

Redesigning the NAV Reasonableness Check for an Asset Management GCC

AI-Assisted Exception Management - Proof of Concept on the iShares Core S&P 500 ETF (IVV)

GCC Project Group-4 | Akshaj Arora · Aviral Joshi · Gayatri Dande · Hrithik Tripathi · Rishab Dave |

Academic Mentor: Prof. Arnab Ghosh

1. Project Overview

A key control activity within fund accounting teams is the **NAV reasonableness check**, which serves as the final validation point before a fund's Net Asset Value (NAV) is published. As asset management operations increasingly scale through Global Capability Centres (GCCs) in India, leading financial institutions are adopting more **standardized and robust control frameworks** to ensure the accuracy, consistency, and timely publication of NAVs across funds and business days.

Traditionally, the NAV reasonableness check relies heavily on an analyst's manual assessment of whether the day's NAV movement is reasonable given prevailing market movements and portfolio activity. While this approach provides an important layer of oversight, it can be **subjective, time-sensitive, and difficult to scale** as the number of funds and complexity of portfolios increase. This creates an opportunity to strengthen the control by making the assessment more **standardized, data-driven, and scalable**, while retaining appropriate analyst oversight.

This project asks a narrow, testable question: can that check be rebuilt as a statistically grounded model, and can AI assist an analyst at the point a flag fires - without ever letting AI make the publication decision itself. Using IVV (a large, liquid, single-benchmark ETF) as the test case, a factor model predicts each day's expected NAV return from the S&P 500 Total Return Index net of expense drag, flags days where actual NAV deviates beyond a calibrated threshold, and routes every flag through four progressively more automated layers of AI assistance - from drafted commentary to an unattended, scheduled exception pipeline.

Key assumptions: This proof of concept holds several factors fixed to isolate whether the core logic works before testing it under harder conditions:

- A single, liquid, single-benchmark ETF (IVV), rather than a multi-asset or actively managed fund
- A constant expense ratio over the backtest window, rather than one that varies with fee changes or fund actions
- No corporate-action noise (splits, special dividends) requiring separate handling in the model
- A 353-day backtest window that does not include a severe market dislocation

These are scoping choices made to test the architecture cleanly, not limitations discovered after the fact. Section 6 addresses what changes when each is relaxed.

353	5	0%	4
-----	---	----	---

trading days back-tested	flags raised, all explained	unexplained-flag / false-positive rate	layers of AI assistance built
--------------------------	-----------------------------	--	-------------------------------

2. GCC Context: Structure, Functions, and Compliance Governance

What is a GCC

A Global Capability Centre (GCC) is an offshore or nearshore operations hub that a global institution – in our example, an asset management firm - operates and staffed to run core middle- and back-office functions on its parent's behalf, at scale and under the parent's own control and risk framework (as distinct from a third-party BPO relationship). Firms such as State Street, BNY Mellon, BlackRock, Northern Trust, and Franklin Templeton run large fund-operations GCCs out of India for exactly this reason.

Core functions typically housed in an asset-management GCC

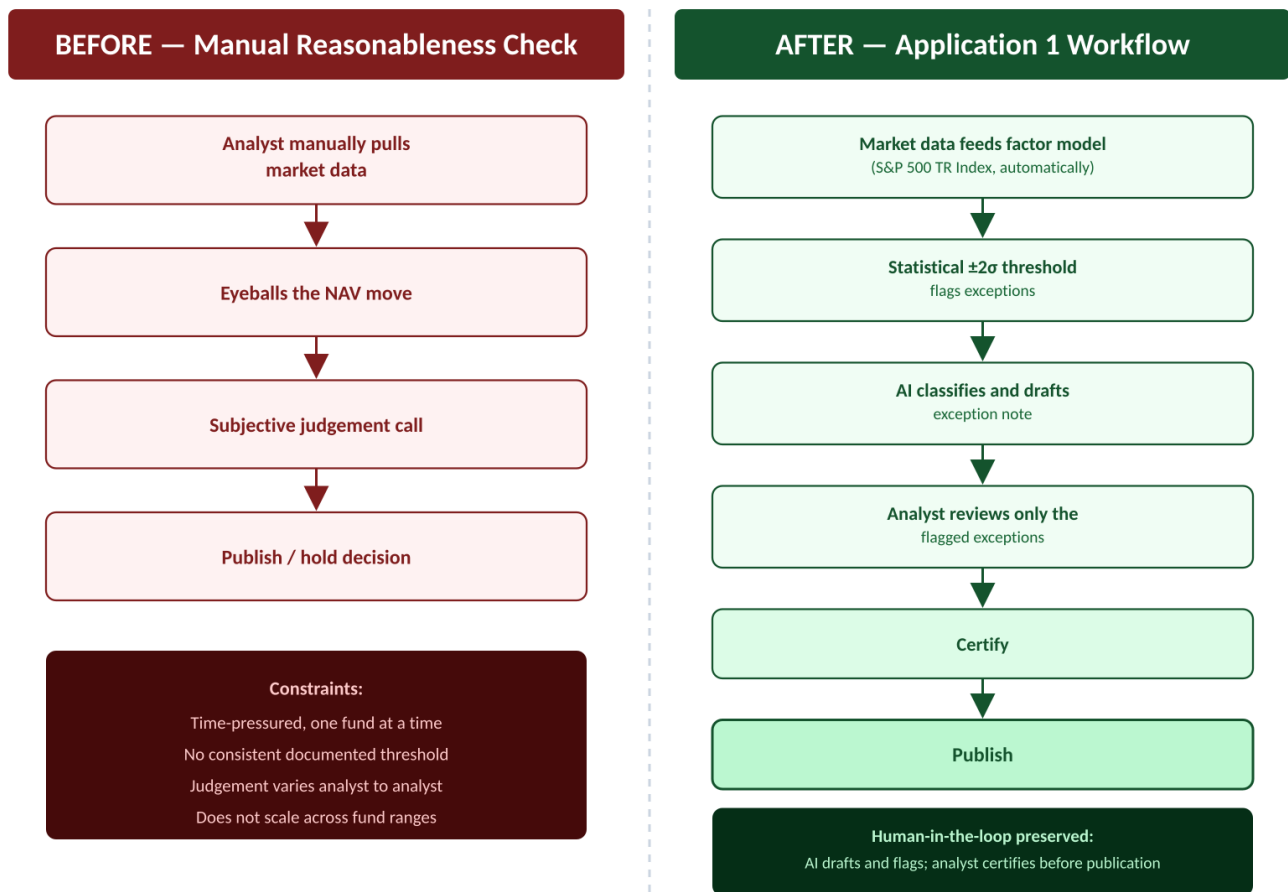
- Fund Accounting & NAV Oversight - striking, checking, and certifying daily NAV; the function this project addresses.
- Investment Operations - trade capture, settlement, reconciliation, corporate actions processing.
- Reporting & Client Servicing - regulatory filings, client and board reporting packs.
- Risk & Compliance - control monitoring, exception governance, regulatory change management.
- Technology & Data - the platforms, pipelines, and (increasingly) the AI/automation layer supporting the above.

This project sits squarely inside the first of those functions - Fund Accounting & NAV Oversight, which is precisely why a GCC, rather than the front office or a third-party vendor, is the natural owner of a controls like this. A GCC already runs NAV oversight across a large and varied fund range, functions as the part of three-lines-of-defence structure described below, and is measured on consistency and control quality across that range rather than on any single fund. A statistically calibrated check with an AI-assisted escalation layer is only worth building if it can be re-calibrated and reused across many funds without rebuilding the workflow each time and a GCC's fund count, repetition, and existing governance structure is exactly the setting where that reuse pays off.

How the compliance/governance operating model runs

Compliance in a GCC is not a single team but an operating model layered across three lines: the first line (the analyst performing the check) owns the control at the point of execution; the second line (Compliance/Risk) sets the policy, thresholds, and escalation rules the control must satisfy and monitors adherence; the third line (Internal Audit) independently tests that the control operated as designed. Every control - including this reasonableness check - must have a named owner, a documented escalation path, and an auditable record of who reviewed and signed off. Regulatory anchors relevant here are SEC Rule 2a-4 (US NAV valuation) and the IOSCO Principles for the Valuation of Collective Investment Schemes, both of which require identifiable human accountability rather than full delegation to an automated system - which is precisely the design boundary this project holds to (Section 5).

3. Process Redesign: Before vs. After



Source: White Paper Section 2 (problem statement) and Section 5.4 (Application 1 / Level 4 workflow)

Dimension	Before (manual check)	After (AI-assisted check)
Basis for decision	Analyst judgement, day-to-day	Statistically calibrated threshold ($\pm 2\sigma$)
Analyst effort	Full review, every fund, every day	Reviews only the ~1.4% of days that flag
Consistency	Varies by analyst and day	Same rule applied identically every day
Documentation	Ad hoc commentary, if any	AI-drafted, consistently formatted exception note
Detection of anomalies	Depends on analyst catching it	Back-tested to catch a 0.5% synthetic pricing error immediately

Dimension	Before (manual check)	After (AI-assisted check)
Escalation	Manual, when someone notices	Automatic email escalation for unexplained flags (Level 4)

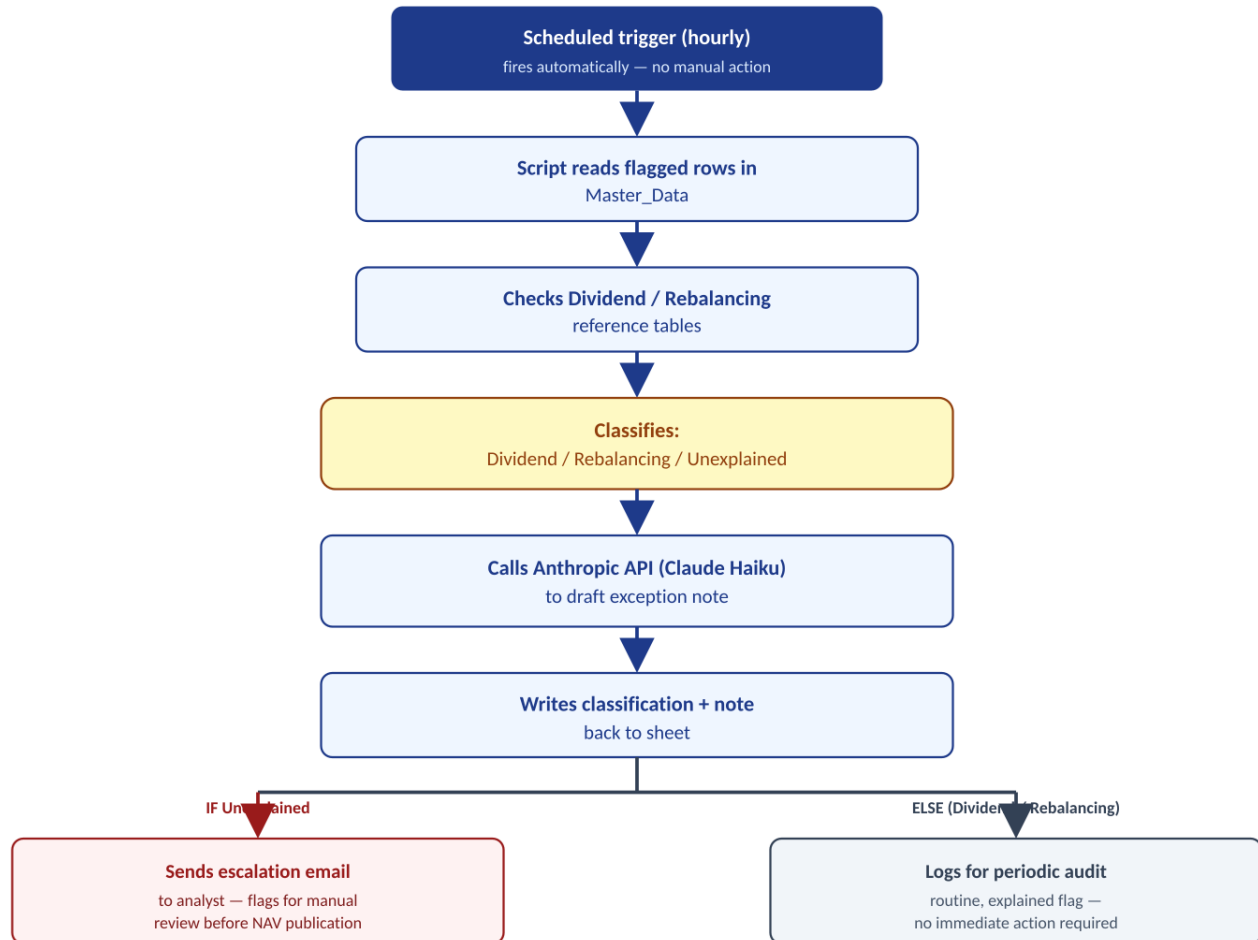
4. Technology Layers, Including the Role of Agentic AI

Four layers of AI assistance were built, each strictly more automated than the last. The first three are AI-assisted: a person opens a tool and prompts it. Only the fourth is agentic in the full sense - it triggers on a schedule, branches its own classification decision from the data, and takes an action (an escalation email), with no person initiating the run.

Layer	What it is	AI's role	What makes it 'agentic' (or not)
1.Exception Commentary	Prompted LLM (Claude)	Drafts a structured exception note for a flag whose classification is supplied by the analyst	Not agentic - no trigger, no branching, human supplies the answer
2.Regression Model	Statistical Regression (OLS, Excel Analysis Tool-Pak)	Estimates the NAV/index relationship ($R^2 = 0.9988$); the fitted residuals define the $\pm 2\sigma$ flagging threshold.	Not agentic - a statistical model, not a workflow
3. Live AI Tool	Browser tool + Claude API	Runs the check in real time and auto-drafts a note the moment a flag fires	Semi-automated - still opened and run by a person
4. Agentic Pipeline	Google Sheets + Apps Script + Anthropic API, hourly trigger	Classifies each flag (Dividend / Rebalancing / Unexplained) by checking reference data itself, drafts the note, writes back to the sheet, and emails an escalation only when Unexplained	Fully agentic - unattended trigger, data-driven branch, automatic downstream action

Level 4 — Agentic Workflow Architecture

Google Sheets + Apps Script + Anthropic API (Claude Haiku)



Human-in-the-loop governance

The workflow never publishes a NAV itself. It classifies, drafts, writes back, and routes — an analyst reviews and certifies every escalated exception.

Evidence captured for submission

- Triggers page screenshot — hourly schedule active (proves genuine automation)
- Apps Script Execution log — shows the run firing on its own schedule
- Before/after screenshot of the sheet — three new columns populated
- Escalation email (Unexplained case) alongside a routine note (Dividend case)
- Synthetic error test — Day-3 injected error correctly triggers escalation branch

Source: White Paper Section 5.4, Level 4 — Agentic Workflow

5. The Reviewer and the Human-in-the-Loop

The design principle held constant across all four layers is that AI produces an output for a human to review; it never authorises NAV publication itself. This is a deliberate governance boundary, not a technical limitation of the tools used - nothing in the Anthropic API or Apps Script prevents fully autonomous action; the boundary is enforced by design because SEC Rule 2a-4 and IOSCO's valuation principles require identifiable human accountability.

A second constraint this design has to hold alongside human accountability is data confidentiality. Every AI-assisted layer in this project sends NAV, benchmark, and expense-ratio figures to an external LLM provider (the Anthropic API) to draft commentary or classify a flag. For this proof of concept, the data involved is limited to a single public ETF's published NAV and a public benchmark index - nothing proprietary or client-identifying. A production rollout across a GCC's actual fund range would need this addressed explicitly before adoption: what data leaves the GCC's environment, where it is processed and retained, and whether that is compatible with the client confidentiality and data-residency obligations a GCC operates under. This project does not resolve that question - it flags it as a precondition for scaling beyond a public single-fund pilot, not an afterthought to be solved post-launch.

Control point	AI's role	Human reviewer's role
Level 1	Drafts commentary from a known classification	Reads, edits if needed, certifies before it enters the record
Level 2	Produces a statistically fitted alternative threshold	Decides whether to adopt, override, or keep the original threshold
Level 3	Calculates the check live and displays pass/flag with a note	Views the result and makes the publish/hold decision manually
Level 4	Runs unattended; classifies, drafts, and escalates Unexplained flags only	Receives the escalation email, investigates, and certifies; routine flags are still logged for periodic audit even with no email sent

A specific point worth making explicit to a reviewer: Level 4 is the only layer that takes an automatic action without prior human sign-off, and that action is deliberately limited to notifying a human - it does not withhold, approve, or alter any figure in the underlying NAV data. Automating the escalation trigger, not the publication decision, is the boundary this project treats as non-negotiable.

6. GCC Operating Model

The NAV reasonableness check is particularly well suited to an Asset Management Global Capability Centre (GCC) because the control combines high-volume, repeatable operational activity with clearly defined governance and escalation requirements. In a typical GCC operating model, the **GCC performs the standardized execution and control activities**, while the **home location retains ownership of the investment, valuation policy, risk appetite, and final governance decisions**. This creates a model in which technology and analytics can improve the efficiency of NAV oversight without transferring accountability away from the asset manager.

What the GCC Performs

Within the GCC, the fund accounting and NAV oversight team can own the day-to-day execution of the reasonableness check across the fund population. This includes:

- Running the NAV reasonableness check for each fund on every business day.
- Applying fund-specific statistical parameters, benchmarks, and thresholds.
- Identifying and classifying exceptions arising from market movements, dividends, rebalancing, or other NAV-moving events.
- Maintaining the exception log and supporting documentation.
- Using AI-assisted tools to draft standardized exception commentary.
- Escalating unexplained exceptions to the appropriate reviewer or control owner.
- Maintaining an auditable record of checks, exceptions, reviewer actions, and outcomes.
- Monitoring control performance and identifying opportunities for further automation or recalibration.

The proposed solution therefore allows the GCC to move from a predominantly manual, analyst-by-analyst assessment toward a standardized control framework that can be consistently applied across a growing fund population

What Remains with the Home Location

The GCC does not replace the asset manager's core ownership or decision-making responsibilities. The home location remains responsible for the broader investment and governance framework, including:

- Ownership of valuation policies and control standards.
- Definition of the overall risk and governance framework.
- Approval of material changes to NAV methodology and control design.
- Oversight of significant or unresolved valuation exceptions.
- Final accountability for the accuracy and publication of the NAV.
- Regulatory and fiduciary accountability for the funds.

This distinction is particularly important for an AI-assisted control. The proposed architecture deliberately keeps **human accountability at the point of publication**: AI can identify, classify, document, and escalate exceptions, but it does not approve, alter, or publish the NAV. This is consistent with the project's human-in-the-loop design principle.

7. Conclusion

For a single, clean, low-turnover index fund like IVV, a relatively simple and low-turnover index ETF, the statistically calibrated reasonableness checks successfully identified **all genuine NAV-moving events** over the 353-day back-test, with **zero unexplained flags and no false positives**. The model also detected a synthetically introduced **0.5% pricing error on its first occurrence**, demonstrating its ability to identify a controlled anomaly. However, these results should be interpreted within the defined scope of the proof of concept. The findings validate the approach for a **single-benchmark, single-factor fund**, but do not yet establish its effectiveness for multi-asset, actively managed, or multi-currency funds. In addition, the back-test period did not include a severe market dislocation, and therefore the model's performance under stressed market conditions remains to be tested.

Implications

- Operationally: a statistically thresholded check reduces analyst review from every fund, every day, to the roughly 1.4% of days that actually flag - without removing the analyst from the publication decision.
- Methodologically: this proof of concept validates the core logic - factor model, threshold, AI-assisted drafting, agentic escalation - for the simplest case in scope. It does not yet validate multi-factor models, multi-currency funds, or performance under a severe market dislocation; those are the next steps, not open questions left unaddressed.
- For the GCC: a reusable, re-parameterizable check like this is what lets NAV oversight scale across a growing fund range without scaling headcount linearly alongside it - which is the case this project was built to make.

That combination - real automation, honestly scoped, with the governance boundary stated rather than assumed - is the standard we tried to hold this project to, and the one we'd want to be judged against.

Way Forward

The proof of concept demonstrates that the NAV reasonableness check can evolve from a manual control into a **standardized, data-driven, and scalable capability within an Asset Management GCC**. The next step is to extend the approach beyond IVV by applying the same control architecture across multiple funds and fund families, with parameters recalibrated to reflect each fund's characteristics. For more complex funds, the statistical layer can be expanded from a single-factor model to a multi-factor framework incorporating relevant benchmarks, sector, currency, and style exposures.

The GCC can serve as the central execution hub, performing daily checks, managing exceptions, maintaining control evidence, and monitoring model performance, while the home location retains policy ownership and final accountability. A structured model-governance framework would enable periodic monitoring and recalibration of thresholds as fund characteristics and market conditions evolve. AI and automation can then be progressively scaled from commentary generation and exception classification to automated escalation, while keeping the final NAV publication decision with a human reviewer.

Before production deployment, the framework would need to be tested across different fund types and market conditions, including severe dislocations, corporate actions, and data-confidentiality requirements. Ultimately, the



opportunity is to build a **reusable GCC-wide NAV control capability** that improves consistency, auditability, and operational efficiency while allowing NAV oversight to scale without increasing manual effort and headcount linearly with fund growth.



Generalising Beyond a Single-Factor Fund

The architecture built here does not depend on IVV's simplicity - only the model that feeds it does. For a single-benchmark equity ETF, one factor (the S&P 500 Total Return Index, net of expense drag) is sufficient to explain daily NAV movement. A multi-asset or actively managed fund would require a multi-factor regression - additional benchmarks, sector or currency exposures, style factors - in place of the single-factor model, and a correspondingly wider or asset-class-specific threshold in place of a flat $\pm 2\sigma$ band. Everything downstream of that model is unchanged: flags still route through the same four layers, from AI-drafted commentary to the agentic escalation pipeline, and the analyst still certifies every publication decision. In other words, added fund complexity raises the statistical difficulty of Level 2, not the operational design of Levels 1, 3, and 4. This is precisely why a GCC is the right place to build and own this capability: it already runs NAV oversight across a large, varied fund range, and a single reasonableness-check workflow - re-parameterised per fund or fund family rather than rebuilt - is what lets that oversight scale without adding headcount linearly to fund count.