*. Small differences in tone, confidence, or speaking order can lead to large differences in outcomes. Two meetings with the same participants and the same information can produce different portfolios. This is not wisdom of crowds; it is occasion noise.

Second, it is *opaque*. Once the discussion ends, individual views cannot by clearly translated into portfolio weights. *Ex post*, every member can plausibly claim that the outcome was "close" to their view. Accountability evaporates precisely when it would be most useful.

Third, it is *status weighted*. Influence is proportional not to the statistical quality of information but to seniority, confidence, and narrative coherence. A weak but well-articulated story can dominate a strong but quietly held view. The committee aggregates personalities, not signals.

In the language of the governance model developed earlier, traditional ICs attempt to raise the information ratio of the central portfolio by pooling signals, but they do so with a mechanism that adds noise and bias. Even when skill exists, discussion-based aggregation can destroy a substantial fraction of this skill. Instead, in this monograph, I argue that algorithmic aggregation is preferable for a host of reasons that I will dissect from the observation that discussion is the wrong place to combine information. If IC members have views, those views should be expressed directly in the portfolio space. In a portfolio-based IC, each member submits a portfolio vector reflecting their beliefs. These vectors are scaled to the same *ex ante* risk to ensure equal influence and then are averaged mechanically. The final portfolio is not the outcome of a negotiation but, rather, is the arithmetic mean of risk-normalised views.

Crucially, this type of algorithmic aggregation does not assume that ICs are wise. It assumes they are noisy. It therefore constrains discretion rather than celebrating it. If the group has skill, averaging preserves it. If the group does not, averaging at least prevents the committee from amplifying its confusion into a single, highly correlated bet.

Exhibit 4 summarizes the weakness of traditional ICs. It is not the idea of centralisation but the way in which individual views are aggregated. Industry-standard ICs rely on qualitative discussion and majority-based consensus. In theory, this should diversify opinions; in practice, it often destroys them.

Traditional ICs suffer from three structural problems. First, they are prone to *group-shift bias*. Information sequencing and social dynamics push the committee toward more extreme positions than the average of individual views would imply. Second, they create *incentive and accountability problems*. Once decisions are taken collectively, it becomes difficult to evaluate individual contributions *ex post*. Third, they face an *aggregation problem*: The final portfolio is not guaranteed to correspond to the arithmetic average of the information contained in individual views.

The algorithmic, portfolio-based IC advocated in this monograph addresses these issues directly. Instead of aggregating opinions verbally at the end of a meeting, each IC member submits a long–short portfolio vector. All vectors are scaled to the same *ex ante* risk to ensure equal influence, are averaged mechanically, and then are rescaled to a target tracking error. Anonymity reduces group-shift bias, while explicit portfolios restore incentives and accountability. Importantly, this change does not require heroic assumptions about skill. If the organisation has no alpha, algorithmic aggregation will not create it. This approach, however, prevents the committee from amplifying noise and behavioural biases. If the organisation does have skill, algorithmic aggregation preserves the benefits of centralisation while also avoiding many of the side effects of discussion-based consensus building. I build on this distinction in the following sections. I first examine why common procedural fixes for traditional ICs help only at the margin. I then develop the case for anonymous portfolio-vector aggregation as a structural governance solution that is robust to noise, agency problems, and human overconfidence.

Exhibit 4. Decision Rules for Investment Committees

	Traditional IC (Majority Based)	Algorithmic IC (Portfolio Based)
Aggregation of opinion	Qualitative: The CIO aggregates investment opinions at the end of a committee meeting and tries to find a consensus among IC members using group discussion.	Algorithmic: Individual portfolios (vectors) are aggregated mechanically by averaging.
Input format	Verbal opinions, qualitative assessments, narratives, and informal votes.	Quantitative long–short portfolio vectors submitted by each IC member.
Equal influence of members	Not guaranteed: Seniority, rhetoric, and speaking order affect outcomes.	Guaranteed by construction through equal risk scaling of portfolios.
Role of CIO	Dominant: The CIO typically chairs the meeting, speaks early, frames the debate, and summarizes the outcome.	Limited: The CIO provides one portfolio vector like any other member and sets the risk budget *ex ante*.
Incentives and accountability	Weak: Individual contributions are hard to evaluate *ex post*; responsibility is shared.	Strong: Individual portfolios are observable and can be evaluated against outcomes.
Susceptibility to group-shift bias	High: Information cascades and group polarisation can lead to extreme consensus positions.	Low: Anonymity and mechanical averaging reduce group dynamics.
Disposition effect	High: Teams tend to delay loss realisation and defend past consensus decisions.	Reduced: Individual portfolios preserve accountability and shorten feedback loops.
Diversification of opinion	Intended but often diluted by discussion dynamics.	Fully preserved through independent portfolio submissions.
Final portfolio	Consensus outcome of discussion; may not reflect the average of individual views.	Exact arithmetic average of individual risk-normalized portfolios, rescaled to a target risk.

Source: Replication of Table 3, "Optimal Design of Investment Committees," in Scherer (2023).

Note: The table highlights that the key distinction is not centralisation per se but the rule by which individual views are aggregated into a portfolio.

Group Dynamics, Noise, and Decision Hygiene

The social psychology literature has been warning for decades that groups are not frictionless averaging machines. Davis and Hinsz (1982) reviewed evidence showing that groups are often no better than the average member and sometimes worse in judgement tasks. High-status members tend to dominate, information that is already widely shared in the group gets far more airtime than distinct information, and after a majority direction becomes visible, groups tend to drift in that direction.

Mulvey and Klein (1998) studied social loafing and collective efficacy in groups. When members perceive others as free riding, their effort declines; the marginal benefit from extra work is small when the outcome is shared, and the marginal cost is real. Collective efficacy (the belief that the group as a whole is capable) helps somewhat but not enough to neutralise the free-rider logic. This is precisely the structural problem many ICs have: Everyone's bonus depends on performance, but nobody's bonus depends much on marginal effort in the meeting.

Cooper and Kagel (2005) used laboratory signalling games to compare individuals and small teams. Teams sometimes outperform individuals because discussion helps avoid very naive strategies, but this is a highly process-dependent approach. Without structured procedures, teams easily get stuck in unproductive arguments or follow a confident member's incorrect view.

Kahneman et al. (2021) pushed these themes into organisations under the label of *noise*: unwanted variability in judgements of essentially identical cases. They showed that when firms run noise audits—having multiple experts independently judge the same case—the dispersion in judgements is far greater than managers expect. Unstructured group deliberation does not reliably reduce this noise. Instead, early and high-status voices anchor the discussion, and occasion noise (e.g., time of day, recent news, mood) drives outcomes. The remedies they advocate are "decision hygiene": independent assessments before any discussion, structured inputs, and algorithmic aggregation of inputs—and only then a group meeting focused on understanding rather than horse trading.

Lovallo and Sibony (2010) applied similar ideas to corporate strategy decisions and argued that the main gains come from better process design, not from exhorting people to be less biased. In the context of ICs, this is an argument for changing how decisions are elicited and aggregated, not a call for yet another training on cognitive biases.

Disposition Effect and Team Behaviour

The disposition effect is the tendency to sell winners too early and hold losers too long. Odean (1998) documented this pattern for retail investors: Realised gains greatly exceed realised losses, even though the sold winners underperform the held losers on average. Standard prospect theory preferences and regret avoidance explain the result.

Cici (2012) showed that professional managers are not immune. Using mutual fund trades and holdings, he found that many funds exhibit a disposition effect, although its prevalence has declined somewhat over time. More worrying for ICs, he reported that team-managed funds often display *stronger* disposition effects than individually managed funds.

Rau (2015) provided experimental evidence that teams can exacerbate the disposition effect. In laboratory asset markets, two-person teams are more reluctant than individuals to realise losses. The social context amplifies the emotional cost of admitting mistakes; closing a losing position in front of a colleague is worse than doing so privately.

For ICs, the implication is awkward. Committees do not merely fail to correct individual biases; under plausible conditions, they amplify them, particularly when the mechanism is embarrassment rather than pure miscalculation. This fits the informal observation that ICs tend to cling to losing positions with elaborate narratives about "giving the thesis time to play out," even when a cold look at the numbers would argue for cutting risk.

Noise in Team Decisions: The Basic Indictment

I was heavily influenced by Kahneman et al.'s (2021) book on *noise* in decision making. Although their work is not about ICs per se, it might as well be. Their object of study is any organisation that relies on human judgement—judges, underwriters, hiring panels, or M&A committees. An IC is just another version of the same thing: a group of people around a table pretending that "deliberation" somehow converts disagreement and ignorance into wisdom. The phrase *noise* is defined as undesirable variability in repeated (across time or individuals) judgements of the same problem. The point is not that people are biased (we knew that already) but that even when bias is held fixed, dispersion in judgements is far greater than organisations realise. Three features are relevant for ICs:

- **Level noise and pattern noise:** Different decision makers systematically sit at different "levels" (harsh versus lenient, risk-on versus risk-off) and react differently to the same cues. Translated into IC language: Some members are chronically bullish on equities, some are chronically scared of drawdowns, and some are in love with whatever factor back-test they saw last week.
- **Occasion noise:** The *same* individual (or committee) will give different answers to the same problem depending on when you ask, in what order the information arrives, who speaks first, and what happened in markets yesterday. In other words, a large part of what passes for so-called macro insight in an IC is just occasion noise with a narrative bolted on afterward.
- **Group amplification of noise:** When groups deliberate, early and high-status speakers (e.g., CIOs and senior strategists) anchor the discussion; others adjust toward them, privately disagreeing but publicly converging. The visible dispersion of views shrinks, but the underlying randomness in which view happens to anchor the meeting remains. This creates a noisy process covered with a thin varnish of apparent unanimity.

My uncomfortable empirical realisation is that whenever you actually run a "noise audit"—giving many judges the same case and measuring dispersion—organisations are shocked by how little agreement they achieved. We have no reason to believe ICs are magically exempt; if anything, the mix of macro storytelling, politics, and career risk should make these factors worse.

In terms of the simple model given in Chapter 2, decision noise is a reminder that our neat mapping from information ratios to hit ratios, $p(\text{IR}) = \Phi(\text{IR})$, is optimistic for real-world committees. A traditional IC is not a frictionless machine that converts N independent alpha streams into $\text{IR}_C = \sqrt{N}\,\text{IR}_D$. This noisy, path-dependent process can easily waste a good part of whatever skill the firm has.

Kahneman's Remedies: Decision Hygiene, Not Hero CIOs

Kahneman et al.'s (2021) book *Noise* has the decency not to end with "train people to be less biased" or "hire geniuses." Instead, it pushes organisational plumbing under the heading of *decision hygiene*. The main prescriptions are as follows:

- **Run noise audits.** Before anyone pretends that the IC "has a view," have members independently assess the same case—in our context, the same macro scenario and risk constraints—and measure how far apart they actually are from reaching consensus.

If dispersion is huge, the governance problem is not "too little debate"; it is "we do not even agree on the sign of expected return."

- **Structure the decision into mediating assessments.** For complex decisions (M&A, hiring, credit), break down the problem into a handful of explicit dimensions—valuation, downside risk, and management quality. Each assessor scores these separately, *before* any group discussion.
- **Aggregate judgements algorithmically.** Combine the separate scores with a simple, agreed-on rule. Only after this structured aggregation do you allow the group to discuss the overall conclusion. Intuition is not banned, but it is delayed until it is at least informed by the decomposed evidence.
- **Use the Mediating Assessments Protocol instead of free-form meetings.** Kahneman and Sibony (2021) call the entire bundle the Mediating Assessments Protocol (MAP): independent assessments, outside views, common scales, aggregation rules, and only then a decision meeting. This protocol is, frankly, everything most ICs do not do.

If one takes *noise* seriously, the standard IC—that is, the CIO monologue, a few supportive comments, some ritual "challenge," and a consensus tilt—is almost the worst possible design. This scenario creates maximum scope for status, anchoring, and occasion noise; it provides minimum structure.

Summary of Committee Pathologies

Drawing on the behavioural literature, we can highlight the empirically relevant distortions as follows:

- **Information sequencing and cascades:** If the CIO or a dominant member states a strong view early in the meeting, other members rationally or irrationally update toward that view because they either infer information from seniority or want to avoid visible disagreement.
- **Group polarisation (group shift):** Teams tend to end up with more extreme positions than the average of *ex ante* individual opinions. Once the perceived majority direction is clear, members have an incentive to move toward it and often beyond their initial position.
- **Free-rider problem:** Individual effort in forming a view is hard to observe. The marginal impact of one member's analysis on the final portfolio is small, yet the time cost is real. This weakens incentives to produce independent research.
- **Disposition effect in teams.** Evidence suggests that teams are even more reluctant than individuals to realise losses, presumably because admitting mistakes is reputationally costlier in a public forum.

All these mechanisms undermine the theoretical benefits of centralisation. Instead of aggregating independent signals, the committee compresses disagreement into an apparently unanimous but noisy and sometimes unstable consensus.

4. ANONYMOUS AVERAGING AS AN INSTITUTIONAL FIX

To address these committee distortions, my proposed remedy is institutionally modest but conceptually sharp. Rather than asking IC members to reach a verbal consensus at the end of a meeting, the CIO requires each member to submit an *individual long–short portfolio vector* over the relevant asset universe. Each vector is scaled to a common level of *ex ante* risk so that all members have equal influence. Most important, all vectors are provided anonymously to ensure that IC members are not subject to loss of face or ridicule. These vectors are averaged and rescaled to a fictious target tracking error. Each member remains accountable for the performance of their portfolios (year-end numbers determine the year-end bonus to incentivise information sharing), but the official IC portfolio is a simple anonymous average. In my experience with this setup, the design attacks several pathologies at once: It reduces group shift, limits CIO dominance, eliminates free riding in the collective decision, and preserves diversification across views.

Anonymous Portfolio Averaging

ICs typically behave like a seminar: Whoever speaks first anchors the discussion, whoever speaks last writes the minutes, and whoever speaks loudest gets the risk budget. That is not information aggregation; it is a social process with a Bloomberg terminal. The proposed anonymous portfolio-vector approach is the opposite. It forces each member to express their view in the only language that matters—positions—and then aggregates those views mechanically.

Let $w^{(j)}$ be the portfolio vector proposed by member j (over the chosen asset universe), and let $\tilde{w}^{(j)}$ be the *risk-normalised* version of that vector (same *ex ante* tracking error for every member). The IC portfolio is simply the average of these normalised vectors:

$$w^{IC} = \frac{1}{M}\sum_{j=1}^{M}\tilde{w}^{(j)}. \tag{4.1}$$

That is the whole design. It looks trivial because it is. The important part is what that triviality buys you:

- **Equal voice by construction:** The normalisation step ensures that "one member, one vote" is implemented in portfolio space rather than in rhetorical space. Without it, the most aggressive member dominates the average just by taking more risk.
- **No CIO capture:** The CIO's portfolio is just one of the $\tilde{w}^{(j)}$. The average prevents a single view—or a single ego—from becoming the house portfolio purely because of that person's status.
- **Noise reduction without pretending to be wise:** Averaging independent signals reduces noise. It does not create alpha out of thin air. If the group has no skill, the average has no skill either; at least it will not amplify the committee's worst behavioural tendencies.

In short, Equation 4.1 illustrates a governance choice: It is a way to extract the diversification benefit of multiple opinions while capping the damage from group dynamics. If you insist on having an IC, this is what it should look like: a dull algorithm with no charisma.

What Anonymous Portfolio Aggregation Adds

Relative to Kahneman's prescriptions, the anonymous portfolio-vector IC introduced in this monograph is both narrower and more radical.

First, it is deliberately *domain specific*. Kahneman et al. (2021) wanted a generic noise-reduction protocol that works for hiring, sentencing, medical diagnosis, and corporate strategy. That forced them to stop at relatively abstract process advice (i.e., MAP). In asset management, we are luckier: The decision we care most about is a portfolio weight vector. We do not need to guess at mediating assessments; the portfolio *is* the mediating assessment.

Second, anonymous aggregation pushes the logic of decision hygiene further than *noise* is willing to go:

- **From scores to portfolios:** Instead of asking IC members for numerical scores on subcriteria and then "using them as inputs" to a still somewhat discretionary debate, we ask for full portfolio vectors, on equal risk, upfront. That automatically embeds their view on all relevant attributes (expected return, risk, diversification, liquidity) without the committee having to agree on a factor model.
- **From structured discussion to minimal discussion.** MAP still expects a decision meeting in which senior people can, in effect, lean on the structured assessments and tilt the outcome. Anonymous vector aggregation removes that last step: After the vectors are in and scaled, the official portfolio is the average. Meetings are for risk-budget housekeeping, not for overriding the algorithm.
- **From noise reduction to incentive alignment:** *Noise* largely treats judges as exogenous; it wants to make their judgements less variable, not to change their incentives. The anonymous IC design does both. It eliminates the incentive to grandstand in meetings (because rhetoric does not move the weights) and restores accountability at the individual level (each member's private portfolio can be tracked *ex post*). That is precisely the case in which traditional ICs are weakest.
- **From "of course we centralise" to "should we centralise at all?":** Finally, Kahneman et al. (2021) assumed the organisation is right to have a committee and focused on making it less bad. This governance model adds an outer loop: Only firms with IR_D above a threshold and sufficient risk-bearing capacity should centralise in the first place. For weak firms, the optimal "decision hygiene" is not MAP; it is to avoid running a single IC portfolio at all.

Stated differently, the approach of Kahneman et al. tells you that your IC is noisier than you think and suggests washing its hands before it touches capital. The anonymous aggregation scheme in this monograph takes the next step: It takes most of the sharp objects out of the committee's hands, lets each member write down their best shot in private, and then averages them. For an industry that claims to believe in diversification, that seems like the least we should do.

Third, the provision of anonymous vectors improves IC governance. We can measure the contribution of each IC member, we can measure the degree of IC member diversification by clustering members with respect to their weight vectors, and we can analyse the decision making of each member and identify and correct potential behavioural biases.

"Why Not Overweight the Best Decision Makers?"

If some committee members are better than others, shouldn't we overweight them? Why give equal weight to mediocre views when we could identify the best forecasters and scale their influence? This argument sounds sensible and is often invoked to justify CIO dominance or informal hierarchies within ICs. Unfortunately, it collapses under minimal scrutiny.

First, statistical confidence is lacking. In realistic settings, the sample size of truly independent decisions is small, regimes change, and noise is large. Apparent skill differences are rarely distinguishable from luck. The *ex post* identification of "the best decision maker" is almost always hindsight bias wearing a spreadsheet.

Second, even if skill differences exist, they are unstable. The member who was right in the last cycle is not guaranteed to be right in the next. Overweighting based on recent success introduces procyclicality and regime dependence—exactly the behaviour ICs claim to want to avoid.

Third—and most important—the argument quietly reintroduces the agency problem. Overweighting the "best" decision maker almost always means overweighting the CIO or the most senior voice. The claim of superior skill is rarely subjected to the same evidentiary standards we would apply to an external manager. Averaging is resisted not because it ignores skill but because it refuses to accept authority as a proxy for it.

From the perspective of the model given in Chapter 3, equal-weight averaging is the default precisely because it is robust to mismeasurement of skill. It sacrifices the theoretical upside of perfect skill identification in exchange for protection against overconfidence and false precision. That trade-off is economically sensible in environments in which true skill is hard to detect.

Risk Normalisation and Tracking-Error Scaling in IC Aggregation

I next discuss the role of tracking-error scaling in the portfolio aggregation protocol and clarify the economic interpretation of the normalisation steps. At first glance, scaling individual IC submissions to a common tracking error and subsequently scaling the final committee portfolio to a target tracking error may appear redundant. In fact, the two steps address conceptually distinct questions. One governs how *individual views are weighted* in the aggregation process; the other governs the *risk appetite of the institution*. Understanding this distinction is important for interpreting group-shift results and for assessing the governance implications of alternative IC designs.

Two Distinct Normalisation Problems

Let $a^{(j)}$ denote the active portfolio vector submitted by committee member j relative to a benchmark, and let Σ denote the covariance matrix used to measure active risk. The *ex-ante* tracking error (TE) of member j is

$$\mathrm{TE}_j \equiv \sqrt{(a^{(j)})^\top \Sigma a^{(j)}}.$$

We envisage two normalisation steps:

- **Member-level scaling (*ex ante*):** Each member's active portfolio is scaled to a common tracking error σ_0:

$$\tilde{a}^{(j)} = \frac{\sigma_0}{\mathrm{TE}_j} a^{(j)}.$$

- **Committee-level scaling (*ex post*):** After aggregation, the resulting active portfolio $\bar{a}$ is scaled to the institution's target tracking error τ^*.

The first step determines the relative influence of individual submissions; the second determines the absolute level of risk taken by the institution. Treating these as separate design choices makes explicit a distinction that is often blurred in traditional IC processes.

Member-Level Scaling: Equal Influence Versus Conviction

Scaling individual submissions to a common *ex ante* tracking error enforces a strong form of symmetry across committee members. Each member contributes an active portfolio with the same risk magnitude, so differences in the aggregated portfolio reflect differences in *directional beliefs* rather than differences in aggressiveness.

From a governance perspective, scaling offers several advantages. First, equal influence implements a literal "one member, one vote" principle in portfolio space. Without scaling, an aggressive member can dominate the average simply by submitting a higher-risk portfolio. In that case, aggregation confounds beliefs with risk tolerance. Second, *ex ante* scaling prevents the use of risk as a political instrument. In real ICs, conviction is often expressed by taking larger positions rather than by persuading others. Normalisation explicitly removes this channel and forces disagreement to be expressed directionally.

Member-level scaling is also important for interpretation. When group shift is measured relative to the arithmetic mean of scaled submissions, deviations can be interpreted as *directional distortions induced by discussion*, not as changes in aggregate risk appetite. This strengthens identification: Group shift captures narrative and sequencing effects rather than heterogeneity in risk aversion.

The cost of this symmetry is that information about conviction is suppressed. If some members genuinely have stronger signals and would optimally take more risk, equal-TE scaling discards that information. Implicitly, the procedure assumes that differences in conviction are not reliable indicators of superior information. This assumption is defensible in many institutional settings, particularly for asset owners with a single balance sheet, but it should be acknowledged.

Committee-Level Scaling: Institutional Risk Choice

Scaling the final portfolio to a target tracking error addresses a different problem. Institutions typically set risk budgets at the portfolio level, not at the level of individual IC members. Regardless of how beliefs are aggregated, the resulting portfolio must satisfy an overall risk constraint to be implementable.

Committee-level scaling therefore reflects institutional practice. It separates the question "what direction should the portfolio tilt?" from the question "how much risk should the institution take?" This separation is often implicit in IC discussions but rarely enforced mechanically. Making it explicit clarifies accountability: The IC determines direction, and the institution (or the board) determines risk appetite.

If member-level scaling is not applied, committee-level scaling preserves relative conviction across members. Members who submit higher-risk portfolios have a greater influence on the final outcome. This approach may be appropriate in some contexts, particularly in asset management firms in which PMs are explicitly evaluated on risk-adjusted performance and differences in conviction are meant to matter. In such settings, however, deviations from the mean must be interpreted carefully because they conflate directional bias with differences in risk taking.

Using Both Steps: Separation of Roles

Applying both member-level and committee-level scaling is not redundant. These two steps control different margins. Member-level scaling fixes relative influence; committee-level scaling fixes absolute risk. Together, they implement a clean separation between belief aggregation and risk choice:

$$\text{belief aggregation} \rightarrow \text{institutional risk budget}$$

This separation is rarely achieved in traditional ICs. In discussion-based settings, members implicitly negotiate both direction and magnitude at the same time, often without realising it. This protocol enforces the separation by construction, making governance choices transparent.

From the perspective of the group-shift analysis, using both steps has a clear interpretive advantage. When individual submissions are risk-normalised, group shift can be interpreted as a distortion in *directional aggregation* rather than as a change in overall risk appetite. This normalisation aligns closely with the concept of group shift in the behavioural literature, which concerns movement toward more extreme positions along a salient dimension, not arbitrary changes in variance.

Implications for Different Institutional Settings

The appropriateness of *ex ante* member-level scaling depends on the institutional context. For asset owners with a single balance sheet, equal influence across IC members is often a governance objective. In such cases, member-level scaling is natural: The committee is meant to pool views, not to allocate risk budgets across individuals.

For asset managers with multiple products and heterogeneous mandates, allowing differences in aggressiveness may be realistic. In that case, one may choose to omit member-level scaling and rely on committee-level scaling instead. However, the interpretation of aggregation outcomes then changes. Deviations from the mean reflect a mixture of belief distortion and risk heterogeneity, and group shift must be analysed in risk-adjusted terms.

Scaling individual IC submissions to a common tracking error and scaling the final portfolio to a target tracking error address different questions and serve different governance purposes. *Ex ante* scaling enforces equal influence and clarifies the interpretation of group shift as a directional bias. *Ex post* scaling reflects institutional risk budgeting and implementation constraints. Using both steps to apply member-level and committee-level scaling separates belief aggregation from risk choice and strengthens identification of governance effects. Relaxing member-level scaling is a useful robustness check, but it weakens causal interpretation by conflating aggregation distortions with differences in risk appetite.

5. INCREMENTAL FIXES TO INVESTMENT COMMITTEES

Traditional ICs are not beyond repair. On the contrary, a number of practical, low-cost interventions are commonly proposed to improve committee outcomes, including hiring higher-quality members, rotating the order of speakers to weaken anchoring, encouraging dissent, formalising agendas, and separating information presentation from decision making. These interventions are sensible, easy to implement, and better than doing nothing.

In this chapter, I evaluate these incremental remedies in the context of the governance model developed in Chapters 3 and 4. The conclusion is deliberately unsentimental: Most of these fixes operate on the *symptoms* of poor IC performance, not on its structural causes. They can reduce noise at the margin, but they do not fundamentally alter the central trade-offs that determine whether an IC is a value-enhancing governance device or a liability.

Hiring Better Committee Members

The most common response to disappointing IC outcomes is to upgrade the committee: Hire more experienced investors, add "star" members with strong track records, or import external experts to raise the average level of skill. In the language of the model, this is an attempt to increase IR_D, the baseline information ratio of individual decision makers.

From a purely economic perspective, this is unobjectionable. If aggregation works at all, then aggregating higher-quality signals should improve the information ratio of the central IC portfolio, raising p_C and therefore the expected number of winning mandates. In the mean–variance expression for firm value, this strengthens the positive term in ΔV.

The problem is that hiring stars does nothing to address the *variance* side of the trade-off. Even a committee of talented people, if organised as a traditional discussion-based IC, still produces a *single* portfolio. If that portfolio is wrong, everyone is wrong together. The variance and bankruptcy-risk terms in the objective remain unchanged. In other words, better inputs raise the mean, but they do not change the fact that centralisation synchronises outcomes.

Moreover, star-heavy committees often worsen CIO dominance rather than mitigate it. Senior, confident, well-known investors are precisely the people whose views anchor discussions. Without a mechanism that limits the influence of any single member on the final portfolio weights, hiring stars risks replacing "average noise" with "high-status noise."

Rotating Who Speaks First

Another widely recommended fix is procedural: By rotating the order in which members speak or requiring the CIO to speak last, it is possible to weaken anchoring and information cascades. This recommendation is consistent with the behavioural literature and with Scherer (2023): Early speakers disproportionately shape the perceived consensus, so changing the order should change the outcome.

In the model, rotating speakers is an attempt to reduce *occasion noise*: the dependence of outcomes on arbitrary features of the meeting, such as who happens to speak first that day. This can indeed reduce variance across meetings. But the key word is "can." Rotating anchors does not eliminate anchoring; it randomises it. The outcome is less systematically biased toward one person, but it remains path dependent.

More important, rotating speakers does not change the mapping from discussion to portfolio weights. The committee still produces one central portfolio, and that portfolio still governs all mandates. The variance term in var(Π_C) and the bankruptcy probability, q_C, are unaffected. The procedure may feel fairer and more collegial, but the economics are unchanged.

Encouraging Dissent and Devil's Advocacy

Formal dissent mechanisms—assigning a devil's advocate, asking for explicit counterarguments, or documenting minority views—are another standard fix. The hope is that structured disagreement prevents premature consensus and forces the committee to confront risks it would otherwise ignore.

Again, this helps at the margin. Dissent can surface information that would otherwise be suppressed and can reduce overconfidence in the final decision. In the model, this potentially improves the quality of the aggregated signal, raising IR_C relative to IR_D.

But dissent within a traditional IC has a built-in ceiling. Unless dissenting views translate mechanically into portfolio weights, they remain cheap talk. Once the vote is taken or the CIO summarises the discussion, dissent disappears into the minutes. The final portfolio still reflects a negotiated consensus, often tilted toward the median or the most powerful voice. From the perspective of the profit distribution, nothing fundamental has changed: Outcomes remain perfectly correlated across mandates.

Standardising Materials and Valuation Models

Another widely recommended remedy is to standardise the information that IC members receive and produce: common templates for presentations, common valuation models, common assumptions about discount rates, growth, inflation, and risk premiums. Instead of each member arriving with custom slides and pet models, everyone is asked to speak the same analytical language.

This fix targets a real problem. One important source of noise in committees is *input heterogeneity*: Members disagree not because they have different views on the world but because they are implicitly answering different questions. One member talks in multiples, another in discounted cash flows; one uses spot inflation, another long-run targets; one embeds optimism in growth, another in discount rates. The resulting discussion is not aggregation but translation.

From the perspective of the model, standardised materials can improve the quality of the individual signals by reducing unnecessary dispersion in ε_i. To the extent that common templates and valuation frameworks help members focus on economically meaningful differences rather than irrelevant modelling choices, they can raise the effective IR_D of individual inputs and, through aggregation, the IR_C of the committee portfolio. This is a genuine benefit. As with the other incremental fixes, however, the impact is limited.

First, common models do not eliminate disagreement; they merely relocate it. Members will still differ in parameter choices, scenario weights, and interpretation of uncertainty. A shared spreadsheet does not force a shared belief. If anything, it can give disagreements a false air of precision.

Second, standardisation tends to increase *correlation of errors*. When everyone uses the same model, the same data sources, and the same assumptions, model risk becomes systemic. In the language of Chapter 3, common frameworks can reduce idiosyncratic noise but increase the correlation of mistakes. This is exactly the wrong direction for weak firms, where synchronised errors dominate the governance problem.

Third, common materials do not change how decisions are translated into portfolios. Even if every IC member uses the same valuation model, the final portfolio in a traditional IC is still the outcome of discussion, negotiation, and CIO summarisation. The mapping from individual views to portfolio weights remains opaque and discretionary. The variance and bankruptcy-risk terms in the objective function are unaffected.

In short, standardised materials are best understood as *input hygiene*. They can clarify individual views and reduce accidental disagreement, but they do not solve the aggregation problem. They ensure better conversation, but they do not determine how much that conversation moves the risk budget.

In contrast, anonymous portfolio-vector aggregation is largely orthogonal to the choice of valuation model. Members may use the same model or different models. What matters is that each model produces a portfolio vector that is then risk-normalised and averaged. Disagreement is expressed directly in weights rather than debated in meetings. Standardisation can therefore be a useful complement to anonymous aggregation, but it is not a substitute for it.

Stated differently, common templates improve the *inputs*; anonymous averaging fixes the *plumbing*. Confusing the two is a reliable way to end up with a beautifully formatted slide deck and a badly designed portfolio.

Bottom Line: Anonymous Aggregation Is a Structural Fix

The anonymous portfolio-vector IC proposed in this monograph should be understood as a response to this limitation. Rather than trying to perfect the discussion, it bypasses the discussion for the purpose of portfolio construction.

Hiring better members still matters: Higher-quality inputs raise the expected quality of the average. Rotating speakers and encouraging dissent may still matter for learning and for future submissions. The decisive difference, however, is that individual views enter the final portfolio through a *mechanical* aggregation rule. Influence is proportional to risk-normalised exposure, not to status or rhetorical skill.

In terms of the model, anonymous aggregation targets the only lever that incremental fixes do not touch: It changes how individual signals map into IR_C without amplifying variance through committee dynamics. Although it does not eliminate the fundamental trade-off of centralisation, it moves the institution closer to the efficient frontier—more of the information benefit, less of the behavioural damage.

The uncomfortable implication is that many familiar IC "best practices" are not wrong but are incomplete. They improve manners, not mathematics. If the goal is to change outcomes rather than meeting quality, the governance design must constrain how decisions are translated into portfolios, not just how politely they are discussed.

6. ORGANISATIONAL BARRIERS TO AVERAGING PORTFOLIO VIEWS

Given the evidence reviewed thus far and given how trivial anonymous portfolio-vector averaging is to implement, a natural question arises: *Why do we almost never see it in practice*? If averaging independent views improves signal quality and reduces noise, why do ICs insist on discussion-heavy consensus processes instead?

The short answer is not technical; it is political. Mechanical averaging fails not because it is inefficient but because it is deeply unattractive to the people who run investment organisations.

Loss of CIO Discretion and Agenda Control

The first and most obvious reason that averaging is unattractive is that it sharply reduces the power of the CIO. In a traditional IC, the CIO controls the agenda, the framing of the problem, the sequencing of speakers, and ultimately, the summary that turns a discussion into a portfolio. Even when formal votes exist, the CIO typically sets the range of admissible outcomes.

Anonymous averaging breaks this mechanism. Once portfolios are submitted, risk-normalised, and averaged, the CIO's role is reduced to setting the risk budget and ensuring implementation. Their personal view is just one vector in the average. This is a categorical loss of influence. From a governance perspective, this is a feature. From the CIO's perspective, it is a demotion.

From Collective Responsibility to Individual Measurability

A closely related issue is accountability. Traditional ICs provide excellent cover. When a decision goes wrong, no one made it alone; the committee made it. The minutes record "robust discussion," not whose view actually drove the risk. Performance attribution becomes diffuse and largely meaningless.

Anonymous portfolio submission changes that. Each member's vector is known *ex ante* and can be evaluated *ex post*. The CIO, in particular, becomes measurable: Their submitted portfolio can be compared with the committee average and to outcomes. The comfortable ambiguity of "we decided" disappears.

This is often presented as a cultural problem ("we want teamwork, not internal competition"). In reality it is an agency problem. Organisations that claim to value accountability routinely design decision processes that limit it.

No Unifying Narrative—and That Is the Point

ICs like stories. A single macro narrative that "explains" the portfolio is psychologically appealing, easy to communicate to boards and clients, and easy to defend after the fact. Averaging destroys this.

The averaged portfolio is rarely a clean expression of one story. It is messy, sometimes internally contradictory, and often hard to justify with a single slide. Equities may be overweight because half the committee is bullish on growth and half is bearish on inflation; bonds may survive not because anyone loves them but because nobody hates them enough.

This lack of narrative coherence is often cited as a flaw. In fact, it is the entire point. Markets do not offer a single true story; they offer multiple plausible ones with uncertain weights. A portfolio that reflects this ambiguity is more honest than one that pretends otherwise. Committees resist averaging precisely because it denies them the illusion of intellectual unity.

Why Resistance Is Likely to Persist

Put together, the resistance to averaging is easy to understand:

- It reduces CIO power and agenda control.
- It makes individual views observable and accountable.
- It removes the comforting fiction of a single, coherent investment story.
- It refuses to anoint a “best” decision maker without statistical evidence.

None of these are technical objections. They are organisational ones.

In the language of this monograph, anonymous averaging improves the information aggregation term without worsening the variance term beyond what centralisation implies. Its main cost is not economic but political. It shifts power from personalities to process.

That is precisely why averaging is rarely adopted—and precisely why it deserves more attention than yet another committee charter promising “robust debate.”

7. IC SIMULATION: PROTOCOL DESIGN

Our central challenge is empirical: Real-world IC deliberations are rarely observable in sufficient detail, and researchers never observe the counterfactual in which the *same committee with the same information and beliefs* aggregates decisions under a different governance rule. In particular, it is infeasible to observe both a traditional, discussion-based IC and its algorithmic twin that mechanically aggregates the same *ex ante* views. In this chapter, I describe the simulation protocol used to study decision making in ICs.

I take a simulation-based approach using an LLM, specifically ChatGPT 5.2 in agentic mode. The LLM is not used as a source of alpha or prediction. Instead, it serves as a *controlled narrative and decision generator* that can impersonate multiple committee members with stable roles, produce structured discussions, and output explicit portfolio vectors. This approach is best understood as a substitute for experimental evidence (simulating decision making under both settings as a controlled experiment as it is undertaken in experimental economics). The aim is to isolate governance and aggregation effects while holding beliefs fixed.

Asset Universe

The asset universe is deliberately coarse, reflecting the level at which most ICs operate when discussing top-down allocation. Four broad asset classes are considered: equities (E), bonds (B), gold (G), and commodities (C). Each committee member submits a portfolio vector

$$w_s^{(j)} = (w_E, w_B, w_G, w_C),$$

with weights summing approximately to 1 and allowing for long–short positions. These portfolios are not meant to be implemented directly. They are stylised expressions of directional views and relative risk allocation.

Committee Composition

The simulated IC consists of eight members plus a chair (the CIO). Each member is assigned a stable role with associated biases that persist across scenarios. **Exhibit 5** summarises the committee composition. In the traditional IC regime, the CIO chairs the meeting and summarises the final outcome. In the anonymous regime, the CIO submits a portfolio vector like any other member, and this vector is not given any additional weight in aggregation.

Decision Regimes

The scenario distinguishes between two decision regimes that differ only in how individual views are aggregated. The comparison between these two regimes isolates the effect of discussion and governance from the effect of beliefs (**Exhibit 6**).

Exhibit 5. Committee Composition and Role Definitions

Member	Role	Characteristic Decision Bias
Dr. Rational (CIO)	Macro-oriented chair	Focuses on consistency across growth, inflation, and policy; prefers moderate tilts and portfolio coherence
Ms. Momentum	Quant PM	Overweights assets with strong recent performance; reluctant to fade trends prematurely
Mr. Value	Fundamental PM	Allocates toward cheap, unloved assets; sceptical of low-yield bonds and fashionable trades
Ms. Risk Parity	Multiasset strategist	Targets balanced risk contributions; resists concentrated directional bets
Dr. Fear	Risk manager	Emphasises drawdowns and tail risk; favours hedges in stress scenarios
Ms. Story	Thematic PM	Responds to dominant narratives (AI, energy transition, deglobalisation); willing to take concentrated thematic bets
Mr. Crowd	Behavioural analyst	Tracks sentiment and peer positioning; prefers allocations that are defensible relative to consensus
Prof. Sceptic	Contrarian	Pushes against consensus; highlights historical reversals and base-rate arguments

Standardised Output Structure

To ensure comparability across scenarios, each simulated IC meeting follows a fixed five-part structure, summarised in **Exhibit 7**.

Scenario Specification

Each scenario is defined by a qualitative macroeconomic description and its implications for the four asset classes. For the simulations, I used a fixed catalogue of 10 scenarios commonly discussed in IC settings, ranging from late-cycle soft landings to equity bubbles and sovereign crises. **Exhibit 8** lists the scenarios and their defining features.

Why an LLM Is a Realistic Tool in This Context

Using an LLM as a simulation engine is appropriate for three reasons. First, the object of interest is not to predict accuracy but to *map from beliefs to decisions* under different governance rules. Second, the LLM allows beliefs to be held fixed while only the aggregation mechanism varies, which is impossible with observational (field) data. Third, IC discussions are inherently narrative, qualitative, and role driven; modern LLMs are well suited to generating such structured narratives while remaining internally consistent.

Exhibit 6. Scenario Catalogue Used in IC Simulations

Scenario	Key Macro Characteristics	Risk Regime
Late-cycle soft landing	Slowing growth, easing inflation, cautious policy normalisation	Neutral/mildly risk-on
Overheating	Above-trend growth, rising inflation, policy behind the curve	Risk-on
Stagflation	Weak growth with persistent inflation; pressure on both equities and bonds.	Risk-off
Recession/deflation	Negative output gap, falling inflation, aggressive monetary easing.	Risk-off
Quantitative easing (QE) + fiscal profligacy	Financial repression, asset price inflation, weak real yields	Risk-on
Quantitative tightening (QT) + fiscal consolidation	Liquidity withdrawal, higher real yields, valuation pressure	Neutral/balanced
Developed market (DM) sovereign crisis	Loss of confidence in sovereign credit; banking–sovereign doom loop risk	Risk-off
Emerging market (EM) balance-of-payments (BoP) crisis	Sudden stop in capital flows, currency depreciation, flight to safety	Risk-off
AI/productivity boom	Positive supply shock, higher trend growth, strong equity performance	Risk-on
Equity bubble (narrow leadership)	Concentrated equity rally with stretched valuations and poor breadth	Late-cycle risk-on

Notes: The risk-regime classification is used to distinguish scenarios in which discussion-based committees face asymmetric reputational incentives to appear "risk-on" or "defensive." Neutral scenarios are those in which neither stance is strongly privileged by the prevailing narrative.

Exhibit 7. Standardised Output Structure of Each Simulation Run

Part	Content
1	Individual starting portfolios and brief rationales
2	Transcript of the traditional IC discussion (two to three rounds)
3	Final traditional IC portfolio compared with the equal-weight mean
4	Anonymous IC portfolio (mechanical average)
5	Diagnostic commentary on CIO dominance, group shift, and loss of diversification

Exhibit 8. Scenario Catalogue Used in IC Simulations

Scenario	Key Macro Characteristics
Late-cycle soft landing	Slowing growth, easing inflation, cautious policy normalisation
Overheating	Above-trend growth, rising inflation, policy behind the curve
Stagflation	Weak growth with persistent inflation; pressure on equities and bonds
Recession/deflation	Negative output gap, falling inflation, aggressive easing
QE + fiscal profligacy	Financial repression, asset price inflation, weak real yields
QT + fiscal consolidation	Liquidity withdrawal, higher real yields, valuation pressure
DM sovereign crisis	Loss of confidence in sovereign credit; gold as hedge
EM BoP crisis	Sudden stop, currency depreciation, flight to safety
AI/productivity boom	Positive supply shock, strong equities, rising real rates
Equity bubble (narrow leadership)	Concentrated equity rally with stretched valuations

The LLM is therefore best viewed as a stochastic laboratory that produces plausible IC behaviour under controlled conditions. It does not replace real data, but it enables analysis of institutional failure modes (e.g., group shift and CIO dominance) that are otherwise inaccessible.

The approach has clear limitations. The LLM is not a true agent-based model, and it does not endogenously optimise utility. The simulations do not generate returns or welfare measures. Incentives are imposed by design rather than by contract. These limitations are accepted because the goal is narrow—that is, to understand how governance and aggregation rules shape outcomes, not to estimate alpha.

Despite these caveats, the protocol provides a transparent and reproducible framework to visualise and quantify IC decision dynamics and to compare traditional discussion-based committees with algorithmic aggregation.

8. IC SIMULATIONS: EVIDENCE ON GROUP-SHIFT BIAS

In this chapter, I extend the governance and design arguments developed earlier by asking a deliberately narrow question: *Does a discussion-based IC exhibit group-shift bias—that is, does the final committee portfolio differ systematically from the arithmetic mean of members'* ex ante *portfolios?* The point is not that committees are irrational per se but that discussion is a *mapping* from dispersed private beliefs to a single portfolio vector. That mapping is status weighted, path dependent, and reputationally constrained. It can therefore introduce (1) *anchoring* (first-mover effects), (2) *coalition selection* (some "clusters" of views become focal), and (3) *career-risk compression* (positions that are costly to defend while the narrative persists are discounted). These mechanisms need not generate dramatic swings in a four-asset simplex, but even small, systematic shifts can matter for a total portfolio and, more important, they reveal how committees manufacture correlated bets.

Experimental Objective and Design

The claim we can test is as follows:

> Holding fixed member identities and their *ex ante* portfolios, does the traditional discussion-based IC outcome deviate from the equal-weight mean of member views, and if so, when is the deviation largest and what features of the meeting generate it?

I follow the protocol defined earlier. For each macro scenario *s*, I record (1) explicit portfolio vectors submitted independently by each committee member and (2) a *traditional* meeting transcript with a final IC portfolio. The *anonymous* committee outcome is the equal-weight mean of the same submitted portfolios (no discussion, no sequencing).

How the Traditional IC Produces a Final Vector

The simulations reflect a common institutional regularity: The final portfolio is rarely the literal average of what members privately believe. Instead, discussion produces an *acceptable set* of portfolios, after which the chair (i.e., the CIO) summarises and edits the result into a coherent final vector. In practice, the outcome can be thought of as

- members begin with $w_s^{(j)}$ (their "optimal" portfolios),
- discussion induces convergence toward a publicly defensible range (often anchored by the first speaker), and
- the final vector is a chair-approved compromise—not a Pareto optimum but a feasible political equilibrium.

This decomposition makes clear why group shift can be *directional* even when numerically modest: The committee is optimising blame, not an objective function.

Measures: Group Shift and Direction

Let $w_s^{(j)} \in \mathbb{R}^4$ denote member *j*'s starting portfolio in scenario *s* over (*E*, *B*, *G*, *C*), and define the neutral equal-weight average (anonymous IC outcome) as follows:

$$\bar{w}_s \equiv \frac{1}{M}\sum_{j=1}^{M} w_s^{(j)}. \tag{8.1}$$

Let w_s^{Trad} denote the final portfolio of the traditional IC after discussion and chair summary.

Group-shift bias is summarised by the L^1 distance from the neutral mean:

$$GS_s \equiv \| w_s^{Trad} - \bar{w}_s \|_1 = \sum_{k\in\{E,B,G,C\}} \left| w_{s,k}^{Trad} - \bar{w}_{s,k} \right|. \tag{8.2}$$

Interpretation

GS_s is the total mass of weights moved away from the arithmetic consensus. Because weights sum to 1, $GS_s \in [0, 2]$ in principle, but in a realistic four-asset policy universe, large values are mechanically capped. Therefore, what matters is not whether the shift is "dramatic" in an absolute sense but whether it is *systematic, state dependent, or aligned with reputational incentives*.

The signed drift vector is defined as follows:

$$\Delta w_s \equiv w_s^{Trad} - \bar{w}_s, \tag{8.3}$$

which identifies the direction of drift (risk-on versus defensive).

Scenario Summary and Computed Group Shift

The simulations cover eight scenario runs from the described protocol. **Exhibit 9** reports the anonymous mean, $\bar{w}_s$; the traditional IC final portfolio, w_s^{Trad}; and the corresponding drifts. **Exhibit 10** ranks scenarios by GS_s.

Data for All Calculations

All values are taken directly from the simulated scenario outputs. Computations are mechanical: Δw_s is a difference in weights, and GS_s is the sum of absolute drift components.

Finding 1: Existence but Not Uniform Magnitude

Group shift is present in every scenario ($GS_s > 0$), but it is not uniformly "large." This is unsurprising: In a four-asset universe, the feasible set of portfolios is small and the CIO typically anchors toward moderation. The relevant feature is that drift is systematic and directional, not that it is always dramatic.

Exhibit 9. Group-Shift Bias Across Simulated Scenarios (Data)

Scenario (First Speaker)	$\bar{w}_s$ (Anonymous/Mean)				w_s^{Trad} (Traditional Final)			
	E	*B*	*G*	*C*	*E*	*B*	*G*	*C*
Stagflation (CIO first)	0.20	0.24	0.29	0.27	0.18	0.22	0.31	0.29
Recession/deflation (Fear first)	0.28	0.51	0.15	0.06	0.24	0.54	0.17	0.05
QE + fiscal (Crowd first)	0.41	0.21	0.22	0.16	0.44	0.19	0.23	0.14
QT + Consolidation (Risk Parity first)	0.36	0.44	0.11	0.09	0.34	0.46	0.11	0.09
DM sovereign crisis (CIO first)	0.35	0.19	0.32	0.14	0.34	0.18	0.34	0.14
EM BoP crisis (Sceptic first)	0.30	0.38	0.23	0.09	0.29	0.37	0.25	0.09
AI/productivity boom (Momentum first)	0.49	0.28	0.08	0.15	0.55	0.22	0.06	0.17
Equity bubble (Story first)	0.45	0.31	0.13	0.11	0.50	0.28	0.12	0.10

Exhibit 10. Computed Drift Δw_s and Group-Shift Magnitude $GS_s = \sum|\Delta|$

Scenario (First Speaker)	ΔE	ΔB	ΔG	ΔC	GS_s	Rank
AI/productivity boom (Momentum first)	+0.06	−0.06	−0.02	+0.02	0.16	1
Recession/deflation (Fear first)	−0.04	+0.03	+0.02	−0.01	0.10	2
Equity bubble (Story first)	+0.05	−0.03	−0.01	−0.01	0.10	2
QE + fiscal (Crowd first)	+0.03	−0.02	+0.01	−0.02	0.08	4
Stagflation (CIO first)	−0.02	−0.02	+0.02	+0.02	0.08	4
QT + consolidation (Risk Parity first)	−0.02	+0.02	0.00	0.00	0.04	6
DM sovereign crisis (CIO first)	−0.01	−0.01	+0.02	0.00	0.04	6
EM BoP crisis (Sceptic first)	−0.01	−0.01	+0.02	0.00	0.04	6

Finding 2: State Dependence and Salience

Group shift is largest in scenarios in which (1) a salient one-dimensional narrative maps into an obvious trade and (2) resisting that trade is reputationally costly. The AI/productivity boom with Momentum speaking first produces the largest drift ($GS = 0.16$), which is consistent with benchmark and peer pressure: The committee becomes more pro-equity than its own mean view. The next largest shifts occur in the recession (Fear first) and equity bubble (Story first) regimes ($GS = 0.10$), where the asymmetry of blame is strongest ("not defensive enough" in a recession; "short the benchmark" in a bubble).

Finding 3: Direction Aligns with Career Risk

The sign pattern of Δw_s is regime consistent:

- **Risk-on and career-risk regimes** (AI boom, equity bubble) shift toward equities and away from bonds ($\Delta E > 0$, $\Delta B < 0$).
- **Risk-off and hedging regimes** (recession, crises) shift toward visible hedges, typically bonds or gold ($\Delta B > 0$ and/or $\Delta G > 0$), with $\Delta E \leq 0$.

This finding supports the view that committees do not simply "average"; they tilt toward what is easiest to defend in the current narrative environment.

Exhibits 11 and **12** summarise the results. Exhibit 11 plots GS_s across scenarios (higher bars imply more group shift). Exhibit 12 visualises the signed drift, Δw_s, by asset class, making the state-dependent direction of group shift immediately visible.

In the simulations, each scenario produces two portfolio vectors over the same asset universe: (1) the arithmetic mean of IC members' *ex ante* submissions (the anonymous outcome) and (2) the discussion-based committee outcome. The group-shift statistic, GS_s, measures how far the discussion-based outcome moved away from the arithmetic mean in the portfolio space.

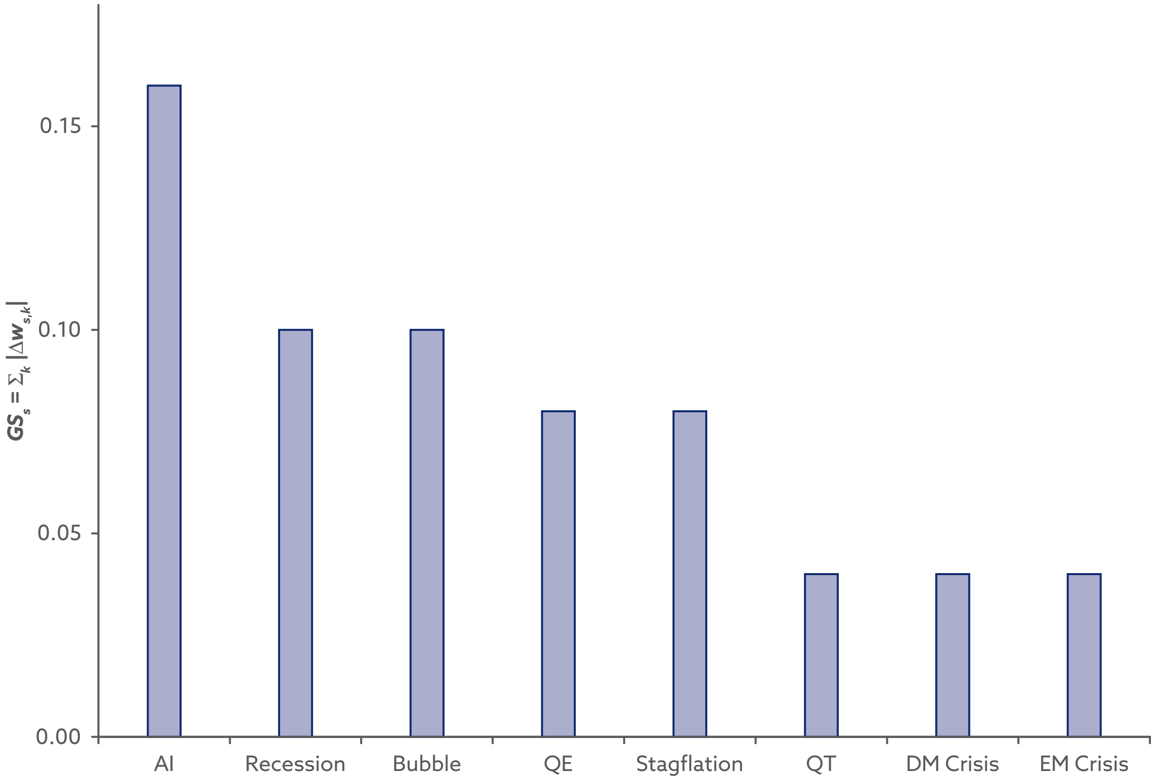
Economic Interpretation

Because both portfolios are weight vectors that (by construction) sum to 1, the drift vector, Δw_s, is self-financing: Increases in some asset weights must be funded by decreases in others. Therefore, GS_s counts the same reallocation twice—once on the buy side and once on the sell side. The economically relevant quantity is the one-way turnover implied by the meeting. Thus, $GS_s/2$ is the fraction of total capital that must be reallocated across asset classes to move from the arithmetic-mean portfolio to the discussion-based portfolio (ignoring transaction costs). If A denotes the total portfolio value, the implied dollar reallocation simply multiplies assets under management by $GS_s/2$.

Two Scenarios with Explicit Mechanism

Two showcase scenarios make this mechanism visible. In both scenarios, the *ex ante* dispersion of views is meaningful, yet the traditional IC outcome shifts toward the salient trade through a combination of first-mover anchoring and career-risk language.

Exhibit 11. Magnitude of Group-Shift Bias GS_s Across Scenarios



Note: The highest group shift occurs in the AI/productivity boom with Momentum speaking first.

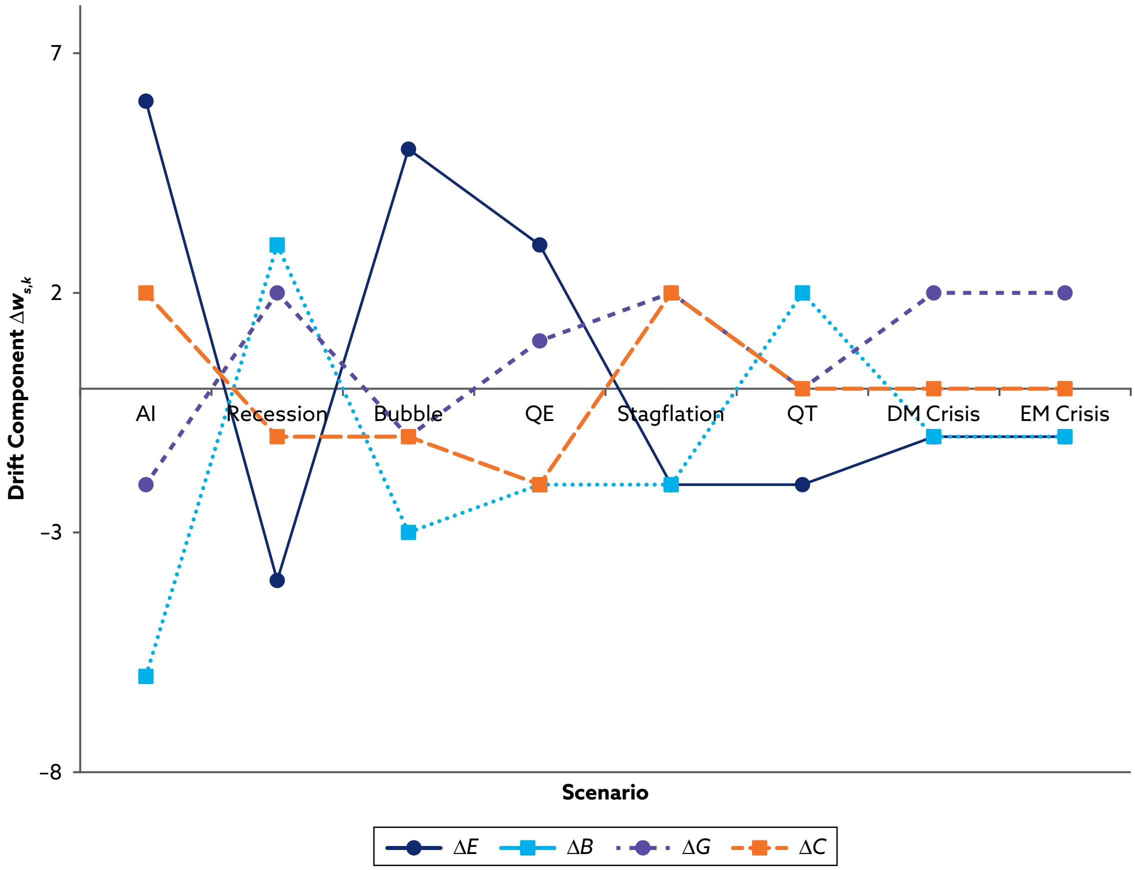
Showcase 1: AI/Productivity Boom (Momentum speaks first)

The anonymous mean is $\bar{w} = (0.49, 0.28, 0.08, 0.15)$, and the traditional IC final is w^{Trad} = (0.55, 0.22, 0.06, 0.17), implying *GS* = 0.16. The transcript points to a three-step mechanism: (1) Momentum and Story anchor an aggressive equity overweight early; (2) Crowd makes underweighting equities reputationally costly ("short the market"); (3) the CIO moderates the extremity but does not restore the arithmetic mean. The outcome is a procyclical drift: A committee that contains explicit caution ends up *more risk-on than its average member*.

Showcase 2: Recession/Deflationary Shock (Fear speaks first)

The anonymous mean is $\bar{w} = (0.28, 0.51, 0.15, 0.06)$, and the traditional IC final is w^{Trad} = (0.24, 0.54, 0.17, 0.05), implying *GS* = 0.10. The mechanism is symmetric: Fear anchors the meeting on drawdown prevention, Momentum confirms the defensive frame, and dissent is compressed because "not defensive enough" is the dominant blame risk. The committee overshoots into a more one-dimensional duration-plus-gold posture than the arithmetic mean.

Exhibit 12. Signed Drift $\Delta w_s = w_s^{Trad} - \bar{w}_s$ by Asset Class



Note: Risk-on narratives (AI, bubble) increase equities and reduce bonds; fear/hedge narratives increase bonds or gold.

What Creates (and limits) Group Shift in This Setup

Three structural feature seem to drive the group split:

- **First-mover anchoring:** The opening speaker shifts the reference point for what is reasonable, so later contributions become negotiations around an anchor rather than an aggregation of independent inputs.
- **Narrative salience (low dimensionality):** Group shift is strongest when the scenario compresses into a single obvious trade (AI → equity overweight; recession → bonds/gold). In more technocratic regimes (QT/consolidation), the narrative is less polarising and *GS* is small.
- **Asymmetric career risk:** Committees move toward positions that are cheap to justify while the regime persists: Being underweight the theme in a bubble/boom or underhedged in a recession is punished more severely than the opposite error.

At the same time, the simulations also clarify why group shift is not always "very pronounced": The asset universe is coarse, the simplex constraint limits extremes, and the CIO role typically imposes moderation. The economically important point is therefore not the absolute size of *GS* in isolation but the fact that it (1) is consistently positive, (2) is regime dependent, and (3) points in the direction of reputationally safe narratives.

Implications for the Governance Model

In the governance model developed earlier, centralisation can be beneficial only if it raises the information ratio enough to justify the increased correlation of outcomes. Group-shift bias provides a concrete mechanism for why traditional ICs may fail to realise that benefit: Discussion introduces narrative-aligned drift, effectively increasing the correlation of errors without adding information. Anonymous portfolio-vector averaging eliminates this channel by construction: It implements $\bar{w}_s$ transparently, preserves *ex ante* dispersion, and prevents the meeting from converting reputational incentives into incremental concentration.

Is It Group-Shift Bias or Just Noise?

The group-shift metric introduced in this chapter measures the deviation of the discussion-based IC outcome from the arithmetic mean of individual submissions. Formally, for scenario *s*,

$$GS_s = \sum_{k \in \{E,B,G,C\}} \left| w_{s,k}^{Trad} - \bar{w}_{s,k} \right|.$$

A natural objection is that GS_s may simply reflect *noise* rather than *bias*. I tackle that objection directly and show that under reasonable definitions, the deviations observed in the simulations cannot be interpreted as mere judgement noise. Instead, they constitute a state-contingent aggregation bias induced by the discussion-based IC process.

Why the Noise Objection Arises

In the sense of Kahneman et al. (2021), *noise* refers to undesirable variability in judgement around a benchmark: Dispersion that is random in sign, is unrelated to incentives, and averages out across repeated decisions of the same type. Because GS_s is defined as a sum of absolute deviations, it is legitimate to ask whether it simply captures random dispersion around the mean portfolio rather than a systematic distortion.

This concern would be well founded if deviations were small, random in sign, and unrelated to features of the meeting or the scenario. In that case, labelling them as "bias" would be inappropriate. Next, I will show that this is not what the simulations produce.

Noise Versus Bias: A Formal Distinction

The distinction between noise and bias can be made precise. Let

$$\Delta w_s \equiv w_s^{Trad} - \bar{w}_s$$

denote the signed drift vector in scenario *s*.

- *Noise* corresponds to Δw_s that is random in sign and magnitude, satisfies $\mathbb{E}[\Delta w_s] = 0$, and is uncorrelated with scenario characteristics or meeting structure.
- *Bias* corresponds to Δw_s that is systematically aligned with state variables (e.g., narrative salience, risk regime, speaking order) and does not average out across similar scenarios.

The L^1 metric GS_s measures the *magnitude* of the distortion, whereas Δw_s captures its *direction*. This separation mirrors standard practice in statistics, where bias is assessed by analysing both the size and sign of deviations.

Why the L^1 Metric is Not Merely Measuring Noise

The L^1 metric offers two advantages in this context. First, it has a transparent economic interpretation: It measures the total portfolio weight reallocated away from the arithmetic consensus. Second, it is invariant to sign cancellations: Offsetting deviations across assets do not mask the extent of the reallocation.

Crucially, the interpretation of GS_s does not stand alone. Throughout this monograph, it is analysed jointly with the signed drift vector, Δw_s. If GS_s reflected noise, we would observe large GS_s values with no systematic pattern in Δw_s. The simulations show the opposite: The sign pattern of Δw_s is highly structured and state dependent.

State Dependence as the Identifying Feature

The most compelling evidence against the noise interpretation is the strong state dependence of the drift. Across scenarios, the direction of Δw_s aligns closely with the dominant narrative and reputational incentives:

- In *risk-on* scenarios (AI/productivity boom, equity bubble), $\Delta E_s > 0$ and $\Delta B_s < 0$: The traditional IC allocates more to equities and less to bonds than implied by the arithmetic mean.
- In *risk-off* scenarios (recession, sovereign, and EM crises), $\Delta B_s > 0$ and/or $\Delta G_s > 0$, typically accompanied by $\Delta E_s < 0$: The IC overshoots toward visible hedges.
- In *technocratic or balanced* scenarios (QT/fiscal consolidation), both GS_s and Δw_s are small.

If deviations were noise, such a consistent sign pattern would be unlikely. Noise would generate positive and negative drifts with roughly equal frequency, independent of scenario type. Instead, the simulations produce predictable, narrative-aligned distortions.

Group-Shift Bias Versus Noise: A Regression Perspective

One might object that the magnitude measure used earlier, $GS_s = \sum_k \left|\Delta w_{s,k}\right|$, may be closer to what the judgement and decision science literature calls *noise* (random variability) than what behavioural finance would call *bias* (systematic distortion). In our setting, GS_s is best interpreted as a *process-induced deviation* from the arithmetic mean: It measures how far the discussion-based IC outcome moves away from the algorithmic average, without yet taking a stance on whether this movement is systematic (bias) or idiosyncratic (noise).

To separate these interpretations, we exploit *direction*. Noise should be approximately mean zero and unrelated to scenario characteristics; group-shift bias, by contrast, should be state contingent and directional. Risk-on narratives should push the committee toward higher equity exposure than the members' mean, whereas risk-off narratives should tilt the committee defensively (i.e., lower equity exposure than the mean).

Regression Specification

For each scenario *s*, define the signed equity drift:

$$\Delta E_s \equiv E_s^{Trad} - \bar{E}_s,$$

where E_s^{Trad} = the equity weight in the discussion-based (traditional) IC portfolio and $\bar{E}_s$ = the equal-weight average equity weight from the same *ex ante* member submissions.

Then, estimate the cross-scenario regression:

$$\Delta E_s = \alpha + \beta \cdot 1\{\text{Risk-on}_s\} + \gamma \cdot 1\{\text{Risk-off}_s\} + \varepsilon_s, \quad (8.4)$$

where the indicators are assigned *ex ante* based on the scenario label. This implementation produces the following:

- Risk-on scenarios: Overheating, AI_Boom, Equity_Bubble
- Risk-off scenarios: Recession, Stagflation, DM_Sov_Crisis, EM_BoP_Crisis
- The omitted (baseline) scenarios are those not classified as either risk-on or risk-off: SoftLanding, QE_Fiscal, QT_Consolidation.

Under the null of pure noise (no state dependence), one expects $\beta = \gamma = 0$.

Results and Interpretation

Exhibit 13 reports the results. In Model 1, the risk-on indicator is associated with a *positive* equity drift of +0.0425 (i.e., +4.25 percentage points of equity weight relative to the arithmetic mean), with a standard error of 0.0181 and a two-sided *p*-value of 0.0516. This result is meaningful given that drifts are bounded by construction and typically arise from marginal adjustments in committee compromise.

The risk-off indicator is negative (–0.0186) but imprecisely estimated ($p = 0.308$), suggesting that this small sample does not provide strong statistical evidence that fear regimes systematically reduce equity weights relative to the mean once the baseline scenarios are accounted for. The overall regression fit is high for a 10-observation design ($R^2 = 0.654$; *F*-test $p = 0.024$); however, given the limited sample size and the stylised nature of scenario classification, we should not overinterpret the result.

Model 2 provides a simple sensitivity check by excluding AI_Boom, the single scenario with the largest measured group shift in the simulation set. The risk-on coefficient falls to +0.0304 and is no longer statistically distinguishable from zero at conventional levels ($p = 0.159$).

Exhibit 13. Scenario-Level Regressions of Equity Drift $\Delta E_s = E_s^{Trad} - \bar{E}_s$

	Model 1 Baseline	Model 2 Excluding AI_Boom
Risk-on indicator	0.0425*	0.0304
	(0.0181)	(0.0189)
Risk-off indicator	−0.0186	−0.0186
	(0.0170)	(0.0158)
Constant	−0.00167	−0.00167
	(0.0128)	(0.0120)
Observations	10	9
R^2	0.654	0.555
Adjusted R^2	0.555	0.407
Residual standard error	0.0222	0.0207
F-statistic (*p*-value)	6.62(0.024)	3.75(0.088)

Notes: Ordinary least squares (OLS) estimates with an intercept. Standard errors in parentheses. Two-sided *p*-values. *denotes $p < 0.10$ (the risk-on coefficient in Model 1 has $p = 0.0516$). Risk-on scenarios: Overheating, AI_Boom, Equity_Bubble. Risk-off scenarios: Recession, Stagflation, DM_Sov_Crisis, EM_BoP_Crisis. Baseline scenarios: SoftLanding, QE_Fiscal, QT_Consolidation.

This sensitivity argues for a conservative reading: *The data are consistent with state-contingent, risk-on directional drift, but the evidence is not stable enough in a 9- or 10-observation cross-section to support strong claims.*

What this Does (and does not) Show

The regression does not prove bias in a causal sense; it is a descriptive check on whether deviations look like scenario-aligned drift rather than purely random noise. The results provide *suggestive* evidence that discussion-based aggregation can induce directional tilts (especially in risk-on regimes), consistent with the mechanisms discussed previously in this chapter (anchoring, narrative salience, and career-risk language). A stronger test would require either (1) more scenarios, (2) repeated Monte Carlo simulations per scenario, or (3) a design that randomises speaking order systematically and estimates its marginal contribution.

Why “Bias” is the Correct Economic Term

In economics, bias does not imply irrationality or error. It denotes a systematic distortion induced by a mechanism. In this assumption, the mechanism is the discussion-based IC: sequential speaking, narrative framing, CIO summarisation, and asymmetric career risk.

The resulting portfolio is not a noisy draw from the mean of individual member beliefs; it is a predictably tilted outcome that reflects which positions are easiest to defend in the prevailing narrative.

Calling these deviations "noise" would therefore be misleading. Noise would average out across comparable decisions. The deviations observed in this scenario do not. They are directional, regime dependent, and induced by the aggregation process. In that sense, group-shift bias is an accurate and conservative description.

Implications and Link to the Governance Model

Clarification of bias strengthens the governance interpretation of our model. The problem with traditional ICs is not that they introduce random error around the consensus; it is that they introduce *correlated, state-contingent distortions*. These distortions increase the effective correlation of errors under centralisation without increasing the information content of the decision.

Anonymous portfolio-vector averaging eliminates this distortion by construction. Because it implements the arithmetic mean of views mechanically, it removes the scope for narrative salience, sequencing, and reputational incentives to tilt the outcome. The resulting portfolio may still be wrong, but it is wrong in the same way as the average member's portfolio would be wrong—not in a systematically amplified direction.

9. INTERNAL AND EXTERNAL VALIDITY OF LLM-BASED IC SIMULATIONS

Using an LLM to simulate an IC naturally raises questions about both internal and external validity. These questions matter. They determine whether the results should be read as identifying a causal mechanism, as providing illustrative narratives, or as offering quantitative guidance. This chapter discusses these issues explicitly and positions the LLM-based simulations within the broader tradition of experimental and quasi-experimental analysis of decision-making institutions.

Internal Validity

Internal validity asks whether the simulation design isolates the effect it claims to measure. In the present context, the object of interest is not forecast accuracy or investment skill but, rather, the mapping from individual beliefs to a collective portfolio choice under different aggregation rules.

Internal validity is strong for three reasons. First, the treatment is cleanly defined. Across simulation runs, the set of committee members, their roles, the macro scenario, and the individual starting portfolios are held fixed. The only elements that vary are (1) the aggregation regime (discussion-based versus anonymous algorithmic averaging) and (2) where relevant, the speaking order in the traditional IC. This isolates the aggregation mechanism as the causal channel.

Second, the simulation produces an observable counterfactual that is effectively unavailable in real-world data. In practice, one never observes both the outcome of a discussion-based IC and the outcome that would have resulted had the same individuals, with the same beliefs, been aggregated mechanically. The LLM protocol delivers both objects by construction, allowing direct comparison between the traditional IC outcome and the equal-weight portfolio-vector aggregate.

Third, outcomes are measured in explicit portfolio space. Each simulation produces numerical weight vectors rather than qualitative statements or votes. This makes it possible to compute deviations, dispersion, and directional drift without interpretative ambiguity. Group-shift bias, dispersion, and coalition structure are therefore not inferred from text but instead are calculated from observable objects.

There are, nevertheless, potential threats to internal validity. LLM outputs are prompt dependent, and poorly specified prompts could mechanically induce the behaviour under study. This risk is mitigated by using stable, minimal prompts and by keeping persona definitions fixed across scenarios. A second concern is stochastic variability: Any single simulation run is an anecdote. This variability is addressed by interpreting results in terms of patterns across scenarios and by focusing on comparative statics rather than on single-point estimates. A third concern is role drift, whereby simulated personas fail to maintain consistent behaviour across scenarios. This inconsistency can be monitored directly by inspecting the distribution of submitted portfolios by role.

Subject to these caveats, the simulations have a clear internally valid interpretation: Conditional on a fixed set of individual portfolio views, discussion-based aggregation induces systematic, state-dependent deviations from the arithmetic mean of those views.

External Validity

External validity concerns whether the results generalise beyond the simulated environment. In this scenario, the limits are increasingly tighter, and they should be stated explicitly.

The simulations are not meant to replicate the decisions of any particular IC nor to deliver quantitatively accurate portfolio weights. LLM agents do not face real compensation contracts, reputational penalties, or career risk. Information is symmetric by design unless asymmetries are explicitly introduced. The asset universe is deliberately simplified, and real-world constraints, such as leverage limits, liquidity management, and regulatory considerations, are abstracted from in our simulation.

These limitations imply that the magnitude of the simulated group shift should not be interpreted literally. A 5 percentage point equity drift in the simulation does not imply that real committees will systematically produce a 5 percentage point bias. The direction of the effect and when it shows up are more likely to generalize. Risk-on narratives tilt discussion-based ICs toward risk-on assets; fear-dominated narratives tilt them defensively; and technocratic, low-salience scenarios produce little deviation from mechanical averaging. These comparative statics are plausible features of real committees, even if their quantitative expression varies across institutions.

The use of an LLM is nevertheless realistic in an important sense. ICs are language-driven institutions. They operate through narratives, slides, framing, and verbal persuasion rather than through explicit optimisation. Modern LLMs are trained on large corpora of institutional and financial language and are therefore well suited to generate plausible forms of argument, conformity pressure, and rhetorical alignment. In this respect, the LLM is not a model of cognition but a model of institutional discourse. The model reproduces the structure of how arguments are made and how consensus is socially constructed, which is precisely the mechanism under investigation.

A helpful analogy is a wind-tunnel test in engineering or a lab experiment in economics. No one claims that a wind-tunnel experiment reproduces the full complexity of flight; its value is found in isolating and visualising specific mechanisms under controlled conditions. Similarly, the LLM-based IC simulation should be read as a laboratory for governance mechanisms. It allows one to observe how sequencing, narrative salience, and aggregation rules interact, while holding beliefs fixed.

The internal validity of the LLM-based simulations is high for the question at hand: How do aggregation rules map individual portfolio views into a collective decision? External validity is necessarily limited, but the exercise is informative about mechanisms rather than magnitudes. The simulations should therefore be interpreted as evidence on plausible institutional failure modes, not as quantitative predictions of real-world portfolio choices.

10. CONCLUSION

ICs are not going away. The economically relevant question is whether they are designed to aggregate information or to provide political cover under the label of "process." In this monograph, I tackled that question in three steps. First, I developed a stripped-down governance model comparing decentralised mandate management to centralised IC decision making. The model made explicit a trade-off that institutional folklore often ignores: Centralisation can improve expected hit rates when there is genuine skill to aggregate, but it also synchronises outcomes across mandates and therefore raises the variance of profits and the probability of joint failure. For weak firms with IR ≈ 0, centralisation is not a cure; it is a way to fail together. For strong firms, centralisation can be rational but only if owners can tolerate a fatter left tail.

Second, I made the argument that even when centralisation is justified, *traditional* committees are a poor aggregation technology. Discussion is nonlinear, status weighted, and path dependent. Centralisation reliably introduces anchoring, information cascades, and free-riding incentives, and it can amplify behavioural biases, such as the disposition effect. The practical implication is that the textbook mapping from N independent alpha streams to $\text{IR}_C = \sqrt{N}\,\text{IR}_D$ is optimistic for real committees. A traditional IC is not a frictionless averaging device; it is a noisy mechanism that can destroy a meaningful fraction of whatever skill the organisation has.

This scenario leads to a simple institutional remedy: anonymous portfolio-vector aggregation. Each member submits a portfolio vector; vectors are risk-normalised to ensure equal influence; the IC portfolio is the arithmetic mean scaled to a target tracking error. This is intentionally dull. The intent is not to construct clever portfolios but, rather, to impose a transparent mapping from dispersed views to implemented positions, while constraining the scope for CIO capture and group dynamics. The rule does not create alpha. It prevents committees from amplifying confusion into a correlated macro bet.

Third—and most important for the behavioural mechanism highlighted in the monograph—I used controlled LLM simulations to test for *group-shift bias*. Because real-world counterfactuals are unobservable, the simulation protocol holds fixed the same cast of committee members and the same *ex ante* portfolio submissions and varies only the aggregation regime. Group shift is measured as the distance between the traditional IC outcome and the equal-weight mean of member portfolios. Across all simulated scenarios, this distance is positive: Discussion-based committees systematically deviate from the arithmetic consensus. Moreover, group shift is not random. It is largest when three conditions coincide: (1) a salient one-dimensional narrative maps into an obvious trade, (2) early speakers anchor the "reasonable" range of portfolios, and (3) reputational or benchmark-relative career risk is invoked. In the simulations, these conditions generate the strongest overshoots in risk-on narrative regimes (AI/productivity boom, equity bubble) and in fear regimes (recession/deflation), with the sign of the shift determined by what is socially costly to resist in that meeting. Appendix A documents each scenario so that readers can trace, line by line, how the narrative and sequencing translate into portfolio drift.

These results sharpen the governance implications. For asset owners (effectively $N = 1$), committees are nearly universal because there is no cross-sectional dispersion of mandates to preserve; only diversification across opinions is available, and process diversifies blame. For asset managers ($N > 1$), the centralisation decision is economically nontrivial because it changes the distribution of profits and the probability of ruin. In both cases, however, the committee's *aggregation rule* is decisive. Traditional meetings introduce directional drift aligned with the salient story of the day, thereby increasing the correlation of errors under centralisation. In the language of the governance model, this mechanism raises var(Π_C) and the left-tail exposure without necessarily raising $\mathbb{E}[\Pi_C]$. Anonymous portfolio averaging shuts down that channel: It preserves the diversification benefit of multiple views while preventing meeting dynamics from manufacturing additional concentration.

A few limitations are worth emphasising, not as apologies but as scope conditions. The LLM simulations are not performance tests and do not claim to predict markets. Their value is methodological: They isolate the mapping from beliefs to portfolios under alternative governance rules in a way that is difficult to observe in practice. The asset universe is deliberately coarse, and risk normalisation is treated implicitly. These simplifications are features, not bugs. They keep the focus on the institutional mechanism.

The institutional takeaway is therefore straightforward. If one insists on having an IC—and most institutions will—then the committee should be prevented from doing too much talking. Collect views in portfolio space, equalise risk, average anonymously, and use meetings for risk budgeting and process review rather than for rhetorical bargaining. Committees will still be imperfect, and averaging will not manufacture skill. This approach will reduce the most damaging failure mode documented here: group-shift bias that turns dispersed uncertainty into a single, highly correlated bet.

ACKNOWLEDGEMENTS

As a young PhD, I thought I could write about any topic. I was clearly wrong. Some papers can be written only after many years of experience; this is one of them. The topic looks deceptively simple from the outside: Put smart people in a room, ask for views, and produce a portfolio. In practice, the hard part is not the arithmetic but the institutional plumbing: who speaks first, what is allowed to be said, what gets written down, what gets quietly ignored, and what is later remembered as "consensus." The real object is not an optimisation problem but a governance mechanism, with incentives, status, career risk, and the very human urge to turn uncertainty into a defensible story.

It takes time to see the recurring failure modes that do not show up in job descriptions or committee charters: how accountability dissolves when inputs are not recorded in portfolio space, how hierarchy substitutes for evidence, how process becomes theatre, and how a committee can manufacture correlated bets precisely when it claims to diversify opinions. It also takes time to learn what tends to work in spite of these frictions—usually the unglamorous procedures that constrain discretion and force clarity.

I am grateful for a career that has allowed me to learn from a wide range of investors, colleagues, and mentors. Many of the ideas in this monograph were shaped, refined, and sometimes challenged through conversations over the years, and I have benefitted enormously from the generosity with which others shared their time and judgement. The most valuable interactions I had were with many friends and peers from the London Quant Group, the Q Group, and Inquire.

For this monograph, I am indebted to ideas and conversations from Dan diBartolemeo, Ed Fiswick, Stefan Hoops, Stephan Kessler, Mark Kritzman, Wai Lee, Sebastian Mallaby, Lionel Martellini, Jose Menchero, Michail Steliaros, Jarrod Wilcox, and Michael Wolf.

APPENDIX A. DETAILED DERIVATIONS FOR THE GOVERNANCE MODEL

This appendix spells out the algebra behind the governance model in Chapter 2. It is intended for readers who would like to see how the proposed model works. Everyone else has the author's blessing to skip it.

From Information Ratios to Hit Ratios

Let r_i be the excess return on mandate i over one year. For simplicity, suppose

$$r_i \sim \mathbb{N}(\mu, \sigma^2),$$

a normal distribution with mean μ and volatility σ. The information ratio is defined as follows:

$$\text{IR} = \frac{\mu}{\sigma}.$$

The *hit ratio*, p, is the probability that the mandate outperforms its benchmark—that is, the probability that excess return is positive:

$$p(\text{IR}) = \text{prob}(r_i > 0).$$

Standardising r_i via

$$Z = \frac{r_i - \mu}{\sigma} \sim \mathcal{N}(0,1),$$

we have

$$\text{prob}(r_i > 0) = \text{prob}\left(Z > \frac{0-\mu}{\sigma}\right) = \text{prob}(Z > -\text{IR}).$$

Because the standard normal distribution is symmetric, prob($Z > -x$) = prob($Z < x$), which equals the standard normal cumulative distribution function, $\Phi(x)$. Hence,

$$p(\text{IR}) = \text{prob}(r_i > 0) = \Phi(\text{IR}). \quad \text{(A.1)}$$

At IR = 0, we get $p(0) = \Phi(0) = 1/2$; as IR increases, so does p.

Distribution of the Number of Outperforming Mandates

There are N mandates. Let I_i be an indicator variable equal to 1 if mandate i outperforms and 0 otherwise:

$$I_i = \begin{cases} 1 \text{ if mandate } i \text{ outperforms,} \\ 0 \text{ otherwise,} \end{cases}$$

Under a given governance regime $g \in \{D, C\}$, we write p_g for the hit ratio. Then,

$$I_i \sim \text{Bernoulli}(p_g),\ \mathbb{E}[I_i] = p_g,\ \text{var}(I_i) = p_g(1 - p_g).$$

Decentralised Regime

Under decentralisation, each mandate is run independently, conditional on the common skill level. The number of winners is

$$n_D = \sum_{i=1}^{N} I_i.$$

Because I_i are independently and identically distributed Bernoulli (p_D) variables, n_D is binomial:

$$n_D \sim \text{Binomial}(N, p_D).$$

Standard results on the binomial distribution give

$$\mathbb{E}[n_D] = Np_D,\ \text{var}(n_D) = Np_D(1 - p_D). \tag{A.2}$$

Centralised Regime

Under centralisation, all mandates hold the same investment committee portfolio. The single indicator variable is as follows:

$$J = \begin{cases} 1 \text{ if the IC portfolio outperforms,} \\ 0 \text{ otherwise,} \end{cases}$$

with $J \sim \text{Bernoulli}(p_c)$. Each of the N mandates is a clone, so

$$n_c = NJ.$$

Thus,

$$n_C = \begin{cases} N \text{ with probability } p_C, \\ 0 \text{ with probability } 1 - p_C. \end{cases}$$

The mean is

$$\mathbb{E}[n_c] = N\mathbb{E}[J] = Np_c,$$

and the variance is

$$\mathrm{var}(n_c) = N^2\ \mathrm{var}(J) = N^2\ p_c(1 - p_c).$$

Summarising,

$$\mathbb{E}[n_c] = Np_c;\ \mathrm{var}(n_c) = N^2\ p_c(1 - p_c). \tag{A.3}$$

Equations A.2 and A.3 mirror Equations 2.2 and 2.3 in the main text.

Profit, Risk Aversion, and the Mean–Variance Objective

Assume fee income is proportional to the number of outperforming mandates:

$$\Pi_g = \pi_0 n_g,$$

where $\pi_0 = fA$ is fee income per winning mandate and $g \in \{D, C\}$.

Using Equations A.2 and A.3,

$$\begin{aligned}
\mathbb{E}[\Pi_D] &= \pi_0\mathbb{E}[n_D] = \pi_0 Np_D,\\
\mathrm{var}(\Pi_D) &= \pi_0^2\mathrm{var}(n_D) = \pi_0^2 Np_D(1-p_D),\\
\mathbb{E}[\Pi_C] &= \pi_0\mathbb{E}[n_C] = \pi_0 Np_C,\ \text{and}\\
\mathrm{var}(\Pi_C) &= \pi_0^2\mathrm{var}(n_C) = \pi_0^2 N^2 p_C(1-p_C).
\end{aligned}$$

We then approximate the owners' expected utility by a mean–variance functional

$$V_g \equiv \mathbb{E}[U(\Pi_g)] \approx \mathbb{E}[\Pi_g] - \frac{\gamma}{2}\mathrm{var}(\Pi_g), \tag{A.4}$$

with $\gamma > 0$ representing a coefficient of absolute risk aversion. Centralisation is preferred if and only if

$$\Delta V = V_C - V_D > 0. \tag{A.5}$$

Substituting,

$$\begin{aligned}
\Delta V &= (\mathbb{E}[\Pi_C] - \mathbb{E}[\Pi_D]) - \frac{\gamma}{2}[\mathrm{var}(\Pi_C) - \mathrm{var}(\Pi_D)]\\
&= N\pi_0(p_C - p_D) - \frac{\gamma\pi_0^2}{2}[N^2 p_C(1-p_C) - Np_D(1-p_D)].
\end{aligned} \tag{A.6}$$

Equation A.6 makes the trade-off explicit: The first term is the gain in expected fee income from a higher hit ratio; the second term is the cost of higher variance.

Skill Threshold Without Bankruptcy

From Equation A.1, we have $p_D = \Phi(\mathrm{IR}_D)$ and $p_C = \Phi(\mathrm{IR}_C)$. To link the two, assume

$$\mathrm{IR}_C = \kappa \mathrm{IR}_D;\ \kappa > 1.$$

For example, $\kappa = \sqrt{N}$ when the IC aggregates N uncorrelated managers of equal quality. Then, Equation A.6 becomes

$$N\pi_0[\Phi(\kappa \mathrm{IR}_D) - \Phi(\mathrm{IR}_D)] > \frac{\gamma\pi_0^2}{2}\{N^2\Phi(\kappa \mathrm{IR}_D)\left[1-\Phi(\kappa \mathrm{IR}_D)\right] - N\Phi(\mathrm{IR}_D)[1-\Phi(\mathrm{IR}_D)]\}. \qquad \text{(A.7)}$$

Zero Skill

At $\mathrm{IR}_D = 0$ and $\mathrm{IR}_C = 0$, we have $p_D = p_C = 1/2$. The left-hand side of Equation A.7 is then zero, and the right-hand side is

$$\frac{\gamma\pi_0^2}{2}(N^2 1/4 - N1/4) = \frac{\gamma\pi_0^2}{8}N(N-1) > 0.$$

Hence, $\Delta V < 0$ at $\mathrm{IR}_D = 0$: A zero-skill firm strictly prefers decentralisation.

Positive Skill and the Threshold

As IR_D increases, the left-hand side of Equation A.7 becomes strictly positive. By continuity, a unique $\mathrm{IR}^* > 0$ exists such that

$$\Delta V(\mathrm{IR}^*) = 0,$$

and

$$\Delta V(\mathrm{IR}_D)\begin{cases} < 0, \mathrm{IR}_D < \mathrm{IR}^*, \\ > 0, \mathrm{IR}_D > \mathrm{IR}^*. \end{cases}$$

Thus, centralisation is optimal if and only if the firm's baseline information ratio exceeds IR^*.

Comparative statics are straightforward: Increasing risk aversion γ or fee scale π_0 enlarges the variance penalty term and raises IR^*; as $\gamma \to 0$ (risk-neutral owners), the right-hand side collapses and $\mathrm{IR}^* \to 0$.

Adding Fixed Costs and Bankruptcy Risk

Real-world managers face fixed costs F (staff, systems, compliance, etc.). Let the firm fail whenever fee income falls below F. Under the simplified profit function $\Pi_g = \pi_0 n_g$, the firm survives as long as

$$\Pi_g = \pi_0 n_g \geq F.$$

Define the minimum number of winners needed to break even as

$$n^* = \left\lceil \frac{F}{\pi_0} \right\rceil.$$

The bankruptcy probability under governance regime g is

$$q_g = \text{prob}(n_g < n^*).$$

Decentralised Regime

With $n_D \sim \text{Binomial}(N, p_D)$,

$$q_D = \sum_{r=0}^{n^*-1} \binom{N}{r} p_D^r (1-p_D)^{N-r}. \tag{A.8}$$

Centralised Regime

With n_C taking values 0 and N,

$$q_C = \begin{cases} 1-p_C, 1 \le n^* \le N, \\ 1, n^* > N. \end{cases} \tag{A.9}$$

Introduce a bankruptcy cost $B > 0$, representing the value lost if the firm fails. The objective becomes

$$V_g^B \approx \mathbb{E}[\Pi_g] - \frac{\gamma}{2}\text{var}(\Pi_g) - Bq_g. \tag{A.10}$$

Centralisation is now preferred if

$$\Delta V^B \equiv V_C^B - V_D^B = \Delta V - B(q_C - q_D) > 0, \tag{A.11}$$

where ΔV is given by Equation A.6.

Equations A.8–A.11 deliver the qualitative conclusions in the main text:

- For weak firms with $p_D \approx p_C \approx 1/2$, we have $\Delta V < 0$ and $q_C - q_D > 0$, so bankruptcy risk strengthens the case against centralisation.
- For moderately strong firms, the mean–variance gain, ΔV, must offset both higher variance and $B(q_C - q_D)$; this raises the skill threshold IR*.

APPENDIX B. REFINING THE PROFIT SPECIFICATION

The analysis in Appendix A deliberately used a stripped-down profit function

$$\Pi_g = \pi_0 n_g,$$

which treated fee income as proportional to the number of outperforming mandates n_g under governance regime $g \in \{D, C\}$. This is best read as modelling the *performance-sensitive* part of the business: flows, performance fees, and the incremental value of mandates that are renewed because they outperformed. In reality, asset managers also collect a more or less stable base fee on all mandates, at least over short horizons. It is therefore natural to ask whether the centralisation results survive a more realistic fee structure.

Affine Fee Structure

Suppose that total yearly profit under governance regime g can be decomposed as

$$\Pi_g = mN + \lambda n_g, \tag{B.1}$$

where

> m is the base management fee per mandate, collected regardless of whether the mandate outperforms in that particular year
>
> N is the number of mandates (constant across regimes)
>
> λ captures the performance-sensitive component of revenues per outperforming mandate: incremental flows, performance fees, survival of the mandate into the next contract cycle, and so on

The previous specification, $\Pi_g = \pi_0 n_g$, implicitly set $m = 0$ and $\lambda = \pi_0$, focusing entirely on the performance-sensitive part. Equation B.1 nests that case but allows for a nonzero base fee.

Given the distributional results for n_g in Equations 2.2 and 2.3, the first and second moments of Π_g are

$$\begin{aligned} \mathbb{E}[\Pi_D] &= mN + \lambda N p_D, \\ \text{var}(\Pi_D) &= \lambda^2 N p_D (1 - p_D), \end{aligned} \tag{B.2}$$

$$\begin{aligned} \mathbb{E}[\Pi_C] &= mN + \lambda N p_C, \\ \text{var}(\Pi_C) &= \lambda^2 N^2 p_C (1 - p_C), \end{aligned} \tag{B.3}$$

where $p_D = \Phi(\text{IR}_D)$ and $p_C = \Phi(\text{IR}_C)$ as before.

With the same mean–variance objective as in Equation 2.4, expected utility under regime g is

$$V_g^{aff} \equiv \mathbb{E}[U(\Pi_g)] \approx \mathbb{E}[\Pi_g] - \frac{\gamma}{2}\text{var}(\Pi_g),$$

and centralisation is preferred if and only if

$$\Delta V^{aff} \equiv V_C^{aff} - V_D^{aff} > 0. \tag{B.4}$$

Using Equations B.2 and B.3, we obtain

$$\begin{aligned} \Delta V^{aff} &= (\mathbb{E}[\Pi_C] - \mathbb{E}[\Pi_D]) - \frac{\gamma}{2}[\text{Var}(\Pi_C) - \text{Var}(\Pi_D)] \\ &= \underbrace{\lambda N(p_C - p_D)}_{\text{gain in performance-sensitive fees}} - \frac{\gamma\lambda^2}{2}[N^2 p_C(1-p_C) - Np_D(1-p_D)]. \end{aligned} \tag{B.5}$$

Two observations are important:

- The base fee term, mN, cancels out of the comparison. It affects the *level* of profits but not the choice between centralised and decentralised governance.
- Equation B.5 is identical to Equation A.6 once we replace π_0 with λ. All the earlier results go through with π_0 interpreted explicitly as the performance-sensitive fee component.

In particular, for a firm with no skill ($\text{IR}_D = \text{IR}_C = 0$ so that $p_D = p_C = 1/2$), the first term in Equation B.5 vanishes and the variance term is strictly negative: Centralisation raises the volatility of profits without raising their mean. For a sufficiently skilful firm ($p_C > p_D > 1/2$), the first term is positive and eventually dominates the variance penalty as IR_D rises, generating the same skill threshold IR* as in the simpler model.

Bankruptcy Threshold with Base Fees

The bankruptcy analysis in the previous section assumed profits of the form $\Pi_g = \pi_0 n_g$ and a fixed cost F that must be covered each period. Under the more realistic affine specification, Equation B.1, the firm fails whenever

$$\Pi_g = mN + \lambda n_g < F.$$

Solving for n_g yields the break-even condition

$$n_g < n^* \equiv \left\lceil \frac{F - mN}{\lambda} \right\rceil, \tag{B.6}$$

with the convention that $n^* \leq 0$ implies zero bankruptcy probability (fixed costs are covered by base fees alone). The bankruptcy probability under governance regime g is therefore

$$q_g^{aff} \equiv \text{prob}(n_g < n^*).$$

Because n_D and n_C have the same distributions as before, the explicit formulas in Equations A.8 and A.9 continue to apply, provided we interpret n^* as in Equation B.6. The only difference is that higher base fees m shift the threshold, n^*, downward: The firm can afford fewer outperforming mandates before it hits the fixed-cost constraint.

The objective with bankruptcy costs becomes

$$V_g^{B,aff} \approx \mathbb{E}[\Pi_g] - \frac{\gamma}{2}\text{var}(\Pi_g) - Bq_g^{aff}, \tag{B.7}$$

and centralisation is optimal if

$$\Delta V^{B,aff} \equiv V_C^{B,aff} - V_D^{B,aff} = \Delta V^{aff} - B(q_C^{aff} - q_D^{aff}) > 0. \tag{B.8}$$

Qualitatively, nothing of substance changes. For weak firms with $\text{IR}_D \approx 0$ and moderate fixed costs, the base fee, mN, reduces the chance of immediate bankruptcy, but it does so under both governance regimes. The *difference* in bankruptcy risk, $q_C^{aff} - q_D^{aff}$, remains large and positive because centralisation still creates highly correlated disaster years. For strong firms with high IR_D and $\text{IR}_C > \text{IR}_D$, the gain in performance-sensitive fee income must still be large enough to offset both the variance penalty and the extra mass in the tail where all mandates underperform at once. The implicit skill threshold, IR*, above which centralisation is optimal now depends on (m, λ, F, B) only through λ and n^*; higher fixed costs F or a larger bankruptcy cost B continue to push IR* upward.

In short, allowing for realistic base fees does not rescue weak firms from the basic dilemma: Centralising investment decisions without underlying skill still means synchronizing bad outcomes. The role of base fees is to move the cliff edge, not to change the direction in which the centralised investment committee pushes the firm.

APPENDIX C. ANONYMOUS PORTFOLIO AVERAGING WITH RISK NORMALISATION

This appendix provides a careful derivation of the anonymous portfolio-vector aggregation rule. The objective is to (1) give each investment committee member equal influence *at a common risk level* and (2) produce a final portfolio with a specified tracking error (active risk) target.

Setup and Notation

Let there be K assets. Let $b \in \mathbb{R}^K$ be the benchmark weight vector, and let $w^{(j)} \in \mathbb{R}^K$ be the portfolio weight vector proposed by IC member $j \in \{1, \ldots, M\}$. The *active* weight vector is defined as

$$a^{(j)} = w^{(j)} - b. \tag{C.1}$$

Let $\Sigma \in \mathbb{R}^{K \times K}$ denote the covariance matrix of asset returns (or, equivalently, the covariance matrix used for active risk). Assume Σ is symmetric and positive semidefinite (in practice, Σ is estimated from data or obtained from a risk model). The *ex ante* tracking error (active risk) of $a^{(j)}$ is the quadratic form

$$\sigma_j \equiv \mathrm{TE}(a^{(j)}) = \sqrt{[a^{(j)}]^{\top} \Sigma a^{(j)}}. \tag{C.2}$$

Risk Normalisation of Member Portfolios

The IC design requires that each member's submitted view enters with an equal risk magnitude. Fix a common per-member tracking error level $\sigma_0 > 0$. A convenient choice is a constant (e.g., 1% tracking error) or the cross-sectional mean of the σ_j; any strictly positive σ_0 yields the same final portfolio after rescaling (shown below).

Define the risk-normalised active vector:

$$\tilde{a}^{(j)} \equiv \begin{cases} \dfrac{\sigma_0}{\sigma_j} a^{(j)} \ \text{if } \sigma_j > 0, \\ 0 \ \text{if } \sigma_j = 0. \end{cases} \tag{C.3}$$

Claim

For any j with $\sigma_j > 0$, the normalised vector $\tilde{a}^{(j)}$ has tracking error of exactly σ_0.

Proof

Using Equations C.2 and C.3,

$$\mathrm{TE}[\tilde{a}^{(j)}] = \sqrt{\left(\frac{\sigma_0}{\sigma_j} a^{(j)}\right)^{\top} \Sigma \left(\frac{\sigma_0}{\sigma_j} a^{(j)}\right)} = \frac{\sigma_0}{\sigma_j}\sqrt{[a^{(j)}]^{\top}\Sigma a^{(j)}} = \frac{\sigma_0}{\sigma_j}\sigma_j = \sigma_0.$$

Thus, each member's portfolio enters the aggregation rule at equal *ex ante* risk.

Equal-Weight Averaging of Normalised Views

The anonymous committee aggregator is the simple average of normalised active vectors:

$$\bar{a} \equiv \frac{1}{M}\sum_{j=1}^{M}\tilde{a}^{(j)}. \tag{C.4}$$

This is the formal implementation of "one member, one vote" in portfolio space. Anonymity matters for incentives and group dynamics, but it does not change the arithmetic in Equation C.4. The *ex ante* tracking error of the average active vector is

$$\sigma_{\bar{a}} \equiv \mathrm{TE}(\bar{a}) = \sqrt{\bar{a}^{\top}\Sigma\bar{a}}. \tag{C.5}$$

Rescaling to the Target Tracking Error

The asset owner specifies a desired tracking error target $\tau^{\star} > 0$ for the final IC portfolio. Define the implemented active portfolio as

$$a^{IC} \equiv \begin{cases} \dfrac{\tau^{\star}}{\sigma_{\bar{a}}}\bar{a} \text{ if } \sigma_{\bar{a}} > 0 \\ 0 \text{ if } \sigma_{\bar{a}} = 0 \end{cases} \tag{C.6}$$

and the implemented portfolio weights as

$$w^{IC} \equiv b + a^{IC}. \tag{C.7}$$

Claim

If $\sigma_{\bar{a}} > 0$, then the final IC active vector, a^{IC}, has tracking error of exactly $\tau^{\star}$.

Proof

Using Equations C.6 and C.5,

$$\mathrm{TE}(a^{IC}) = \sqrt{\left(\frac{\tau^{\star}}{\sigma_{\bar{a}}}\bar{a}\right)^{\top} \Sigma \left(\frac{\tau^{\star}}{\sigma_{\bar{a}}}\bar{a}\right)} = \frac{\tau^{\star}}{\sigma_{\bar{a}}}\sqrt{\bar{a}^{\top}\Sigma\bar{a}} = \frac{\tau^{\star}}{\sigma_{\bar{a}}}\sigma_{\bar{a}} = \tau^{\star}.$$

Invariance to the Choice of the Normalisation Level

The intermediate constant in Equation C.3, σ_0, is a convenient device to equalise member risk *before* averaging. The final portfolio does not depend on the specific value of σ_0, as long as it is strictly positive.

To see this, replace σ_0 with $c\sigma_0$ for some constant $c > 0$. Then, each $\tilde{a}^{(j)}$ in Equation C.3 is multiplied by c, so $\bar{a}$ in Equation C.4 is multiplied by c, and $\sigma_{\bar{a}}$ in Equation C.5 is also multiplied by c. In Equation C.6, the factor c cancels:

$$a^{IC} = \frac{\tau^*}{c\sigma_{\bar{a}}}(c\bar{a}) = \frac{\tau^*}{\sigma_{\bar{a}}}\bar{a}.$$

Hence, the implemented portfolio, w^{IC}, is invariant to the arbitrary choice of σ_0.

Interpretation

The full procedure can be summarised as follows:

1. Convert each proposed portfolio to an active vector $a^{(j)}$.
2. Scale each active vector to a common tracking error σ_0.
3. Average the scaled active vectors.
4. Scale the average to the target tracking error, τ^*.
5. Add back the benchmark to obtain w^{IC}.

This procedure implements equal influence across members at equal risk, avoids chief investment officer dominance by construction, and produces a portfolio with a risk level that is chosen *ex ante* by the asset owner rather than endogenously in a meeting.

APPENDIX D. HISTORICAL INVESTMENT COMMITTEE FAILURES

The governance model in Chapter 2 is intentionally abstract. It treats the firm as a bundle of mandates, a profit function, and a choice between more or less centralisation. To see why this abstraction matters in practice, it is useful to look at a few cases in which investment committees or their functional equivalents failed on a large scale. In each case, the pattern lines up closely with the mechanisms in this monograph: centralisation without sufficient skill, CIO dominance, and an underpricing of variance and bankruptcy risk.

Orange County (1994): One Trade for Everyone

In the early 1990s, Orange County, California, ran a large investment pool for the county and for local public entities. The elected treasurer effectively acted as CIO, concentrating the pool in leveraged bets on falling interest rates using structured notes and repos. When the Federal Reserve raised rates aggressively in 1994, the pool suffered losses of around USD1.5 billion–USD1.6 billion and the county filed for bankruptcy.

Formally, an oversight committee and a board were both signing off on the investment policy. Substantively, they behaved like a rubber-stamp IC: They endorsed or at least did not challenge an approach that violated the spirit of their hierarchy of "safety, liquidity, yield" in favour of carry and leverage.

In the language of this monograph:

- The county had no measurable alpha; it was a weak firm with $\mathrm{IR}_D \approx 0$. Centralising all public cash into one leveraged strategy did not raise expected skill, but it maximised the correlation of outcomes and therefore the variance of Π_g and the bankruptcy probability, q_g.
- The oversight committee implicitly behaved as if the risk-aversion parameter, γ, in the objective

$$V_g^B \approx \mathbb{E}[\Pi_g] - \gamma/2\,\mathrm{Var}(\Pi_g) - Bq_g$$

  were close to zero and as if the cost of failure, B, were negligible. In reality, B was "county bankruptcy."
- An anonymous portfolio-vector IC—forcing several independent portfolios to be written down, scaled, and averaged—would not have guaranteed prudence, but it would have made it harder for a single treasurer to run one highly leveraged macro bet for the entire pool.

Orange County provides a textbook example of what happens when a weak firm centralises aggressively and treats leverage as free.

University Endowments (2008): The Illiquidity Trap

Large US university endowments embraced what was called the Yale model—that is, heavy allocations to private equity, hedge funds, and other illiquid assets, often with substantial unfunded commitments and some explicit leverage. When the 2008 financial crisis hit, endowments suffered large drawdowns and, in several cases, severe liquidity stress. Harvard's endowment, for example, experienced losses of more than 20% and had to raise costly liquidity and cut spending.

Most endowments had a formal IC and a professional CIO. The committee set long-run policy; the CIO and internal team implemented it. In practice, the governance bargain was simple: Committees bought into the CIO's long-horizon illiquidity narrative and did not insist on a granular, quantitative treatment of liquidity and survival constraints.

In the framework of this monograph:

- Many endowments arguably had positive skill, $IR_D > 0$, and could rationally consider some degree of centralisation. The problem is that the governance choice focused almost entirely on increasing $\mathbb{E}[\Pi_g]$ and smoothing the reported volatility of the funding ratio, while underweighting both $\text{var}(\Pi_g)$ and the probability, q_g, of being forced into disruptive budget cuts or distressed asset sales.
- An anonymous vector-style IC, which forces each member to write down and own an explicit portfolio, including liquidity buffers, would not have eliminated the appetite for illiquid assets. It would, however, have increased the chance that more conservative views on liquidity and leverage could have influenced the final allocation.

This endowment experience thus illustrates the case of an informed investor that may justifiably centralise. The IC, however, collectively overlooked the implicit liabilities (cash flow needs). The IC design in this case can do little other than explain the endowment's economic exposure more clearly.

J.P. Morgan's Chief Investment Office (2012): The "London Whale"

In 2012, J.P. Morgan's Chief Investment Office accumulated a large synthetic credit portfolio using credit default swap indices. The positions, associated with the trader dubbed the "London Whale," generated losses in excess of USD6 billion when markets moved against the book. Internal and external reviews of the incident pointed to serious weaknesses in risk governance: rare and superficial risk-committee meetings, model changes that masked risk, and incomplete information flowing up to senior management and the board.

In this case, the Chief Investment Office unit functioned as an internal IC with its own risk committee and oversight channels. On paper, governance looked robust; in practice, the unit had both the mandate and the stature to drive a concentrated, poorly understood risk position with insufficient challenge.

Viewed through the lens of this monograph:

- The unit's book was a fully centralised portfolio: Either the synthetic credit hedge "worked" for the bank's balance sheet or it did not. That structure offered essentially no diversification across independent views. The relevant R_C in Chapter 3 took the values {0, *N*}.
- The supposed hedging skill behind the positions was poorly documented and hard to verify. The bank behaved as if the firm were in the "strong" region, in which case IR_C is genuinely higher than IR_D, but governance and *ex post* losses suggest otherwise.
- Committee structures existed but behaved as classic discussion-based ICs: late, infrequent, and heavily anchored on the Chief Investment Office narrative. They did not behave like independent providers of portfolio vectors that could be benchmarked and averaged.

This example is less about the absence of committee process and more about its capture: centralisation plus Chief Investment Office dominance without the counterweight of independent, quantifiable contributions.

UK Pension LDI Crisis (2022): Leveraged Herding

In 2022, a sharp rise in UK government bond yields following a fiscal policy announcement triggered a collateral and liquidity crisis for many defined benefit pension schemes that had implemented leveraged liability-driven investment (LDI) strategies. To meet margin calls on their interest rate derivatives, funds sold long-dated gilts into a falling market, amplifying the sell-off and prompting central bank intervention.

Most schemes had trustee boards and ICs, often advised by consultants and LDI managers. The committees embraced a single risk management narrative: hedging liabilities through derivatives to stabilise measured funding ratios. Leverage and collateral liquidity risk were underappreciated at the IC level; many trustees did not fully understand the mechanics of collateral waterfalls and stress scenarios.

In terms of the model in Chapter 2:

- At the system level, funds migrated from a variety of bond and cash allocations toward a highly correlated LDI structure. This raised the cross-sectional correlation of outcomes and made system-wide stress more likely. It is, in effect, an increase in the degree of centralisation without a clear increase in underlying IR.
- The governance focus was on reducing the volatility of reported funding ratios, not on the survival condition or avoiding permanent losses. This was a reoccurring mistake layed open by the events.
- Committees outsourced much of the portfolio design to external CIO equivalents (consultants and LDI managers). There was little understanding which independent internal views on leverage and liquidity were elicited, compared, and averaged.

This LDI episode is a reminder that it is insufficient for committees to adopt what looks like a sophisticated consensus. If everyone adopts the same structure, the system behaves like one highly centralised IC portfolio.

Lessons for IC Design

From these cases, three common themes emerge that reinforce this monograph's main arguments:

- **Centralisation without sufficient skill is dangerous.** Weak or mediocre firms that centralise end up with the same expected outcome as they would under decentralisation, but they do so with a much higher variance and a higher probability of ruin. Orange County and many LDI users fit this pattern.
- **CIO dominance and narrative anchoring are structural risks.** Once a single story (illiquidity premia, synthetic hedging, liability matching) dominates the committee, information cascades, group polarisation, and free riding kick in. The committee stops being a diversified provider of independent views and becomes a rhetorical support act for the CIO.
- **Explicitly pricing variance and bankruptcy would change decisions.** Had these institutions formulated their governance choices in terms of V_g^B as in Chapter 3, with nonzero g and B, many of the decisions that now look like obvious mistakes would have been less palatable *ex ante*.

The anonymous portfolio-vector IC is not a cure-all; it does not conjure skill where none exists. It does, however, directly target the failure modes on display in these episodes by forcing independent portfolio construction, reducing the scope for group shift, and capping the influence of any single CIO narrative on the implemented portfolio.

REFERENCES

Bär, M., A. Kempf, and S. Ruenzi. 2011. "Is a Team Different from the Sum of Its Parts? Evidence from Mutual Fund Managers." *Review of Finance* 15 (2): 359–96. doi:10.1093/rof/rfq014.

Bliss, R. T., M. E. Potter, and C. Schwarz. 2008. "Performance Characteristics of Individually Versus Team Managed Funds." *Journal of Portfolio Management* 34 (3): 110–19. doi:10.3905/jpm.2008.706248.

Cici, G. 2012. "The Prevalence of the Disposition Effect in Mutual Fund Trades." *Journal of Financial and Quantitative Analysis* 47 (4): 795–820. doi:10.1017/S0022109012000348.

Cooper, D. J., and J. H. Kagel. 2005. "Are Two Heads Better Than One? Team Versus Individual Play in Signalling Games." *American Economic Review* 95 (3): 477–509. doi:10.1257/0002828054201431.

Davis, J. H., and V. B. Hinsz. 1982. "Current Research Problems in Group Performance and Group Dynamics." In *Group Decision-Making*, edited by H. Brandstätter, J. H. Davis, and G. Stöcker-Kreichgauer. London: Academic Press.

diBartolomeo, D. 2018. "Centralized Portfolio Management (A Twenty Six Year Tale of the Most Expensive Wine We've Ever Served)." *Northfield News* (June): 1, 5–6.

Grinold, R. C., and R. N. Kahn. 2000. *Active Portfolio Management*. 2nd ed. New York: McGraw-Hill.

Kahneman, D., O. Sibony, and C. R. Sunstein. 2021. *Noise: A Flaw in Human Judgment*. London: William Collins.

Lee, W. 2000. *Theory and Methodology of Tactical Asset Allocation*. Chichester, UK: Wiley.

Lovallo, D., and A. Sibony. 2010. "The Case for Behavioral Strategy." *McKinsey Quarterly* (1 March).

Mulvey, P. W., and H. J. Klein. 1998. "The Impact of Perceived Loafing and Collective Efficacy on Group Goal Processes and Group Performance." *Organizational Behavior and Human Decision Processes* 74 (1): 62–87. doi:10.1006/obhd.1998.2753.

Odean, T. 1998. "Are Investors Reluctant to Realize Their Losses?" *Journal of Finance* 53 (5): 1775–98. doi:10.1111/0022-1082.00072.

Rau, H. A. 2015. "The Disposition Effect in Team Investment Decisions: Experimental Evidence." CEGE Discussion Paper No. 256, University of Göttingen.

Scherer, B. 2022. "Portfolio Factory." *Journal of Portfolio Management* 48 (4): 7–13. doi:10.3905/jpm.2022.1.364.

Scherer, B. 2023. "Optimal Design of Investment Committees." *Journal of Asset Management* 25: 129–35. doi:10.1057/s41260-023-00330-3.

Sharpe, W. F. 1981. "Decentralized Investment Management." *Journal of Finance* 36 (2): 217–34. doi:10.1111/j.1540-6261.1981.tb00436.x.

Tetlock, P. E., and D. Gardner. 2015. *Superforecasting: The Art and Science of Prediction*. New York: Crown.

Named Endowments

CFA Institute Research Foundation acknowledges with sincere gratitude the generous contributions of the Named Endowment participants listed below.

Gifts of at least US$100,000 qualify donors for membership in the Named Endowment category, which recognizes in perpetuity the commitment toward unbiased, practitioner-oriented, relevant research that these firms and individuals have expressed through their generous support of CFA Institute Research Foundation.

Ameritech
Anonymous
Theodore R. Aronson, CFA
Asahi Mutual Life Insurance Company
Batterymarch Financial Management
Boston Company
Boston Partners Asset Management, L.P.
Gary P. Brinson, CFA
Brinson Partners, Inc.
Capital Group International, Inc.
Concord Capital Management
Dai-Ichi Life Insurance Company
Daiwa Securities
Mr. and Mrs. Jeffrey Diermeier
Gifford Fong Associates
John A. Gunn, CFA
Investment Counsel Association of America, Inc.
Jacobs Levy Equity Management
Jon L. Hagler Foundation
Long-Term Credit Bank of Japan, Ltd.
Lynch, Jones & Ryan, LLC
Meiji Mutual Life Insurance Company
Miller Anderson & Sherrerd, LLP
John B. Neff, CFA
Nikko Securities Co., Ltd.
Nippon Life Insurance Company of Japan
Nomura Securities Co., Ltd.
Payden & Rygel
Provident National Bank
Frank K. Reilly, CFA
Salomon Brothers
Sassoon Holdings Pte. Ltd.
Scudder Stevens & Clark
Security Analysts Association of Japan
Shaw Data Securities, Inc.
Sit Investment Associates, Inc. Standish, Ayer & Wood, Inc.
State Farm Insurance Company
Sumitomo Life America, Inc.
T. Rowe Price Associates, Inc. Templeton Investment Counsel Inc. Frank Trainer, CFA
Travelers Insurance Co.
USF&G Companies
Yamaichi Securities Co., Ltd.

Senior Research Fellows

Financial Services Analyst Association

For more on upcoming CFA Institute Research Foundation publications and webcasts, please visit www.cfainstitute.org/research/foundation.

Available online at
rpc.cfainstitute.org

ISBN 978-1-952927-78-2