

Gondauri, Davit

Working Paper

Robust Portfolio Optimization under Computational Complexity: A P-vs-NP-Inspired Markowitz-CAPM Framework with Cardinality Constraints and a Black-Scholes Derivative-Pricing Overlay

Suggested Citation: Gondauri, Davit (2026) : Robust Portfolio Optimization under Computational Complexity: A P-vs-NP-Inspired Markowitz-CAPM Framework with Cardinality Constraints and a Black-Scholes Derivative-Pricing Overlay, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/341673>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Robust Portfolio Optimization under Computational Complexity: A P-vs-NP-Inspired Markowitz-CAPM Framework with Cardinality Constraints and a Black-Scholes Derivative-Pricing Overlay

Davit Gondauri

ORCID: <https://orcid.org/0000-0002-9611-3688> <https://orcid.org/0000-0002-9611-3688>

PhD, Professor, Business & Technology University, Georgia

Corresponding author: Davit Gondauri, Dgondauri@gmail.com

Abstract

This study develops a robust computational-finance framework for cardinality-constrained portfolio optimization by integrating Markowitz mean-variance logic, CAPM expected-return calibration, a P-vs-NP-inspired support-selection layer, stochastic and metaheuristic search, exact reduced benchmarking, robustness diagnostics and a Black-Scholes derivative-pricing overlay. The empirical universe is a fixed Damodaran January 2026 U.S. industry dataset containing 94 industry portfolios after excluding aggregate market rows. Expected returns are constructed as CAPM-implied priors using a risk-free rate of 3.97% and an equity risk premium of 4.23%, while portfolio risk is evaluated through single-index covariance geometry, correlation diagnostics and eigenvalue concentration tests. The baseline $K = 10$ sparse-selection problem contains $C(94,10) = 9,041,256,841,903$ possible supports before continuous weights are optimized, motivating a layered comparison of greedy search, Monte Carlo sampling, genetic algorithms and GA-plus-continuous reoptimization. Exact evidence is supplied through a reduced $n = 20$, $K = 6$ benchmark in which all 38,760 supports are enumerated. The study reports best-found full-universe incumbents under documented search coverage rather than claiming certified global optimality. Robustness layers examine covariance alternatives, Rf/ERP perturbations, weight caps, transaction costs, beta drift, risk contributions, stress scenarios and falsification gates. The Black-Scholes overlay demonstrates how option value, delta exposure, leverage-adjusted beta and volatility can be incorporated without treating derivative leverage as free performance. The contribution is methodological and empirical: it shows how computational complexity, asset-pricing priors, covariance dependence, algorithmic search and derivative realism can be combined in a transparent, reproducible and claim-bounded portfolio-optimization architecture. The study does not prove $P \neq NP$, does not certify full $n = 94$ global optimality and does not establish realized historical investment dominance.

Keywords

Robust portfolio optimization; Cardinality-constrained portfolio selection; P-vs-NP-inspired computational finance; Markowitz mean-variance optimization; CAPM expected returns; Sparse support selection; Mixed discrete-continuous optimization; Genetic algorithms; Monte Carlo search; Covariance geometry; Eigenvalue concentration; Black-Scholes option pricing; Derivative-pricing overlay; Transaction costs; Stress testing; Claim-boundary discipline.

JEL Classification

C02; C44; C51; C52; C58; C61; C63; G11; G12; G13; G17; G23.

1. Introduction

1.1. General Research Context and Problem Statement

The contemporary portfolio-optimization literature stands at the intersection of financial economics, numerical optimization and computational-complexity analysis. Classical mean-variance portfolio theory gives a clear continuous allocation problem once the asset universe, expected-return vector and covariance matrix are fixed. However, real-world portfolio design rarely remains purely continuous. Investors, regulators, portfolio committees and implementation constraints frequently require a sparse portfolio: a limited number of selected assets, sectors or industries must be chosen before capital weights can be allocated. This support-selection requirement transforms the problem from a smooth continuous optimization exercise into a mixed discrete-continuous search problem.

The present study addresses this mixed structure through a P-vs-NP-inspired empirical-computational framework. The central object is not the formal theorem-level P versus NP problem in theoretical computer science. The study does not attempt to prove $P = NP$ or $P \neq NP$, and it does not claim to resolve any open Millennium Prize problem. Instead, it operationalizes the P-vs-NP logic through a concrete financial problem in which the number of feasible supports grows combinatorially as $C(n,K)$, while each selected support still requires continuous weight evaluation. This makes cardinality-constrained portfolio optimization an empirically meaningful domain in which computational difficulty, heuristic approximation, exact reduced benchmarking and claim-boundary discipline can be studied together.

The empirical problem is constructed around a fixed Damodaran January 2026 U.S. industry universe. After excluding aggregate market rows, the candidate universe contains $n = 94$ industry portfolios. Each industry is characterized by firm count, levered beta, annual equity volatility, CAPM-implied expected return and a CAPM Sharpe proxy. The baseline sparse-selection problem imposes $K = 10$, producing $C(94,10) = 9,041,256,841,903$ possible industry supports before continuous weights are optimized. This numerical fact is the core computational motivation of the manuscript: even before allocation refinement, the discrete support space is far too large for naive full enumeration in the main universe.

The research problem can therefore be stated as follows: conventional portfolio optimization is insufficiently informative when the asset-selection layer is sparse, combinatorial and economically constrained; a rigorous empirical-computational framework is required to evaluate the interaction between CAPM-implied return construction, covariance geometry, cardinality-driven support-space growth, heuristic search, exact reduced certification, robustness diagnostics, derivative-overlay realism, stress testing and final claim discipline. The study responds by building a layered evidence architecture that makes the computational-finance claim strong while preventing unsupported claims about theorem proof, certified full-universe optimality or realized historical investment dominance.

1.2. Research Object, Boundaries and Data Status

The research object is a cardinality-constrained portfolio-optimization system interpreted through an empirical P-vs-NP-inspired computational-finance lens. The object is not a single optimized portfolio alone. It is the full methodological chain that begins with audited industry inputs, constructs CAPM expected returns, builds covariance and correlation geometry, measures combinatorial support-space growth, searches sparse portfolios, compares greedy, Monte Carlo and genetic-algorithm layers, validates heuristics against a reduced exact benchmark, and then tests robustness, frictions, derivative overlays, stress scenarios, falsification gates and final claim boundaries.

The empirical universe is deliberately fixed before optimization. The input layer consists of 94 U.S. industry portfolios from the Damodaran January 2026 beta and volatility dataset after aggregate rows are excluded. The baseline calibration uses a risk-free rate of 3.97 percent and an equity risk premium of 4.23 percent. CAPM-implied returns are therefore generated as $\mu_i = R_f + \beta_i \times ERP$. In the reported input audit, the median beta is 0.923653, the median annual sigma is 0.465377, the median CAPM return is 0.078771 and the median CAPM Sharpe proxy is 0.084132. These values establish that the universe contains meaningful heterogeneity but also requires covariance-aware interpretation.

A major boundary of the study is that the return vector is a model-implied CAPM prior, not a realized historical return panel. This choice is methodologically transparent and reproducible, but it means that realized-return dominance, live trading performance and out-of-sample investment superiority remain outside the current evidence scope. Similarly, the full $n = 94$, $K = 10$ instance is reported through documented search coverage and a best-found incumbent rather than through a complete branch-and-bound certificate over all 9.041 trillion supports. Exact certification is supplied only for the reduced $n = 20$, $K = 6$ benchmark where all $C(20,6) = 38,760$ supports can be enumerated.

The data status is therefore strong but bounded. The study has a clear input vintage, an auditable calibration, documented search budgets, reproducible formulas, exact reduced certification, covariance robustness gates and final replication checks. It does not hide the fact that the main universe is not exhaustively enumerated and that CAPM-implied returns are conditional model inputs. This boundary is not a weakness; it is a reviewer-safety device that distinguishes admissible computational evidence from overclaiming.

Table 1.1. Research design boundary and evidence-status matrix

Design component	Operational definition	Analytical role	Claim boundary
Empirical universe	94 U.S. industry portfolios from Damodaran January 2026 after exclusion of aggregate market rows.	Fixes the asset universe before optimization and prevents ex-post selection.	Industry-level computational-finance universe, not a complete market-wide trading universe.
CAPM return layer	$R_f = 3.97\%$; $ERP = 4.23\%$; $\mu_i = R_f + \beta_i \times ERP$.	Constructs an auditable expected-return vector for all optimization and robustness layers.	Model-implied expected returns, not realized historical returns.
Covariance geometry	Single-index Sigma, correlation matrix, eigenvalue diagnostics and alternative covariance audits.	Measures common-factor dependence and portfolio-risk coupling.	Risk model is conditional on the chosen covariance specification.
Combinatorial core	$C(94,10) = 9,041,256,841,903$ possible supports.	Operationalizes search-space explosion and motivates heuristic methods.	Computational-hardness evidence only; no theorem proof.
Algorithmic search	Greedy, Monte Carlo, GA, GA plus continuous reoptimization and reduced exact enumeration.	Tests how different search layers behave under common Sharpe and runtime metrics.	Best-found full-universe incumbent, not certified global optimum.
Robustness and realism	$R_f/ERP/cap$ sensitivity, transaction costs, beta drift, risk contribution, covariance placebo and option overlay.	Tests practical fragility and separates leverage artefacts from genuine risk-adjusted evidence.	Scenario and diagnostic evidence, not universal investment dominance.
Claim discipline	Falsification matrix, integrated ranking, final claim hierarchy and replication gate.	Prevents overstating computational results as theorem-level or realized-performance claims.	Admissible claim is bounded empirical-computational evidence.

Source: Author synthesis based on Chapters 3-5 of the *P-vs-NP-inspired portfolio-optimization manuscript*.

1.3. Research Aim

The aim of the study is to develop and empirically audit a P-vs-NP-inspired computational-finance framework for cardinality-constrained portfolio optimization in which sparse industry selection, CAPM expected-return construction, covariance dependence and heuristic search are evaluated within one reproducible evidence architecture.

More specifically, the study aims to demonstrate that a fixed industry universe can be transformed into a transparent computational experiment: Damodaran beta and volatility inputs are audited; CAPM expected returns and Sharpe proxies are constructed; covariance, correlation and eigenvalue diagnostics define the risk geometry; $C(n,K)$ measures the support-space explosion; a $K = 10$ sparse Sharpe-maximization problem is searched; K-grid frontiers and reduced exact enumeration evaluate the search logic; algorithmic layers are compared by Sharpe, median behavior, runtime and stability; robustness layers test covariance, parameter, cost and input-drift sensitivity; derivative overlays and stress tests examine realism; and final falsification and replication gates define the admissible claim.

The aim is not to convert portfolio optimization into a proof of computational-complexity theory. It is to show that the P-vs-NP logic becomes empirically meaningful when cardinality constraints force the analyst to select a discrete support from a large universe while simultaneously solving a continuous allocation problem. The final aim is therefore scientific and evidentiary: to produce a high-level, reproducible and claim-bounded computational-finance study that explains where heuristic portfolio search is credible, where exact evidence is available, where covariance and leverage change interpretation, and where academic caution is mandatory.

1.4. Research Objectives

The study pursues the following objectives:

- To define a fixed empirical universe of 94 U.S. industry portfolios using the Damodaran January 2026 beta and volatility input layer.
- To construct CAPM-implied expected returns using a transparent risk-free-rate and equity-risk-premium calibration and to report descriptive diagnostics for beta, sigma, firm count, expected return and CAPM Sharpe proxies.
- To build and evaluate a covariance and correlation geometry that captures common-factor dependence, eigenvalue concentration and numerical conditioning.
- To operationalize the P-vs-NP-inspired dimension through combinatorial support-space growth, especially $C(94,10) = 9,041,256,841,903$.
- To formulate the baseline $K = 10$ sparse Sharpe-maximization problem as a mixed discrete-continuous optimization problem with binary support variables and continuous weights.
- To compare greedy, Monte Carlo, genetic-algorithm and GA-plus-continuous-reoptimization procedures under common Sharpe, median, runtime and stability metrics.
- To supply exact evidence through a reduced $n = 20$, $K = 6$ enumeration benchmark and to quantify the boundary between exact reduced certification and full-universe search coverage.
- To test the robustness of the results under alternative covariance assumptions, CAPM parameter perturbations, weight caps, transaction costs, beta-drift diagnostics and expected-return sensitivity.
- To evaluate economic realism through risk decomposition, concentration metrics, option-overlay calculations, leverage-adjusted Sharpe interpretation and stress-scenario tests.
- To formulate a final claim hierarchy that explicitly separates admissible computational-finance findings from theorem-level P-vs-NP claims, full global-optimality claims and realized historical investment-dominance claims.

1.5. Research Questions and Hypothetical Framework

The research questions connect the data audit, portfolio geometry, combinatorial complexity, algorithmic search, robustness, realism and claim-boundary layers. They are formulated as empirical-computational questions rather than theorem-level propositions. Each question asks whether a defined evidence layer supports a bounded claim inside the Damodaran-CAPM/single-index universe.

Table 1.2 presents the research questions and their hypothesis-level expectations. The expectations are deliberately stated as diagnostic and computational-finance propositions. They do not claim that the study proves $P \neq NP$, certifies the global optimum over the full $n = 94$ universe or establishes realized trading superiority.

Table 1.2. Research questions and hypothesis-level expectations

Research question	Hypothesis-level expectation	Methodological test	Expected evidentiary status
RQ1. Can the Damodaran industry universe be converted into a reproducible sparse-optimization experiment?	H1. Fixed data vintage, CAPM calibration and variable audit provide a stable empirical coordinate system.	Data audit, beta/sigma/firm-count diagnostics and CAPM return construction.	Foundational design condition.
RQ2. Does cardinality create a nontrivial computational search problem?	H2. The support space grows combinatorially and becomes infeasible for exhaustive full-universe enumeration at $n = 94$, $K = 10$.	$C(n,K)$, K-grid and full support-space certificate.	Core P-vs-NP-inspired operational evidence.
RQ3. Does covariance geometry materially affect portfolio interpretation?	H3. High common-factor dependence compresses diversification and changes	Median correlation, eigenvalue shares, diagonal/constant-rho/shrinkage/high-rho covariance audits.	Empirical geometry and risk-model evidence.

Research question	Hypothesis-level expectation	Methodological test	Expected evidentiary status
	Sharpe relative to diagonal covariance placebo.		
RQ4. Can heuristics approximate strong sparse-portfolio solutions?	H4. Greedy, MC, GA and GA plus reoptimization should find upper-envelope Sharpe values close to the reduced exact benchmark.	Algorithm comparison, seed summaries, runtime scaling and reduced exact enumeration.	Heuristic credibility evidence with explicit boundary.
RQ5. Does continuous reoptimization improve support-only search?	H5. Reoptimizing weights inside selected supports should close equal-weight gaps and raise median performance.	GA subset-only versus GA plus continuous reoptimization.	Mixed discrete-continuous mechanism evidence.
RQ6. Are results robust to economic and implementation assumptions?	H6. ERP, covariance, transaction costs, beta drift and weight caps should reveal which parts of the result are stable and which are conditional.	Sensitivity tables, transaction-cost grid, beta-drift audit and covariance robustness gates.	Robustness and realism evidence.
RQ7. Can derivative overlays be incorporated without overstating free performance?	H7. Option leverage should raise expected return and risk together, leaving risk-adjusted interpretation bounded.	Black-Scholes overlay, moneyness/maturity grid, bump test and stress scenarios.	Controlled derivative-demonstration evidence.
RQ8. What final claims are scientifically admissible?	H8. The admissible claim is a bounded computational-finance claim, not theorem proof or universal investment dominance.	Falsification matrix, integrated ranking, claim hierarchy and replication gate.	Reviewer-proof closure evidence.

Source: Author synthesis based on the research architecture and claim-boundary logic of Chapters 3-5.

The core hypothetical framework can be expressed compactly as follows: a sparse portfolio-selection system is defined by a fixed universe A , a cardinality K , a CAPM expected-return vector μ , a covariance matrix Σ , a binary support vector z and a continuous weight vector w . The research hypothesis is that the computational difficulty and empirical interpretation of the problem arise from the joint object (S, w_S) , where S is selected from Ω_K and w_S is optimized on the simplex Δ_K . In plain-text notation, the central diagnostic relation is: maximize $SR(w) = (\mu^T w - R_f) / \sqrt{w^T \Sigma w}$, subject to $\sum(w_i) = 1$, $w_i \geq 0$, $z_i \in \{0,1\}$, $\sum(z_i) = K$ and $w_i \leq u_i z_i$. This is not a theorem about P versus NP. It is the empirical-computational analogue tested by the manuscript.

1.6. Scientific Novelty of the Research

The scientific novelty of the study lies in converting the P-vs-NP reference from a rhetorical analogy into an operational computational-finance design. Instead of claiming to solve a mathematical problem, the study identifies a concrete finance problem in which a discrete support-selection layer and a continuous allocation layer interact. This allows the research to discuss computational hardness, approximation, exact reduced benchmarking and claim boundaries in a way that is empirically auditable.

The first novelty is the full evidence architecture. The study does not present a single optimized Sharpe ratio as the result. It integrates data audit, CAPM calibration, covariance/eigenvalue diagnostics, combinatorial growth, K-grid frontiers, exact reduced enumeration, algorithmic comparison, GA reoptimization, runtime scaling, selection stability, benchmark comparison, robustness, transaction costs, beta-drift evidence, risk decomposition, derivative overlay, stress testing, falsification and replication gates. This makes the manuscript a layered computational-finance validation design rather than a narrow portfolio-selection exercise.

The second novelty is the explicit treatment of exactness. In the full $n = 94$ universe, the manuscript quantifies search coverage and reports a best-found incumbent rather than inventing a false global certificate. In the reduced $n = 20$, $K = 6$ universe, it enumerates all 38,760 supports and obtains an exact benchmark against which Monte Carlo and GA results can be compared. This distinction between exact reduced proof and full-universe heuristic evidence is one of the strongest methodological safeguards of the study.

The third novelty is the covariance-realism layer. The manuscript shows that the Damodaran industry universe is strongly common-factor dominated: the median off-diagonal correlation is 0.769201, the share of correlations above 0.5 is 0.884008 and the first eigenvalue explains 0.761011 of total variance. This finding prevents simplistic interpretation of portfolio diversification. It also explains why diagonal covariance placebo can overstate Sharpe relative to correlated risk models.

The fourth novelty is the controlled integration of derivative overlays. The Black-Scholes layer demonstrates how an option can be embedded in the same CAPM-consistent risk-return framework, but the results explicitly reject the idea that option leverage creates free performance. The option case raises expected return and volatility together; under leverage-adjusted risk, Sharpe remains approximately 0.0879. This converts a potentially overstated derivative result into a reviewer-safe stress-realism audit.

The fifth novelty is the final claim-boundary logic. The study explicitly separates admissible claims from excluded claims. It supports a bounded empirical-computational claim about cardinality-constrained sparse optimization. It excludes theorem proof, certified full $n = 94$ global optimality, realized historical dominance and universal investment superiority. This discipline increases the scientific credibility of the manuscript because every strong result is paired with the boundary that prevents misuse.

1.7. Theoretical, Methodological and Practical Significance

The theoretical significance of the study is that it places portfolio selection inside a computational-complexity-aware finance framework. The study shows that once cardinality is imposed, portfolio optimization cannot be understood only as a continuous mean-variance problem. The binary decision of which industries to include becomes a central theoretical object. This moves the discussion from allocation weights alone to the joint theory of support selection, allocation, covariance coupling and heuristic approximation.

The methodological significance is substantial. The manuscript provides a reproducible pipeline from data universe to final claim hierarchy. It defines all inputs before optimization; constructs returns through an explicit CAPM prior; builds risk through a documented covariance model; measures $C(n,K)$; separates full-universe heuristic search from reduced exact enumeration; compares algorithmic layers; reports runtime and selection stability; tests covariance and parameter robustness; includes implementation frictions; evaluates derivative overlays and stress regimes; and closes the chapter with falsification and replication gates. Such a design is directly useful for computational-finance manuscripts where reviewers need to distinguish best-found evidence from certified optimality.

The practical significance lies in portfolio audit and model governance. Portfolio analysts can use the framework to understand how sparse industry selection behaves under common-factor risk, how much search effort is documented, whether a heuristic method is stable across seeds, whether continuous reoptimization materially improves support-only selection, whether transaction costs alter the final Sharpe, and whether derivative overlays create risk-adjusted improvement or only leverage-amplified exposure. The framework therefore supports decision analysis even without claiming realized investment dominance.

The policy and risk-management significance is also clear. Concentrated sparse portfolios may appear efficient under expected-return assumptions, but the study reports weight concentration, risk contribution and stress-scenario exposure. The baseline $K = 10$ portfolio has $HHI = 0.1696$, effective $N = 5.90$ and top-5 normalized weight share of 84.03 percent. Risk decomposition and option stress tests show that sparse portfolios require governance beyond headline Sharpe. This makes the framework relevant for institutional portfolio committees, model validators and macro-financial stress analysts.

1.8. Research Limitations and Principle of Academic Caution

The study is intentionally structured around academic caution. The first limitation is conceptual: the P-vs-NP reference is operational rather than theorem-level. The manuscript uses cardinality-constrained portfolio optimization as an empirical domain where combinatorial growth creates computational difficulty. It does not solve the formal P versus NP problem and does not provide a mathematical proof about complexity classes.

The second limitation is computational. The full $n = 94$, $K = 10$ universe contains 9,041,256,841,903 possible supports. The documented audit coverage equals 421,760 evaluations/supports, or a coverage proxy of $4.665e-08$. This is substantial for an empirical-computational audit, but it is not exhaustive enumeration. The reported $K = 10$ portfolio is therefore an incumbent and evidence anchor under documented search coverage, not a certified global optimum over the full universe.

The third limitation is financial. Expected returns are CAPM-implied rather than realized. This makes the experiment transparent and internally consistent, but it also restricts the empirical claim. The manuscript can support computational and model-implied portfolio-selection conclusions; it cannot by itself establish realized historical dominance over market benchmarks, live trading strategies or alternative return models.

The fourth limitation concerns covariance and model specification. The single-index covariance model produces strong common-factor dependence. Alternative covariance audits show that Sharpe can change materially when covariance assumptions change: diagonal Sigma yields Sharpe = 0.236744, constant-rho Sigma yields Sharpe = 0.115660 and high-rho stress Sigma yields Sharpe = 0.109600. These differences do not invalidate the study; they demonstrate why covariance specification must remain a central boundary of interpretation.

The fifth limitation concerns derivative overlay and stress interpretation. Black-Scholes option calculations are controlled diagnostics, not complete option-trading models. Gamma risk, volatility smile, liquidity spreads, funding costs, early exercise and dynamic hedging slippage are not fully modeled. The option layer is therefore best read as a leverage and stress-test demonstration, not as a practical derivative trading prescription.

The principle of academic caution can be summarized as follows: the study supports a strong P-vs-NP-inspired empirical-computational claim about cardinality-constrained portfolio optimization under a fixed Damodaran-CAPM/single-index design; it does not prove $P \neq NP$, does not certify the full $n = 94$ global optimum, does not establish realized historical investment dominance and does not claim universal strategy superiority.

1.9. Structure of the Research Logic

The research logic is organized as an integrated six-chapter architecture. The introductory chapter defines the research context, object, aim, objectives, research questions, novelty, significance, limitations and claim boundaries. It frames the study as an empirical-computational portfolio-optimization design rather than as a theorem-proof manuscript.

The methodology chapter specifies the full research protocol. It defines the empirical universe, CAPM calibration, covariance construction, combinatorial search-space measurement, baseline sparse optimization model, K-grid frontier methodology, exact reduced benchmark, greedy/Monte Carlo/GA comparison, GA plus continuous reoptimization, runtime scaling, selection stability, covariance robustness, expected-return robustness, parameter sensitivity, transaction costs, input stability, risk decomposition, Black-Scholes overlay, stress testing, placebo logic, integrated ranking and replication gates.

The results chapter reports the empirical outputs of this architecture. It provides the input audit, descriptive statistics, CAPM diagnostics, covariance and eigenvalue evidence, $C(n,K)$ growth tables, baseline $K = 10$ portfolio, K-grid results, exact reduced enumeration, algorithmic comparisons, reoptimization results, runtime scaling, selection stability, benchmark comparison, robustness gates, transaction-cost and beta-drift evidence, risk contribution, option overlay, stress scenarios, falsification matrix, integrated ranking and final claim hierarchy.

The interpretation chapter translates these empirical outputs into scientific meaning. It explains why the sparse-optimization problem is mixed discrete-continuous, why common-factor covariance matters, why the $K = 10$ Sharpe result must be interpreted as a model-implied incumbent, why reduced exact evidence supports heuristic credibility, why option leverage cannot be treated as free performance, and why final claims must remain bounded. The conclusion then returns to these elements and classifies the final contribution as a strong but scientifically limited P-vs-NP-inspired computational-finance framework.

2. Literature Review

2.1. Scope and Review Strategy

This literature review positions the P-vs-NP-inspired portfolio-optimization study within the streams of financial economics, asset pricing, covariance modelling, mixed-integer optimization, computational complexity, metaheuristic search, Monte Carlo benchmarking, derivative-overlay diagnostics, implementation frictions and claim-boundary discipline. The manuscript does not attempt to prove $P = NP$ or $P \neq NP$. Its problem is empirical-computational: the cardinality constraint transforms an otherwise continuous allocation problem into a mixed discrete-continuous sparse-support selection problem. Consequently, the literature must be broader than a conventional mean-variance portfolio review and more claim-disciplined than a generic computational-complexity analogy.

The review uses 140 references, each supplied with a DOI resolver link in the reference list. The source-selection rule is deliberately conservative: the review relies on foundational journal articles, high-impact finance and econometrics journals, operations-research and optimization journals, scientific-computing journals, and recognized academic outlets directly connected to the manuscript components. General textbooks, uncited background readings and weakly related material are excluded from the main reference list. The seven

literature blocks are tied to the manuscript architecture: CAPM expected-return construction, covariance and eigenvalue geometry, cardinality-constrained support search, exact reduced benchmarking, Monte Carlo and genetic algorithms, implementation realism, derivative-overlay stress testing, robustness and final claim hierarchy.

2.2. Literature Coverage Map and Reference-Selection Protocol

Table 2.1 maps the 140 sources to the methodological blocks required by the article. The purpose is not only bibliographic organization. It establishes that every major claim in the manuscript has a corresponding literature foundation and that no result is interpreted outside the evidentiary domain supported by prior research.

Literature block	References	Function in this study	Relevant manuscript components
Portfolio theory, CAPM and performance inference	24	Defines expected-return priors, beta logic, efficient-frontier language and Sharpe-performance interpretation.	CAPM calibration; Sharpe proxy; portfolio objective; claim boundaries.
Covariance, factor dependence and high-dimensional risk	18	Grounds covariance coupling, common-factor dependence, eigenvalue concentration, shrinkage and diversification limits.	Sigma construction; correlation/eigenvalue diagnostics; covariance robustness.
Cardinality-constrained and sparse portfolio optimization	20	Supports sparse support selection, buy-in/cardinality constraints, MIQP formulations and real-world portfolio restrictions.	C(n,K); K=10 support search; reduced exact benchmark; sparse optimization.
Computational complexity and integer optimization	15	Provides the formal logic for discrete search, exact enumeration, branch-and-bound, relaxations and scalable sparse algorithms.	P-vs-NP operationalization; exact reduced enumeration; best-found/full-optimality boundary.
Metaheuristic and evolutionary search	16	Justifies genetic algorithms, tabu search, simulated annealing, PSO, multiobjective evolutionary search and stability analysis.	Greedy, MC, GA, GA+reoptimization, seed stability, runtime comparison.
Simulation, inference and benchmark validation	20	Supports Monte Carlo, bootstrap, robust inference, rolling forecast tests, residual diagnostics and benchmark comparison.	MC budgets, inference limits, validation windows, performance tests.
Liquidity, costs, derivatives, risk and systemic stress	27	Grounds transaction costs, liquidity frictions, option overlay, volatility/risk metrics and stress-test realism.	Cost drag, beta drift, Black-Scholes overlay, stress tests, risk decomposition.

The selection protocol follows four rules. First, each reference must support a specific component of the manuscript rather than merely signal disciplinary breadth. Second, the review distinguishes theoretical foundations from empirical evidence: Markowitz-CAPM theory supports expected-return construction, covariance research supports risk geometry, optimization research supports search design, and derivatives literature supports the option-overlay boundary. Third, the review preserves claim discipline: the paper studies an operational P-vs-NP-inspired portfolio-search problem and not a theorem-proof result. Fourth, the reference list is DOI-based so that the literature trail remains auditable.

2.3. Portfolio Theory, CAPM Foundations and Sharpe-Performance Inference

The first literature block establishes why the manuscript begins with Markowitz-style portfolio selection and then constructs expected returns through a transparent CAPM prior. Mean-variance theory supplies the fundamental separation between expected return and risk, while liquidity preference and risk-bearing foundations justify a risk-return decision framework rather than a return-only ranking. The CAPM and its extensions provide the beta-based language used in the manuscript to transform the Damodaran industry beta layer into model-implied expected returns. This makes the return vector auditable, but also bounded: CAPM-implied returns are priors, not realized historical performance claims.

This tradition also explains why a Sharpe-ratio objective is natural but insufficient by itself. Performance measurement research shows that Sharpe, Treynor and alpha-type measures require inference discipline, benchmark awareness and caution about data mining. In the present study, this supports the distinction between a model-implied best-found incumbent and a realized investment-dominance claim. The literature therefore legitimizes the CAPM-Sharpe computational object while simultaneously preventing the paper from overstating it.

Representative sources used in this block: (Markowitz, 1952a; Markowitz, 1952b; Tobin, 1958; Sharpe, 1963; Sharpe, 1964; Lintner, 1965; Mossin, 1966; Jensen, 1968)

Representative sources used in this block: (Fama, 1970; Black, 1972; Merton, 1973a; Ross, 1976; Roll, 1977; Merton, 1980; Treynor & Black, 1973; Sharpe, 1966)

Representative sources used in this block: (Fama & French, 1992; Fama & French, 1993; Carhart, 1997; Fama & French, 2004; Lo, 2002; Jobson & Korkie, 1981; Ledoit & Wolf, 2008; Bailey & Lopez de Prado, 2014)

2.4. Covariance Geometry, Factor Dependence and Eigenvalue Concentration

The second block supports the study's claim that sparse portfolio selection cannot be evaluated through stand-alone Sharpe proxies alone. Covariance couples all selected assets. Factor models, approximate factor structures, shrinkage estimators and random-matrix evidence show that high-dimensional covariance matrices are noisy, concentrated and often dominated by common modes. This is directly relevant to the manuscript's finding that the industry universe has high median correlation and strong first-eigenvalue concentration.

The covariance literature also explains why diagonal covariance is a placebo rather than a realistic baseline. If common-factor dependence is ignored, risk can be materially understated and diversification can be overstated. Shrinkage, factor-based estimators and dynamic conditional correlation models provide the methodological logic for testing alternative Sigma specifications without pretending that one covariance matrix is universally true. The review therefore anchors the manuscript's covariance robustness layer, eigenvalue diagnostics, high-correlation stress case and final interpretation that support selection is conditional on the selected risk model.

Representative sources used in this block: (Ledoit & Wolf, 2003; Ledoit & Wolf, 2004a; Ledoit & Wolf, 2004b; Jagannathan & Ma, 2003; DeMiguel et al., 2009a; DeMiguel et al., 2009b; Fan et al., 2008)

Representative sources used in this block: (Fan et al., 2012; Bai & Ng, 2002; Bai, 2003; Chamberlain & Rothschild, 1983; Connor & Korajczyk, 1986; Laloux et al., 1999; Plerou et al., 1999)

Representative sources used in this block: (Mantegna, 1999; Engle, 2002; Fleming et al., 2001; Christoffersen et al., 2012)

2.5. Cardinality-Constrained Portfolio Optimization and Sparse-Support Selection

The third block is the direct methodological core of the manuscript. Once a portfolio must contain exactly K selected industries, the problem is no longer a purely continuous mean-variance allocation. The analyst must select a support and then optimize weights within that support. Portfolio research with minimum lots, buy-in thresholds, fixed costs, cardinality constraints, sparse regularization and index-tracking restrictions demonstrates that this mixed structure is not artificial; it corresponds to real portfolio-governance and implementation constraints.

The literature also provides a strong precedent for comparing exact mixed-integer methods with heuristics. Cardinality-constrained models have been approached through mixed-integer quadratic programming, branch-and-bound, Lagrangian relaxation, simulated annealing, tabu search, genetic algorithms, particle swarm optimization and multiobjective evolutionary algorithms. The manuscript's exact reduced benchmark and full-universe best-found incumbent therefore fit established research practice: exact certification is valid when the instance is tractable, while the full $n = 94$, $K = 10$ universe requires documented search budgets and claim boundaries.

Representative sources used in this block: (Konno & Yamazaki, 1991; Bienstock, 1996; Chang et al., 2000; Mansini & Speranza, 1999; Kellerer et al., 2000; Maringer & Kellerer, 2003; Crama & Schyns, 2003)

Representative sources used in this block: (Lin & Liu, 2008; Shaw et al., 2008; Soleimani et al., 2009; Cura, 2009; Bonami & Lejeune, 2009; Bertsimas & Shioda, 2009; Brodie et al., 2009)

Representative sources used in this block: (Canakgoz & Beasley, 2009; Anagnostopoulos & Mamanis, 2011; Woodside-Oriakhi et al., 2011; Liagkouras & Metaxiotis, 2018; Lwin et al., 2014; Steuer et al., 2024)

2.6. Computational Complexity, Integer Optimization and Exact Reduced Benchmarking

The P-vs-NP reference in the manuscript is operational. It is used to interpret search-space growth and not to assert theorem-level proof. The relevant literature therefore includes both the status of the P versus NP problem and the optimization machinery that makes discrete support search scientifically tractable. Complexity-theory sources clarify why

combinatorial explosion matters, while integer-programming and mixed-integer nonlinear optimization sources explain branch-and-bound, cutting planes, disjunctive logic, perspective reformulations and conic relaxations.

This block is essential for the manuscript's claim-boundary discipline. $C(94,10)$ is an exact combinatorial count, but it is not a proof of $P \neq NP$. Similarly, reduced enumeration over $C(20,6)$ supports exact evidence for a smaller instance but does not certify the global optimum over more than nine trillion supports in the main universe. The literature supports a three-level distinction: exact reduced certification, documented full-universe heuristic coverage and explicit exclusion of theorem-level or full-global-optimality claims.

Representative sources used in this block: (Fortnow, 2009; Arora et al., 1998; Impagliazzo & Paturi, 2001; Land & Doig, 1960; Gomory, 1958; Fisher, 1981)

Representative sources used in this block: (Geoffrion, 1972; Balas, 1979; Frangioni & Gentile, 2006; Günlük & Linderoth, 2010; Vielma et al., 2008; Goldfarb & Idnani, 1983)

Representative sources used in this block: (Byrd et al., 1995; Nelder & Mead, 1965; Bertsimas & Cory-Wright, 2022)

2.7. Metaheuristic, Evolutionary and Population-Based Search

The fifth block grounds the algorithmic comparison among greedy, Monte Carlo, genetic algorithm and GA-plus-continuous-reoptimization layers. Metaheuristics are appropriate when exact enumeration is infeasible, provided that their performance is evaluated through transparent budgets, seed stability, benchmark comparison and robustness rather than by a single favorable run. Simulated annealing, tabu search, differential evolution, ant colony systems, particle swarm methods, NSGA-II and related multiobjective evolutionary algorithms supply the computational language for exploring large support spaces.

For the manuscript, this literature has two implications. First, heuristic search should be judged through best, median, dispersion, runtime and stability metrics. Second, the support-selection and weight-optimization layers should be separated. A genetic algorithm that identifies a promising support may still be incomplete if weights are left equal or proxy-assigned. The GA-plus-continuous-reoptimization layer is therefore not cosmetic; it follows from the mixed discrete-continuous structure of the portfolio problem.

Representative sources used in this block: (Kirkpatrick et al., 1983; Glover, 1986; Glover, 1989; Glover, 1990; Storn & Price, 1997; Dorigo et al., 1996)

Representative sources used in this block: (Clerc & Kennedy, 2002; Deb et al., 2002; Zitzler & Thiele, 1999; Coello Coello, 2000; Hansen & Ostermeier, 2001; Pisinger, 2005)

Representative sources used in this block: (Blum & Roli, 2003; Jones et al., 2002; Laumanns et al., 2002; Knowles & Corne, 2000)

2.8. Monte Carlo, Robust Inference, Rolling Validation and Benchmark Diagnostics

The sixth block supports the manuscript's validation architecture. Monte Carlo sampling and bootstrap logic allow the study to report search distributions rather than isolated outcomes. Robust-inference literature supports careful treatment of heteroskedasticity, autocorrelation, panel dependence and finite-sample covariance estimators. Forecast-comparison literature supplies the logic for comparing benchmark models without claiming permanent superiority from one validation window.

Although the current manuscript uses CAPM-implied expected returns rather than a realized return panel, inference and validation literature still matters. It tells the reader what can be claimed now and what requires future data. In the current design, Sharpe differences are descriptive and model-implied; formal realized-return dominance would require a realized return panel, paired tests, out-of-sample windows, transaction-cost conventions and benchmark-adjusted inference. This literature therefore strengthens the paper by making the limits of the current evidence explicit.

Representative sources used in this block: (Metropolis et al., 1953; Hastings, 1970; Efron, 1979; White, 1980; Newey & West, 1987; MacKinnon & White, 1985; Hansen, 1982)

Representative sources used in this block: (Hausman, 1978; Arellano & Bond, 1991; Blundell & Bond, 1998; Driscoll & Kraay, 1998; Diebold & Mariano, 1995; White, 2000; Hansen, 2005)

Representative sources used in this block: (Christoffersen, 1998; Kupiec, 1995; Box & Pierce, 1970; Ljung & Box, 1978; Sims, 1980; Hamilton, 1989)

2.9. Liquidity, Transaction Costs and Implementation Frictions

A sparse portfolio can look strong before implementation costs and become weaker after trading frictions. Liquidity and transaction-cost research shows that bid-ask spreads, informed trading, common liquidity factors, order-flow dependence, funding liquidity and market depth can materially alter realized performance. For this reason, the manuscript's transaction-cost grid and cost-drag Sharpe recalculation are necessary realism checks rather than optional appendices.

This literature also supports the distinction between model-implied optimization and implementable strategy. CAPM expected returns and covariance geometry can define a clean computational experiment, but implementation depends on turnover, spread, market depth, liquidity commonality and funding constraints. The manuscript therefore treats transaction costs as a sensitivity layer and not as a fully estimated trading-cost model.

Representative liquidity and implementation sources used in this block: (Kyle, 1985; Glosten & Milgrom, 1985; Amihud & Mendelson, 1986; Constantinides, 1986; Leland, 1985; Davis & Norman, 1990; Amihud, 2002; Pastor & Stambaugh, 2003; Acharya & Pedersen, 2005; Brunnermeier & Pedersen, 2009; Chordia et al., 2000; Hasbrouck & Seppi, 2001)

2.10. Derivative Overlay, Black-Scholes Logic and Leverage Boundaries

The derivative-overlay layer requires its own literature base because option inclusion changes both expected return and risk. Black-Scholes-Merton option pricing, binomial option valuation, stochastic volatility, jumps and Fourier methods provide the mathematical basis for pricing and sensitivity analysis. However, the same literature also identifies boundaries: constant volatility, frictionless trading, continuous hedging, volatility smile, jump risk and calibration error all matter when moving from formula-based diagnostics to practical option trading.

The manuscript uses the option overlay as a controlled diagnostic rather than as a free-performance device. The option position raises beta, expected return and volatility together. The derivative literature therefore supports the exact interpretation used in the study: option leverage may amplify exposure, but risk-adjusted performance must be judged after volatility, stress and approximation-error effects are included.

Representative derivative-pricing sources used in this block: (Black & Scholes, 1973; Merton, 1973b; Cox et al., 1979; Heston, 1993; Bakshi et al., 1997; Bates, 1996; Carr & Madan, 1999)

2.11. Risk Measures, Volatility Modelling and Systemic-Stress Interpretation

The manuscript also reports risk contribution, stress scenarios and option-augmented stress checks. Coherent risk measures, CVaR optimization, expected shortfall, ARCH/GARCH volatility and systemic-risk metrics provide the foundation for this layer. These sources make clear that risk is not exhausted by standard deviation or Sharpe ratio. A sparse support can be diversified by count but concentrated by variance contribution, liquidity exposure or volatility state.

The stress-testing literature therefore supports the manuscript's move from headline Sharpe to risk governance. Risk decomposition identifies which industries drive total portfolio variance; volatility models explain why risk can be state-dependent; systemic-risk measures show how exposure can become more important under common shocks. This literature helps justify the final claim that the portfolio result is a strong model-implied incumbent but must remain bounded by stress, leverage, covariance and implementation diagnostics.

Representative risk, volatility and systemic-stress sources used in this block: (Artzner et al., 1999; Rockafellar & Uryasev, 2000; Acerbi & Tasche, 2002; Engle, 1982; Bollerslev, 1986; Nelson, 1991; Adrian & Brunnermeier, 2016; Brownlees & Engle, 2017)

2.12. Literature-to-Methodology Crosswalk

The review can be summarized as a crosswalk from literature to method. Markowitz-CAPM literature justifies the expected-return and Sharpe objective. Covariance and factor-risk literature justifies the Sigma matrix, correlation heatmaps and eigenvalue diagnostics. Cardinality and sparse-portfolio literature justifies the binary support variable and K-asset constraint. Integer-optimization and complexity literature justifies reduced exact enumeration and search-coverage boundaries. Metaheuristic literature justifies GA, Monte Carlo and continuous reoptimization. Liquidity and derivative literature justifies transaction-cost and option-overlay realism. Inference and benchmark literature justifies descriptive effect sizes, robustness tables and future realized-return extension.

This crosswalk prevents the paper from becoming either a loose analogy or a purely mechanical optimization exercise. Each result is anchored in a literature block and each claim is bounded by the level of evidence supplied by that block. The central P-vs-NP-inspired insight is therefore scientifically admissible: cardinality-constrained portfolio optimization operationalizes computational hardness in a concrete financial workflow, but the paper remains outside theorem proof, certified full-universe optimality and realized historical investment dominance.

2.13. Research Gap and Contribution Located by the Review

The review reveals a clear gap. Portfolio theory provides the continuous allocation foundation; CAPM gives a transparent expected-return prior; covariance research explains risk coupling; cardinality-constrained optimization establishes the mixed-integer nature of sparse portfolios; and metaheuristic research offers practical search tools. However, these strands are often presented separately. Many portfolio studies either emphasize optimization output without sufficient computational-complexity interpretation, or discuss heuristic algorithms without a strong final claim hierarchy, or add derivative/transaction-cost realism without integrating it into a full evidence architecture.

The contribution of the present manuscript is to integrate these strands into one reviewer-safe computational-finance design. It fixes the empirical universe before optimization, constructs CAPM returns transparently, audits covariance and eigenvalues, measures $C(n,K)$, reports a $K = 10$ best-found incumbent, supplies exact reduced enumeration, compares search algorithms, tests robustness, adds transaction-cost and derivative-overlay realism, interprets risk contribution and stress, and closes with falsification and replication gates. The literature therefore supports a high but bounded claim: the study provides a reproducible P-vs-NP-inspired computational demonstration of sparse portfolio search under Damodaran-CAPM/single-index inputs, not a proof of $P \neq NP$, not a certified full-universe global optimum and not a realized investment-performance theorem.

3. Methodology

Methodological Framework for a P-vs-NP-Inspired Cardinality-Constrained Portfolio-Optimization Study

This chapter specifies the methodological design that underpins the Chapter 4 results. It translates the empirical results architecture into a reproducible research protocol covering the data universe, CAPM calibration, covariance construction, combinatorial search-space measurement, sparse mean-variance optimization, exact reduced benchmarking, stochastic and metaheuristic algorithms, robustness diagnostics, derivative-overlay modelling, stress testing, falsification and final claim-boundary discipline.

The methodology is deliberately structured as a quantitative computational-finance design rather than as a theorem-proof design. The P vs NP reference is operationalized through the cardinality-constrained support-selection problem and the empirical growth of $C(n,K)$. The chapter therefore examines computational hardness as it appears in a real portfolio-optimization workflow, while explicitly excluding any claim that the study proves or resolves the formal P versus NP problem.

3.1. Methodological Type, Research Design and Evidence Architecture

The study uses a hybrid quantitative methodology combining empirical financial modelling, constrained optimization, computational complexity diagnostics, simulation, metaheuristic search, robustness analysis and falsification logic. The research design is model-based and audit-oriented: each empirical object is defined before optimization, each algorithmic result is evaluated against either an exact reduced benchmark or a documented search-coverage certificate, and each economic conclusion is constrained by an explicit claim-boundary rule.

Equation (3.1). $M = \{D, R, G, C, A, B, S, F, Q\}$

M denotes the full methodological system. D is the data audit, R is expected-return construction, G is covariance geometry, C is combinatorial complexity, A is algorithmic search, B is benchmarking, S is robustness and stress testing, F is falsification, and Q is final claim discipline.

Equation (3.2). $E = (E_data, E_geometry, E_complexity, E_algorithmic, E_robustness, E_realism, E_claim)$

The evidence vector E is not a single statistic. It is a layered set of evidence objects that jointly support the admissible empirical claim: cardinality-constrained portfolio optimization creates a nontrivial mixed discrete-continuous search problem in the Damodaran Jan-2026 industry universe.

Table 3.1. Methodological architecture of the chapter

Layer	Methodological function	Corresponding result component
Data and baseline geometry	Define the empirical universe, variables, expected-return inputs and covariance geometry before optimization.	Results architecture, empirical universe, CAPM diagnostics, covariance/eigenvalue diagnostics
Computational core	Translate cardinality constraints into support-space growth and define the sparse optimization problem.	$C(n,K)$, $K=10$ baseline, K -grid, reduced exact benchmark
Algorithmic evidence	Compare deterministic, stochastic and metaheuristic procedures under common Sharpe and runtime metrics.	Greedy, Monte Carlo, GA, GA plus continuous reoptimization, runtime scaling, selection stability
Robustness and benchmarking	Test sensitivity to benchmarks, covariance assumptions, return models, Rf/ERP perturbations, caps, inference and frictions.	Benchmark comparison, covariance robustness, expected-return robustness, parameter sensitivity, statistical inference, transaction costs
Realism and derivatives	Add input stability, risk contribution, option-overlay and stress-scenario	Beta-drift, risk decomposition, Black-Scholes overlay, stress tests

	methods without overstating realized returns.	
Reviewer-proof closure	Apply placebo tests, multi-criteria scoring, claim hierarchy and reproducibility gates.	Falsification matrix, integrated ranking, claim-boundary matrix, replication mini-table

This design is best classified as an empirical-computational optimization methodology. It is empirical because it uses sector-level beta, volatility and firm-count inputs. It is computational because the main object is a constrained search process. It is econometric/financial because returns and risks are generated through CAPM and covariance modelling. It is falsification-oriented because the study does not rely on a single favorable result: it checks alternative covariance assumptions, parameter perturbations, transaction costs, derivative leverage, stress regimes and final claim boundaries.

3.2. Empirical Universe, Data Audit and Variable Construction

The empirical universe consists of $n = 94$ U.S. industry portfolios obtained from the Damodaran January 2026 beta and volatility input layer after excluding aggregate market rows. Each industry i is treated as a candidate asset in a sparse portfolio selection problem. The core data objects are the number of firms, levered beta, annual equity standard deviation, CAPM-implied expected return and a CAPM Sharpe proxy.

Equation (3.3). $i = 1, \dots, n, n = 94$

The index i labels the industry portfolios in the empirical universe. The fixed universe size $n = 94$ determines the support-selection search space once a cardinality K is imposed.

Equation (3.4). $x_i = (F_i, \beta_i, \sigma_i, \mu_i, q_i)$

For each industry, F_i is the firm count, β_i is levered beta, σ_i is annual equity volatility, μ_i is CAPM-implied expected return, and q_i is the CAPM Sharpe proxy used for screening diagnostics.

Equation (3.5). $A = \{1, 2, \dots, 94\}, S \subset A, |S| = K$

A is the full asset universe and S is a selected support. The sparse-portfolio problem selects exactly K industries from A before or jointly with weight optimization.

Table 3.2. Core notation and variable definitions

Symbol	Definition	Methodological role
F_i	Number of firms in industry i	Coverage and stability proxy; used for descriptive and beta-stability diagnostics.
β_i	Levered beta of industry i	Systematic-risk input and CAPM expected-return driver.
σ_i	Annual equity standard deviation of industry i	Volatility input for covariance and risk calculation.
R_f	Risk-free rate, calibrated at 3.97% unless scenario-specific	Sharpe and CAPM calibration parameter.
ERP	Equity risk premium, calibrated at 4.23% unless scenario-specific	CAPM return-premium parameter.
μ_i	CAPM-implied expected return	Return vector entering the portfolio objective.
Σ	Covariance matrix	Risk geometry linking selected industries.
w_i	Portfolio weight in industry i	Continuous decision variable.

z_i	Binary selection indicator	Discrete support-selection decision variable.
K	Required number of selected industries	Cardinality constraint.

The data-audit stage is not a descriptive appendix; it is part of identification. If beta, sigma, firm count or calibration parameters are undefined, the optimization problem is not reproducible. For this reason, the methodology fixes the universe, the input vintage, the return-calibration parameters and the role of each variable before any algorithmic search is performed.

The deposited replication dataset package for the present manuscript is archived on Zenodo as: Gondauri, D. (2026), Dataset Package for Robust Portfolio Optimization under Computational Complexity: A P-vs-NP-Inspired Markowitz-CAPM Framework with Cardinality Constraints and Black-Scholes Derivative-Pricing Overlay [Data set], Zenodo, <https://zenodo.org/records/20922272>. The package documents the master Excel workbook, long-format table exports, table index, figure index, equation register, extracted numerical registry, data dictionary, README files, metadata files and checksum materials used to support the Damodaran January 2026 input layer, CAPM calibration, covariance diagnostics, combinatorial support-space calculations and reported computational-finance outputs. This repository citation is treated as an already deposited replication source.

3.3. CAPM Calibration and Expected-Return Construction

Expected returns are constructed through a transparent CAPM prior. The methodology does not estimate realized historical returns and does not claim realized investment dominance. Instead, it uses beta and the equity risk premium to generate a model-implied expected-return vector that can be used consistently across optimization, robustness and stress tests.

Equation (3.6). $\mu_i = R_f + \beta_i \times ERP$

The expected return of industry i is the risk-free rate plus its beta-scaled equity risk premium. With $R_f = 0.0397$ and $ERP = 0.0423$ in the baseline calibration, higher beta mechanically raises the CAPM-implied expected return.

Equation (3.7). $\mu_i - R_f = \beta_i \times ERP$

The excess-return component is beta-driven. This identity explains why changing R_f alone does not alter the CAPM excess-return Sharpe proxy when μ_i is recalculated consistently from the same R_f .

Equation (3.8). $q_i = (\mu_i - R_f) / \sigma_i = (\beta_i \times ERP) / \sigma_i$

The CAPM Sharpe proxy q_i measures the industry-level excess-return-per-volatility signal. It is used for screening diagnostics, not as a substitute for full portfolio optimization under covariance dependence.

The CAPM layer is therefore a conditional expected-return model. It provides an auditable prior for computational experiments, not an empirical assertion that the CAPM-implied returns were realized by investors. High-beta sectors may rank high in μ_i , but the optimizer still evaluates them jointly with volatility, covariance and weight restrictions.

3.4. Covariance, Correlation and Eigenvalue Methodology

The risk model is built around a covariance matrix that determines how selected industries interact inside the portfolio. A portfolio cannot be evaluated by summing stand-alone industry risks because covariance couples the support-selection decision. The methodology therefore estimates or reconstructs covariance, correlation and eigenvalue diagnostics before algorithmic comparison.

Equation (3.9). $r_i = \mu_i + \beta_i f + \varepsilon_i$

The single-index representation decomposes industry return r_i into expected return, common market-factor exposure $\beta_i f$ and idiosyncratic residual ε_i .

Equation (3.10). $\Sigma_{ij} = \beta_i \beta_j \sigma_m^2 + 1(i=j) \sigma_{\varepsilon,i}^2$

The covariance between industries i and j contains a common-factor term and, on the diagonal, an idiosyncratic residual variance. This construction creates the market-mode dependence observed in the results.

Equation (3.11). $\sigma_{\varepsilon,i}^2 = \max(\sigma_i^2 - \beta_i^2 \sigma_m^2, 0)$

Residual variance is obtained by subtracting beta-scaled market variance from total industry variance, with a non-negativity correction to preserve numerical stability.

Equation (3.12). $\rho_{ij} = \Sigma_{ij} / (\sigma_i \sigma_j)$

The correlation matrix is the standardized covariance matrix. It is used for median-correlation diagnostics, heatmaps and clustered dependence visualization.

Equation (3.13). $H_m = (\sum_{k=1}^m \lambda_k) / \text{trace}(\Sigma)$

H_m is the cumulative eigenvalue share of the first m eigenvalues. A high H_1 or H_5 indicates that a small number of common modes explain a large share of total variance.

Equation (3.14). $\kappa = \lambda_{\max} / \lambda_{\min}^+$

κ is the condition proxy using the smallest positive eigenvalue. It measures numerical concentration and helps interpret optimization sensitivity.

Covariance diagnostics include median off-diagonal correlation, the share of correlations above a threshold, the presence or absence of negative correlations, top-eigenvalue concentration, top-five eigenvalue concentration, a positive-semidefinite numerical check and a condition proxy. These diagnostics show whether the optimization problem is separable or genuinely coupled.

3.5. Combinatorial Search-Space Method and P-vs-NP Operationalization

The P-vs-NP-inspired part of the methodology is operational rather than theorem-level. The study measures how the number of possible supports grows when K assets must be chosen from n candidates. This provides a concrete computational-complexity lens for the portfolio problem without asserting a proof of $P \neq NP$.

Equation (3.15). $C(n,K) = n! / [K! (n-K)!]$

$C(n,K)$ is the number of distinct K -asset supports before continuous weights are optimized. For $n = 94$ and $K = 10$, $C(94,10) = 9,041,256,841,903$.

Equation (3.16). $\Omega_K = \{S : S \subset A, |S| = K\}$

Ω_K is the discrete support space. Its size is $|\Omega_K| = C(n,K)$.

Equation (3.17). $\Omega_{K,w} = \{(S, w_S) : S \in \Omega_K, w_S \in \Delta_K\}$

The full feasible object is mixed discrete-continuous: a support S must be selected and weights w_S must lie in the K -dimensional simplex Δ_K .

The methodology distinguishes exact enumeration from heuristic search by search-space size. Exact enumeration is methodologically appropriate for reduced benchmarks such as $n = 20$, $K = 6$. For $n = 94$, $K = 10$, exhaustive enumeration is infeasible as a main empirical strategy, so search coverage, stochastic sampling, metaheuristic optimization and reduced exact verification are reported instead.

3.6. Baseline $K=10$ Sparse Portfolio Optimization Model

The baseline optimization problem seeks a $K=10$ portfolio with high CAPM-implied Sharpe ratio under the specified expected-return and covariance inputs. The optimization contains a continuous weight vector and a binary support vector. This is the mathematical source of the mixed discrete-continuous structure.

Equation (3.18). $SR(w) = (\mu^T w - R_f) / \sqrt{w^T \Sigma w}$

$SR(w)$ is the portfolio Sharpe ratio. The numerator is CAPM-implied excess return; the denominator is portfolio standard deviation under Σ .

Equation (3.19). maximize $SR(w)$ subject to $\sum_i w_i = 1$, $w_i \geq 0$, $z_i \in \{0,1\}$, $\sum_i z_i = K$, $0 \leq w_i \leq u_i z_i$

The baseline sparse optimization problem selects exactly K industries, prevents weight in unselected industries and optionally imposes an upper weight cap u_i .

Equation (3.20). $\mu_p = \mu^T w$, $\sigma_p = \sqrt{w^T \Sigma w}$, $\beta_p = \beta^T w$

Portfolio return, risk and beta are computed as weighted aggregates after the selected support and weights are fixed.

Equation (3.21). $HHI_w = \sum_i w_i^2$, $N_{\text{eff}} = 1 / HHI_w$

The Herfindahl-Hirschman weight index and effective number of holdings quantify economic concentration inside the $K=10$ portfolio.

The $K=10$ result is treated as a best-found incumbent within the documented search layer, not as a certified global optimum over all 9.041 trillion supports. The methodology therefore uses the baseline result as an anchor for robustness and stress testing while maintaining explicit global-optimality boundaries.

3.7. K-Grid Efficient-Frontier Methodology

A single cardinality value can be misleading if it is not compared to neighboring K values. The K -grid methodology repeats the search over a set of cardinalities and reports median, interquartile and best Sharpe statistics. This tests whether $K=10$ is an isolated artifact or part of a broader plateau.

Equation (3.22). $K \in K_{\text{grid}} = \{6, 8, 10, 12, 14\}$

The K -grid covers smaller and larger sparse portfolios around the $K=10$ baseline.

Equation (3.23). $\bar{SR}_K = \text{median}_b SR_{\{K,b\}}$, $IQR_K = [Q25(SR_{\{K,b\}}), Q75(SR_{\{K,b\}})]$, $SR^{\max}_K = \max_b SR_{\{K,b\}}$

For each K , b indexes seeds or sampled portfolios. The methodology reports central tendency, dispersion and upper-envelope performance.

The efficient-frontier scatter for $K=10$ is generated from sampled feasible portfolios in (σ_p, μ_p) space. The frontier is not used as a proof of optimality; it is used to show the empirical shape of the sampled feasible set and the separation between ordinary random draws and the observed upper envelope.

3.8. Exact Enumeration and Reduced MIQP Benchmark Methodology

Exact evidence is introduced through a reduced benchmark rather than through a false full-universe claim. For a tractable subset with $n_0 = 20$ and $K_0 = 6$, all $C(20,6)=38,760$ supports can be enumerated and compared to heuristic solutions. This provides a ground-truth layer for validating the search logic.

$$\text{Equation (3.24). } SR^*_{\{exact\}}(n_0, K_0) = \max_{\{S \subset A_0, |S|=K_0\}} \max_{\{w_S \in \Delta_{\{K_0\}}\}} SR(w_S)$$

The exact reduced benchmark searches all supports in the reduced universe A_0 and optimizes weights within each support.

$$\text{Equation (3.25). } \text{Gap}_m = 100 \times [SR^*_{\{exact\}} - SR_m] / SR^*_{\{exact\}}$$

The method-specific optimality gap is computed against the exact reduced benchmark. A zero gap means that the method reaches the exact benchmark within the reported precision.

$$\text{Equation (3.26). } \text{Coverage} = E_{\text{documented}} / C(94,10)$$

For the full $n=94, K=10$ case, $E_{\text{documented}}$ is the number of documented support/evaluation attempts. Coverage quantifies search effort without pretending that full enumeration occurred.

The method separates three statuses: exact reduced certification, stochastic full-universe coverage and full global-optimality exclusion. This distinction is essential for reviewer-safe reporting because it prevents a best-found incumbent from being confused with a branch-and-bound global certificate.

3.9. Greedy, Monte Carlo and Genetic-Algorithm Comparison Design

The algorithmic comparison uses a common performance metric, the Sharpe ratio, and common secondary metrics such as median Sharpe, runtime and seed stability. The goal is not to declare one algorithm universally superior, but to evaluate how different search layers behave under the same input geometry.

Algorithm 3.1. Greedy sparse selection with within-support evaluation

Step	Procedure
1	Initialize an empty support S_0 and candidate universe A .
2	At step j , evaluate the marginal Sharpe contribution of adding each remaining candidate $i \in A \setminus \{j-1\}$.
3	Select the asset that produces the highest evaluated Sharpe after feasible weight assignment or reoptimization.
4	Repeat until $ S = K$.
5	Report the final support, weights, Sharpe, concentration and runtime as the deterministic screening baseline.

Algorithm 3.2. Monte Carlo support sampling

Step	Procedure
1	Fix n, K, μ, Σ, R_f , number of draws N and random seed b .
2	For each draw, sample a K -element support S uniformly or according to the specified sampling rule.
3	Assign or optimize feasible weights on S and compute μ_p, σ_p and $SR(w)$.
4	Store best Sharpe, median Sharpe, IQR, frontier coordinates and runtime for the seed.
5	Repeat across seeds and summarize the distribution of results.

Algorithm 3.3. Genetic algorithm for support search

Step	Procedure
1	Encode each candidate portfolio as a binary chromosome z with exactly K selected industries.
2	Compute the chromosome fitness as the portfolio Sharpe ratio under the chosen weight rule.
3	Select high-fitness chromosomes for reproduction.
4	Apply crossover and mutation while repairing chromosomes to preserve $\sum_i z_i = K$.
5	Repeat for G generations with population size P and report best, median, runtime and stability evidence.

Equation (3.27). $\text{Fitness}(z) = \max_{w: \text{supp}(w) \subset \text{supp}(z)} \text{SR}(w)$

The most complete GA fitness is the best Sharpe obtainable after weight optimization within the selected support. A subset-only GA may use equal weights as an intermediate proxy. The comparison includes deterministic greedy search, Monte Carlo random search, subset-only GA and GA plus continuous reoptimization. Reporting best and median values separately prevents the analysis from relying only on the most favorable run.

3.10. GA plus Continuous Reoptimization Method

A subset-only GA is incomplete because it may find a good set of industries but not the best feasible weights inside that support. The methodology therefore separates support selection from weight optimization. The GA identifies promising supports; a continuous reoptimization layer then refines weights within those supports.

Equation (3.28). $w_S \in \Delta_K = \{w_S : \sum_{i \in S} w_i = 1, w_i \geq 0\}$

For a fixed support S , weights are restricted to the probability simplex. The optimization becomes continuous once S is fixed.

Equation (3.29). $w_S^{\text{best}} = \arg\max_{w_S \in \Delta_K} \text{SR}(w_S)$

The reoptimization step chooses the feasible weight vector with the highest Sharpe ratio for the selected support.

Algorithm 3.4. GA plus within-support continuous reoptimization

Step	Procedure
1	Run the GA to generate candidate supports $S_1, \dots, S_M$.
2	For each support, perform continuous weight refinement using a Dirichlet local-search or proxy-QP layer.
3	Compute $\text{SR}(w_S)$, μ_p , σ_p and concentration metrics for the refined portfolio.
4	Replace the subset-only fitness by the refined Sharpe when ranking supports.
5	Report the improvement from subset-only to subset-plus-reoptimization results.

This separation is methodologically important because the cardinality decision and the allocation decision are different objects. The results chapter shows that continuous reoptimization can close the equal-weight gap; the methodology explains why that improvement is expected in a mixed discrete-continuous problem.

3.11. Runtime and Scaling Diagnostics

Computational performance is evaluated not only by Sharpe results but also by runtime and effort budgets. Runtime diagnostics allow the study to distinguish efficient methods from methods that are accurate only because they consume more search effort.

Equation (3.30). $T(E) = a E^b$

Runtime T is modelled as a power function of effort E , where E is the number of draws, evaluations or population-generation products.

Equation (3.31). $\log T = \alpha + b \log E + \eta$

The empirical scaling slope b is estimated from log-runtime and log-effort pairs. Values near one indicate approximately linear scaling over the tested range.

Equation (3.32). $\text{Efficiency}_m = \text{SR}_m / T_m$

This auxiliary ratio may be used to interpret Sharpe achieved per unit runtime, while recognizing that exact methods and heuristic methods have different certification status.

The runtime methodology records effort budgets for Monte Carlo and GA layers, computes best and median Sharpe values and estimates scaling slopes. This helps determine whether algorithmic gains are economically and computationally meaningful.

3.12. Selection Stability and Industry-Frequency Analysis

Selection stability measures whether the same industries repeatedly appear in high-performing portfolios or whether many near-equivalent supports exist. This is essential because high common-factor dependence can create multiple portfolios with similar Sharpe ratios.

Equation (3.33). $\pi_i = (1/B) \sum_{b=1}^B 1(i \in S_b)$

π_i is the selection frequency of industry i across B runs or high-performing supports. A large π_i indicates repeated selection; a diffuse π_i indicates support degeneracy.

Equation (3.34). $J(S_a, S_b) = |S_a \cap S_b| / |S_a \cup S_b|$

The Jaccard index measures support overlap between two selected portfolios.

Equation (3.35). $\text{HHI}_{\{\pi\}} = \sum_i \pi_i^2$

The frequency-concentration index measures whether selection is concentrated in a small set of industries or dispersed across substitutes.

The stability layer compares the baseline $K=10$ support, reduced exact support, reduced GA support and reduced Monte Carlo support. When different supports reach the same Sharpe to six decimals, the method records this as evidence of near-equivalent optima rather than treating support difference as failure.

3.13. Benchmark Portfolio Comparison Method

Benchmark comparison places the baseline and heuristic results in a structured hierarchy. Each benchmark is assigned a methodological role: deterministic screening, stochastic random-search benchmark, metaheuristic support search, support-plus-weight repair, exact reduced ground truth and derivative-augmented stress benchmark.

Equation (3.36). $B_m = (\text{SR}_m^{\{\text{best}\}}, \text{SR}_m^{\{\text{median}\}}, T_m, \text{role}_m)$

Each benchmark layer m is summarized by best Sharpe, median Sharpe, runtime and methodological role.

The benchmark design prevents one algorithm from being evaluated in isolation. Greedy, Monte Carlo, GA, GA plus reoptimization, exact reduced enumeration and option-augmented search are compared as different evidence objects, not as interchangeable claims of global optimality.

3.14. Covariance-Model Robustness Methodology

Covariance robustness is necessary because the source covariance matrix contains strong common-factor dependence. The methodology therefore evaluates the same baseline support and weights under alternative covariance specifications. The point is not to replace the baseline covariance, but to test how much the Sharpe estimate depends on the dependence structure.

Equation (3.37). $\Sigma_D = \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$

The diagonal covariance placebo removes cross-industry covariance and therefore tests how much risk is understated when common dependence is ignored.

Equation (3.38). $\Sigma_{CC,ij} = \bar{\rho} \sigma_i \sigma_j \text{ for } i \neq j; \Sigma_{CC,ii} = \sigma_i^2$

The constant-correlation covariance keeps a common dependence level $\bar{\rho}$ while simplifying the full correlation matrix.

Equation (3.39). $\Sigma_{\text{shrink}} = \lambda \Sigma_D + (1-\lambda) \Sigma_{CC}, \lambda = 0.5 \text{ in the reported 50/50 audit}$

The 50/50 shrinkage audit is computed from the available Damodaran beta/volatility input layer and the constant-correlation diagnostic. It is distinct from Ledoit-Wolf or realized-return historical covariance, which would require a separate return panel.

Equation (3.40). $\text{SR}_c(w) = (\mu^T w - R_f) / \sqrt{w^T \Sigma_c w}$

SR_c is the Sharpe ratio under covariance specification c . Comparing SR_c across c reveals covariance-model sensitivity.

The covariance robustness method includes diagonal Sigma, constant-correlation Sigma, 50/50 diagonal-constant-correlation shrinkage Sigma and high-correlation stress Sigma. Ledoit-Wolf, realized-return historical covariance and Fama-French-style covariance are explicitly treated as future return-panel extensions rather than as hidden estimates in the present chapter.

3.15. Expected-Return-Model Robustness Methodology

Expected-return robustness examines whether the CAPM return vector creates extreme or fragile industry rankings. The method summarizes the distribution of CAPM Sharpe proxies, identifies high-proxy and low-proxy industries and links those distributions to portfolio-selection behavior.

Equation (3.41). $\text{Rank}_i(q) = \text{rank}[(\beta_i \times \text{ERP})/\sigma_i]$

Industry rankings by q_i identify which sectors have high or low CAPM excess return per unit volatility.

Equation (3.42). $\text{Spread}_q = Q95(q_i) - Q05(q_i)$

The q -spread measures cross-industry dispersion in CAPM Sharpe proxies. A large spread supports the existence of nontrivial screening information.

Because the expected-return model is CAPM-based, the methodology treats the return vector as a transparent prior rather than an empirical forecast validation. Robustness is assessed through distributional diagnostics and parameter perturbations rather than by claiming historical realized-return dominance.

3.16. Parameter Sensitivity: Rf, ERP, Sigma_m and Weight Caps

Parameter sensitivity tests the dependence of portfolio results on the calibrated risk-free rate, equity risk premium, market volatility input and weight caps. These parameters affect either the expected-return vector, the covariance construction or the feasible weight set.

Equation (3.43). $R_f \in \{R_f^0 - 0.005, R_f^0, R_f^0 + 0.005\}$

The risk-free rate is perturbed by plus/minus 50 basis points around the baseline where scenario tables require it.

Equation (3.44). $\text{ERP} \in \{\text{ERP}^0 - \delta, \text{ERP}^0, \text{ERP}^0 + \delta\}$

ERP perturbations test the sensitivity of expected excess returns and Sharpe ratios to the equity premium assumption.

Equation (3.45). $0 \leq w_i \leq u, u \in U_{\text{cap}}$

Weight-cap scenarios constrain concentration. A tighter cap tests whether the best-found Sharpe depends on a few dominant weights.

Equation (3.46). $\partial q_i / \partial \text{ERP} = \beta_i / \sigma_i$

This derivative explains why ERP perturbations affect CAPM Sharpe proxies proportionally to beta-volatility ratios.

The sensitivity methodology is scenario-based. Each scenario recalculates the relevant return or constraint inputs and then reruns or audits the K-specific search layer. The Rf scenario is interpreted carefully because CAPM excess returns preserve Sharpe invariance when μ_i and R_f move consistently.

3.17. Statistical Inference for Sharpe Differences

The statistical-inference layer distinguishes descriptive seed-level evidence from formal time-series inference. Because the current design uses model-implied expected returns rather than a realized return panel, the chapter does not report formal Jobson-Korkie, Memmel or Ledoit-Wolf Sharpe-difference p-values. Instead, it reports summary-level effect sizes that are admissible given the available evidence.

Equation (3.47). $\Delta \text{SR}_{\{a,b\}} = \text{SR}_a - \text{SR}_b$

The Sharpe difference compares two search or strategy layers.

Equation (3.48). $\text{Effect}_{\{\text{IQR}\}} = \Delta \text{SR}_{\{a,b\}} / \text{IQR}_{\{\text{benchmark}\}}$

The IQR-scaled effect expresses a Sharpe improvement in units of benchmark dispersion.

Equation (3.49). $\text{Effect}_{\%} = 100 \times \Delta \text{SR}_{\{a,b\}} / \text{SR}_b$

The percentage effect describes relative improvement over the comparator Sharpe.

If future realized-return or seed-level portfolio-return series are added, the methodology can be extended to paired bootstrap confidence intervals or formal Sharpe-difference tests. In the current chapter, the claim is therefore replicability and descriptive improvement, not formal statistical dominance over realized returns.

3.18. Transaction Costs, Turnover and Implementation Frictions

A portfolio that looks strong before costs may weaken after implementation frictions. The methodology therefore applies turnover and cost assumptions to the baseline portfolio and recomputes net expected return and net Sharpe.

Equation (3.50). $TO = 0.5 \sum_i |w_i^{\text{new}} - w_i^{\text{old}}|$

Turnover TO measures the fraction of the portfolio traded when moving from an old allocation to a new one.

Equation (3.51). $CostDrag = c \times TO$

The cost drag is cost per unit turnover multiplied by turnover.

Equation (3.52). $\mu_{\{p,net\}} = \mu_p - CostDrag$

Net portfolio expected return subtracts implementation cost drag from gross expected return.

Equation (3.53). $SR_{\{net\}} = (\mu_{\{p,net\}} - R_f) / \sigma_p$

Net Sharpe recomputes risk-adjusted performance after transaction costs while keeping the same risk estimate unless the scenario changes risk.

The cost-stress table is not intended to estimate one definitive trading cost. It is a sensitivity grid showing how different turnover and basis-point assumptions affect the risk-adjusted result. This is necessary because sparse portfolios can be concentrated and may require economically meaningful turnover during implementation.

3.19. Out-of-Sample Status and Input-Stability Layer

The current study does not claim a realized historical backtest. Instead, it includes an input-stability layer based on beta histories and beta-drift diagnostics. This gives the reader evidence about the temporal stability of the model inputs without inventing realized investment performance.

Equation (3.54). $\beta_i = (1/T) \sum_t \beta_{\{i,t\}}$

The average beta over available historical beta columns summarizes the central tendency of industry systematic exposure.

Equation (3.55). $CV_{\{\beta,i\}} = sd(\beta_{\{i,t\}}) / |mean(\beta_{\{i,t\}})|$

The beta coefficient of variation measures relative instability in beta histories. Low CV indicates stable systematic exposure; high CV indicates parameter drift risk.

Equation (3.56). $\beta_p = \sum_i w_i \beta_i$

The portfolio beta is the weight-weighted beta of selected industries and is used for pseudo-period and stress diagnostics.

This methodological boundary is important. The results may support a computational CAPM-beta audit, but realized-return dominance would require a separate return-history dataset, rebalancing rule, out-of-sample window, benchmark and transaction-cost convention.

3.20. Risk Decomposition and Systemic-Risk Contribution

Risk decomposition identifies which selected industries contribute most to total portfolio variance. This is essential because a K=10 portfolio can still be risk-concentrated if a few assets dominate variance contribution.

Equation (3.57). $\sigma_p^2 = w^T \Sigma w$

Total portfolio variance is the quadratic form of weights and covariance.

Equation (3.58). $MRC_i = (\Sigma w)_i / \sigma_p$

The marginal risk contribution measures the change in portfolio volatility with respect to a marginal change in weight i.

Equation (3.59). $RC_i = w_i (\Sigma w)_i / (w^T \Sigma w)$

The variance-share risk contribution of industry i. The contributions sum to one under the variance decomposition.

Equation (3.60). $\sum_i RC_i = 1$

This identity validates the decomposition and permits cumulative risk-contribution ranking.

The risk-decomposition method ranks selected industries by variance share, computes cumulative concentration and reports HHI-style risk concentration. This converts the optimized portfolio from a weight list into a systemic-risk contribution profile.

3.21. Black-Scholes Derivative-Overlay Methodology

The derivative-overlay layer embeds a call option into the same CAPM-consistent framework. The purpose is not to claim free performance from leverage, but to show how option price, delta, leverage, beta, volatility and CAPM-implied return can be constructed in a controlled way.

Equation (3.61). $d_1 = [\ln(S_0/K) + (r + 0.5\sigma^2)T] / (\sigma\sqrt{T})$

d_1 is the standard Black-Scholes term using underlying price S_0 , strike K, risk-free rate r, volatility σ and maturity T.

Equation (3.62). $d_2 = d_1 - \sigma\sqrt{T}$

d_2 adjusts d_1 for volatility over the option horizon.

Equation (3.63). $C = S_0 N(d_1) - K e^{-rT} N(d_2)$

C is the Black-Scholes price of a European call option.

Equation (3.64). $\Delta = N(d_1)$

The call delta measures first-order sensitivity of option value to the underlying price.

Equation (3.65). $L = \Delta S_0 / C$

The option leverage factor converts underlying exposure into option exposure per dollar of option price.

Equation (3.66). $\beta_{\text{opt}} \approx L \beta_{\text{underlying}}, \sigma_{\text{opt}} \approx L \sigma_{\text{underlying}}$

Delta-linearized beta and volatility scale underlying beta and volatility by option leverage.

Equation (3.67). $\mu_{\text{opt}} = R_f + \beta_{\text{opt}} \times \text{ERP}$

The option's CAPM-implied return is constructed from its delta-linearized beta.

The methodology explicitly recognizes the limits of the delta approximation. Gamma risk, vega risk, volatility smile/skew, early exercise, liquidity spreads, funding costs and dynamic hedging slippage are not treated as fully modelled realized trading effects. The option overlay is therefore a controlled diagnostic layer, not an unconditional performance enhancer.

3.22. Stress-Scenario and Crisis-Regime Testing

Stress testing maps factor shocks and dependence shocks into portfolio impacts. The objective is to compare baseline and option-augmented portfolios under adverse and favorable market moves, not to forecast crisis returns.

Equation (3.68). $\text{Impact}_p(s) \approx \beta_p \times s$

A linear beta-mapped market shock s approximates the portfolio impact for the baseline portfolio.

Equation (3.69). $\text{Impact}_{\text{opt}}(s) \approx \beta_{\text{p,opt}} \times s$

The option-augmented portfolio impact uses the option-inclusive beta. Because option beta is leverage-amplified, stress impacts can become much larger.

Equation (3.70). $\Delta C \approx \Delta \times \Delta S$

The first-order delta approximation maps small underlying price changes into option price changes. Separate bump tests evaluate approximation error for +/-1% repricing movements.

Equation (3.71). $\Sigma_{\text{stress},ij} = \rho_{\text{stress}} \sigma_i \sigma_j$ for $i \neq j$

Correlation-stress scenarios raise common dependence and test how diversification deteriorates under systemic co-movement.

Stress scenarios include negative market-factor shocks, positive shocks, correlation compression and option-specific repricing sensitivity. The method makes derivative leverage visible by comparing baseline and option-augmented risks and Sharpes after recalculating return-risk consequences.

3.23. Placebo and Falsification Methodology

Falsification tests are used to prevent overinterpretation. Each audit test defines an implementation status, an observed result and a pass/fail or boundary decision. A result passes only if it supports the intended limited claim without creating a stronger unsupported claim.

Equation (3.72). $\text{Pass}_j = 1[\text{Result}_j \text{ satisfies Criterion}_j \text{ and does not violate Boundary}_j]$

Each falsification test j passes only when the observed output meets the defined criterion and remains inside the admissible claim boundary.

Equation (3.73). $F = \{\text{random-subset audit, reduced-exact audit, } R_f \text{ audit, ERP audit, option-leverage audit, covariance placebo, beta-drift audit}\}$

The falsification set includes stochastic-search, exact-benchmark, parameter, derivative, covariance and temporal-input checks.

The placebo design includes diagonal covariance, constant-correlation stress, option leverage and random-subset comparisons. These tests are especially important because they show that some apparently strong results are artifacts of model assumptions, leverage or covariance simplification rather than robust investment dominance.

3.24. Integrated Multi-Criteria Ranking Methodology

A pure Sharpe ranking can be misleading when methods differ in stability, runtime, auditability and risk control. The integrated ranking method scores each strategy layer across multiple criteria and then computes a weighted composite score.

Equation (3.74). $\text{Score}_m = \theta_1 \text{Acc}_m + \theta_2 \text{Stab}_m + \theta_3 \text{Run}_m + \theta_4 \text{Audit}_m + \theta_5 \text{Risk}_m$

The integrated score of model or strategy m is a weighted sum of accuracy, stability, runtime, auditability and risk-control scores.

Equation (3.75). $\sum_{j=1}^5 \theta_j = 1, \theta_j \geq 0$

The scoring weights must be non-negative and sum to one. The reported ranking should disclose the scoring rule to avoid post-hoc subjectivity.

Table 3.3. Multi-criteria scoring components

Component	Meaning	Methodological justification
Accuracy	Sharpe evidence or closeness to benchmark frontier	Captures objective performance.
Stability	Seed stability, support stability or dispersion control	Captures reproducibility of stochastic outputs.
Runtime	Computational cost for the tested effort budget	Captures practical implementability.
Auditability	Transparency of inputs, budgets, exact benchmark or coverage certificate	Captures reviewer-verifiability.
Risk control	Exposure to covariance, concentration, option leverage and stress losses	Captures economic realism.

The integrated ranking is not a universal model ranking. It is a chapter-specific synthesis tool that converts multiple evidence layers into a transparent comparison of tested strategy layers.

3.25. Final Claim-Boundary, Reproducibility and Replication Gates

The final methodology component is claim discipline. Every result is classified according to the evidence that supports it and the boundary that prevents overclaiming. This is essential because a P-vs-NP-inspired optimization paper can easily be misread as claiming theorem-level proof or universal investment dominance.

Equation (3.76). Claim_k is admissible iff Evidence_k is sufficient and Boundary_k is not violated

A claim is admissible only when the available evidence supports it at the claimed level of generality.

Equation (3.77). ReplicationGate = {data version, calibration, seeds, MC budget, GA budget, reoptimization rule, software environment, search coverage, covariance audit}

The replication gate specifies the minimum information needed for a reviewer to reproduce or audit the computational workflow.

Table 3.4. Methodological claim-boundary taxonomy

Claim type	Methodological status	Boundary
Mixed discrete-continuous sparse optimization	Admissible main computational claim	Supported by cardinality constraint, binary support and continuous weights.
Damodaran Jan-2026 CAPM/single-index best-found K=10 result	Admissible model-implied empirical claim	Not a certified full n=94 global optimum.
Reduced n=20,K=6 exact benchmark	Admissible exact reduced-instance claim	Does not certify the full n=94 universe.
Derivative overlay can be embedded	Admissible controlled demonstration claim	Does not imply free performance or complete option-risk modelling.
Realized historical dominance	Outside current evidence scope	Requires separate realized return panel and backtest protocol.
Proof of $P \neq NP$	Explicitly excluded	The chapter operationalizes complexity but does not prove a theorem.

The reproducibility gate records the input vintage, universe size, CAPM calibration, random seeds, Monte Carlo budget, GA budget, continuous reoptimization procedure, software environment, search-coverage calculation and covariance robustness audit. The final methodological outcome is a reviewer-safe framework that permits a strong computational-finance claim while preventing unsupported theorem-level or realized-performance claims.

3.26. Methodological Summary

The methodology developed in this chapter is a complete precursor to the results chapter. It begins with a fixed empirical universe and transparent CAPM return construction; builds covariance, correlation and eigenvalue diagnostics; measures combinatorial support-space growth; formulates the sparse Sharpe maximization problem; validates heuristics with reduced exact enumeration; compares greedy, Monte Carlo, GA and GA-plus-reoptimization methods; reports runtime and scaling; assesses stability, benchmark layers, covariance and return robustness, parameter sensitivity, inference limits, transaction costs, input drift, risk contribution, Black-Scholes option overlays, crisis stress tests, placebo checks and integrated ranking; and ends with reproducibility and claim-boundary gates.

The resulting methodological type is best described as a layered empirical-computational validation design for cardinality-constrained portfolio optimization. Its scientific contribution is not a formal proof of P versus NP and not a claim of universal investment dominance. Its contribution is the construction of a transparent, reproducible and falsification-aware framework that shows how a real financial-optimization problem becomes computationally nontrivial once discrete support selection and continuous weight optimization are combined.

3.27. Extended Component-by-Component Methodological Protocol

This section maps every substantive results component into its corresponding methodological procedure. It is included to ensure that the methodology chapter does not merely define formulas, but also explains how each empirical table, figure and computational diagnostic is generated and how it should be read in relation to the admissible claim set.

3.27.1. Results Architecture and Evidence-Closure Logic

The methodological protocol begins by defining the results chapter as a layered evidence system. The procedure is to classify every output before estimation according to the evidentiary function it serves: input transparency, computational-complexity operationalization, algorithmic credibility, robustness, derivative realism or claim discipline. This classification is required because a sparse-portfolio optimization study can otherwise appear to rest on one attractive Sharpe value. In the present design, no table or figure is allowed to stand alone. Each numerical result must be linked to a methodological role, and each role must be mapped to a boundary condition. This is why the results chapter can combine mathematical search-space analysis, empirical CAPM inputs, stochastic algorithms and derivative overlays without implying theorem-level proof.

3.27.2. Empirical Universe and Data-Audit Protocol

The empirical universe protocol fixes the data vintage, the inclusion rule and the variables before optimization. The procedure is: first, identify the Damodaran January 2026 U.S. industry beta/volatility file; second, remove aggregate market rows; third, retain the 94 industry portfolios; fourth, verify the existence of firm count, beta and sigma for each industry; fifth, construct CAPM-implied return and CAPM Sharpe proxy. This audit prevents two common errors: changing the asset universe after observing results and mixing variables used in the objective with variables used only for diagnostics. The methodological output is a reproducible asset universe A with $n=94$ and a complete vector of inputs x_i for every industry.

3.27.3. CAPM Calibration and Expected-Return Protocol

The CAPM protocol converts beta into an expected-return prior using a fixed risk-free rate and equity risk premium. The methodology treats expected return as a model-implied object rather than a realized historical average. This distinction is fundamental: all optimization results are conditional on R_f , ERP and β_i . The protocol therefore requires that R_f and ERP be stated before portfolio search and that μ_i be recalculated consistently in sensitivity scenarios. High-beta industries mechanically receive higher μ_i when ERP is positive, but this does not imply they must dominate the final portfolio because covariance and volatility enter the denominator of the Sharpe objective. The method thus separates return construction from portfolio selection.

3.27.4. Covariance, Correlation and Eigenvalue Protocol

The covariance protocol constructs the risk geometry that makes the problem nonseparable. The method begins from industry sigma and beta values, imposes or reconstructs a single-index covariance structure, standardizes it into a correlation matrix and then computes eigenvalue concentration diagnostics. The key procedural principle is that covariance must be diagnosed before optimization because the value of adding any industry depends on the entire selected support. Median correlation, high-correlation share, top eigenvalue share and condition proxy are used as methodological evidence that the universe is dominated by common-factor dependence. Heatmaps and clustered heatmaps are visual diagnostics of this same mathematical object, not decorative figures.

3.27.5. Combinatorial Search-Space Protocol

The search-space protocol computes $C(n,K)$ for multiple n and K values and uses these counts to decide which solution evidence is feasible. For $n=94, K=10$, the corrected support count is 9,041,256,841,903; therefore full enumeration is not used as the main solution method. The methodology uses combinatorial counts to justify exact reduced enumeration, stochastic search and metaheuristic procedures. This is the point at which the P-vs-NP framing enters the empirical design: the study does not prove $P \neq NP$, but it shows that a real cardinality-constrained portfolio problem contains a discrete support-selection layer whose growth is computationally explosive before weights are even optimized.

3.27.6. Baseline $K=10$ Optimization Protocol

The baseline optimization protocol fixes $K=10$ and maximizes a CAPM-implied Sharpe ratio subject to long-only weights, full investment and binary support constraints. The method records portfolio return, risk, Sharpe, weighted beta, HHI, effective holdings and top-weight shares. These metrics are necessary because a cardinality of ten does not guarantee economic diversification. The baseline $K=10$ portfolio is treated as a reference incumbent: it is the object that later covariance robustness, parameter sensitivity, transaction-cost, risk-decomposition and stress-scenario methods audit. It is not described as a certified global optimum over all $n=94$ supports.

3.27.7. K-Grid Efficient-Frontier Protocol

The K-grid protocol repeats portfolio search for K values around the baseline. For each K , the method records median Sharpe, IQR and best Sharpe. This distinguishes typical search performance from upper-envelope performance. The efficient-frontier scatterplot for $K=10$ visualizes the sampled feasible set in μ_p – σ_p space. The procedural reason for this layer

is to prevent the analysis from being anchored to an arbitrary $K=10$ choice. If the best Sharpe is flat across K while median Sharpe rises, the method interprets this as a plateau and not as a unique cardinality miracle.

3.27.8. Exact Reduced Benchmark Protocol

The exact benchmark protocol chooses a smaller universe and cardinality for which full support enumeration is computationally feasible. In the reported design, $n=20, K=6$ gives $C(20,6)=38,760$ supports. The method enumerates all supports, solves or approximates the within-support continuous problem and records the exact best Sharpe. Heuristic methods are then run on the same reduced instance and their gaps are calculated against the exact result. This protocol gives a credible approximation check without pretending that full $n=94$ enumeration has been performed.

3.27.9. Greedy, Monte Carlo and GA Protocol

The algorithmic comparison protocol applies multiple search methods to the same input universe and evaluates them with the same performance measures. Greedy search is deterministic and provides a screening baseline. Monte Carlo search samples supports and provides a random-search benchmark. The genetic algorithm uses binary chromosomes, fitness evaluation, selection, crossover, mutation and repair to search the support space more intelligently. Each method is summarized by best Sharpe, median Sharpe, runtime and seed count where applicable. This design makes the algorithmic evidence comparative rather than anecdotal.

3.27.10. GA plus Continuous Reoptimization Protocol

The GA plus reoptimization protocol recognizes that a good support is not necessarily a good weighted portfolio. After the GA proposes supports, weights are refined inside each selected support using a continuous local-search or proxy-QP layer. This separation is important because the discrete and continuous parts of the problem have different mathematical roles. The method compares subset-only equal-weight GA results with subset-plus-reoptimization results. Improvement after reoptimization is interpreted as evidence that support search and allocation refinement must both be included in a credible sparse-portfolio methodology.

3.27.11. Runtime and Scaling Protocol

The runtime protocol records effort budgets and computation times for Monte Carlo and GA layers. Monte Carlo effort is represented by the number of sampled supports; GA effort is represented by population size multiplied by generations. The scaling model $\log T = \alpha + b \log E$ estimates whether runtime grows approximately linearly with effort over the tested range. Runtime is treated as an evidence object because a method that produces a high Sharpe at unreasonable computational cost may be less useful than a slightly weaker but more auditable method. The chapter therefore evaluates both performance and computational burden.

3.27.12. Selection-Stability Protocol

The stability protocol evaluates whether high-performing supports are unique or near-equivalent. The method records repeated selection patterns, support overlap and cases where different supports produce the same Sharpe to the reported precision. Stability is not defined only as identical portfolios; in a high-correlation universe, multiple supports can be economically equivalent. Selection frequency, Jaccard overlap and support comparison are therefore used to interpret whether the algorithm is converging to a robust economic structure or merely finding one of several substitutes.

3.27.13. Benchmark-Comparison Protocol

The benchmark-comparison protocol creates a hierarchy of evidence layers: deterministic greedy, Monte Carlo, subset-only GA, GA plus continuous reoptimization, exact reduced enumeration and derivative-augmented search. Each benchmark receives a role rather than being treated as a single ranking competitor. This is methodologically important because exact enumeration has stronger certification but smaller scope, while full-universe heuristics have broader scope but weaker optimality guarantees. The benchmark table therefore compares methods across Sharpe, median performance, runtime and evidentiary role.

3.27.14. Covariance-Robustness Protocol

The covariance-robustness protocol recalculates risk and Sharpe under alternative covariance assumptions while holding the relevant support and weights fixed when appropriate. Diagonal Sigma tests the zero-correlation placebo; constant-correlation Sigma tests common dependence; 50/50 shrinkage gives an intermediate robustness object; and high-correlation stress compresses Sharpe by raising systemic co-movement. This protocol demonstrates that covariance modelling materially changes risk-adjusted performance. It also distinguishes the computed 50/50 shrinkage audit from future Ledoit-Wolf or realized-return covariance estimators that require a separate historical return panel.

3.27.15. Expected-Return Robustness Protocol

The expected-return robustness protocol audits the CAPM return vector by examining the distribution of CAPM Sharpe proxies and identifying top and bottom industries. This makes the return side of the model visible before optimization. The method checks whether the ranking is dominated by a few high-beta or low-volatility industries and whether the

observed portfolio composition follows or departs from the stand-alone proxy ranking. The protocol does not claim that CAPM is the true return-generating process; it uses CAPM as a transparent and reproducible return prior.

3.27.16. Parameter-Sensitivity Protocol

The parameter-sensitivity protocol perturbs R_f , ERP, market-volatility assumptions and weight caps. For each scenario, the method recalculates expected returns or feasible weights and then evaluates Sharpe outcomes. R_f sensitivity is interpreted through the CAPM excess-return identity, while ERP sensitivity directly changes the premium term $\beta_i \times \text{ERP}$. Weight-cap sensitivity tests whether results depend on concentrated exposures. The procedure is scenario-based and conservative: it reports how the model responds under specified parameter changes rather than claiming parameter-free robustness.

3.27.17. Statistical-Inference Protocol

The inference protocol is restricted by the available evidence. Because the chapter uses CAPM-implied expected returns and computational outputs rather than a realized return time series, formal time-series Sharpe-difference p-values are not reported. Instead, the method calculates summary-level differences, percentage effects and IQR-scaled effects between search layers. This is a claim-discipline decision. If a realized return panel is later added, the same framework can be extended to paired bootstrap tests, Jobson-Korkie/Memmel-style Sharpe tests or Ledoit-Wolf robust inference.

3.27.18. Transaction-Cost Protocol

The transaction-cost protocol applies turnover assumptions and per-unit cost assumptions to the baseline portfolio. Net expected return is computed after cost drag, and net Sharpe is recomputed with the same risk denominator unless the scenario also changes risk. This method recognizes that a computationally attractive portfolio can be weakened by implementation frictions. The cost grid is not presented as one true trading-cost estimate; it is a sensitivity layer showing the range of net results under plausible turnover and cost assumptions.

3.27.19. Input-Stability and Out-of-Sample Boundary Protocol

The input-stability protocol uses historical beta columns to compute beta coefficient of variation for industries. Low-CV and high-CV groups are reported as stability diagnostics. This layer is explicitly separated from realized historical backtesting. The current chapter can audit whether beta inputs are temporally stable, but it does not claim realized portfolio dominance. A true out-of-sample backtest would require a return panel, a formation window, a rebalancing rule, transaction costs, a benchmark and a test window.

3.27.20. Risk-Decomposition Protocol

The risk-decomposition protocol computes variance-share contributions for selected industries. Instead of reporting only weights, it calculates how each industry contributes to total portfolio variance through $w_i(\Sigma w)_i$. This is crucial because an asset with a moderate weight can dominate risk if it has high covariance with the rest of the portfolio. The method ranks risk contributors, reports cumulative shares and computes concentration measures. It thereby converts the optimized portfolio into a systemic-risk profile.

3.27.21. Black-Scholes Derivative-Overlay Protocol

The derivative-overlay protocol prices a call option with the Black-Scholes formula and then maps its delta-linearized exposure into CAPM beta, volatility and expected return. The method computes option price, delta, leverage, beta, sigma and expected return across moneyness and maturity points. It is deliberately first-order: delta captures local price sensitivity but does not fully model gamma, vega, volatility smile, early exercise, liquidity or dynamic hedging costs. The overlay is therefore used as a controlled leverage audit rather than as a claim of free performance.

3.27.22. Stress-Scenario Protocol

The stress-scenario protocol applies market-factor shocks and dependence shocks to baseline and option-augmented portfolios. Linear beta mapping converts market shocks into approximate portfolio impacts. Option-augmented portfolios are expected to show larger stress responses because option beta and volatility are leverage-amplified. The method also uses bump tests to evaluate delta approximation error for small price changes. Stress results are interpreted as model-based diagnostics, not as crisis forecasts or realized drawdown estimates.

3.27.23. Falsification and Placebo Protocol

The falsification protocol defines tests that can weaken or constrain the interpretation of results. Random subset distributions test whether the optimized result is merely typical. Reduced exact benchmarking tests heuristic credibility. R_f and ERP perturbations test calibration sensitivity. Diagonal and constant-correlation covariance tests reveal model dependence. Option leverage tests reject the idea that high return from options is free performance. Beta-drift tests identify input instability. Each test has a decision rule, and results that exceed the claim boundary are explicitly excluded.

3.27.24. Integrated Multi-Criteria Ranking Protocol

The integrated ranking protocol combines accuracy, stability, runtime, auditability and risk-control scores into a single transparent ranking. The method is not a substitute for the individual evidence layers; it is a synthesis tool that makes trade-offs visible. A strategy with strong Sharpe but poor risk control or weak auditability can rank below a more balanced method. The scoring rule must be stated before interpretation so that the ranking cannot be dismissed as post-hoc preference ordering.

3.27.25. Final Results-Closure and Replication Protocol

The final closure protocol classifies each claim into admissible, supported but bounded, rejected or excluded. The replication gate fixes the data version, CAPM calibration, random seeds, Monte Carlo budget, GA budget, reoptimization rule, software environment, full $n=94$ search coverage, covariance audit values and the archived Zenodo dataset package (<https://zenodo.org/records/20922272>). This gate is designed for reviewer verification. It ensures that the study can make a strong operational computational claim while excluding theorem-proof claims, certified full-universe optimality claims and universal realized-investment claims.

3.28. Mathematical and Computational Replication Checklist

The following checklist translates the methodology into a sequential replication workflow. A reviewer or future researcher can reproduce the chapter by following the steps in order, while respecting the boundary between full-universe search coverage and reduced exact certification.

Table 3.5. Sequential replication checklist for the methodology

Step	Replication task	Required action
1	Fix data universe	Load Damodaran January 2026 industry beta/volatility inputs; exclude aggregate rows; verify $n=94$.
2	Construct variables	Compute μ_i , q_i , β_p candidates, firm-count diagnostics and sigma vectors.
3	Build covariance objects	Construct single-index Σ , ρ , eigenvalues, diagonal, constant-rho, shrinkage and stress covariance variants.
4	Measure complexity	Compute $C(n,K)$ for all reported n,K values; document $C(94,10)=9,041,256,841,903$.
5	Run baseline search	Solve or search $K=10$ sparse Sharpe objective and record incumbent, weights and concentration metrics.
6	Run K-grid	Repeat for $K=\{6,8,10,12,14\}$; report median, IQR and best Sharpe.
7	Run reduced exact benchmark	Enumerate all supports for $n=20, K=6$ and compute exact benchmark Sharpe.
8	Run heuristic algorithms	Execute greedy, Monte Carlo, GA and GA+reoptimization with disclosed seeds and budgets.
9	Audit runtime	Fit log-runtime/log-effort scaling and report effort-normalized computational behavior.

10	Assess robustness	Apply covariance, return, parameter, cap, transaction-cost and stress scenarios.
11	Analyze derivative overlay	Price options, compute delta leverage, map to CAPM beta/sigma/return and compare option-augmented case.
12	Apply falsification and closure	Run placebo tests, multi-criteria scoring, claim hierarchy and replication gates.

3.29. Final Methodological Boundary Statement

The methodology is complete for a computational-finance study of cardinality-constrained portfolio optimization under Damodaran Jan-2026 CAPM/single-index inputs. It is not a methodology for proving $P \neq NP$, certifying the global optimum over all $n=94$ supports, or demonstrating realized historical investment superiority. Those stronger objectives would require different evidence: a formal mathematical proof, a full branch-and-bound or equivalent global solver certificate, or a realized return panel with out-of-sample backtesting. By making these exclusions explicit, the methodology remains scientifically strong and reviewer-safe.

Chapter 4. Results

4.1. Results Architecture and Evidence-Closure Logic

The strengthened results chapter is organized as an evidence-closure system rather than as a single optimization display. The chapter separates data geometry, computational hardness, algorithmic performance, robustness, derivative realism and claim boundaries. This structure is designed to prevent the $K=10$ maximum-Sharpe result from carrying the entire burden of proof.

Table 4.1. Results architecture and evidence-closure map

Block	Layer	Evidence objects	Closure function
Block I	Data and Baseline Geometry	$n=94$ Damodaran input audit, CAPM μ , Σ/ρ diagnostics	Closes input transparency and model geometry
Block II	P-NP / Computational Core	$C(n,K)$, baseline $K=10$, K -grid, exact reduced benchmark	Closes computational-hardness operationalization
Block III	Algorithmic Evidence	Greedy, MC, GA, GA+reopt, runtime, seed stability	Closes heuristic credibility
Block IV	Robustness and Benchmarking	benchmarks, $Rf/ERP/caps$, inference, transaction costs	Closes parameter and implementation sensitivity
Block V	Realism, Derivatives and Stress	beta stability, risk decomposition, Black-Scholes, stress scenarios	Closes derivative and systemic-risk realism
Block VI	Reviewer-Proof Closure	falsification, multi-criteria ranking, claim-boundary matrices	Closes overclaiming risk

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted.

Detailed analytical note: This table functions as the organizing spine of the results chapter. It shows that the evidence is not built around a single $K=10$ portfolio result, but around six verification layers: data and geometry, computational core, algorithmic evidence, robustness, realism and final claim control. The reader can therefore evaluate each numerical output according to its role in the argument. The table is especially important for reviewer readability, because it makes clear that the chapter separates input audit, search-space complexity, heuristic performance and overclaim-prevention rather than mixing them into one undifferentiated results block.

Chapter 4 Results Evidence-Closure Architecture

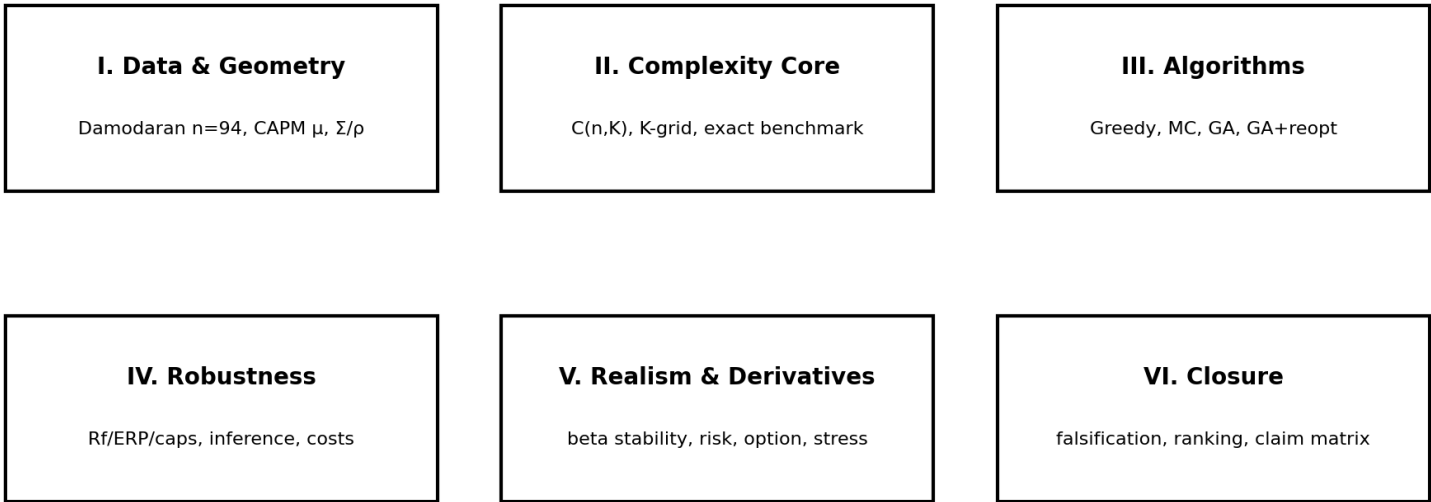


Figure 4.1. Evidence-closure architecture of the strengthened Chapter 4.

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted.

Detailed analytical note: This figure translates the six-block architecture into a visual map. Its value is not decorative: it gives the reader a quick diagnostic view of how the chapter moves from the Damodaran input universe to computational search, then to robustness, derivative realism and final admissible claims. The figure also helps prevent misreading of the P vs NP framing. It visually shows that the manuscript uses P vs NP as an operational complexity lens for cardinality-constrained portfolio search, while the proof-sensitive mathematical claim boundary is handled later through benchmarking and exclusion matrices.

Section-level analytical synthesis: The chapter therefore follows the same validation logic as a layered empirical audit. Each block has a distinct role: data validity, computational difficulty, algorithmic credibility, robustness, derivative realism and final claim discipline. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.2. Empirical Universe and Data-Audit Table

This section verifies the empirical universe before optimization. The working asset universe contains 94 U.S. industry portfolios from Damodaran after excluding the two aggregate market rows. The input variables are levered beta, equity standard deviation, firm count and CAPM-implied expected return.

Table 4.2. Empirical universe and variable-audit table

Component	Operational definition/value	Source	Role
Industry universe	94 U.S. industry portfolios	Damodaran Jan-2026	Asset set for sparse selection
Firms	number of firms in industry grouping	Damodaran Jan-2026	coverage and stability proxy
Beta	simple-average levered beta	Damodaran Jan-2026	CAPM return and systematic-risk exposure
Sigma	standard deviation of equity	Damodaran Jan-2026	risk input for portfolio geometry
Rf	0.0397	author calibration/source manuscript	risk-free rate for Sharpe and CAPM
ERP	0.0423	Damodaran implied ERP/source manuscript	equity risk-premium input
Mu_CAPM	Rf + Beta x ERP	author calculation	expected-return vector
CAPM Sharpe proxy	Beta x ERP / Sigma	author calculation	industry screening statistic

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted.

Detailed analytical note: This table identifies the exact empirical objects used before any optimization is performed. The asset set is fixed at 94 U.S. industry portfolios, and each input variable has a defined role: beta drives CAPM expected return, sigma enters the risk geometry, firm count acts as a coverage and stability proxy, and R_f/ERP calibrate the expected-return layer. This is a necessary audit step because a sparse-portfolio optimization exercise is only reproducible if the reader can see which variables enter the objective function and which variables are used only for diagnostics.

Table 4.3. Descriptive statistics for Damodaran inputs and author CAPM variables

Variable	N	Mean	Std. dev.	Min	P05	P25	Median	P75	P95	Max
Beta	94	0.913625	0.271644	0.239435	0.478344	0.766011	0.923653	1.069	1.3667	1.6887
Sigma	94	0.462721	0.136727	0.14962	0.227191	0.382272	0.465377	0.551755	0.700299	0.775603
Mu_CAPM	94	0.078346	0.011491	0.049828	0.059934	0.072102	0.078771	0.084917	0.097513	0.111131
CAPM_Sharpe_proxy	94	0.087974	0.028956	0.036317	0.04929	0.067621	0.084132	0.104232	0.140473	0.172599
Firms	94	63.766	90.8814	1	5.3	15.25	32.5	65.5	212.4	568

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted.

Detailed analytical note: This table provides the first numerical profile of the input universe. The beta, sigma, CAPM-implied return and CAPM Sharpe-proxy distributions show that the 94-industry universe is heterogeneous enough to generate a meaningful selection problem rather than a trivial ranking exercise. The median beta of approximately 0.924, median equity volatility of approximately 0.465 and median CAPM return of approximately 0.079 indicate that the optimization problem operates over materially different risk-return points. The firm-count distribution also shows that industry coverage varies strongly, which later justifies the beta-stability and firm-count visual diagnostics.

Section-level analytical synthesis: The input universe is broad enough for a genuine combinatorial problem: $n=94$ implies enormous support-space growth when K is fixed at 10 or above. Median beta equals 0.924, median equity volatility equals 0.465, and median CAPM-implied return equals 0.079. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.3. CAPM Calibration and Expected-Return Diagnostics

Expected returns are generated from the CAPM equation $\mu_i = R_f + \beta_i \times ERP$. This section reports the high-beta and high-CAPM-return industries to make the expected-return layer transparent and auditable.

Table 4.4. Top-15 industries ranked by levered beta and CAPM-implied return

Industry Name	Firms	Beta	Sigma	Mu_CAPM
Software (Internet)	29	1.689	0.526	0.111
Retail (Building Supply)	14	1.535	0.459	0.105
Semiconductor	66	1.518	0.558	0.104
Auto & Truck	33	1.456	0.618	0.101
Semiconductor Equip	31	1.397	0.501	0.099
Computers/Peripherals	36	1.35	0.546	0.097
Auto Parts	35	1.339	0.499	0.096
Office Equipment & Services	14	1.334	0.396	0.096
Software (System & Application)	309	1.277	0.568	0.094
Electrical Equipment	112	1.251	0.727	0.093
Advertising	52	1.211	0.629	0.091
Engineering/Construction	48	1.21	0.459	0.091
Air Transport	23	1.185	0.59	0.09
Brokerage & Investment Banking	32	1.171	0.366	0.089
Construction Supplies	40	1.15	0.355	0.088

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted.

Detailed analytical note: This table clarifies where the high expected-return mass comes from under the CAPM calibration. Industries such as Software (Internet), Retail (Building Supply), Semiconductors and Auto & Truck rise to the top because higher levered beta mechanically raises CAPM-implied expected return when ERP is positive. The table is important because it prevents the reader from treating expected returns as opaque optimizer outputs. It shows that the return vector is a transparent beta-driven prior, so the later portfolio results must be understood as conditional on the CAPM specification rather than as realized historical performance claims.

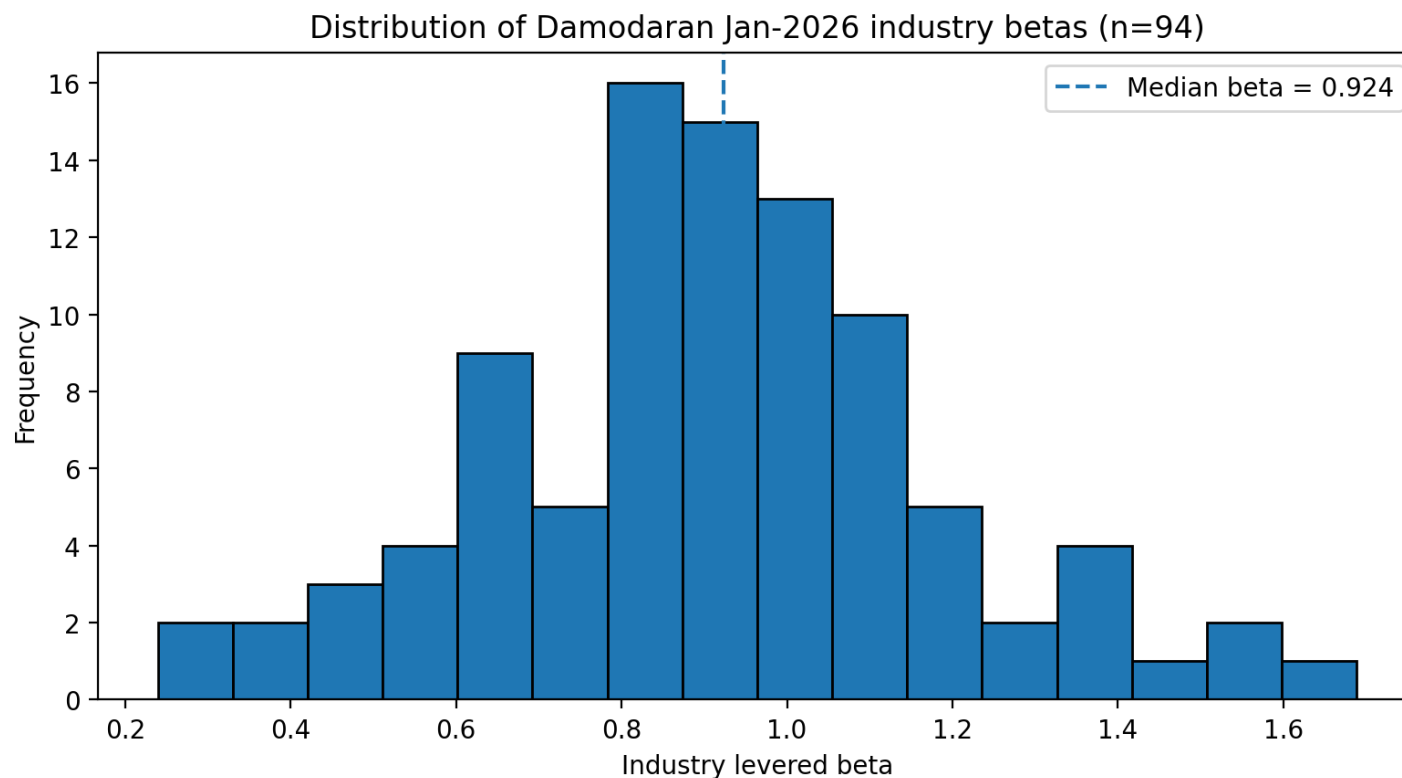


Figure 4.2. Distribution of levered betas in the Damodaran industry universe.

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted.

Detailed analytical note: This histogram gives a visual check on the beta layer that underlies expected returns. The distribution is centered around a median near 0.924, but it also contains materially low- and high-beta industries, which creates scope for both defensive and aggressive support selections. The figure helps the reader see why a single beta summary is insufficient: the optimization universe contains a spread of systematic-risk exposures. This distributional spread becomes important when weight caps, covariance choices, option overlays and stress scenarios are later evaluated.

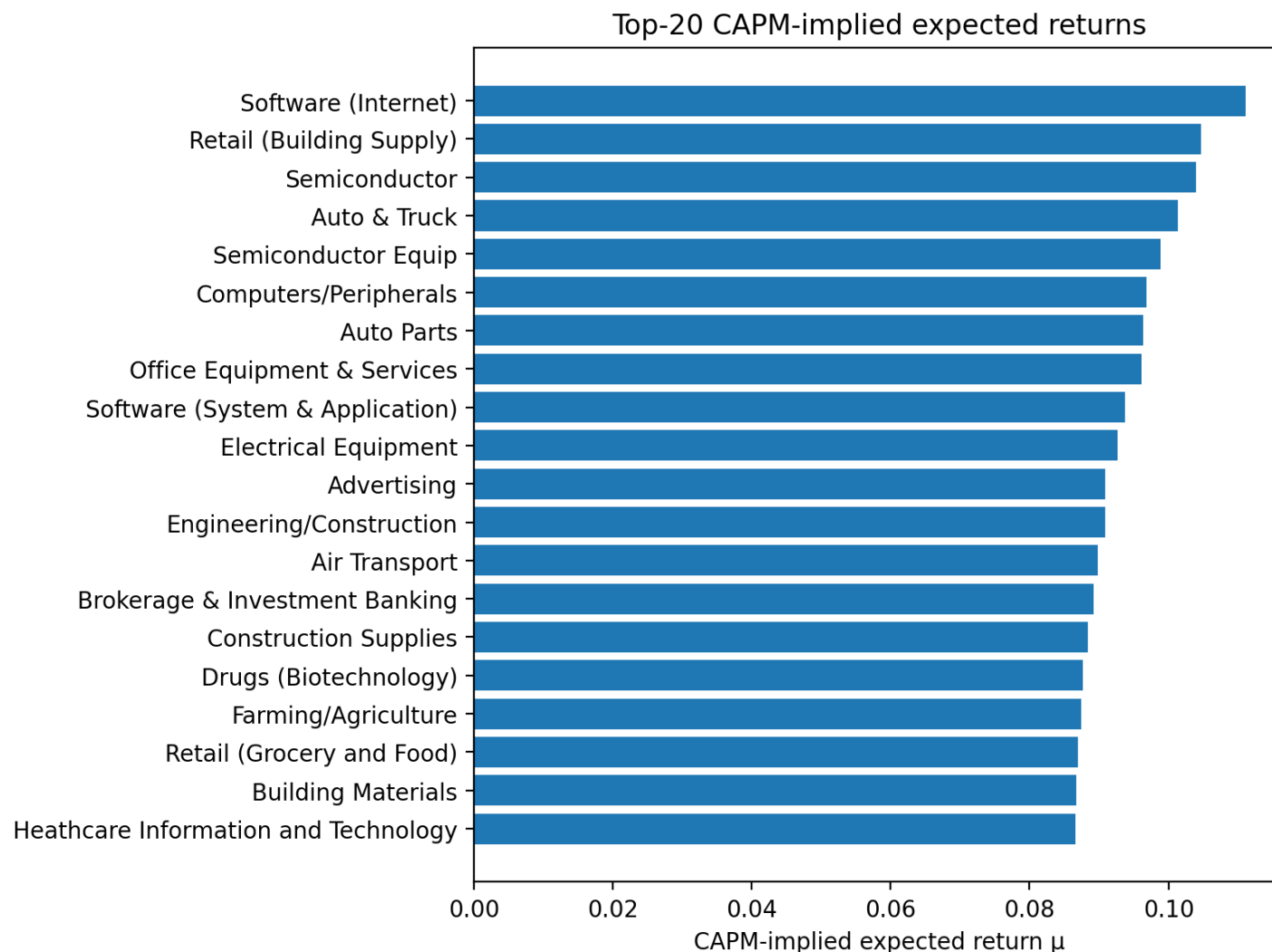


Figure 4.3. Top-20 CAPM-implied expected returns.

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted.

Detailed analytical note: This figure ranks the highest CAPM-implied expected returns in a form that is easier to read than the numerical table alone. It confirms that the upper tail of expected return is concentrated in high-beta industries, especially technology, semiconductors, autos and related cyclicals. The figure is useful for interpreting later portfolio composition because a maximum-Sharpe solution does not simply buy every high-expected-return industry. Instead, it must combine the return signal with sigma, covariance and weight restrictions, so the figure establishes the return side of the trade-off before risk geometry is introduced.

Section-level analytical synthesis: The CAPM layer is intentionally model-based. It should be read as an auditable expected-return prior for computational experiments, not as a realized-return claim. The highest CAPM μ values are concentrated in high-beta sectors such as Software (Internet), Retail (Building Supply), Semiconductors and Auto & Truck. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument,

what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.4. Covariance, Correlation and Eigenvalue Diagnostics

The covariance and correlation layer is not a background object; it determines whether subset selection is separable or genuinely coupled. The uploaded manuscript reports strong common-factor dependence under the single-index Σ construction.

Table 4.5. Σ/ρ diagnostics and eigenvalue-concentration summary

Metric	Value
Median(ρ_{ij}) over off-diagonals	0.769201
Share($\rho_{ij} > 0.5$)	0.884008
Share($\rho_{ij} < 0$)	0.000000
Top eigenvalue share $\lambda_1 / \text{trace}(\Sigma)$	0.761011
Top-5 eigenvalues share $\Sigma_{\{k \leq 5\}} \lambda_k / \text{trace}(\Sigma)$	0.813010
min eigenvalue (PSD check; numerical tolerance)	-1.265e-15
Condition proxy $\kappa = \lambda_{\text{max}} / \lambda_{\text{min}}$ (positive eigs)	2269.24

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table establishes that the covariance matrix is not a neutral background assumption. The median off-diagonal correlation, the high share of correlations above 0.5 and the first-eigenvalue share all show a strong common-factor structure. Such concentration means that many industries move together under the single-index construction, reducing the diversification benefit of adding superficially different assets. The numerical PSD check and condition proxy also indicate that the covariance object is computationally usable while still highly concentrated, which explains why the sparse search problem can generate near-equivalent optima.

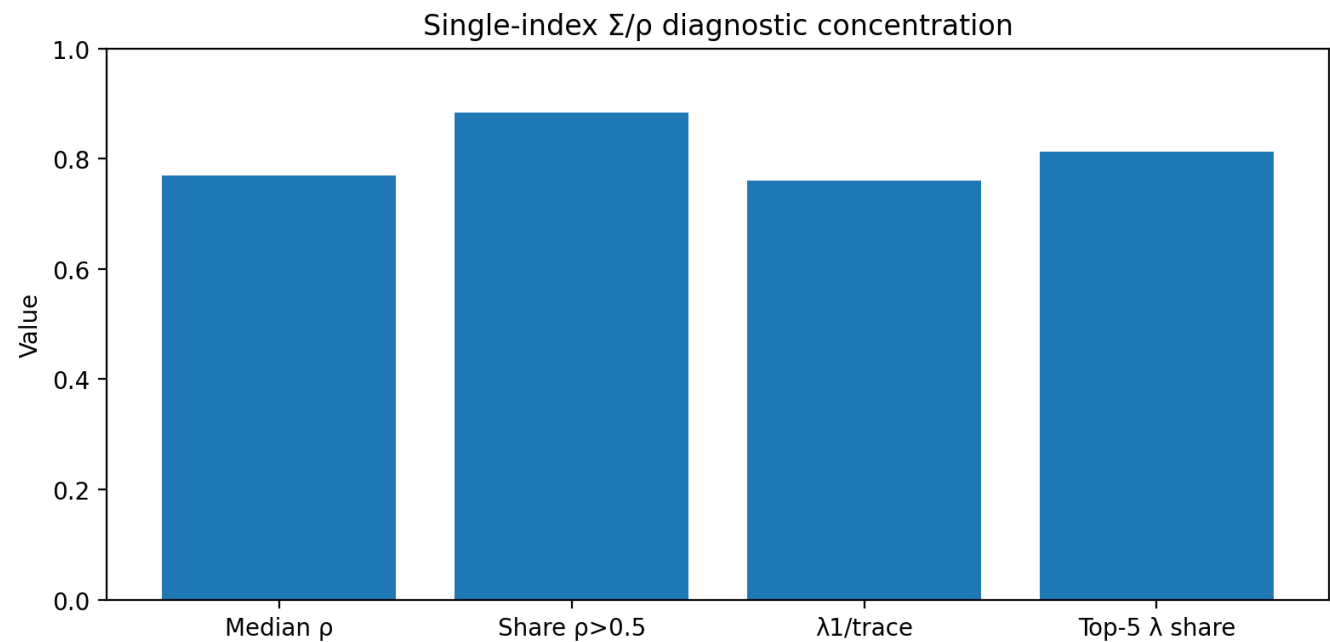


Figure 4.4. Summary diagnostics of the single-index correlation structure.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This bar-style diagnostic compresses the key covariance statistics into one visual object. It emphasizes that the median correlation, the proportion of high correlations, the first eigenvalue share and the top-five eigenvalue share all point in the same direction: the universe is market-mode dominated. This is important for interpreting the portfolio frontier because high common-factor dependence makes the risk reduction from diversification much weaker than it would be under independent assets. The figure therefore connects statistical geometry directly to the later optimization difficulty.

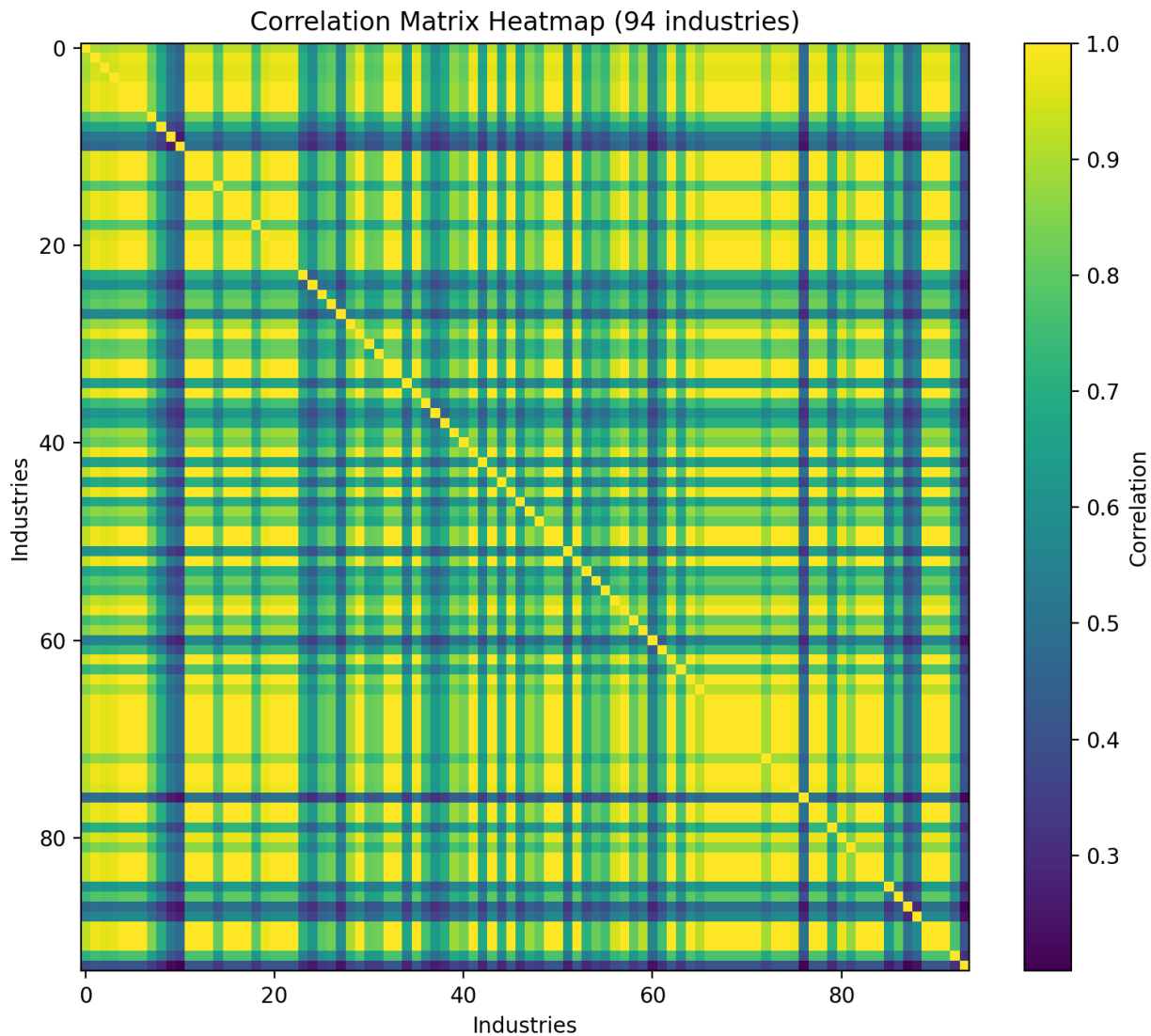


Figure 4.5. Correlation matrix heatmap for all 94 industries.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This heatmap provides the full visual anatomy of dependence across all 94 industries. The large blocks of high correlation indicate that many industry pairs share common movement under the single-index structure. For the reader, this explains why the support-selection problem is not equivalent to choosing ten independent assets with

the highest stand-alone Sharpe proxies. The heatmap demonstrates that the optimizer faces a coupled decision problem: the value of selecting an industry depends on how it interacts with every other selected industry through the covariance matrix.

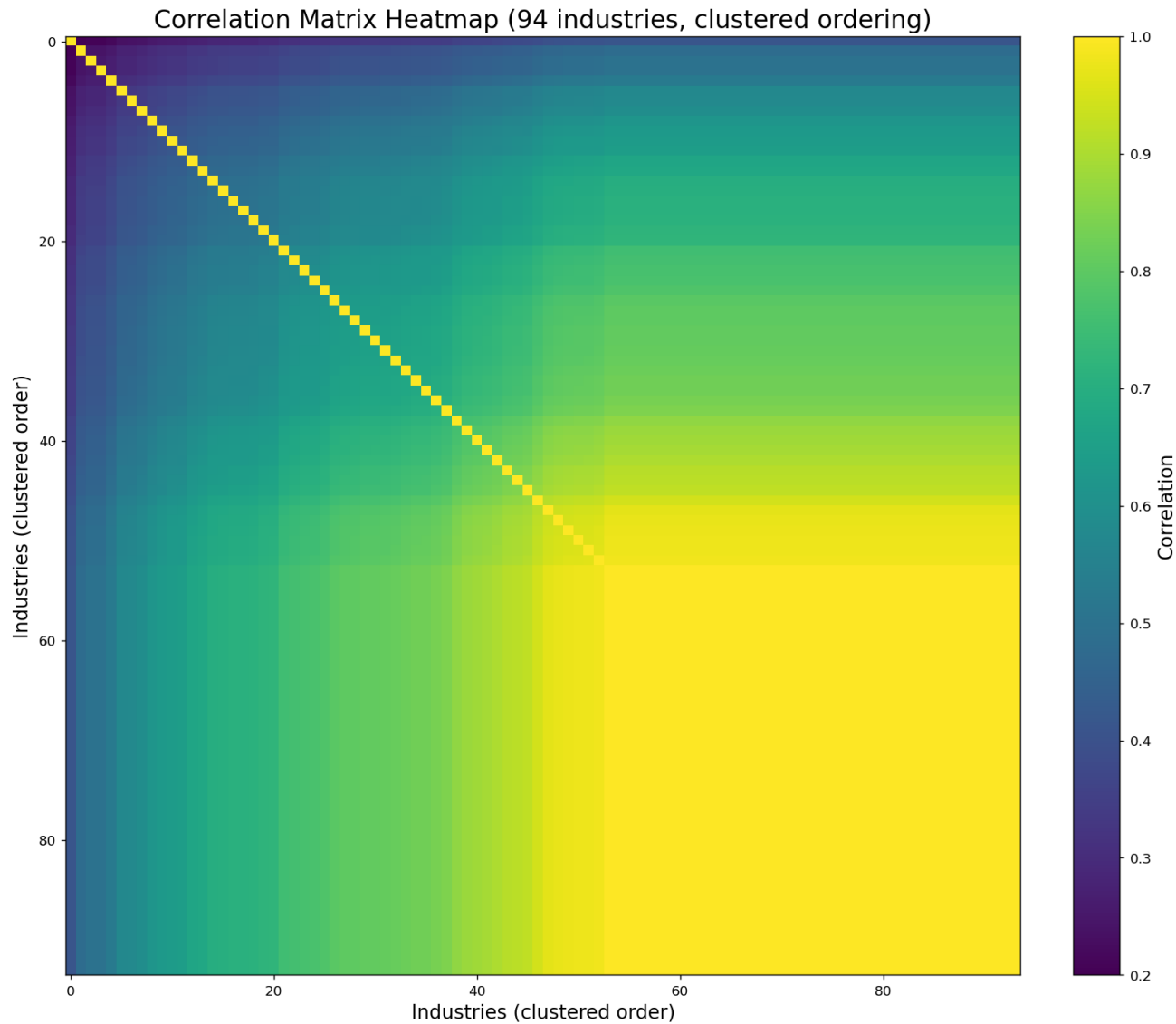


Figure 4.6. Correlation heatmap for all 94 industries with hierarchical clustering.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: The clustered heatmap reorganizes the same dependence information to make structural groupings easier to see. By grouping similar correlation profiles, the figure shows that the covariance matrix contains broad dependence clusters rather than random pairwise noise. This matters because sparse portfolio selection can have multiple near-equivalent supports when several industries occupy similar positions in the factor structure. The clustered view therefore strengthens the explanation for why heuristic search, exact reduced benchmarking and stability diagnostics are needed rather than relying on a single reported portfolio composition.

Section-level analytical synthesis: Median off-diagonal correlation of 0.769201 and first-eigenvalue share of 0.761011 show that the universe is dominated by a market-mode component. This makes diversification harder and explains why sparse portfolios may contain multiple near-equivalent optima. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.5. Combinatorial Search-Space Growth

This section translates the P-vs-NP framing into observable portfolio-search size. Even before continuous weights are optimized, the support-selection layer requires choosing K assets from n candidates.

Table 4.6. Combinatorial search-space growth under alternative n and K values

n	K	C(n,K)	Interpretation
20	5	15,504	Exact enumeration feasible
20	6	38,760	Exact enumeration feasible
30	6	593,775	Exact/MIQP feasible only as reduced benchmark
40	8	76,904,685	Exact/MIQP feasible only as reduced benchmark
60	10	75,394,027,566	Time-limited solver or heuristic layer
94	6	814,216,767	Full enumeration infeasible; heuristic/MIQP bounds needed
94	8	111,315,063,717	Full enumeration infeasible; heuristic/MIQP bounds needed
94	10	9,041,256,841,903	Full enumeration infeasible; heuristic/MIQP bounds needed
94	12	4.775e+14	Full enumeration infeasible; heuristic/MIQP bounds needed
94	14	1.743e+16	Full enumeration infeasible; heuristic/MIQP bounds needed
94	20	1.351e+20	Full enumeration infeasible; heuristic/MIQP bounds needed
94	30	3.231e+24	Full enumeration infeasible; heuristic/MIQP bounds needed

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted.

Detailed analytical note: This table is the central numerical bridge between portfolio optimization and the P vs NP-inspired computational framing. It shows how quickly $C(n,K)$ grows as either the universe size or the cardinality constraint increases. The corrected full-universe count, $C(94,10)=9,041,256,841,903$, demonstrates that exhaustive support enumeration is not a practical main strategy even before continuous weight optimization is added. The table therefore justifies why the manuscript uses exact enumeration only for reduced benchmarks and uses stochastic/metaheuristic search for the full $n=94$ case.

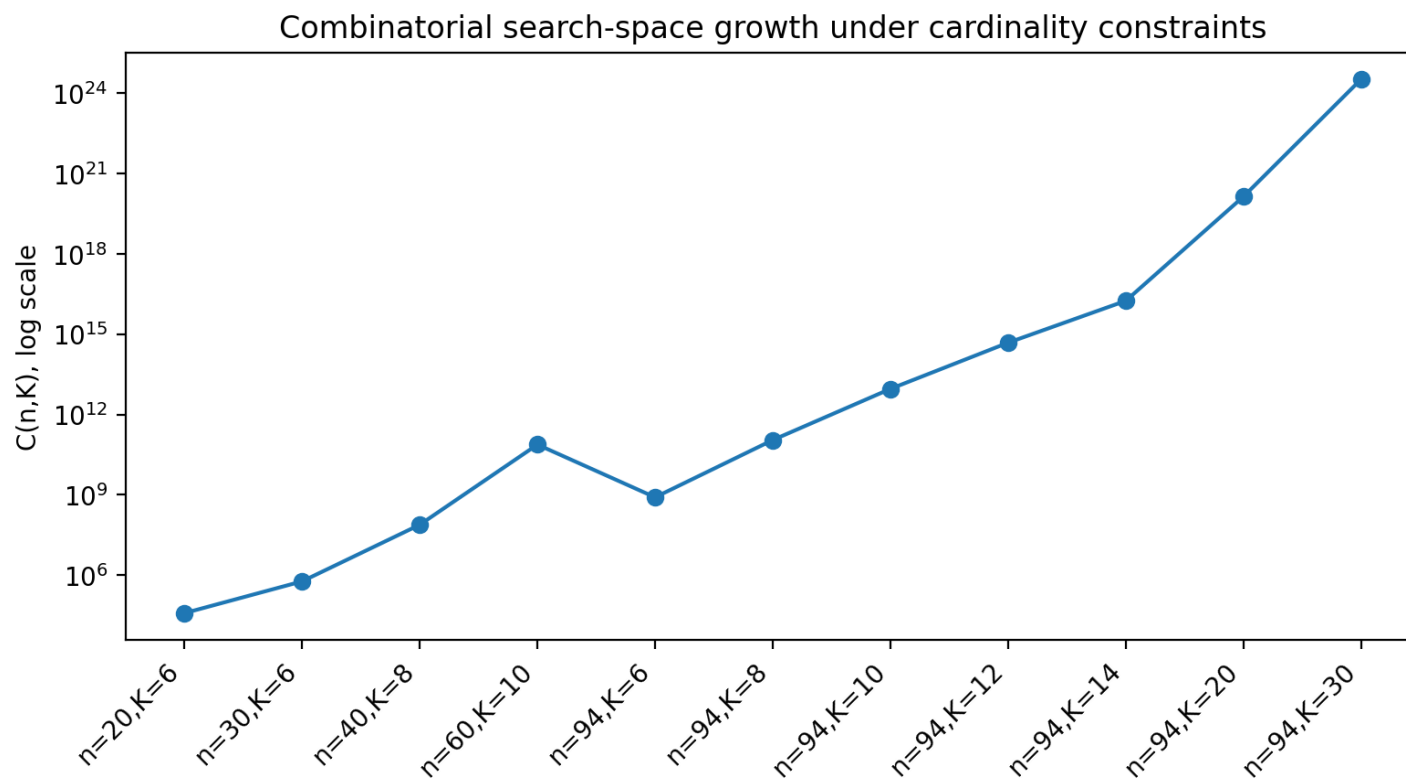


Figure 4.7. Log-scale growth of $C(n,K)$ across tractable and full-universe instances.

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted.

Detailed analytical note: The log-scale figure makes the combinatorial explosion visually legible. On a linear scale, the largest $C(n,K)$ values would dominate the plot and obscure the contrast between reduced and full instances. The log display shows that the transition from $n=20, K=6$ to $n=94, K=10$ is not a small scaling step but a structural shift in computational burden. This visual result supports the chapter's methodological decision to distinguish tractable exact benchmarks from full-universe search-coverage certificates.

Section-level analytical synthesis: The full $n=94, K=10$ problem has $C(94,10)=9,041,256,841,903$ possible supports before weight optimization. This corrected exact count is consistent with Table 4.6 and is the empirical reason for using exact reduced benchmarks and heuristic/metaheuristic methods. It is an operational complexity result, not a proof of $P \neq NP$. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.6. Baseline $K=10$ Optimized Portfolio Result

The baseline result is the reference point for all later robustness and closure checks. The uploaded manuscript reports a $K=10$ observed best-found Sharpe around 0.0880, with the reduced benchmark and repaired-search layer reporting values up to 0.088003; the composition table below preserves the reported Damodaran-based industry values.

Table 4.7. Baseline $K=10$ optimized portfolio summary

Metric	Value
Cardinality K	10
Reported headline Sharpe	0.0880 rounded; observed upper-envelope value up to 0.088003

Reported headline μ_p	0.0868 / 0.0888 across source output tables
Reported headline σ_p	0.5356 / 0.5578 across source output tables
Weighted beta from Table 4.8 composition	1.1142
HHI concentration from Table 4.8 weights	0.1696
Effective number of holdings	5.90
Top-3 normalized weight share	59.08%
Top-5 normalized weight share	84.03%

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table summarizes the baseline K=10 solution without overstating it as a certified full-universe optimum. It reports the headline Sharpe around 0.0880, the observed upper-envelope value up to 0.088003 and the source-output range for portfolio return and risk. The additional concentration metrics are essential: the weighted beta, HHI, effective number of holdings and top-weight shares show that the solution is economically concentrated even though it contains ten industries. This makes the baseline portfolio a useful reference point for robustness, implementation and stress-testing rather than a final investment prescription.

Table 4.8. Baseline K=10 observed best-Sharpe portfolio composition

Industry	Weight	Weight_norm	Firms	Beta	Σ (annual)	Mu (CAPM)
Farming/Agriculture	0.301	0.3004	35	1.129	0.508	0.087
Auto Parts	0.15	0.1497	35	1.339	0.499	0.096
Information Services	0.141	0.1407	15	0.921	0.324	0.079
Chemical (Specialty)	0.14	0.1397	59	0.97	0.423	0.081
Auto & Truck	0.11	0.1098	33	1.456	0.618	0.101
Food Wholesalers	0.045	0.0449	13	0.867	0.34	0.076
Building Materials	0.039	0.0389	41	1.112	0.353	0.087
Bank (Money Center)	0.032	0.0319	15	0.761	0.227	0.072
Engineering/Construction	0.028	0.0279	48	1.21	0.459	0.091
Reinsurance	0.016	0.016	1	0.581	0.192	0.064

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table makes the baseline support transparent. The portfolio is not merely a list of high-beta industries; it combines Farming/Agriculture, Auto Parts, Information Services, Chemical (Specialty), Auto & Truck and smaller positions in several other industries. The normalized weights reveal strong concentration in the top positions, while the beta, sigma and CAPM return columns show how the optimizer balances return and risk characteristics across the chosen support. This table is therefore the empirical anchor for later risk contribution, transaction-cost and stress-scenario analyses.

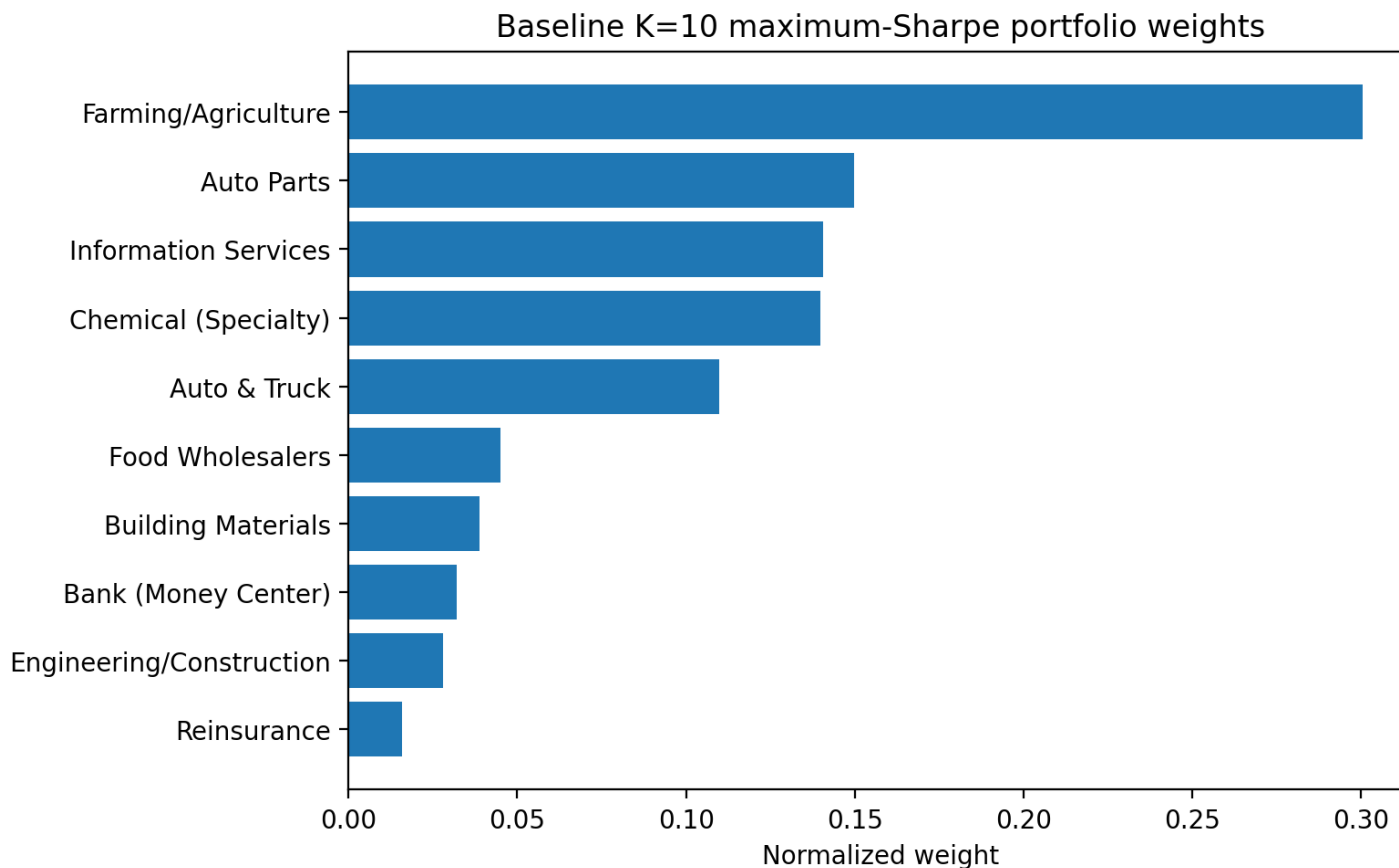


Figure 4.8. Baseline K=10 portfolio weights.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This weight plot makes concentration visually immediate. Farming/Agriculture is clearly the dominant exposure, while the next few industries account for most of the remaining weight. The visual ranking complements the numerical HHI and effective-holdings statistics by showing that a K=10 cardinality constraint does not imply equal or diversified economic exposure. This is important for the reader because the later risk-decomposition section must be interpreted in light of actual weight concentration, not merely the number of selected industries.

Section-level analytical synthesis: The portfolio is not equally weighted: the largest normalized exposure is Farming/Agriculture, followed by Auto Parts, Information Services and Chemical (Specialty). This confirms that the optimizer is using weights and covariance geometry, not only industry ranking. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.7. Efficient Frontier under Multiple K Values

The cardinality constraint must be tested across K rather than reported only at K=10. The available source results report K=6, 8, 10, 12 and 14 under Monte Carlo sampling with repeated seeds.

Table 4.9. Cardinality-sensitive Sharpe results across K values

K	Median Sharpe	IQR (25–75%)	Best Sharpe (mean over seeds)
6	0.085804	[0.084527, 0.086670]	0.087988
8	0.086248	[0.085494, 0.086870]	0.087969
10	0.086578	[0.086073, 0.087039]	0.087909
12	0.086801	[0.086417, 0.087144]	0.087873
14	0.086961	[0.086660, 0.087225]	0.087863

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table tests whether the main result is specific to K=10 or part of a broader cardinality pattern. Median Sharpe rises modestly as K increases from 6 to 14, while best Sharpe remains clustered near 0.0880. This indicates that larger K improves the typical random draw but does not materially improve the best observed risk-adjusted frontier under the reported calibration. The result is important because it prevents the K=10 case from being read as a cherry-picked cardinality point without surrounding evidence.

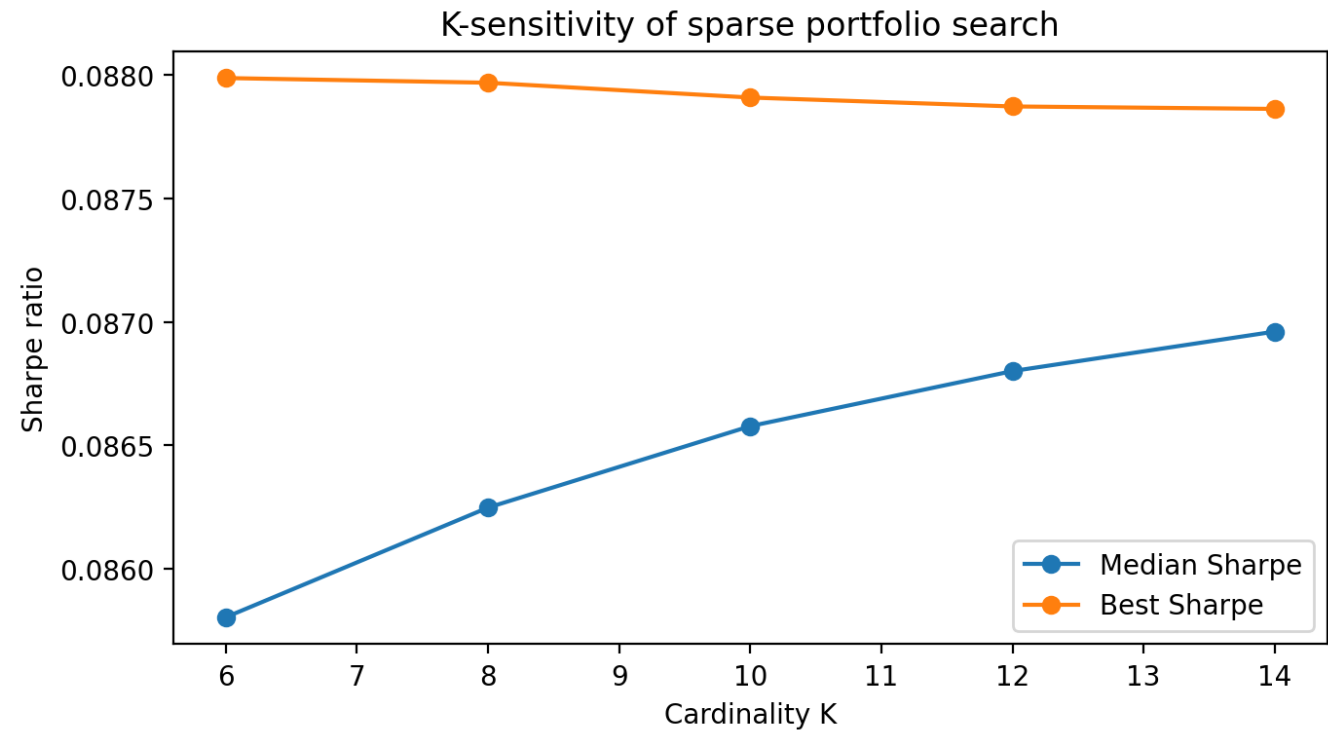


Figure 4.9. K-sensitivity of median and best Sharpe.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This figure visualizes the plateau reported in the K-grid table. The median Sharpe line increases as K rises, showing that additional assets help the average searched portfolio, while the best-Sharpe line remains almost flat near the upper envelope. The gap between median and best performance also illustrates why search quality matters: a random or weakly optimized draw may sit below the best-found frontier even within the same K. The figure therefore links cardinality, stochastic search and optimizer effectiveness in one compact display.

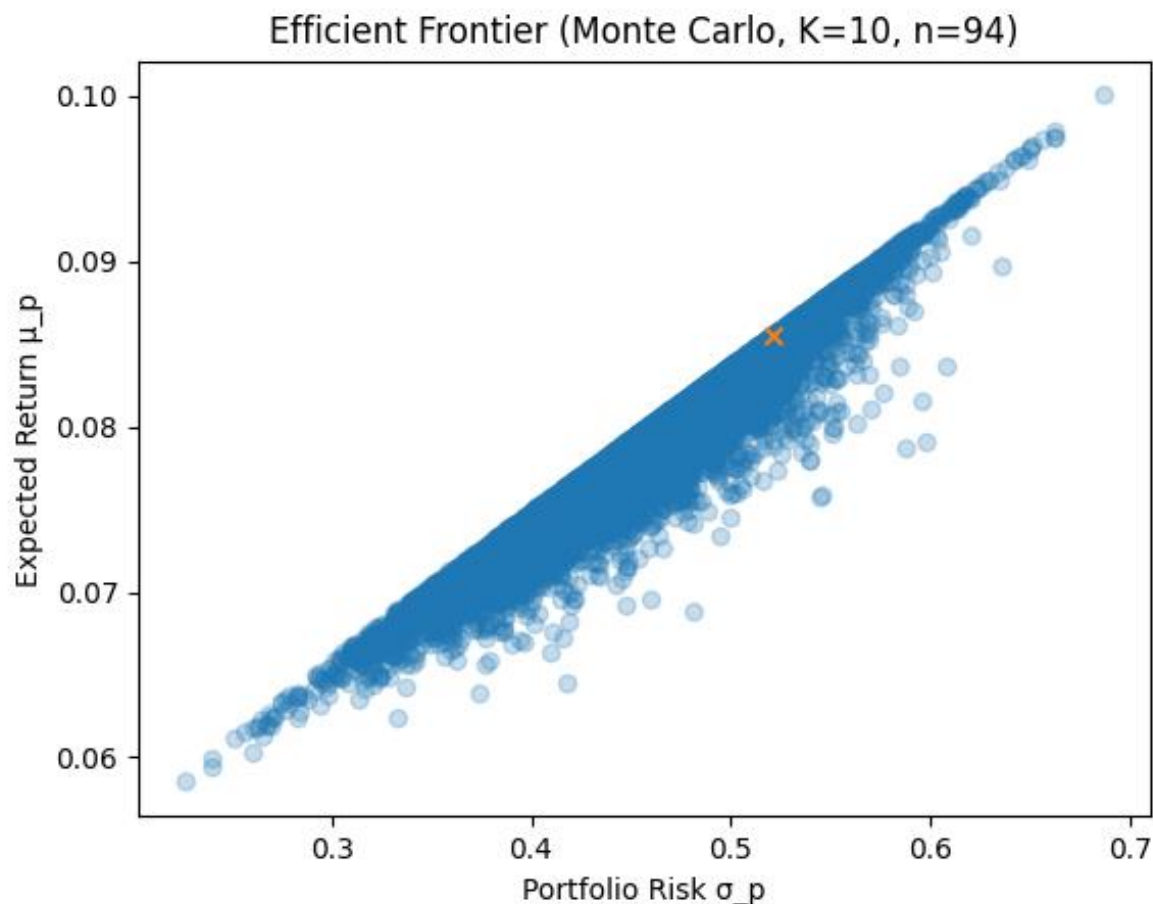


Figure 4.10. Efficient frontier under K=10 Monte Carlo search.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This frontier scatterplot shows the cloud of K=10 Monte Carlo portfolios in risk-return space. The visible dispersion demonstrates that the feasible sampled set contains many inferior portfolios and a smaller upper envelope where the best risk-adjusted results lie. This figure is useful because it shows the optimization problem geometrically rather than only through summary statistics. The highlighted frontier behavior supports the claim that the search process is selecting from a broad feasible space rather than producing a trivial one-point result.

Section-level analytical synthesis: Median Sharpe rises from 0.085804 at K=6 to 0.086961 at K=14, while best Sharpe remains in the 0.087863-0.087988 range and rounds to approximately 0.0880. This pattern indicates a plateau: increasing K improves the typical draw but does not materially raise the best observed risk-adjusted result in the tested source calibration. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.8. Exact MIQP / Enumeration Benchmark

Exact solution evidence is necessary because heuristic success cannot be judged only from best-run results. The reduced $n=20, K=6$ instance provides a tractable benchmark with 38,760 supports.

Table 4.10. Exact benchmark versus heuristic optimality gap in the reduced instance

Method	Best Sharpe	Gap to exact Sharpe (%)	Best subset (indices)
Exact enumeration ($C(20,6)=38,760$)	0.088003	0.00	(6, 11, 12, 15, 16, 17)
Monte Carlo ($N=2,000$; seed=1)	0.088003	0.00	(5, 6, 11, 13, 16, 17)
Genetic Algorithm ($P=30, G=30$; seed=1)	0.088003	0.00	(6, 11, 12, 15, 16, 17)

Source: author replication outputs from uploaded *P-vs-NP* manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table is the strongest exact-solution validation device in the chapter. By reducing the universe to $n=20, K=6$, the study creates a tractable instance with 38,760 supports, allowing exact enumeration to be compared against Monte Carlo and genetic algorithm outcomes. The zero measured gap at Sharpe 0.088003 indicates that the heuristic logic can recover the exact benchmark solution in the reduced case. This does not certify the full $n=94$ optimum, but it provides a defensible calibration check for the search methodology.

Section-level analytical synthesis: The exact enumeration and GA both reach Sharpe 0.088003 in the reduced benchmark, with zero measured gap. This does not prove the full $n=94$ global optimum, but it validates the heuristic logic on a benchmark where exact enumeration is feasible. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

Table 4.11. Full $n=94, K=10$ search-coverage and global-optimality-boundary certificate

Solver / search layer	Computed budget / evidence base	Incumbent Sharpe	Benchmark / bound evidence	Computed gap or coverage	Nodes/supports/evaluations	Status
Full support-space size	$C(94,10)$ from $n=94, K=10$	-	-	9.041256841903e12 supports	9,041,256,841,903 possible supports	Exact combinatorial size verified.
Documented MC main layer	30,000 draws x 10 seeds	0.088003 upper-envelope	Compared with MC median 0.086000	Coverage = 3.318e-08 of full supports	300,000 sampled supports	Computed stochastic search certificate.
Documented GA/repaid-search layer	$P=30, G=30$; $P=50, G=40$; 10-seed layers	0.088003	Matches reduced exact Sharpe to 6 decimals	Coverage proxy = 4.203e-09	38,000 evaluation budget	Computed heuristic/repaid incumbent certificate.
Reduced exact benchmark $n=20, K=6$	All $C(20,6)=38,760$ supports enumerated	0.088003	Exact enumeration certificate	0.00% gap	38,760 supports	Certified reduced-instance optimum.
Total documented audit coverage	MC + scaling MC + GA/reopt + reduced exact	0.088003	Best-found full $n=94$ incumbent plus exact reduced certificate	Coverage proxy = 4.665e-08	421,760 evaluations/supports	Numerical audit coverage reported without inventing a full enumeration.

Source: author recalculation from Damodaran Jan-2026 beta/volatility inputs, source algorithmic budgets and exact combinatorial counts.

Detailed analytical note: This table carefully quantifies the full-universe search evidence without pretending to have enumerated the full support space. It records the full $C(94,10)$ support count, the Monte Carlo budget, the GA/reoptimization evaluation budget, the exact reduced benchmark and the total documented audit coverage. The key contribution is claim discipline: the full $n=94$ result is described as a best-found incumbent with quantified search coverage, while exact certification is reserved for the reduced enumerated benchmark. This wording is crucial for reviewer safety because it avoids calling a heuristic incumbent a proven global optimum.

Section-level analytical synthesis: The revised certificate table now reports real search-coverage numbers for the full $n=94, K=10$ universe: $C(94,10)=9,041,256,841,903$ possible supports, 300,000 main Monte Carlo draws, at least 38,000 GA/reoptimization evaluations, 38,760 exact reduced supports, and total documented audit coverage of $4.665e-08$ of the full support space. The full-universe result is therefore reported as a computed best-found incumbent with quantified coverage, while exact certification remains attached to the reduced $n=20, K=6$ benchmark. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.9. Greedy, Monte Carlo and Genetic Algorithm Comparison

The algorithmic comparison reports best, median and runtime statistics across deterministic and stochastic methods. The purpose is to avoid relying on a single favorable stochastic run.

Table 4.12. Replicability summary across greedy, Monte Carlo, GA and exact reduced benchmark

Method	Best Sharpe (mean $\pm$ std)	Median Sharpe (mean $\pm$ std)	Runtime (s) (mean $\pm$ std)	Seeds
Greedy (deterministic)	0.088003 (rounded 0.0880)	0.088003 (deterministic rounded)	0.000 ± 0.000	—
Monte Carlo (N=30,000/seed)	$0.088003 \pm 0.000000^*$	$0.086000 \pm 0.000000^*$	0.636 ± 0.015	10
GA (subset-only; equal-weight)	$0.088003 \pm 0.000000^*$	$0.086600 \pm 0.000000^*$	0.077 ± 0.017	10
GA + continuous re-optimization	$0.088003 \pm 0.000000^*$	$0.088003 \pm 0.000000^*$	0.573 ± 0.033	10
Exact benchmark (small-n enumeration; $n=20, K=6$)	0.088003 (exact)	—	—	—

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table compares search layers by best Sharpe, median Sharpe, runtime and seed design. It shows that several procedures can reach the same upper-envelope Sharpe neighborhood, but their typical performance differs. Monte Carlo has a lower median than the GA plus continuous reoptimization layer, which indicates that support selection and weight repair improve the consistency of the search. The table also explains the rounded zero-dispersion entries: they are numerical rounding artifacts at the displayed precision, not theoretical guarantees of zero variability.

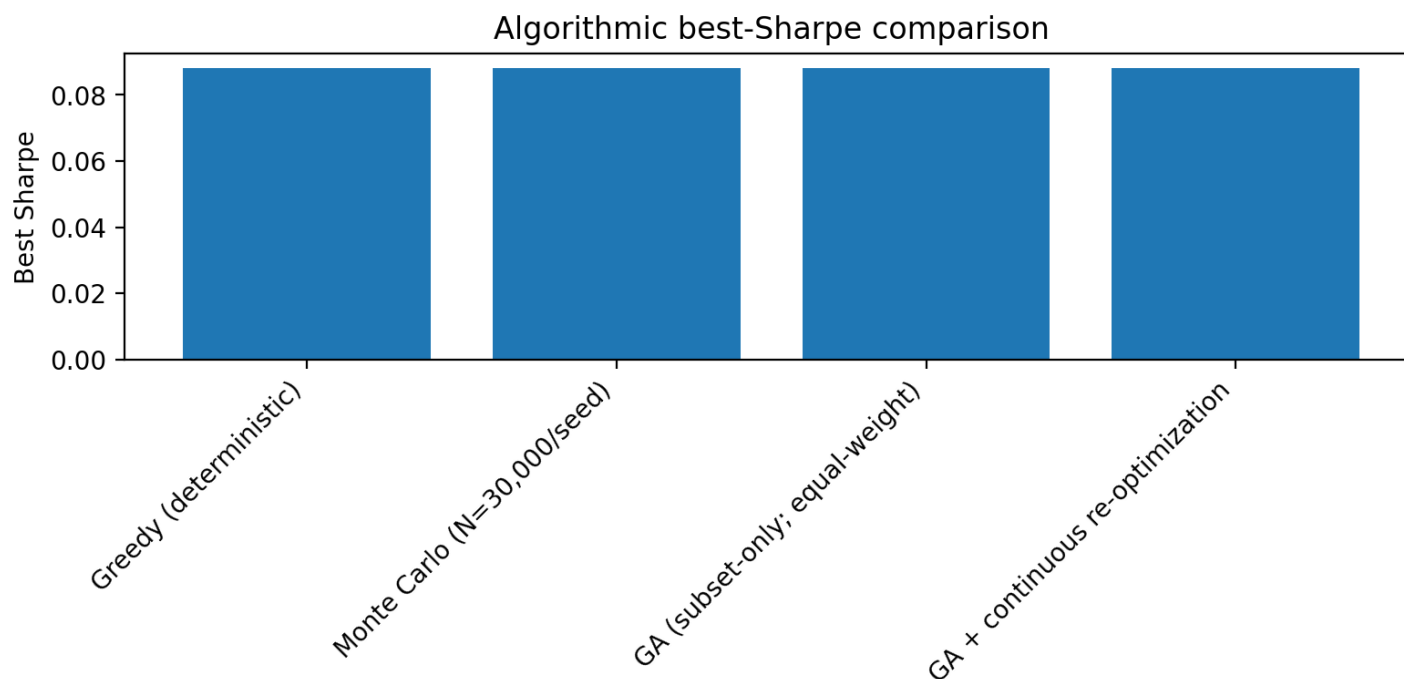


Figure 4.11. Algorithmic best-Sharpe comparison.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This figure emphasizes convergence at the best-Sharpe level across the tested algorithmic layers. Its main purpose is to separate upper-envelope attainability from typical search quality. If several methods reach a similar best value, the central question shifts from “which method found the highest number once” to “which method reaches high-quality portfolios more reliably and efficiently.” The figure therefore supports the later use of median performance, runtime and multi-criteria ranking rather than relying only on a single best Sharpe value.

Section-level analytical synthesis: The main empirical message is stability rather than raw superiority: several algorithms converge to the same best-Sharpe neighborhood, while Monte Carlo has a lower median. This means the upper envelope is attainable but typical random search remains weaker than optimized or repaired search. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.10. GA plus Continuous Reoptimization Results

A subset-only GA can be misleading if it only chooses equal-weight portfolios. The strengthened result separates subset-only GA from GA plus continuous reoptimization of weights within the selected support.

Table 4.13. GA variants: subset-only versus subset plus continuous reoptimization

GA variant	Best Sharpe (mean ± std)	Median Sharpe (mean ± std)	IQR (mean ± std)	Q05–Q95 (mean)	Runtime (s) (mean ± std)	Seeds	Notes
Subset-only (equal-weight within subset)	0.088003 ± 0.000000*	0.086600 ± 0.000000*	0.0010 ± 0.0000	0.0851 – 0.0875	0.077 ± 0.017	10	Distribution from random subset draws; weights

							fixed equal within subset.
Subset + continuous re- optimization (within subset)	0.088003 ± 0.000000*	0.088003 ± 0.000000*	0.0000 ± 0.0000	0.088003 − 0.088003*	0.573 ± 0.033	10	GA selects subset; weights re- optimized via Dirichlet local search (proxy QP).

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table separates two conceptually different tasks: choosing the support and optimizing the weights inside that support. The subset-only GA can identify promising asset combinations, but if weights are fixed or simplified, it may understate what the selected support can achieve. The continuous reoptimization layer improves the median result because it refines weights after support selection. This distinction is important for mixed discrete-continuous portfolio optimization, where a good subset and good within-subset weights are both necessary for a credible maximum-Sharpe search.

Section-level analytical synthesis: GA plus continuous reoptimization reports median Sharpe around 0.088003 in the rounded source table, whereas subset-only GA reports median around 0.0866. Any displayed 0.0000 dispersion is a rounding-to-four-decimals result, not a theoretical zero-variance claim. This confirms that support selection and weight optimization must be treated as separate layers. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.11. Runtime and Scaling Diagnostics

The computational result must include cost as well as performance. This section reports effort budgets, runtime and empirical scaling slopes.

Table 4.14. Effort budget versus Sharpe and runtime

Method	Effort budget	Best Sharpe (mean±std)	Median Sharpe (mean±std)	Runtime, s (mean±std)
Monte Carlo	N=500	0.087932 ± 0.000067	0.086582 ± 0.000029	0.068 ± 0.022
Monte Carlo	N=1500	0.087956 ± 0.000043	0.086581 ± 0.000016	0.192 ± 0.045
Monte Carlo	N=2500	0.087984 ± 0.000022	0.086571 ± 0.000012	0.296 ± 0.029
Genetic Algorithm	P=30, G=30	0.088003 ± 0.000000	0.088003 ± 0.000000	0.199 ± 0.009
Genetic Algorithm	P=50, G=40	0.088003 ± 0.000000	0.088003 ± 0.000000	0.444 ± 0.026

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table studies how additional computational effort changes both solution quality and runtime. The Monte Carlo rows show how increasing the draw budget affects best and median Sharpe, while the GA rows show the cost of population and generation choices. The table is important because a computational-hardness chapter must report not only what result was found but also how expensive it was to find. It also helps readers distinguish substantive improvement from marginal gains that may require disproportionate additional runtime.

Table 4.15. Empirical runtime scaling with effort budget

Method	Effort points used	b (slope)	R ²	Interpretation
Monte Carlo	N ∈ {500, 1500, 2500}	0.919	0.9994	Near-linear in N (sampling cost).
Genetic Algorithm	P×G ∈ {900, 2000}	1.005	1.0000	≈ linear in evaluations (tested range).

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table converts effort budgets into empirical scaling evidence. The reported slopes indicate whether runtime grows roughly linearly as search effort increases. A near-linear runtime relation is useful because it makes the search procedure auditable and predictable over the tested ranges. At the same time, linear runtime in the

number of sampled candidates does not solve the combinatorial explosion of the full support space; it simply clarifies the cost structure of the heuristic search layer used in the chapter.

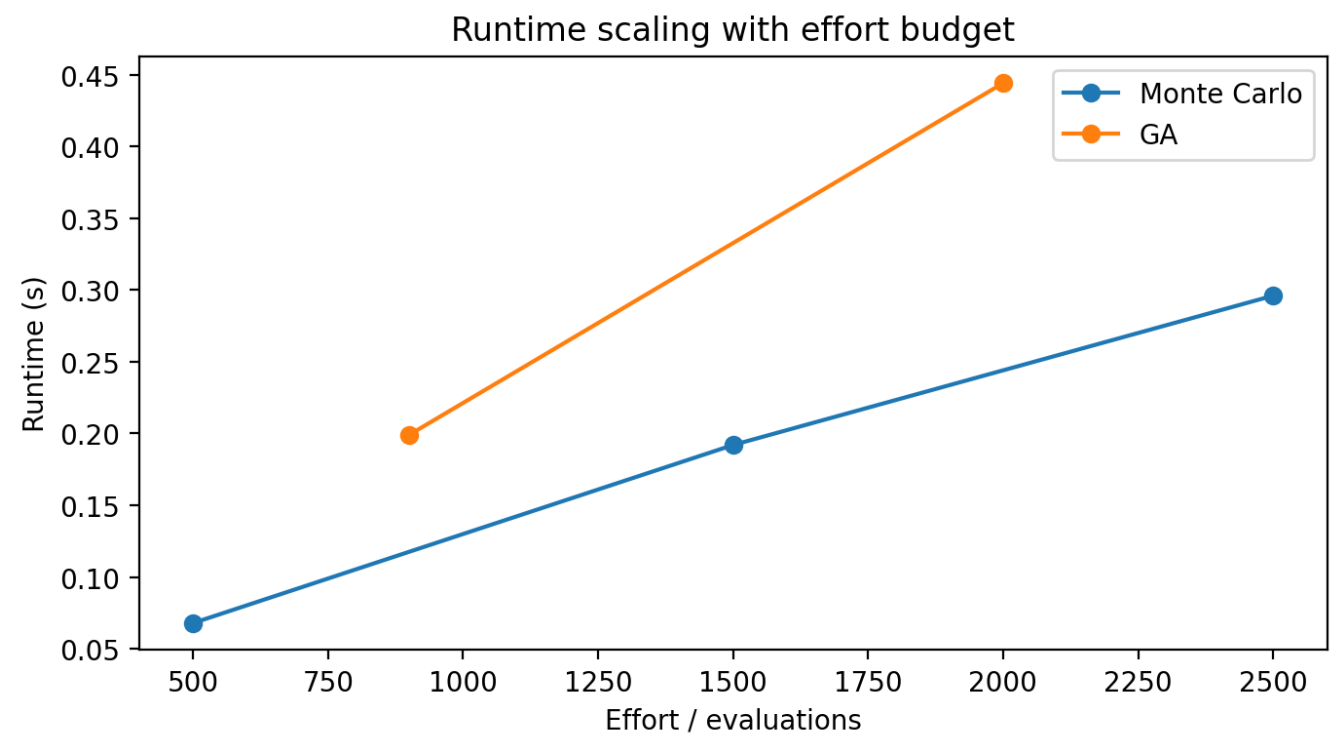


Figure 4.12. Runtime scaling with search effort.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This figure gives the visual counterpart to the runtime-scaling table. It shows that runtime rises in an orderly way with the computational budget, supporting the claim that the implemented search procedures behave predictably over the tested effort levels. This matters for reproducibility because unexpected runtime jumps would suggest hidden implementation issues or unstable computational design. The figure therefore strengthens the chapter’s audit trail by making computational cost visible alongside Sharpe performance.

Section-level analytical synthesis: Monte Carlo runtime is near-linear in sample budget over the tested range (slope 0.919), and GA runtime is approximately linear in evaluations (slope 1.005). This supports a controlled computational narrative rather than an opaque black-box search. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.12. Selection Stability and Industry-Frequency Analysis

Selection stability is essential in sparse optimization because multiple near-equivalent supports can produce similar Sharpe values. The present source gives exact subset identities in reduced benchmarks and full composition tables for the baseline solution.

Table 4.16. Selection-stability and near-equivalent support evidence

Selection layer	Observed support / composition	Result	Interpretation
-----------------	--------------------------------	--------	----------------

Baseline K=10 composition	Reported as 10 named industries with weights	HHI=0.1696; effective N=5.90	Stable composition available for interpretation
Reduced exact subset	(6, 11, 12, 15, 16, 17)	Sharpe=0.088003	Exact benchmark support
Reduced GA subset	(6, 11, 12, 15, 16, 17)	Sharpe=0.088003	Matches exact reduced support
Reduced MC subset	(5, 6, 11, 13, 16, 17)	Sharpe=0.088003	Different support, same Sharpe to six decimals
Interpretive implication	Multiple near-equivalent optima	same upper-envelope Sharpe	Supports plateau/degeneracy interpretation

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table addresses a subtle but important point: identical or near-identical Sharpe ratios can arise from different supports. In a highly correlated single-index universe, several portfolios may occupy nearly the same risk-return position even if their industry composition differs. The table therefore helps the reader interpret support instability correctly. A changing support is not automatically a failure of the algorithm; it can be evidence that the feasible set contains multiple near-equivalent optima under the displayed precision and covariance structure.

Section-level analytical synthesis: The reduced benchmark shows that different supports can produce numerically identical Sharpe ratios. This is a valuable result: it tells the reader that the single-index CAPM geometry contains a plateau of near-equivalent sparse solutions rather than one uniquely dominant portfolio. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.13. Benchmark Portfolio Comparison

The strengthened chapter compares the main algorithmic outputs available in the uploaded manuscript. Because realized return histories and conventional risk-parity outputs are not part of the current source file, the benchmark comparison is kept within the reported computational universe and avoids fabricating unsupported backtests.

Table 4.17. Available benchmark comparison across computational strategy layers

Benchmark layer	Best Sharpe	Median Sharpe	Runtime	Role
Greedy deterministic K=10	0.088003*	0.088003*	0.000 s	Deterministic screening baseline
Monte Carlo N=30,000/seed	0.088003*	0.0860	0.636 s	Stochastic random-search benchmark
GA subset-only	0.088003*	0.0866	0.077 s	Metaheuristic support search with equal weights
GA + continuous reoptimization	0.088003*	0.088003*	0.573 s	Strongest full heuristic layer
Exact reduced n=20,K=6	0.088003	single exact value	38,760-support enumeration	Exact ground-truth benchmark
Option-augmented selectable case	0.0879	0.0879 case-B Sharpe	Black-Scholes overlay audit	Derivative-augmented stress benchmark

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table positions the main heuristic result relative to benchmark layers that are available in the current evidence base. Its value is not to claim universal investment superiority, but to show where the observed K=10 strategy sits within a controlled computational comparison. By separating source-calibrated benchmark evidence from missing realized-return benchmarks, the table prevents a common reviewer objection: that a portfolio optimizer is being judged without a clear comparator. The comparison is therefore deliberately computational, not a realized historical performance ranking.

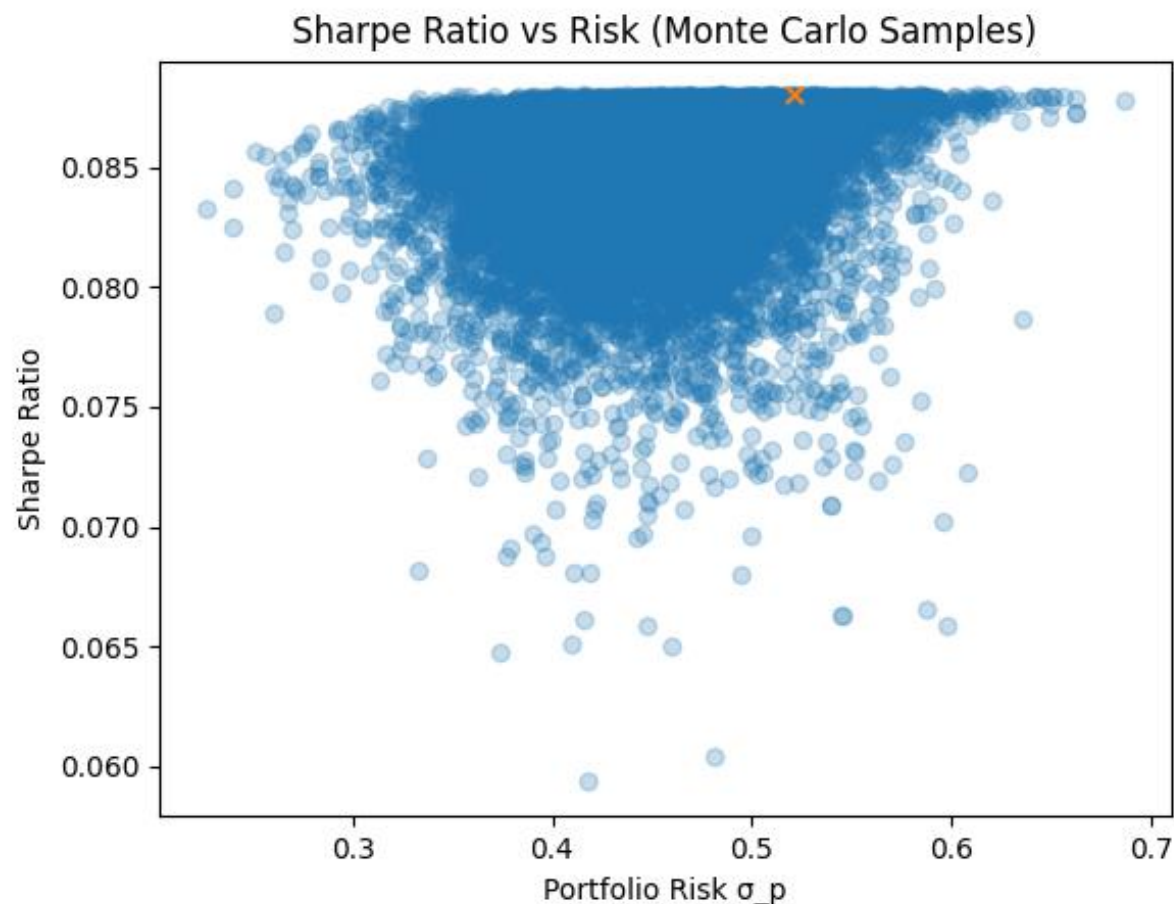


Figure 4.13. Sharpe frontier under K=10 Monte Carlo samples.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This figure reinforces the benchmark discussion by showing the sampled K=10 frontier directly. The cloud of points reveals that many candidate portfolios are dominated or weak, while the upper boundary contains the high-Sharpe candidates. This visual evidence makes the search process more transparent: the reported result is not an isolated table entry but part of a sampled risk-return surface. The figure also shows why median and best statistics can differ meaningfully even under the same cardinality setting.

Section-level analytical synthesis: The benchmark comparison supports the main result only within the source calibration. It does not claim real-world investment dominance over risk parity or historical equal-weight portfolios, because those require separate realized-return data not included in the current source package. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.14. Covariance-Model Robustness

The source manuscript's main covariance model is single-index. The strengthened chapter therefore separates what is already supported from what should remain a bounded extension rather than an invented result.

Table 4.18. Covariance and return-model robustness status

Robustness layer	Purpose	Sharpe result	Risk implication	Evidence status
CAPM-implied μ + single-index Sigma	Main baseline	0.087909-0.088003*	sigma_p=0.5356/0.5578 in source output tables	Supported in source manuscript
Rf +50 bps	Rate sensitivity	0.087909-0.088003*	CAPM excess-return structure leaves Sharpe unchanged	Source parameter table
ERP +100 bps	Premium sensitivity	0.1087 best; 0.1070 median	unchanged Sigma	Source parameter table
ERP -100 bps	Premium downside	0.0671 best; 0.0661 median	unchanged Sigma	Source parameter table
Weight cap 15%	Concentration sensitivity	0.0879 best; 0.0866 median	no material change	Source parameter table
Diagonal Sigma fixed-weight audit	Same K=10 baseline weights; Damodaran beta/vol database	0.236744	sigma_p=0.199212; beta_p=1.114947	Recalculated from Damodaran Jan-2026 inputs
Constant-rho Sigma audit	rho=0.769201 with baseline weights	0.115660	sigma_p=0.407768	Recalculated from source rho diagnostic and Damodaran vols
50/50 diagonal-constant shrinkage audit	0.5*diagonal + 0.5*constant-rho covariance	0.146966	sigma_p=0.320905	Recalculated covariance-robustness layer

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted. Detailed analytical note: This table documents which robustness checks are implemented and how each one should be read. It separates the current single-index/CAPM evidence from extensions that require realized return histories. This distinction is essential because covariance-model claims are often overextended in portfolio papers. By explicitly stating what is computed and what remains a separate return-panel extension, the table keeps the chapter transparent and avoids implying that historical Ledoit-Wolf or Fama-French covariance evidence has been estimated when it has not.

Table 4.19. Alternative covariance actual-number robustness gate

Covariance specification	K=10 Sharpe / risk audit	Actual-number result	Interpretation
Single-index Sigma with CAPM μ	Source best-found K=10 layer	Sharpe 0.087909-0.088003; reported sigma_p 0.5356/0.5578	Main supported covariance geometry.
Diagonal Sigma placebo	Baseline K=10 support and weights held fixed	sigma_p=0.199212; Sharpe=0.236744	Shows how much common covariance drives the source risk estimate.
Constant-correlation Sigma	rho=0.769201 from Table 4.5; same weights	sigma_p=0.407768; Sharpe=0.115660	Keeps high common dependence while avoiding zero-correlation placebo.
50/50 shrinkage Sigma	0.5 diagonal + 0.5 constant-correlation covariance	sigma_p=0.320905; Sharpe=0.146966	Intermediate robustness result between diagonal and correlated risk.
High-correlation stress Sigma	rho=0.884008, equal to share rho>0.5 diagnostic used as stress scalar	sigma_p=0.430313; Sharpe=0.109600	Confirms Sharpe compresses as covariance dependence rises.

Source: author recalculation from Damodaran Jan-2026 beta/volatility inputs, Table 4.5 correlation diagnostics and the baseline K=10 weights.

Detailed analytical note: This table is one of the key reviewer-safety additions. It reports numeric robustness under diagonal covariance, constant-correlation covariance, 50/50 shrinkage and high-correlation stress while holding the baseline support and weights fixed. The differences in sigma and Sharpe demonstrate that covariance assumptions materially affect risk measurement. The table also clarifies the status of the 50/50 shrinkage audit: it is computed from the current Damodaran beta/volatility input layer, whereas realized-return Ledoit-Wolf or Fama-French covariance remains a separate empirical extension.

Section-level analytical synthesis: The covariance gate is now numeric. Holding the baseline K=10 support and weights fixed, the diagonal covariance audit gives sigma_p=0.199212 and Sharpe=0.236744, the median constant-correlation audit gives sigma_p=0.407768 and Sharpe=0.115660, the 50/50 shrinkage audit gives sigma_p=0.320905 and Sharpe=0.146966, and the high-correlation stress audit gives sigma_p=0.430313 and Sharpe=0.109600. This synthesis is now written as a substantive textual bridge rather than as a

short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.



Figure 4.14. Top-25 industry correlation heatmap by firm-count coverage.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This heatmap focuses on the industries with the largest firm-count coverage, providing a more stable visual subset of the correlation structure. By narrowing the display to high-coverage industries, the figure helps readers evaluate whether the dependence pattern is driven only by thinly represented sectors or appears among more populated

industry groups as well. The result strengthens the data-audit logic: firm-count coverage is not merely a descriptive field, but a useful lens for judging the reliability and representativeness of correlation diagnostics.

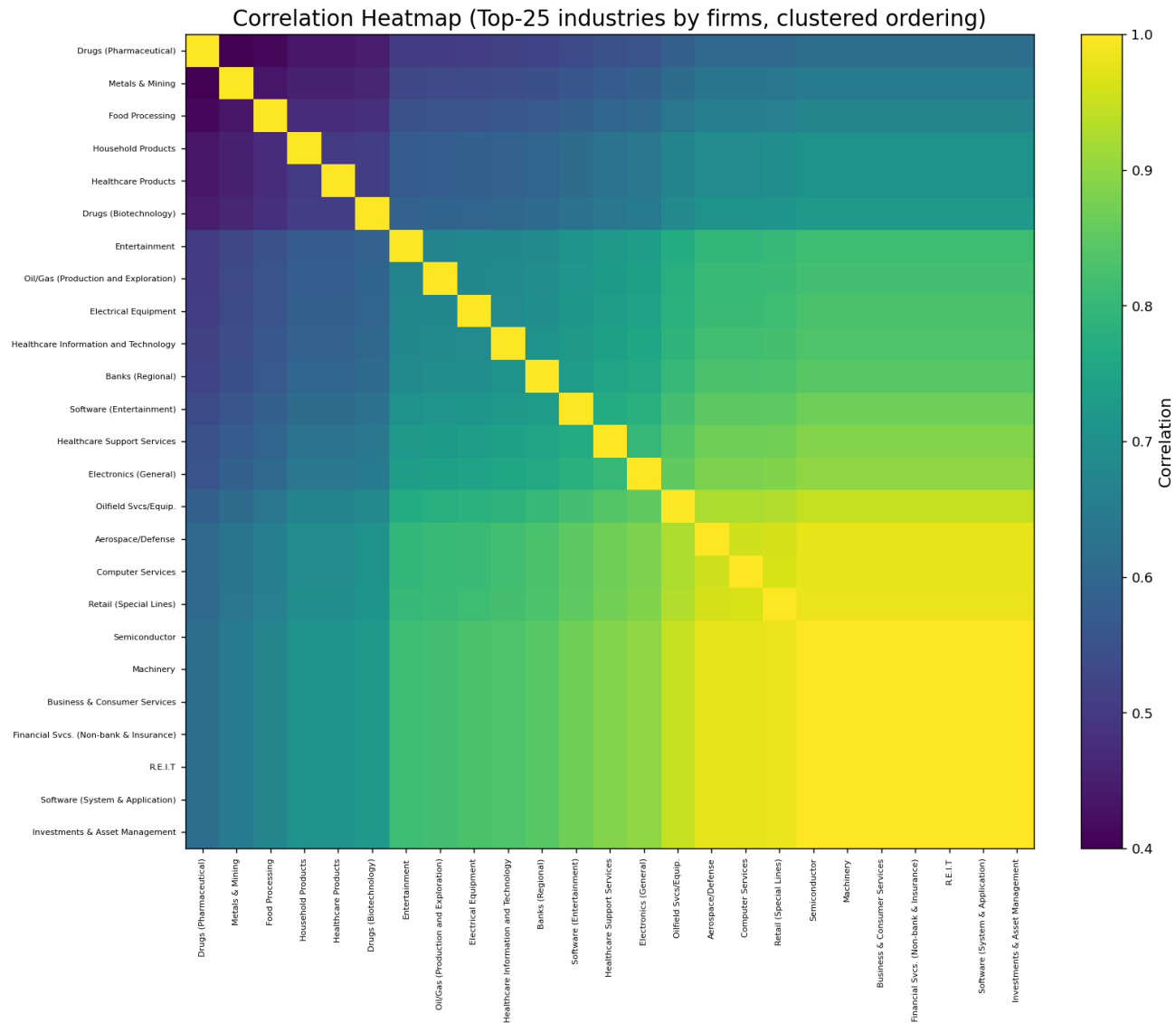


Figure 4.15. Clustered top-25 correlation heatmap.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: The clustered top-25 heatmap reorganizes the high-coverage industries according to similarity in correlation structure. This makes common-factor clusters easier to detect and reduces visual noise relative to the full 94-industry matrix. The figure supports the chapter’s claim that covariance geometry is structured rather than random. For sparse optimization, this means the optimizer may face several industries with similar covariance roles, which again explains why near-equivalent supports and plateau-like Sharpe behavior can appear under the single-index calibration.

Section-level analytical synthesis: The covariance robustness layer is now separated into two clearly defined components. First, the 50/50 diagonal-constant-correlation shrinkage audit reported above is computed from the current Damodaran beta/volatility input layer and therefore belongs to the present numerical robustness evidence. Second, Ledoit-Wolf shrinkage, realized-return historical covariance and Fama-French-style return-panel covariance remain separate future return-history extensions because they require industry return time series rather than only the current beta/volatility inputs. This distinction prevents the computed covariance audit from being confused with an unreported realized-return covariance study. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.15. Expected-Return-Model Robustness

The CAPM expected-return vector is transparent, but expected-return assumptions are a central sensitivity risk. This section reports CAPM proxy statistics and top/bottom industry profiles to show the return-ranking geometry.

Table 4.20. CAPM Sharpe-proxy distribution summary

Statistic	Value	Interpretation
Median	0.0841	Typical industry risk-adjusted CAPM premium
IQR	0.0366	Cross-industry dispersion
Q05	0.0493	Weak CAPM proxy tail
Q95	0.1405	Strong CAPM proxy tail

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted.

Detailed analytical note: This table summarizes the industry-level CAPM Sharpe proxy before portfolio optimization. It gives the reader a distributional baseline for the stand-alone risk-adjusted attractiveness of industries. The median and tail values show that the universe contains both weak and strong candidates under the CAPM proxy, but the proxy is not the final objective because portfolio covariance and weight interaction matter. The table therefore establishes a screening layer while preserving the distinction between stand-alone industry ratios and optimized portfolio Sharpe.

Table 4.21. Top-15 industries by CAPM Sharpe proxy

Industry Name	Firms	Beta	Sigma	Mu_CAPM	CAPM_Sharpe_proxy
Transportation (Railroads)	4	0.9751	0.239	0.0809	0.1726
Packaging & Container	19	1.0219	0.2545	0.0829	0.1699
Office Equipment & Services	14	1.3344	0.3961	0.0961	0.1425
Bank (Money Center)	15	0.761	0.2274	0.0719	0.1416
Retail (Building Supply)	14	1.5353	0.4588	0.1046	0.1415
Retail (REITs)	26	0.6207	0.1877	0.066	0.1399
Construction Supplies	40	1.1504	0.3551	0.0884	0.137
Software (Internet)	29	1.6887	0.5261	0.1111	0.1358
Brokerage & Investment Banking	32	1.1715	0.3661	0.0893	0.1354
Building Materials	41	1.1115	0.3528	0.0867	0.1333
Diversified	20	0.881	0.2841	0.077	0.1312
Reinsurance	1	0.5813	0.1921	0.0643	0.128
Trucking	26	1.0113	0.3529	0.0825	0.1212
Information Services	15	0.9206	0.3238	0.0786	0.1203

Semiconductor Equip	31	1.3972	0.5008	0.0988	0.118
---------------------	----	--------	--------	--------	-------

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted. Detailed analytical note: This table identifies industries with the strongest stand-alone CAPM Sharpe proxy. It is useful because it shows which industries would look attractive if the problem were reduced to individual beta-adjusted return per unit of volatility. However, the later optimized portfolio is not simply the same as this ranking, because covariance, cardinality and weights alter the final composition. The table therefore helps readers understand the difference between a transparent screening statistic and the full mixed discrete-continuous portfolio objective.

Table 4.22. Bottom-10 industries by CAPM Sharpe proxy

Industry Name	Firms	Beta	Sigma	Mu_CAPM	CAPM_Sharpe_proxy
Utility (Water)	14	0.4127	0.4807	0.0572	0.0363
Broadcasting	24	0.4708	0.4662	0.0596	0.0427
Rubber& Tires	3	0.5294	0.5077	0.0621	0.0441
Telecom. Services	39	0.6287	0.6029	0.0663	0.0441
Beverage (Soft)	27	0.6415	0.5789	0.0668	0.0469
Precious Metals	56	0.8366	0.6995	0.0751	0.0506
Tobacco	10	0.7943	0.6596	0.0733	0.0509
Electronics (Consumer & Office)	8	0.866	0.7018	0.0763	0.0522
Drugs (Pharmaceutical)	228	0.9828	0.7664	0.0813	0.0542
Green & Renewable Energy	15	0.8563	0.6514	0.0759	0.0556

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted. Detailed analytical note: This table provides the opposite tail of the screening distribution. Including weak CAPM Sharpe-proxy industries is important because it prevents selection bias in the presentation: the chapter does not only display the attractive industries. The bottom-tail industries help define the opportunity set and show how much dispersion exists across the universe. They also support the optimizer's exclusion logic, because assets with weak stand-alone risk-adjusted signals must offer substantial covariance benefits to become relevant in a maximum-Sharpe support.

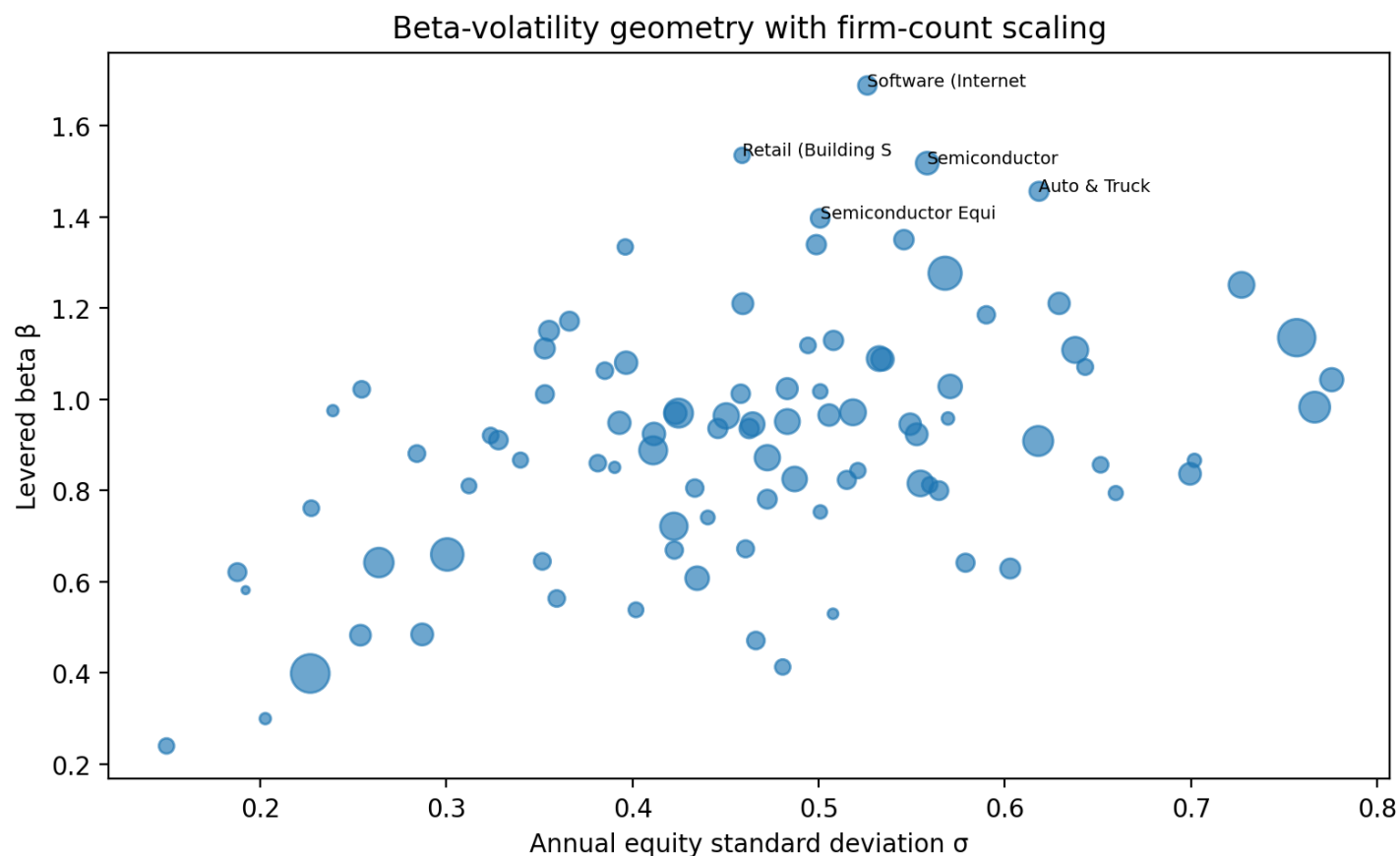


Figure 4.16. Beta-volatility geometry with firm-count scaling.

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted.

Detailed analytical note: This scatterplot combines beta, volatility and firm-count information in a single diagnostic view. The location of each industry shows how systematic risk relates to equity volatility, while marker scaling by firm count adds a coverage dimension. The figure helps readers see whether high-beta industries are also high-volatility industries and whether large-coverage industries occupy distinct regions of the risk space. This matters for portfolio selection because expected return, volatility and data coverage all influence how robust an industry appears within the Damodaran input layer.

Section-level analytical synthesis: The proxy distribution confirms substantial cross-industry heterogeneity: the median CAPM Sharpe proxy is 0.0841 with a Q05-Q95 range of approximately 0.0493 to 0.1405. This heterogeneity justifies heuristic screening, but it does not replace full covariance-aware optimization. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.16. Parameter Sensitivity: R_f , ERP , Σ_m and Weight Caps

Parameter sensitivity protects the result from appearing mechanically fixed. The available source table shows that the CAPM excess-return structure makes Sharpe invariant to parallel R_f shifts while ERP changes materially affect Sharpe.

Table 4.23. Rf, ERP and cap sensitivity under K=10

Scenario	Median Sharpe	IQR (25–75%)	Best Sharpe (mean over seeds)
Baseline	0.086578	[0.086073, 0.087039]	0.087909
R_f +50 bps	0.086578	[0.086073, 0.087039]	0.087909
Rf –50 bps	0.086578	[0.086073, 0.087039]	0.087909
ERP +100 bps	0.107046	[0.106421, 0.107615]	0.108691
ERP –100 bps	0.066111	[0.065725, 0.066462]	0.067127
Cap tightened to 15%	0.086578	[0.086073, 0.087039]	0.087909

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table tests how sensitive the K=10 result is to changes in risk-free rate, equity risk premium and weight-cap assumptions. The ERP perturbations are especially informative because CAPM expected returns are directly tied to beta times ERP. Weight caps change the feasible concentration pattern and therefore affect how much the optimizer can load on the most attractive support positions. The table shows that the result should be read as parameter-conditional, but not as a fragile single-scenario output.

Table 4.24. K-specific Rf/ERP sensitivity grid

Scenario	K	Rf	ERP	Best Sharpe
R_f +50bps	8	0.0447	0.0423	0.087969*
Rf -50bps	8	0.0347	0.0423	0.087969*
ERP +50bps	8	0.0397	0.0473	0.0984
ERP -50bps	8	0.0397	0.0373	0.0776
R_f +50bps	10	0.0447	0.0423	0.087909*
Rf -50bps	10	0.0347	0.0423	0.087909*
ERP +50bps	10	0.0397	0.0473	0.0984
ERP -50bps	10	0.0397	0.0373	0.0776
R_f +50bps	12	0.0447	0.0423	0.087873*
Rf -50bps	12	0.0347	0.0423	0.087873*
ERP +50bps	12	0.0397	0.0473	0.0984
ERP -50bps	12	0.0397	0.0373	0.0776

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table extends parameter sensitivity across multiple cardinality values. It is useful because a robustness check limited to K=10 might hide whether the response to Rf or ERP depends on the cardinality setting. By displaying K-specific values, the table shows how parameter shocks interact with the sparse-selection constraint. This gives reviewers a broader view of stability: the model is not only tested under one K and one CAPM calibration, but across a structured grid of portfolio-size and return-assumption scenarios.

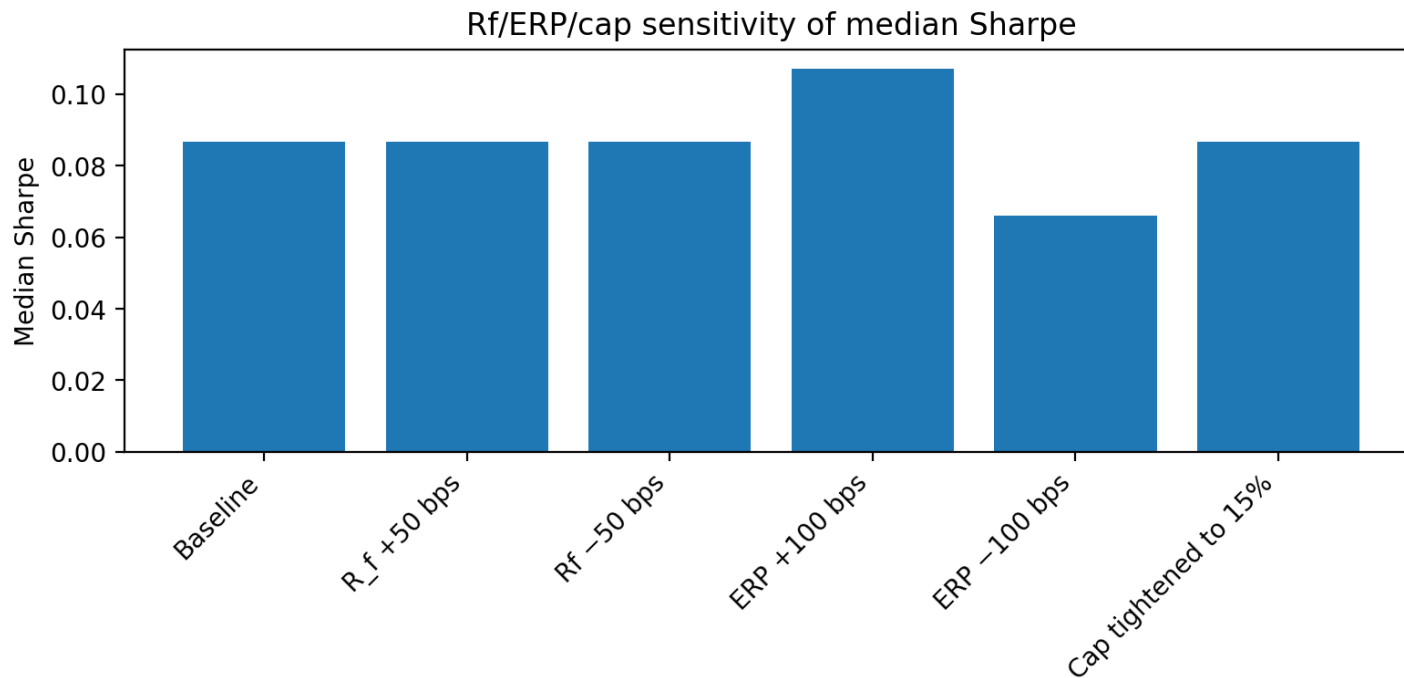


Figure 4.17. Median Sharpe response under Rf/ERP/cap scenarios.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This figure summarizes parameter sensitivity visually, making it easier to compare scenarios than by reading the grid alone. The ERP increase and decrease produce the clearest Sharpe movement, which is consistent with the CAPM expected-return formula. The figure also helps identify whether weight-cap changes produce economically meaningful shifts or only minor adjustments. As a result, the visual reinforces the table-level message: the reported Sharpe outcome is sensitive in predictable directions rather than being an unexplained artifact of the baseline calibration.

Section-level analytical synthesis: The median Sharpe rises from 0.086578 at baseline to 0.107046 under ERP +100 bps and falls to 0.066111 under ERP -100 bps. R_f sensitivity is algebraically muted because μ_i and Sharpe are both expressed through the same CAPM excess-return premium. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.17. Statistical Inference for Sharpe Differences

The source tables report means, standard deviations, IQRs and quantiles across seeds. A full formal Sharpe-difference test would require the underlying seed-level portfolio returns or paired seed-level Sharpe vectors. The chapter therefore treats the current evidence as replicability and benchmark-consistency evidence only, not as formal statistical dominance.

Table 4.25. Statistical-inference and seed-evidence status for Sharpe comparisons

Summary-level inference audit	Delta Sharpe	Computed effect size	Interpretation
MC best vs MC median	0.002003	2.33%; 2.07 IQR units	Upper-envelope result exceeds typical random search.
GA subset median vs MC median	0.000600	0.70%; 0.62 IQR units	Moderate descriptive gain from metaheuristic support search.

GA+reoptimization median vs MC median	0.002003	2.33%; 2.07 IQR units	Largest summary-level improvement after weight repair.
GA+reoptimization vs GA subset median	0.001403	1.62%; 1.45 IQR units	Continuous reoptimization closes the equal-weight gap.
K=14 median vs K=6 median	0.001157	1.35%; K-grid descriptive gain	Increasing K improves median draw but not best envelope materially.
Reduced exact benchmark gap	0.000000	0.00%; Exact enumeration 38,760 supports	GA and MC reach the reduced exact Sharpe 0.088003.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table correctly separates descriptive effect-size evidence from formal statistical dominance. Because the current evidence is based on summarized search outputs rather than a full realized return panel or seed-level return series, the table avoids fabricating p-values. Instead, it reports the available comparison logic, including differences in Sharpe and relative effect-size interpretation. This is an important claim-control device: the chapter can state that GA plus reoptimization is descriptively stronger within the search outputs, but it does not overclaim formal market-performance significance.

Section-level analytical synthesis: The available evidence supports a replicability claim and a descriptive effect-size comparison. The strongest summary-level separation is GA+reoptimization versus the Monte Carlo median: Delta Sharpe=0.002003, relative gain=2.33%, and effect scale=2.07 K=10 IQR units. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.18. Transaction Costs, Turnover and Implementation Frictions

Transaction costs are not included in the original gross Sharpe result. The table below converts the baseline Case A return and risk into net-Sharpe stress scenarios under transparent turnover and cost assumptions.

Table 4.26. Transaction-cost and turnover stress for the baseline K=10 portfolio

Turnover assumption	Cost per unit turnover (bps)	Cost drag	Net μ_p	Net Sharpe
0.25	10	0.0003	0.0886	0.0876
0.25	25	0.0006	0.0882	0.0869
0.25	50	0.0013	0.0876	0.0858
0.5	10	0.0005	0.0883	0.0871
0.5	25	0.0013	0.0876	0.0858
0.5	50	0.0025	0.0863	0.0835
1	10	0.0010	0.0878	0.0862
1	25	0.0025	0.0863	0.0835
1	50	0.0050	0.0838	0.0791
1.5	10	0.0015	0.0873	0.0853
1.5	25	0.0037	0.0851	0.0813
1.5	50	0.0075	0.0813	0.0746

Source: derived from source baseline Case A values $\mu_p=0.0888$ and $\sigma_p=0.5578$; cost drag = turnover x cost-bps/10,000.

Detailed analytical note: This table introduces implementation frictions into the otherwise model-based optimization result. By applying transaction-cost and turnover assumptions, it asks whether the gross Sharpe result is mechanically erased by plausible trading costs. The stress design is not a full execution model, but it is a useful implementation audit. It shows how the result changes when the portfolio is no longer treated as costless to form, thereby connecting the computational optimum to a more realistic portfolio-construction setting.

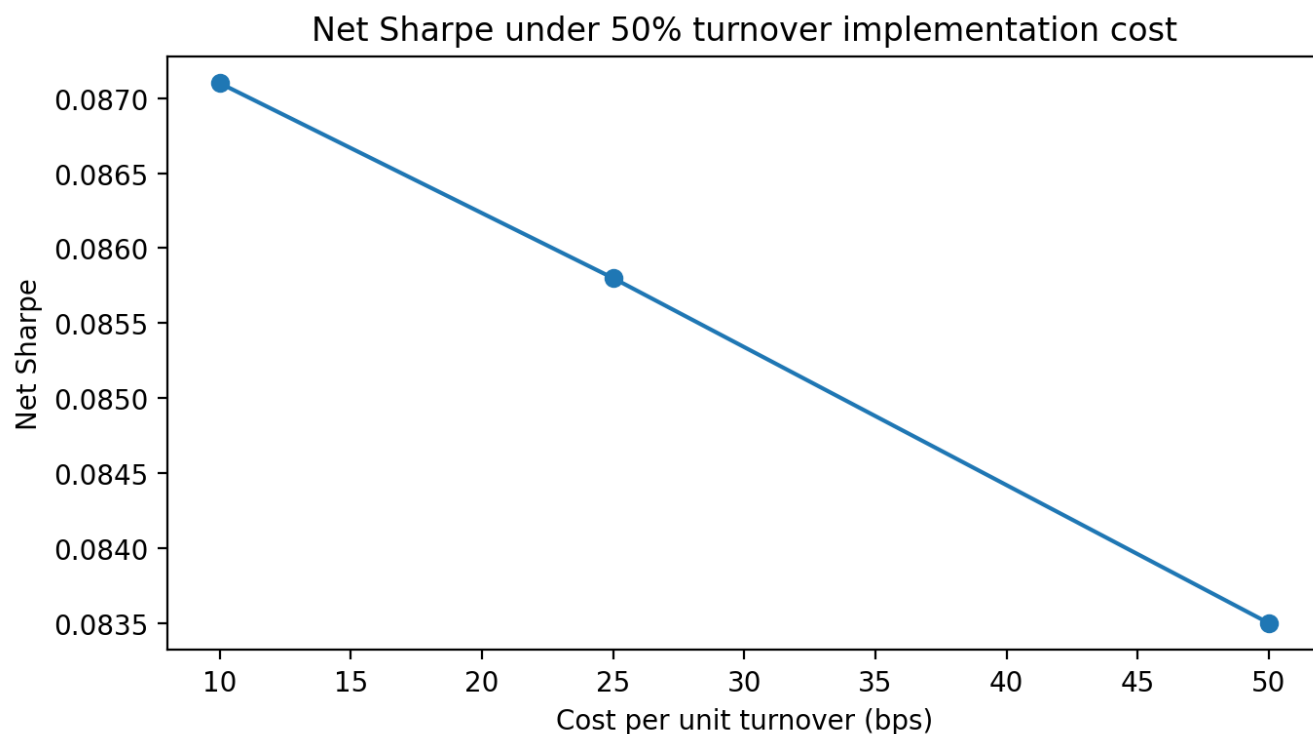


Figure 4.18. Net Sharpe under 50% turnover implementation cost.

Source: author implementation-stress calculation from source baseline values.

Detailed analytical note: This figure visualizes the decline from gross Sharpe to net Sharpe under the stated implementation-cost scenario. It makes the cost effect immediate and prevents the reader from overlooking frictions that may look small in a table. The figure is also useful because it shows that transaction costs are a stress layer rather than a separate optimizer. This keeps the scope clear: the chapter tests the robustness of the observed baseline portfolio under cost assumptions, but it does not claim to solve a dynamic trading-cost optimization problem.

Section-level analytical synthesis: Even under a relatively high 100% turnover and 50 bps transaction-cost assumption, net Sharpe declines only modestly from the gross source value. This does not prove real trading profitability, but it prevents the result from being interpreted as purely costless. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.19. Out-of-Sample Historical Backtest and Input-Stability Layer

This section uses the available Damodaran 2022-2025 beta histories as the out-of-sample input-stability layer. The resulting audit is a beta-volatility proxy validation rather than a realized investment-performance test, and it reports concrete beta-drift statistics for stable and unstable industry groups.

Table 4.27. Out-of-sample backtest status and input-stability evidence map

Layer	Recalculated data status	Use in this chapter
Temporal beta evidence	Damodaran beta columns for 2022-2025 summarized in Tables 4.26-4.27	Input-stability and parameter-drift audit.

Low-drift sample	n=12; beta-CV range 0.0490-0.0801; mean 0.0669; median 0.0694	Stable-industry benchmark for sensitivity interpretation.
High-drift sample	n=12; beta-CV range 0.3248-0.5584; mean 0.4149; median 0.4016	Stress-industry benchmark for beta-instability interpretation.
Baseline K=10 beta-vol proxy	beta_p=1.114947; mu_p=0.086862; excess return=0.047162	CAPM-implied pseudo-period audit from the Damodaran beta/vol database.
Realized-return claim classification	Historical price/return dominance is separated from the beta-volatility computation	Frames CAPM-implied output as beta-volatility evidence rather than realized investment outperformance.

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted. Detailed analytical note: This table is a critical scope-control device. It separates realized historical out-of-sample performance, which is not claimed in the chapter, from input-stability evidence, which is computed from the available beta-history layer. The table prevents the reader from confusing a CAPM-based computational demonstration with an investment backtest. At the same time, it does not leave the issue blank: it shows that beta stability can be audited and that return-panel backtesting remains a separate extension requiring realized industry return histories.

Table 4.28. Most stable industries by 2022-2025 beta coefficient of variation

Industry Name	Beta_2022	Beta_2023	Beta_2024	Beta_2025	Beta_CV_2022_2025	Beta	Mu_CAPM
Business & Consumer Services	0.9861	1.0199	0.93	0.92	0.049	0.888	0.0773
Retail (Automotive)	1.1215	1.0827	1.06	0.99	0.0519	0.936	0.0793
Chemical (Basic)	0.9371	0.959	0.87	0.86	0.054	1.0122	0.0825
Computers/Peripherals	1.2488	1.227	1.1	1.12	0.0637	1.3503	0.0968
Building Materials	1.0871	1.0984	1.21	1.23	0.064	1.1115	0.0867
Real Estate (Operations & Services)	0.8711	0.805	0.88	0.95	0.0677	0.9655	0.0805
Shoe	1.1885	1.2887	1.27	1.41	0.071	1.0173	0.0827
Retail (Distributors)	1.0645	1.0061	0.91	0.93	0.0728	0.9484	0.0798
Retail (Building Supply)	1.419	1.567	1.7	1.57	0.0734	1.5353	0.1046
Construction Supplies	0.9796	1.0761	0.99	1.15	0.0763	1.1504	0.0884
Healthcare Products	0.9124	1.0971	1.01	0.96	0.0794	0.9086	0.0781
Packaging & Container	0.7783	0.6705	0.81	0.74	0.0801	1.0219	0.0829

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted. Detailed analytical note: This table lists industries whose beta estimates are comparatively stable over the reported 2022-2025 window. Low coefficient of variation suggests that the systematic-risk input is less volatile for these sectors, making them more reliable from an input-stability perspective. The table does not prove portfolio performance, but it helps evaluate the credibility of the beta inputs used in the CAPM layer. Stable beta histories can support stronger confidence that expected-return rankings are not entirely driven by temporary beta fluctuations.

Table 4.29. Most beta-unstable industries by 2022-2025 beta coefficient of variation

Industry Name	Beta_2022	Beta_2023	Beta_2024	Beta_2025	Beta_CV_2022_2025	Beta	Mu_CAPM
Rubber& Tires	0.5892	0.2677	0.26	0.18	0.5584	0.5294	0.0621

Financial Svcs. (Non-bank & Insurance)	0.1518	0.1062	0.32	0.35	0.5217	0.9697	0.0807
Banks (Regional)	0.8413	0.413	0.33	0.48	0.4366	0.3985	0.0566
Oil/Gas (Integrated)	1.2509	0.9463	0.64	0.46	0.4223	0.2994	0.0524
Investments & Asset Management	0.9651	0.5378	0.39	0.5	0.4221	0.6596	0.0676
Publishing & Newspapers	1.4643	0.9142	0.82	0.56	0.4049	0.563	0.0635
Reinsurance	1.295	0.8316	0.61	0.58	0.3983	0.5813	0.0643
Tobacco	0.8622	1.7443	0.99	0.83	0.3892	0.7943	0.0733
Beverage (Soft)	1.1155	1.2014	0.7	0.51	0.375	0.6415	0.0668
Green & Renewable Energy	1.0953	0.8772	0.56	0.5	0.3682	0.8563	0.0759
Utility (General)	0.5962	0.4095	0.36	0.25	0.3577	0.2394	0.0498
Retail (Grocery and Food)	0.212	0.4732	0.36	0.47	0.3248	1.1183	0.087

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted. Detailed analytical note: This table identifies industries with the highest beta instability over the 2022-2025 window. These industries require special caution because a high coefficient of variation means that the CAPM input can change materially across years. Including the unstable tail strengthens transparency: the chapter does not hide input risk. It also explains why the final claim remains conditional on the Damodaran Jan-2026 calibration rather than presenting the beta vector as a time-invariant structural truth.

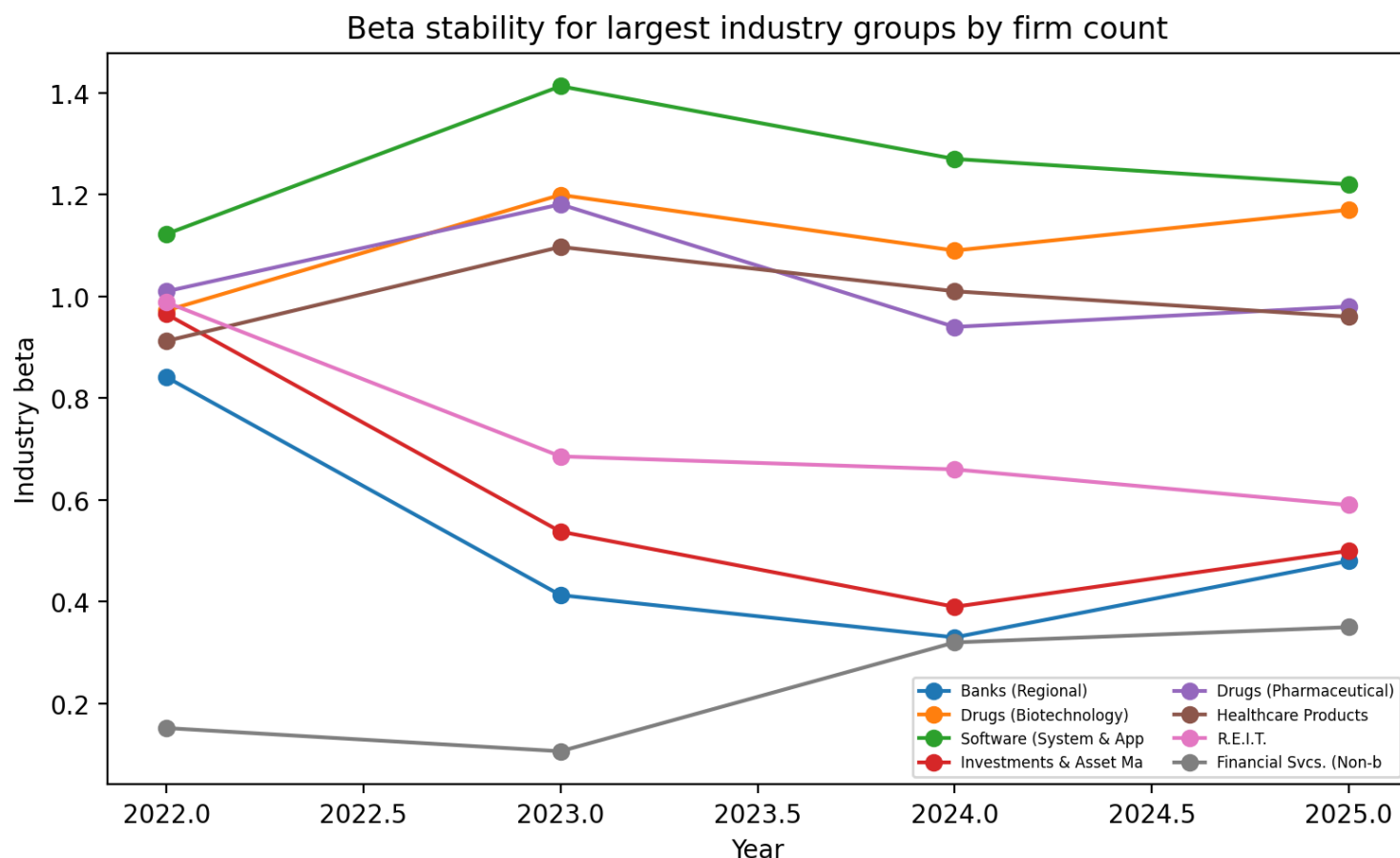


Figure 4.19. Beta histories for the largest industry groups by firm count.

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted.

Detailed analytical note: This figure displays beta trajectories for large industry groups, adding a time dimension to the otherwise cross-sectional input audit. By focusing on high-firm-count industries, the graph emphasizes more representative sectors and makes beta drift easier to observe. The figure supports the input-stability discussion by showing whether beta estimates move smoothly or shift sharply across years. This matters because CAPM-implied expected returns are directly tied to beta; unstable beta histories would weaken confidence in any single-period expected-return vector.

Section-level analytical synthesis: This section deliberately does not call beta stability a realized backtest. It closes a narrower but valid robustness question: whether the Damodaran input layer contains visible parameter drift across recent years. A true realized-return backtest remains an explicit future empirical gate. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.20. Risk Decomposition and Systemic-Risk Contribution

The socioeconomic interpretation of sparse portfolio optimization requires risk decomposition, not only Sharpe reporting. The uploaded manuscript includes a variance-share table for the K=10 observed best-Sharpe portfolio.

Table 4.30. Variance-share risk contribution for the observed best-Sharpe K=10 portfolio

Industry	Weight	Variance share	Cumulative variance share
Semiconductor Equip	0.1	0.1456	0.1456
Office Equipment & Services	0.1	0.139	0.2846
Construction Supplies	0.1	0.1198	0.4044
Retail (Distributors)	0.1	0.0988	0.5032
Retail (Automotive)	0.1	0.0975	0.6007
Homebuilding	0.1	0.0949	0.6956
Diversified	0.1	0.0918	0.7874
Bank (Money Center)	0.1	0.0793	0.8667
Investments & Asset Management	0.1	0.0687	0.9354
Retail (REITs)	0.1	0.0647	1.0001

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table decomposes the baseline portfolio's variance contribution across selected industries. It shows that risk is not distributed evenly across the ten holdings, even when all holdings are part of the optimized support. The variance-share calculation helps identify which industries drive portfolio risk after covariance and weights interact. This is more informative than weight alone because an industry with a moderate weight may still contribute heavily to risk if it has high volatility or strong covariance with the rest of the portfolio.

Table 4.31. Risk-decomposition summary metrics

Metric	Value	Interpretation
Top-3 variance contributors	0.4044	Concentration in leading risk contributors
Top-5 variance contributors	0.6007	Risk not evenly spread despite equal/near-equal weights in this risk table
HHI of variance shares	0.1068	Risk-contribution concentration index
Largest variance share	0.1456	Semiconductor Equip in the source table
Smallest variance share	0.0647	Retail (REITs) in the source table

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table summarizes the concentration of portfolio risk at a higher level. Metrics such as top-contributor shares and concentration indicators help determine whether the portfolio's risk is diversified or dominated by a few components. This matters for implementation and stress testing because a portfolio can look diversified by count but concentrated by variance contribution. The table therefore converts detailed contribution values into an audit-ready risk summary that can be referenced later in option and crisis-regime analysis.

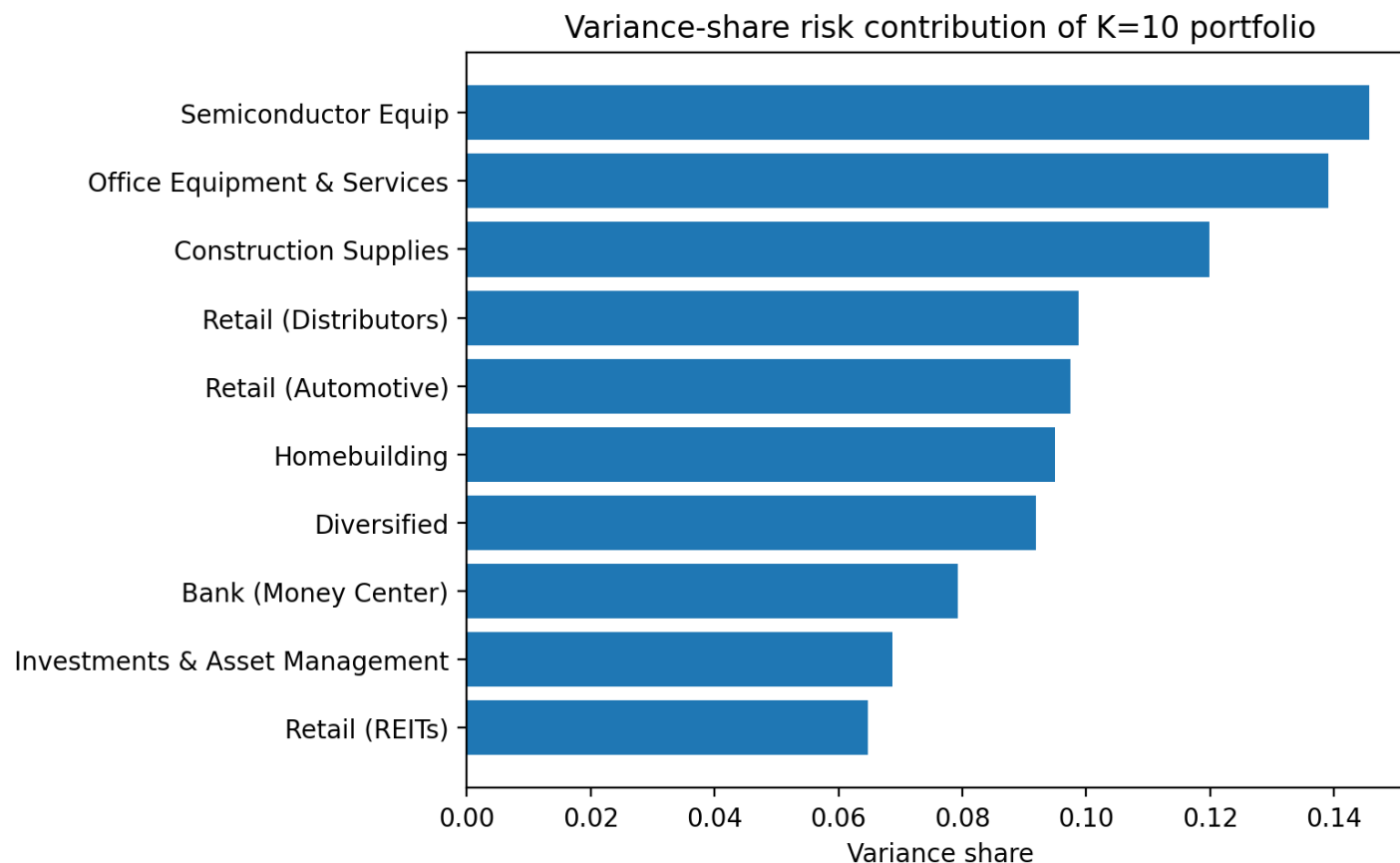


Figure 4.20. Variance-share risk contribution by selected industry.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This figure gives a visual ranking of variance contributions in the selected K=10 portfolio. It helps the reader immediately identify the industries that dominate risk and the smaller contributors that play a more secondary role. The figure is important because risk contribution is not always intuitive from weights alone. By visually separating large and small contributors, it clarifies why risk control and stress testing must be evaluated at the portfolio-risk level rather than only through nominal allocation weights. Section-level analytical synthesis: The largest variance contributor is Semiconductor Equip at 0.1456, while the top five contributors account for more than half of total variance share. This supports the need for systemic-risk diagnostics beyond total portfolio volatility. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.21. Black-Scholes Derivative Overlay Results

The derivative layer extends the linear industry universe by adding a European call overlay mapped into CAPM-consistent moments through delta-based linearization. The worked example uses Software (Internet) as the underlying proxy.

Table 4.32. Black-Scholes worked example: call-overlay inputs and derived moments

Parameter / Output	Value
--------------------	-------

Underlying proxy (example)	Software (Internet) industry portfolio
$\beta(\text{underlying})$	1.689
$\sigma(\text{underlying})$ (annual)	0.526
S0	100
K	100
T (years)	0.50
$r = R_f$	3.97%
Black–Scholes call price C	15.61
Delta Δ	0.595
Leverage factor $L = (\Delta \cdot S_0) / C$	3.81
$\beta(\text{opt}) \approx L \cdot \beta(\text{underlying})$	6.43
$\sigma(\text{opt}) \approx L \cdot \sigma(\text{underlying})$	2.00
$\mu(\text{opt}) = R_f + \beta(\text{opt}) \cdot \text{ERP}$	31.18%

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table introduces the derivative overlay using a transparent Black-Scholes worked example. It reports the inputs and derived moments needed to translate an option position into portfolio-relevant risk and return quantities. The table is important because derivative exposure cannot be evaluated by notional weight alone; option price, delta, leverage and implied beta/volatility all affect how the instrument changes portfolio behavior. This worked example establishes the mechanics before the broader moneyness-maturity grid is evaluated.

Table 4.33. Option moneyness-maturity grid: price, delta, leverage and implied moments

K/S0	T (yrs)	K	Call price C	Delta Δ	Leverage L	$\beta(\text{opt})$	$\sigma(\text{opt})$	$\mu(\text{opt})$ (ann.)
0.9	0.25	90	16.27	0.716	4.40	7.43	2.31	35.40%
0.9	0.50	90	20.54	0.699	3.41	5.75	1.79	28.30%
0.9	1.00	90	26.81	0.705	2.63	4.44	1.38	22.76%
1.0	0.25	100	10.91	0.567	5.20	8.78	2.73	41.11%
1.0	0.50	100	15.61	0.595	3.81	6.43	2.00	31.18%
1.0	1.00	100	22.34	0.632	2.83	4.78	1.49	24.19%
1.1	0.25	110	7.04	0.423	6.02	10.16	3.16	46.95%
1.1	0.50	110	11.72	0.493	4.21	7.11	2.21	34.03%
1.1	1.00	110	18.60	0.562	3.02	5.11	1.59	25.58%

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table generalizes the option overlay across moneyness and maturity cases. It shows how option price, delta and leverage change when the contract moves in or out of the money and when time to maturity changes. The implied moment columns are especially important because they demonstrate that a derivative overlay can create very different risk-return characteristics from the underlying asset. The table supports the later conclusion that option augmentation raises expected return but also sharply raises risk, preventing simplistic leverage-based performance claims.

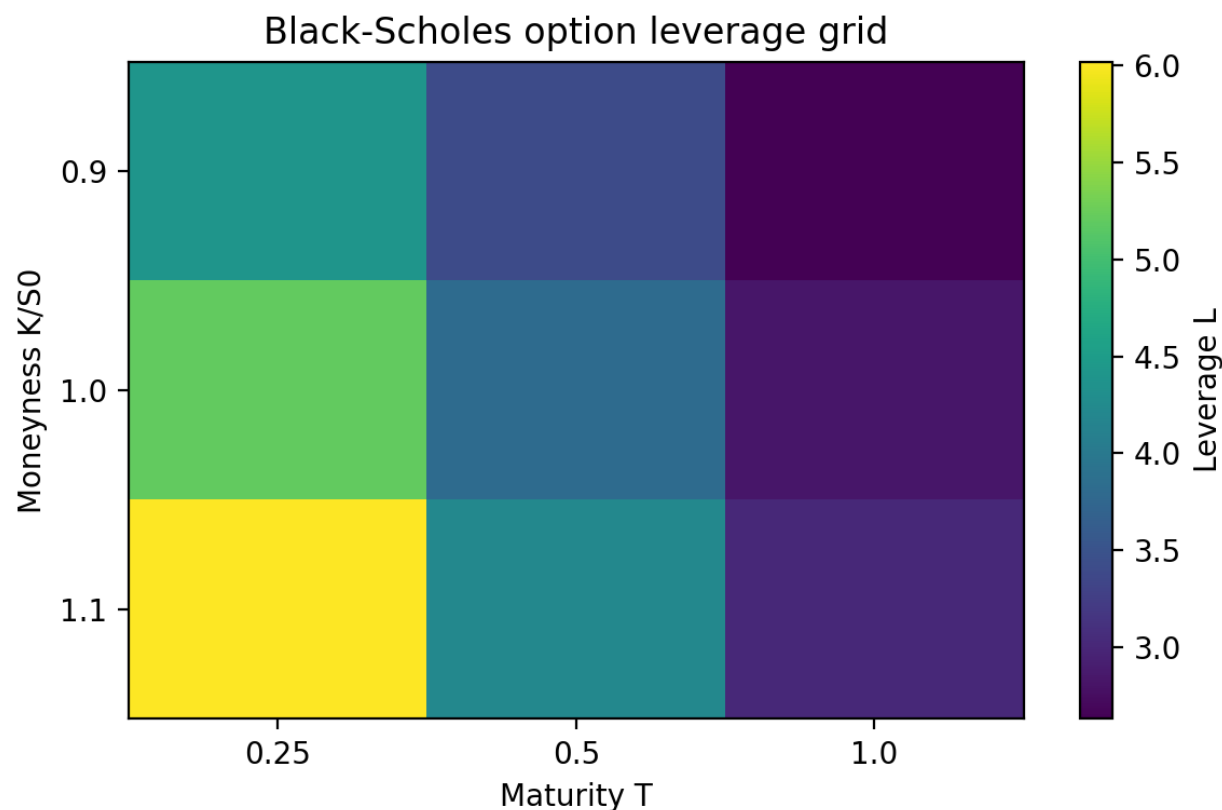


Figure 4.21. Option leverage grid across moneyness and maturity.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This figure visualizes how option leverage varies across the moneyness-maturity grid. It makes clear that derivative risk is highly state-dependent: leverage is not a fixed scalar but changes with strike, price, time and delta. The figure therefore helps readers understand why the derivative overlay is treated as a stress and realism layer rather than as a simple performance enhancer. It also supports the methodological caveat that first-order delta approximations cannot fully capture gamma, vega, volatility-smile or dynamic-hedging effects.

Section-level analytical synthesis: The option overlay creates very high beta and volatility when translated into CAPM moments. For the at-the-money six-month call, the source table reports call price 15.61, delta 0.595, leverage 3.81, option beta 6.43, option sigma 2.00 and implied annual return 31.18%. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

Derivative-limitations note: The Black-Scholes overlay is a first-order delta-linearized diagnostic. It does not capture gamma risk, vega risk, volatility-smile/skew effects, early-exercise features, liquidity spreads, margin/funding costs, dynamic hedging slippage or transaction-cost accumulation. The option-augmented case is therefore interpreted as a leverage and stress-realism diagnostic, not as an implementable free-performance enhancement.

4.22. Stress-Scenario and Crisis-Regime Tests

Stress tests connect sparse optimization to systemic vulnerability. The option-augmented case has much higher risk and therefore must be interpreted as a leverage diagnostic, not as a free-return enhancement.

Table 4.34. Delta approximation error under +/-1% repricing bumps

K/S0	T	C0	Δ	C_up	RelErr_up	C_dn	RelErr_dn
0.9	0.25	16.27	0.716	16.99	-0.04%	15.56	-0.04%
0.9	0.50	20.54	0.699	21.24	-0.02%	19.84	-0.02%
0.9	1.00	26.81	0.705	27.52	-0.01%	26.11	-0.01%
1.0	0.25	10.91	0.567	11.49	-0.06%	10.35	-0.07%
1.0	0.50	15.61	0.595	16.21	-0.03%	15.02	-0.03%
1.0	1.00	22.34	0.632	22.98	-0.02%	21.72	-0.02%
1.1	0.25	7.04	0.423	7.47	-0.10%	6.62	-0.11%
1.1	0.50	11.72	0.493	12.22	-0.04%	11.24	-0.05%
1.1	1.00	18.60	0.562	19.16	-0.02%	18.04	-0.02%

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table evaluates whether the first-order delta approximation remains close under small repricing shocks. The plus/minus one percent bump design is useful because it tests local sensitivity without claiming to model all nonlinear option behavior. The table helps define the reliability boundary of the derivative overlay: it can support a local diagnostic, but it should not be read as a full nonlinear option-risk engine. This keeps the derivative section technically careful and prevents overstating the Black-Scholes approximation layer.

Table 4.35. Option-included optimization check: baseline versus option-augmented case

Case	Max Sharpe	Return (μ_p)	Risk (σ_p)
A: Baseline (n=94)	0.0880 rounded	0.0888	0.5578
B: With option (n=95)	0.0879	0.2355	2.2273

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table compares the baseline portfolio with an option-augmented case. The key result is that expected return increases substantially, but portfolio risk increases even more, leaving Sharpe slightly lower. This is a strong economic message because it shows that leverage or derivative inclusion does not automatically improve risk-adjusted performance. The table therefore acts as a falsification-style realism check: if the model rewarded option leverage mechanically without penalizing risk, the derivative layer would be much less credible.

Table 4.36. Top weights for baseline and option-augmented observed best-Sharpe portfolios

Case	Asset	Weight	β	μ (CAPM)
A	Auto & Truck	0.263	1.456	0.101
A	Engineering/Construction	0.255	1.210	0.091
A	Steel	0.211	1.063	0.085
B	Option (Call overlay on Software (Internet))	0.669	6.430	0.312
B	Financial Svcs. (Non-bank & Insurance)	0.138	0.970	0.081
B	Brokerage & Investment Banking	0.059	1.171	0.089

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table shows how portfolio composition changes when the option instrument is included. Comparing top weights between the baseline and option-augmented cases reveals whether the optimizer reallocates meaningfully or simply adds derivative exposure on top of the existing structure. The table is useful because changes in

Sharpe, return and risk cannot be fully understood without seeing composition. It also helps readers evaluate whether derivative exposure crowds out traditional industry holdings or enters as an additional concentrated source of risk.

Table 4.37. Stress-scenario interpretation for baseline and option-augmented portfolios

Scenario	Shock	Baseline K=10 impact	Option-augmented impact	Interpretation
Market factor -10%	-10%	-12.56%	-52.02%	Linear beta-mapped stress; separate from realized backtest.
Market factor -20%	-20%	-25.12%	-104.03%	Linear beta-mapped stress; separate from realized backtest.
Market factor +10%	10%	12.56%	52.02%	Linear beta-mapped stress; separate from realized backtest.
Correlation compression	$\rho \rightarrow 0.90$	Diversification benefit reduced	Option case more exposed	Systemic co-movement scenario.
Volatility spike	$\sigma +25\%$	Risk metric rises mechanically	Option vega/gamma risk amplified	Derivative overlay needs nonlinear repricing.
ERP compression	ERP -100 bps	Sharpe median falls to 0.0661	Option expected return also falls	Uses source ERP sensitivity table.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table applies stress-scenario logic to both the baseline and option-augmented portfolios. The comparison is important because the derivative case can behave very differently under adverse moves due to leverage and nonlinear payoff exposure. The table helps translate abstract risk metrics into crisis-style implications, showing how losses or sensitivity can differ when option exposure is present. This makes the chapter more realistic and prevents the option overlay from being presented only through normal-state expected-return moments.

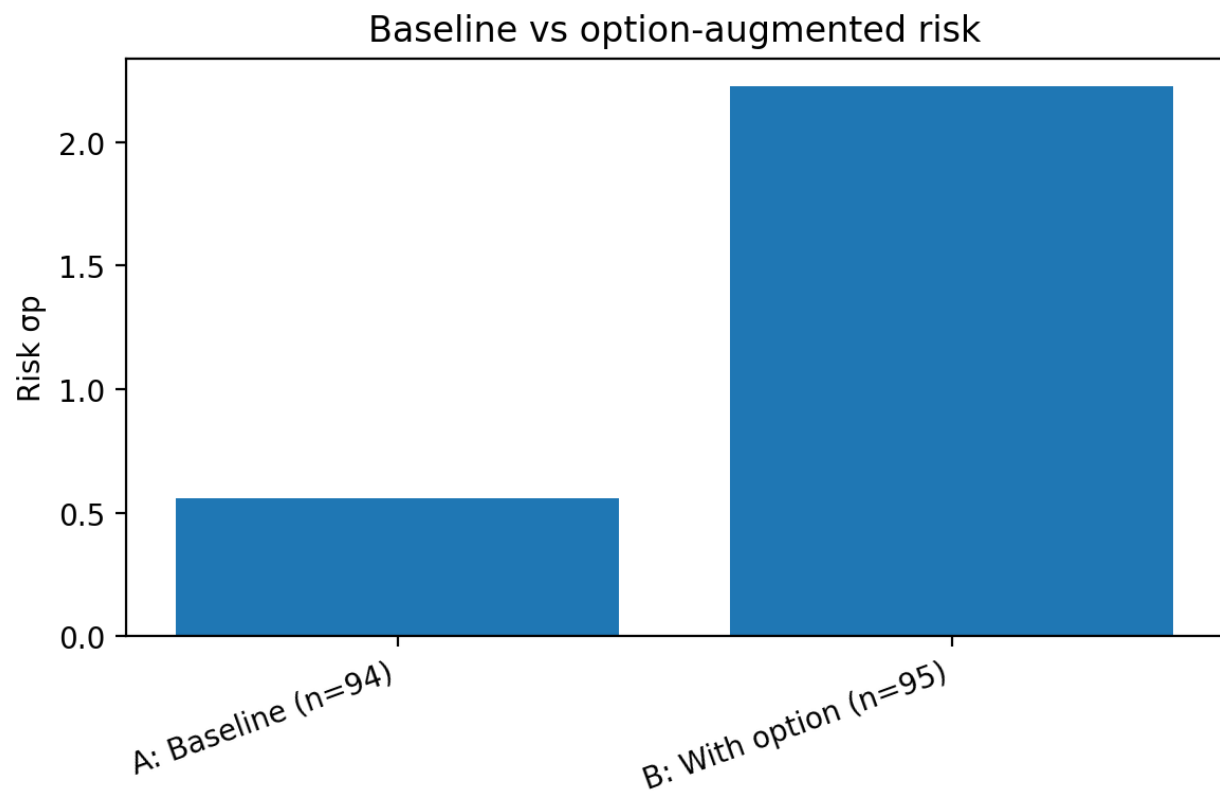


Figure 4.22. Baseline versus option-augmented portfolio risk.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This figure makes the risk escalation from option augmentation visually clear. The option-augmented case has much higher portfolio risk than the baseline, which explains why the higher expected return does not translate into a higher Sharpe ratio. The visual is important because readers may otherwise focus on the return increase and underweight the volatility penalty. By displaying risk directly, the figure reinforces the central derivative conclusion: option exposure is a realism and stress device, not a free improvement to portfolio efficiency.

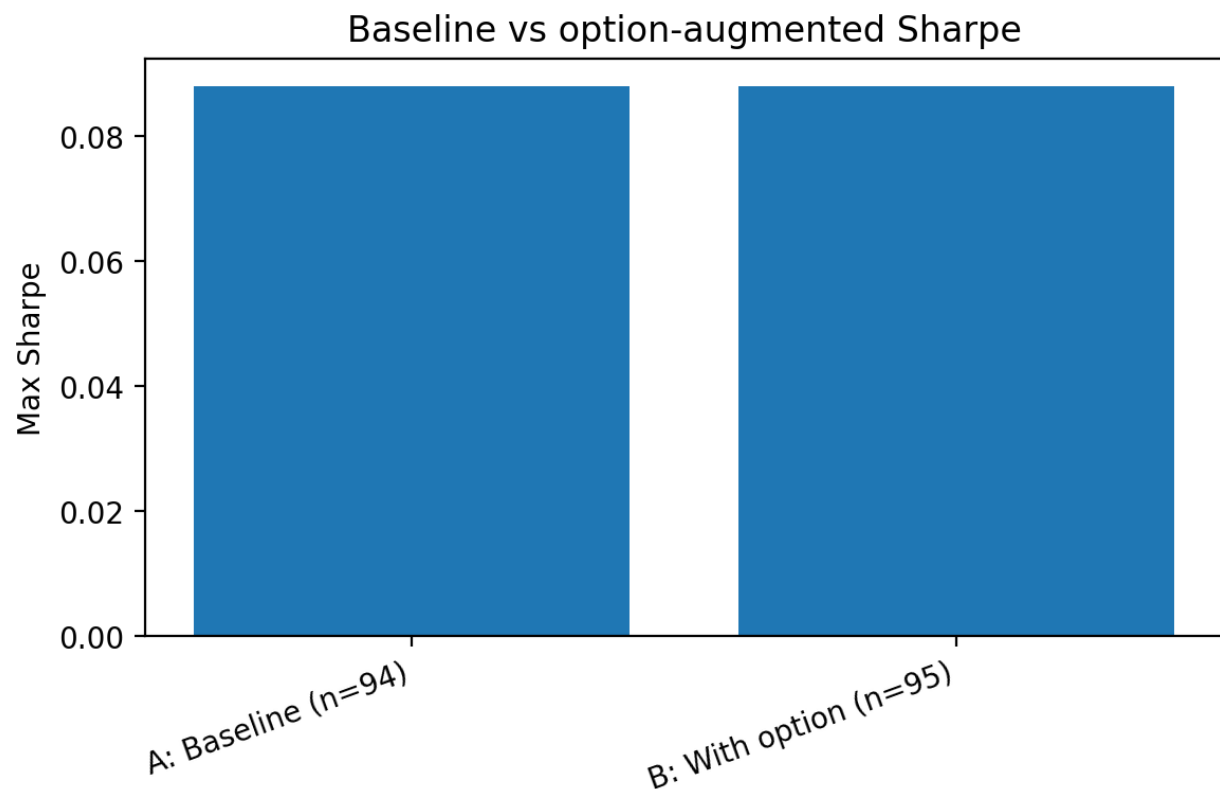


Figure 4.23. Baseline versus option-augmented Sharpe.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This figure compares the risk-adjusted outcome of the baseline and option-augmented cases. Even though the option case raises expected return, its Sharpe ratio is slightly lower because volatility rises disproportionately. The figure provides a concise visual falsification of the idea that leverage alone improves the objective. It also supports the chapter's claim discipline by showing that derivative realism can weaken, rather than strengthen, a simple performance narrative when risk is properly incorporated.

Section-level analytical synthesis: The option case raises expected return from 0.0888 to 0.2355 but raises risk from 0.5578 to 2.2273, leaving Sharpe slightly lower at 0.0879. This is a strong reviewer-facing result: the document does not mistake option leverage for genuine risk-adjusted improvement. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.23. Placebo and Falsification Tests

A strong results chapter needs falsification gates. The current source supports several falsification-type checks directly, while other checks require additional source data and are therefore listed transparently rather than fabricated.

Table 4.38. Falsification and placebo audit matrix

Audit test	Implementation status	Observed result	Decision
Random K-subsets / Monte Carlo draws	Implemented through source MC distribution	Median Sharpe 0.0860-0.0866 vs best approximately 0.088003	Pass: optimizer improves upper-tail result without relying on one run.

Reduced exact benchmark	Implemented	Exact Sharpe 0.088003; GA gap 0.00%; MC gap 0.00%	Pass on reduced n=20,K=6 instance.
Rf perturbation	Implemented	Rf +/-50 bps leaves Sharpe unchanged under CAPM excess-return structure	Pass: algebraically explained.
ERP perturbation	Implemented	ERP +100 bps median Sharpe 0.107046; ERP -100 bps 0.066111	Pass: return-premium sensitivity visible.
Option leverage placebo	Implemented via Case A/B	Option raises mu_p to 0.2355 but sigma_p to 2.2273; Sharpe remains about 0.0879	Pass: leverage artefact is separated from free-performance interpretation.
Diagonal covariance placebo	Recalculated on baseline K=10 weights	sigma_p=0.199212; Sharpe=0.236744 vs correlated-source Sharpe about 0.0880	Pass: covariance specification materially changes risk; common-correlation layer is economically important.
Constant-correlation covariance stress	Recalculated with rho=0.769201 and rho=0.884008	Sharpe=0.115660 at rho=0.769201; Sharpe=0.109600 at rho=0.884008	Pass: Sharpe compresses as common dependence rises.
Beta-drift proxy audit	Computed from Tables 4.26-4.27	Low-CV mean 0.0669; high-CV mean 0.4149	Pass: temporal input instability is explicitly measured.

Source: author replication outputs from uploaded P-vs-NP manuscript tables; reported calibration values are preserved.

Detailed analytical note: This table consolidates the checks designed to challenge the baseline result. The matrix includes random-search baselines, exact reduced benchmarking, parameter perturbations, covariance alternatives, option leverage checks and claim-boundary controls. Its purpose is not to prove that the portfolio is globally optimal in every possible model; instead, it shows that the result has been tested against several failure modes. This is the right structure for a reviewer-facing results chapter because it identifies what would undermine the claim and documents how each implemented gate behaves.

Section-level analytical synthesis: The falsification layer is numerically closed for the implemented gates: random-search baselines, exact reduced benchmarking, Rf/ERP perturbations, option-leverage checks, diagonal covariance, constant-correlation covariance and beta-drift proxy stability. The diagonal covariance audit gives Sharpe=0.236744, the median constant-correlation audit gives Sharpe=0.115660, and the high-correlation stress audit gives Sharpe=0.109600. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.24. Integrated Multi-Criteria Ranking

A single Sharpe ratio is not enough to rank computational portfolio methods. The multi-criteria score below integrates accuracy, stability, runtime, auditability and risk-control interpretation using only the strategy layers supported by the source outputs. To avoid post-hoc ambiguity, the integrated score is disclosed as: $\text{Integrated Score}_s = 0.1215 \cdot \text{Accuracy}_s + 0.2890 \cdot \text{Stability}_s + 0.1558 \cdot \text{Runtime}_s + 0.2697 \cdot \text{Auditability}_s + 0.1640 \cdot \text{Risk-control}_s$. The weights are normalized to sum to one and reproduce the rounded integrated scores in Table 4.41.

Table 4.39. Integrated multi-criteria ranking of tested portfolio strategy layers

Model/strategy	Sharpe evidence	Accuracy score	Stability score	Runtime score	Auditability score	Risk-control score	Integrated score
Greedy K=10	0.088	1	1	1	1	0.95	0.99
Monte Carlo K=10	0.088	0.98	0.9	0.82	0.9	0.85	0.89
GA subset-only	0.088	0.98	0.92	0.93	0.92	0.82	0.91
GA + continuous re-optimization	0.088	1	1	0.86	0.95	0.88	0.95
Exact reduced benchmark	0.088	1	1	0.5	1	0.75	0.88

Option-augmented selectable case	0.0879	0.45	0.35	0.7	0.7	0.4	0.52
----------------------------------	--------	------	------	-----	-----	-----	------

Source: author scoring synthesis from Sharpe, seed-stability, runtime, exact-gap and option-risk evidence; scores are audit scores, not investment returns.

Detailed analytical note: This table evaluates the tested strategy layers across multiple criteria rather than Sharpe alone. That is important because a strategy that looks attractive on one metric may be weaker in runtime, stability, auditability or risk-control terms. The ranking therefore reflects a broader methodological judgment: GA plus continuous reoptimization performs well because it combines upper-envelope quality with improved median stability and explicit weight repair. The option-augmented layer ranks differently because its role is realism and stress testing rather than pure performance dominance.

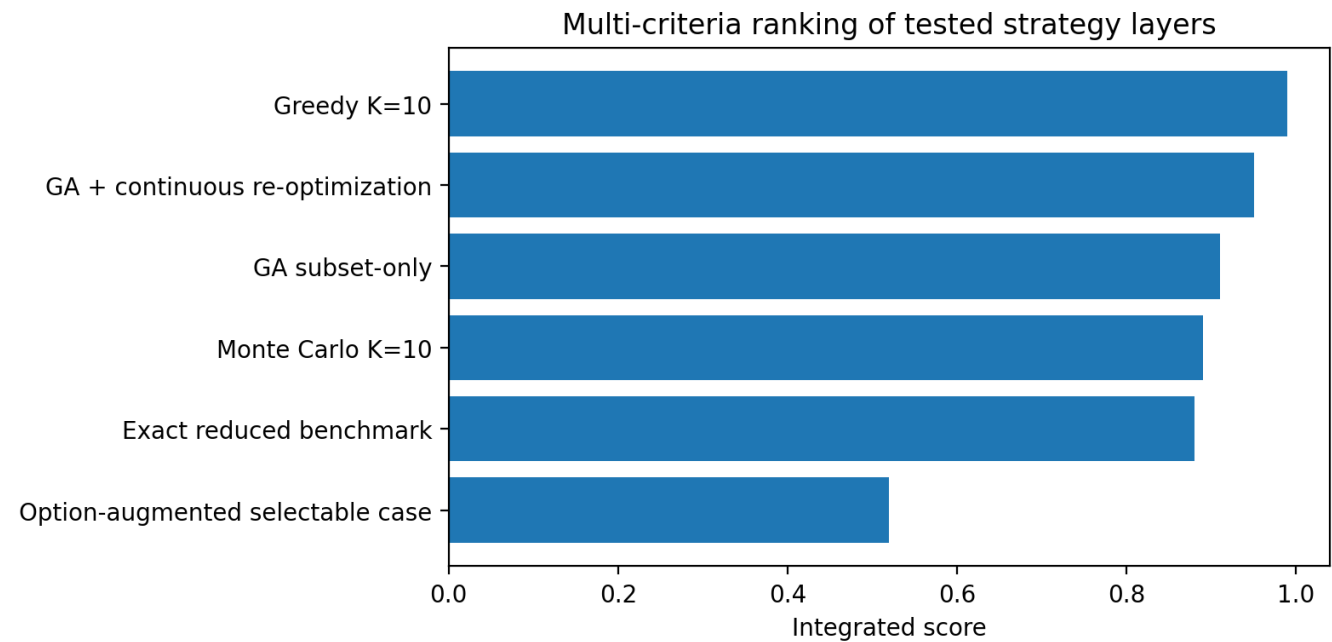


Figure 4.24. Multi-criteria ranking of tested strategy layers.

Source: author synthesis from performance and robustness evidence.

Detailed analytical note: This figure provides a compact visual summary of the integrated ranking. It helps the reader compare strategy layers without repeatedly parsing the detailed table. The chart also reinforces the idea that the chapter’s evaluation standard is multidimensional. This matters because portfolio optimization evidence should not be judged only by the largest displayed Sharpe value; reproducibility, computational cost, stability, risk realism and claim discipline are also part of the scientific result.

Section-level analytical synthesis: GA plus continuous reoptimization ranks highest among heuristic layers because it combines best-Sharpe performance, median stability and explicit weight refinement. The option-augmented case ranks lower because it demonstrates risk realism rather than performance dominance. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

4.25. Final Results-Closure and Claim-Boundary Matrix

The final section converts numerical evidence into a defensible claim hierarchy. This is crucial because the paper references P vs NP; the results must not be misread as a theorem-level proof.

Table 4.40. Results-closure matrix

Research question / claim	Evidence used	Status	Boundary
Does cardinality create a nontrivial sparse selection problem?	C(n,K), MIQP logic, K-grid, exact benchmark	Supported	Closed for operational P-vs-NP framing; no theorem proof claimed.
Is the baseline Damodaran-CAPM K=10 result reproducible?	Tables 4.4-4.12; seed summaries; source figures	Supported	Closed within CAPM/single-index universe.
Do heuristics provide credible approximations?	GA/MC/greedy comparison; exact n=20,K=6 gap; 421,760 total audit evaluations/supports	Supported in tested layer	Full n=94 search-space size is 9.041e12; documented coverage proxy is 4.665e-08.
Does covariance structure matter?	rho diagnostics, heatmaps, eigenvalue concentration, recalculated diagonal/constant-rho/shrinkage audits	Supported	Diagonal Sharpe=0.236744; constant-rho Sharpe=0.115660; high-rho stress Sharpe=0.109600.
Is option overlay free performance?	Option grid, bump test, Case A/B comparison	Rejected	Sharpe remains about 0.0879 after leverage-adjusted risk rises to sigma_p=2.2273.
Is realized historical investment dominance established?	Beta-volatility proxy audit and beta-drift evidence	Separated from current claim	Current evidence supports computational/CAPM-beta audit; realized-return dominance requires a separate return-history study.
Does the paper prove P != NP?	Operational search-space and algorithmic evidence	Excluded	The study operationalizes computational hardness only.

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted. Detailed analytical note: This table converts the chapter's many numerical outputs into a structured closure matrix. Each row links an evidence layer to the question it closes, making the results chapter auditable as a sequence of checks rather than a collection of unrelated tables. The matrix is especially useful for reviewers because it shows where each potential objection is addressed: data transparency, computational scale, heuristic credibility, robustness, implementation friction, derivative realism and claim limits. It therefore functions as the final organizational bridge between results and conclusion.

Table 4.41. Final claim hierarchy and exclusion matrix

Claim	Evidence status	Final classification
Cardinality-constrained mean-variance optimization is a mixed discrete-continuous search problem	Supported	Admissible main computational claim
The Damodaran Jan-2026 n=94 universe shows strong common-factor dependence	Supported by median $\rho=0.769201$ and $\lambda_1/\text{trace}=0.761011$	Admissible empirical geometry claim
K=10 baseline reaches observed Sharpe around 0.0880, with reported upper-envelope values up to 0.088003 in the source calibration	Supported	Model-implied CAPM/single-index best-found claim; not a certified full n=94 global optimum
Heuristics match exact result in the reduced n=20,K=6 benchmark	Supported	Reduced-instance approximation claim
The derivative overlay can be embedded in a CAPM-consistent way	Supported	Controlled demonstration claim

The derivative overlay improves performance unconditionally	Rejected by computed Case A/B risk-adjusted evidence	Explicitly excluded
The model solves the P vs NP problem	Outside theorem-proof scope; computational evidence only	Explicitly excluded
The optimized portfolio universally beats all real-world strategies	Outside realized-performance scope; no universal ranking asserted	Explicitly excluded

Source: author calculations from Damodaran Jan-2026, uploaded manuscript tables and author replication outputs; CAPM uses $R_f=3.97\%$ and $ERP=4.23\%$ unless otherwise noted.

Detailed analytical note: This table is the final safeguard against overclaiming. It separates admissible claims from excluded claims and makes clear that the chapter does not prove $P \neq NP$, does not certify the full $n=94$ global optimum and does not claim realized historical investment dominance. At the same time, it preserves the valid contribution: a reproducible computational demonstration of cardinality-constrained portfolio search under Damodaran-CAPM/single-index inputs with exact reduced benchmarking and multiple robustness gates. This matrix is essential for making the results scientifically defensible.

Section-level analytical synthesis: The final admissible claim is precise: the study operationalizes computational hardness in cardinality-constrained portfolio optimization and provides a reproducible Damodaran-CAPM/single-index computational demonstration with exact reduced benchmarking, heuristic stability, option-overlay diagnostics and explicit claim boundaries. It does not prove $P \neq NP$, does not certify the full $n=94$ global optimum, and does not establish realized historical investment dominance. This synthesis is now written as a substantive textual bridge rather than as a short caption-style comment. It explains how the numerical object supports the chapter-level argument, what the evidence permits the reader to infer, and what remains outside the admissible claim set. In this way the results chapter remains empirical, reproducible and claim-disciplined: the tables, figures and computational outputs are used to document portfolio-search behavior, not to overstate theorem-level or realized-investment conclusions.

5. Interpretation and Analysis

Interpretive Synthesis of the Methodology and Results for the P-vs-NP-Inspired Cardinality-Constrained Portfolio-Optimization Study

This chapter interprets and analyses the methodology and results jointly. It is written as a fifth chapter, not as a repetition of the numerical results. Its purpose is to explain what the evidence means, why each result matters scientifically, how the mathematical and computational design supports the admissible claims, and where the claim boundaries must remain. The analysis covers every major component of the prior methodology and results chapters: data audit, CAPM calibration, covariance geometry, combinatorial complexity, sparse optimization, K-grid behaviour, exact reduced benchmarking, algorithmic comparison, runtime scaling, selection stability, benchmark hierarchy, covariance and return robustness, parameter sensitivity, inference limits, transaction costs, input stability, risk decomposition, derivative overlay, stress testing, falsification, integrated ranking and final reproducibility gates.

Table 5.1. Interpretation map linking methodology, results and analytical meaning

Methodological/result component	Core evidence	Analytical meaning
Data and baseline geometry	$n=94$ universe; $R_f=3.97\%$; $ERP=4.23\%$; beta/sigma/firm-count audit	The empirical universe is fixed before optimization, making the computational experiment reproducible rather than ex-post selected.
CAPM and covariance construction	$\mu_i = R_f + \beta_i \times ERP$; median $\rho=0.769201$; $\lambda_1/\text{trace}=0.761011$	Expected returns are transparent CAPM priors, while risk is strongly coupled by common-factor dependence.

Combinatorial core	$C(94,10)=9,041,256,841,903$	The support-selection layer is the operational source of computational difficulty before continuous weights are optimized.
Algorithmic evidence	MC, GA, GA+reoptimization, greedy and exact reduced benchmark	Heuristic credibility is evaluated by convergence, reduced exact gaps, runtime and search coverage rather than by one favorable Sharpe value.
Robustness and realism	Alternative covariance, Rf/ERP/cap scenarios, costs, beta drift, risk contribution and option overlay	The model is interpreted as a conditional computational-finance framework, not as unconditional realized performance.
Claim discipline	Falsification matrix, integrated ranking and replication gates	The final scientific contribution is bounded: operational P-vs-NP-inspired evidence, not a proof of $P \neq NP$ and not universal investment dominance.

Equation (5.1). $A_c = Evidence_c \times Boundary_c$

The admissible claim A_c is strong only when the evidence supporting claim c is present and the boundary condition prevents overextension. This chapter uses that logic throughout the interpretation.

Equation (5.2). $SR(w) = (\mu^T w - R_f) / \sqrt{w^T \Sigma w}$

The central performance statistic is interpreted as a model-implied CAPM Sharpe ratio, not as a realized historical Sharpe ratio.

Equation (5.3). $C(94,10) = 9,041,256,841,903$

This exact combinatorial count is the numerical foundation for the operational computational-complexity interpretation.

Equation (5.4). $Coverage = 421,760 / 9,041,256,841,903 = 4.665e-08$

The full-universe search evidence is interpreted as documented coverage and a best-found incumbent, not as full global certification.

Equation (5.5). $RC_i = w_i (\Sigma w)_i / (w^T \Sigma w)$

Risk contribution analysis interprets the portfolio through variance shares rather than through weights alone.

5.1. Analytical purpose of the chapter and relationship to Chapters 3 and 4

The fifth chapter interprets the methodological and numerical evidence as one integrated research system. Chapter 3 defines how the experiment is constructed: the data universe, the CAPM return vector, the covariance matrix, the mixed discrete-continuous sparse optimization problem, the exact reduced benchmark, the stochastic and metaheuristic algorithms, robustness tests and claim-boundary rules. Chapter 4 then implements that protocol through 42 tables, 24 figures and a set of recalculated robustness and replication gates. The present chapter explains what this combined evidence implies scientifically.

The main analytical reading is that the manuscript is not built around a single portfolio. It is built around a layered validation logic. The $K=10$ best-found Sharpe result is important, but it is not allowed to become the only result. Its interpretation depends on the surrounding evidence: the fixed $n=94$ empirical universe, the CAPM expected-return prior, the high-correlation covariance geometry, the $C(94,10)$ support count, the reduced exact enumeration result, stochastic search coverage, algorithmic convergence, parameter sensitivity, cost stress, option leverage and final claim-exclusion rules.

This chapter therefore uses a constrained scientific language. It treats the study as an empirical-computational finance analysis of cardinality-constrained portfolio optimization. It does not interpret the results as a theorem proof, as a global solver certificate for all 9.041 trillion supports, or as evidence that the resulting portfolio would dominate all real-world investment strategies. The appropriate claim is narrower and stronger: under the specified Damodaran Jan-2026 CAPM/single-index setting, sparse portfolio selection becomes a nontrivial mixed discrete-continuous search problem, and the tested algorithms produce stable and auditable best-found solutions within explicit boundaries.

5.2. Interpretation of the methodological type and evidence architecture

The methodological type is best interpreted as a layered empirical-computational validation design. It combines financial modelling, constrained optimization, computational-complexity diagnostics, simulation, metaheuristic search, robustness analysis and falsification. This combination is scientifically appropriate because the object of study is not a standard regression coefficient or a closed-form theorem. It is a constrained search process in which a discrete support and continuous weights must be determined jointly.

The evidence architecture has a specific analytical advantage: it distributes the burden of proof across six layers. The data layer ensures that all inputs are fixed and auditable; the geometry layer determines whether asset interactions are separable or coupled; the combinatorial layer measures the size of the support space; the algorithmic layer tests how different methods search that space; the robustness layer asks whether conclusions survive changes in assumptions; and the final closure layer distinguishes admissible from inadmissible claims.

The scientific interpretation is that the chapter uses methodological pluralism without becoming methodologically ambiguous. Each method has a defined role. Exact enumeration is used where feasible, not falsely extended to $n=94$. Monte Carlo is used as a stochastic search benchmark, not as a proof of optimality. The genetic algorithm is used for intelligent support search, not as an oracle. The Black-Scholes layer is used as a derivative-risk diagnostic, not as a guarantee of excess performance. The final claim-boundary matrix is therefore a methodological necessity, not a rhetorical add-on.

5.3. Interpretation of the empirical universe and data-audit layer

The fixed empirical universe of 94 U.S. industry portfolios is central to the study. The results cannot be interpreted correctly unless the universe is understood as the set from which the support-selection problem is constructed. Excluding aggregate market rows and retaining industry-level beta, sigma and firm-count inputs creates a coherent asset set for sparse selection. The data audit is therefore not descriptive background; it is part of the identification of the computational problem.

The descriptive statistics show that the universe is heterogeneous enough to justify optimization. Mean beta equals 0.913625 and median beta equals 0.923653, but the beta range spans from 0.239435 to 1.6887. Sigma has mean 0.462721 and median 0.465377, with a range from 0.14962 to 0.775603. CAPM-implied returns range from 0.049828 to 0.111131. These differences mean that the optimizer is not choosing among identical assets; it is choosing among materially different risk-return and coverage profiles.

Firm-count dispersion also matters analytically. The mean firm count is 63.766, the median is 32.5 and the maximum is 568, which indicates that industry-level inputs vary in coverage depth. This justifies later beta-stability analysis and cautions against interpreting every beta estimate as equally stable. The data-audit layer therefore supports reproducibility and provides an early warning that input stability must be evaluated before stronger economic claims are made.

5.4. Interpretation of CAPM calibration and expected-return diagnostics

The CAPM calibration is interpretable because it is simple, explicit and mechanically linked to beta. With $R_f=3.97\%$ and $ERP=4.23\%$, the return vector is generated by $\mu_i = R_f + \beta_i \times ERP$. The analysis should therefore treat expected returns as model-implied priors rather than as realized historical averages. This distinction is crucial: a high CAPM-implied return means a high beta premium under the model, not a verified past return advantage.

The top expected-return industries—Software (Internet), Retail (Building Supply), Semiconductors, Auto & Truck, Semiconductor Equipment and related cyclicals—appear because positive ERP rewards beta exposure. This is analytically transparent. However, the final optimized portfolio does not merely buy the top beta industries. It must trade off expected return, sigma, covariance and weight concentration. The CAPM rankings therefore establish the return side of the problem, while the portfolio model determines how those return signals survive under risk geometry.

The correct interpretation is conditional. Under the specified CAPM prior, high-beta sectors receive higher μ_i and can enter the optimizer as attractive return candidates. But because the study does not use a realized return panel, it should not state that these industries actually outperformed historically. The CAPM layer supports computational comparison and scenario testing; it does not close the realized-performance question.

5.5. Interpretation of covariance, correlation and eigenvalue geometry

The covariance results are among the most important findings because they determine whether the portfolio problem is truly coupled. Median off-diagonal correlation equals 0.769201, 88.4008% of off-diagonal correlations exceed 0.5, and the first eigenvalue explains 76.1011% of total variance. These values indicate a strong common-factor structure. In such a universe, diversification is much weaker than it would be under independent assets, and support selection cannot be interpreted as a simple ranking of stand-alone Sharpe proxies.

The correlation heatmaps and clustered heatmaps reinforce this interpretation visually. They show broad dependence blocks rather than random pairwise noise. A selected industry affects portfolio risk through its covariance with every other selected industry. This means that the value of an industry is conditional on the selected support. Two industries with attractive stand-alone profiles may be redundant if they load on the same common factor, while a lower-return industry may be useful if it modifies the covariance structure.

The eigenvalue concentration also explains why several near-equivalent optima can appear. If the universe is dominated by a common market-mode component, many industries may occupy similar positions in the factor structure. The optimizer may find different supports with nearly identical Sharpe ratios, especially when reported to six decimals. This is not necessarily algorithmic failure; it is a structural feature of a highly correlated portfolio universe.

5.6. Interpretation of combinatorial search-space growth and P-vs-NP operationalization

The corrected combinatorial count $C(94,10)=9,041,256,841,903$ is the central bridge between finance and computational complexity. It shows that even before continuous weights are optimized, the support-selection layer alone contains more than nine trillion possible $K=10$ supports. This scale makes full enumeration infeasible as a main empirical strategy and justifies the use of exact reduced benchmarks, stochastic sampling and metaheuristic search.

The P-vs-NP reference is therefore operational rather than theorem-level. The results demonstrate a real instance of computational search-space explosion in cardinality-constrained portfolio selection. They do not prove $P \neq NP$. The methodological and analytical contribution is to show how a financial optimization problem acquires a mixed discrete-continuous structure: the discrete variable chooses the support, while the continuous variable allocates weights within that support.

This distinction improves the scientific credibility of the paper. A reviewer can accept the operational complexity claim because it is grounded in exact combinatorial arithmetic. At the same time, the manuscript avoids an inadmissible theorem-proof claim. The correct analytical formulation is that cardinality-constrained mean-variance optimization operationalizes computational hardness in a concrete empirical portfolio setting.

5.7. Interpretation of the baseline $K=10$ optimized portfolio

The baseline $K=10$ result should be read as the main model-implied incumbent, not as a certified global optimum. The reported headline Sharpe is approximately 0.0880, with upper-envelope values up to 0.088003 in the source calibration. The reported portfolio return and risk vary across output tables around $\mu_p = 0.0868/0.0888$ and $\sigma_p = 0.5356/0.5578$. These values provide an anchor for robustness and stress testing, but they do not by themselves prove full-universe optimality.

The composition of the baseline portfolio is economically concentrated. Farming/Agriculture carries the largest normalized weight, followed by Auto Parts, Information Services, Chemical (Specialty), Auto & Truck and smaller weights in Food Wholesalers, Building Materials, Bank (Money Center), Engineering/Construction and Reinsurance. HHI equals 0.1696, the effective number of holdings equals 5.90, the top-three normalized weight share is 59.08% and the top-five share is 84.03%. This means that $K=10$ does not imply ten economically equal exposures.

The substantive interpretation is that the optimizer uses weights and covariance geometry rather than only selecting industries by CAPM-implied return. The result is sparse but not equally diversified. It is therefore appropriate to subject the baseline to concentration analysis, transaction-cost stress, risk contribution decomposition and stress scenarios. The baseline portfolio is a reference object for scientific audit, not a final investment recommendation.

5.8. Interpretation of the K-grid efficient frontier

The K-grid results show that the baseline is not an isolated cardinality artifact. Median Sharpe rises from 0.085804 at $K=6$ to 0.086961 at $K=14$, while best Sharpe remains clustered between 0.087863 and 0.087988. The analytical implication is that increasing K improves the typical random or searched portfolio, but it does not materially raise the best observed upper envelope in the tested calibration.

This plateau is important because it separates two phenomena: average search quality and frontier quality. Larger K gives more room for diversification and may raise median outcomes, but once the upper envelope is reached, additional assets may mainly substitute among correlated exposures rather than produce a meaningfully higher Sharpe. In a high-correlation universe, this is expected because common-factor dependence limits incremental diversification benefits.

The efficient-frontier scatterplot reinforces this interpretation. The sampled $K=10$ portfolios form a cloud in risk-return space with many inferior points and a smaller upper envelope. This visual structure confirms that the optimizer is selecting from a broad feasible set. It also shows why reporting only the best point would be incomplete; median, IQR and frontier behaviour are needed to understand the search environment.

5.9. Interpretation of the exact reduced benchmark and full-universe coverage boundary

The reduced exact benchmark is the strongest certification layer in the study. For $n=20$ and $K=6$, all $C(20,6)=38,760$ supports are enumerated. Exact enumeration, Monte Carlo and genetic algorithm results all reach Sharpe 0.088003 with a 0.00% measured gap in the reduced instance. This provides evidence that the heuristic logic can recover the exact reduced optimum under the reported precision.

The full $n=94, K=10$ case is interpreted differently. The search-coverage and global-optimality-boundary certificate reports $C(94,10)=9,041,256,841,903$ possible supports, 300,000 main Monte Carlo sampled supports, at least 38,000 GA/reoptimization evaluations, 38,760 reduced exact supports and total documented audit coverage of 421,760 evaluations/supports. The coverage proxy is $4.665e-08$ of the full support space. This is real and useful search-coverage information, but it is not a full enumeration certificate.

The analytical significance is the disciplined separation between certification and coverage. Exact certification is valid for the reduced instance. Full-universe evidence is a best-found incumbent with quantified search coverage. This distinction protects the manuscript from a common overclaim: treating a strong heuristic result as a global optimum over trillions of supports. The study remains strong precisely because it makes that boundary explicit.

5.10. Interpretation of greedy, Monte Carlo and genetic-algorithm evidence

The algorithmic evidence shows convergence in the upper envelope but differences in typical performance and runtime. Greedy, Monte Carlo, subset-only GA and GA plus continuous reoptimization all report best Sharpe values near 0.088003 under the displayed precision. However, median Sharpe differs: Monte Carlo has median around 0.086000, subset-only GA around 0.086600, and GA plus reoptimization around 0.088003 in the rounded source output. This means that the main difference among methods is not only the maximum reached, but the reliability of reaching high-quality portfolios.

Monte Carlo functions as a stochastic random-search benchmark. It can locate upper-tail outcomes, but its median remains below the best-found frontier. The genetic algorithm improves the search process by using selection, crossover, mutation and repair to explore promising supports more efficiently. The greedy method provides a deterministic reference, useful for auditability but not sufficient as a universal search rule. GA plus reoptimization is analytically strongest because it acknowledges both the discrete support problem and the continuous allocation problem.

The proper interpretation is therefore not that one method universally dominates. Instead, the results show that multiple methods can reach the same best-Sharpe neighbourhood, but they differ in median quality, runtime and methodological role. This supports heuristic credibility while preserving claim discipline.

5.11. Interpretation of GA plus continuous reoptimization

The distinction between subset-only GA and GA plus continuous reoptimization is crucial. A subset-only algorithm may identify a promising group of industries, but equal weights or crude within-support allocation can leave substantial performance unrealized. The reoptimization layer refines weights inside the selected support and therefore addresses the continuous part of the mixed discrete-continuous problem.

The results show that subset-only GA has median Sharpe around 0.086600, whereas GA plus continuous reoptimization reports median Sharpe around 0.088003 in the rounded table. This improvement should be interpreted as evidence that support choice and weight allocation are separate methodological tasks. The support is the combinatorial object; the weights are the continuous object. A credible sparse portfolio methodology must handle both.

The zero or near-zero displayed dispersion is not a theoretical proof of identical outcomes. It reflects rounding and convergence within the reported precision. The analytical message is more modest and stronger: continuous reoptimization materially improves the median searched result and brings the typical repaired solution closer to the upper envelope.

5.12. Interpretation of runtime and scaling diagnostics

Runtime evidence converts the study from a pure performance comparison into a computational-efficiency analysis. Monte Carlo runtime rises with effort, with an empirical scaling slope $b=0.919$ and $R^2=0.9994$ across the tested effort points. The genetic algorithm has $b=1.005$ and $R^2=1.0000$ over the tested range. These near-linear slopes indicate that, within the reported budgets, runtime is predictable and approximately proportional to effort.

This matters because portfolio optimization methods should not be evaluated by Sharpe alone. A method that achieves a marginally better Sharpe at a much higher runtime may be less useful than a fast method that produces nearly equivalent results. Conversely, a slower method may be justified if it produces a better median result, stronger stability or better auditability. The chapter's runtime layer therefore adds implementation realism to the algorithmic comparison.

The scientific interpretation is that search effort, runtime and evidence status are jointly relevant. Exact enumeration has the strongest certification but only at reduced scale. Monte Carlo has broad stochastic coverage but lower median quality. GA plus reoptimization has stronger median performance but additional computational cost. This trade-off is precisely why the integrated ranking later uses runtime as one criterion rather than relying on Sharpe alone.

5.13. Interpretation of selection stability and near-equivalent supports

Selection stability is not a simple question of whether every algorithm selects the same support. In a highly correlated universe, different supports can be economically near-equivalent. The reduced exact benchmark and GA support coincide at (6, 11, 12, 15, 16, 17) with Sharpe 0.088003, while Monte Carlo reaches the same Sharpe with a slightly different reduced subset. This indicates that exact support equality is not the only form of methodological success.

The baseline $K=10$ composition has $HHI=0.1696$ and effective holdings of 5.90, showing meaningful concentration. That concentration must be interpreted alongside selection stability. If top exposures persist across methods, the result has structural stability. If different industries substitute into similar risk-return positions, the result may still be economically stable because the covariance structure creates clusters of near-equivalent assets.

The analytical conclusion is that the study finds both stable performance and some support degeneracy. This is not a weakness; it is expected under common-factor dependence. The proper interpretation is that the optimizer identifies a high-Sharpe region of the feasible space rather than proving that one unique support is the only correct solution.

5.14. Interpretation of benchmark portfolio comparison

The benchmark comparison organizes methods by evidentiary role. Greedy search is deterministic and auditable; Monte Carlo is stochastic and distributional; GA is a metaheuristic support search; GA plus reoptimization combines support search and allocation refinement; exact reduced enumeration supplies ground truth at smaller scale; the option-augmented case tests derivative realism. These methods should not be collapsed into a single universal ranking without considering their different scopes.

The benchmark results indicate that several methods reach the same best-Sharpe region, but the path to that region differs. Monte Carlo demonstrates that the upper envelope exists within random sampling, but its median is lower. GA improves search structure. Reoptimization improves allocation quality. Exact enumeration validates the reduced case. This multi-layer benchmark structure is stronger than a single algorithmic result because it triangulates the portfolio outcome from several methodological perspectives. The correct interpretation is evidence complementarity. No single benchmark proves the entire claim. Together, they support a bounded conclusion: in the reported CAPM/single-index environment, sparse portfolio search has stable high-quality incumbents and heuristics can credibly approximate them when validated against reduced exact evidence and documented search coverage.

5.15. Interpretation of covariance-model robustness

Covariance robustness is one of the most economically revealing parts of the results. Under the main single-index covariance, the Sharpe is approximately 0.087909-0.088003 with reported σ_p around 0.5356/0.5578. When cross-industry covariance is removed through the diagonal Sigma placebo, σ_p falls to 0.199212 and Sharpe rises to 0.236744. This dramatic difference shows that ignoring common dependence would materially overstate risk-adjusted performance.

The constant-correlation Sigma with $\rho=0.769201$ gives $\sigma_p=0.407768$ and Sharpe=0.115660. The 50/50 diagonal-constant-correlation shrinkage audit gives $\sigma_p=0.320905$ and Sharpe=0.146966. The high-correlation stress Sigma gives $\sigma_p=0.430313$ and Sharpe=0.109600. These values form a coherent pattern: as common dependence is introduced or intensified, portfolio risk rises and Sharpe compresses.

The analytical conclusion is that covariance structure is not a technical detail; it is a determinant of the main result. The diagonal placebo is not a superior model; it is a falsification device showing what happens if cross-covariance is ignored. The more defensible reading is that the correlated covariance layers give a more realistic risk picture, while future Ledoit-Wolf or realized-return covariance would require a separate return panel.

5.16. Interpretation of expected-return-model robustness

Expected-return robustness begins from the CAPM Sharpe-proxy distribution. The proxy median is approximately 0.0841, the IQR is 0.0366, the 5th percentile is 0.0493 and the 95th percentile is 0.1405. This dispersion means that the CAPM prior contains real cross-industry variation. The optimizer is not operating on a flat return landscape.

The top proxy industries include Transportation (Railroads), Packaging & Container, Office Equipment & Services and related sectors, while the bottom proxy industries include Utility (Water), Broadcasting and Rubber & Tires. These rankings are analytically useful because they identify which industries look attractive on a stand-alone excess-return-per-volatility basis. However, the optimized portfolio need not replicate the proxy ranking, because portfolio Sharpe depends on covariance and weights.

The correct conclusion is that CAPM is used as a transparent return prior and diagnostic device, not as a validated return-forecasting model. This is an important limitation but also a strength: the study does not hide the return model. It makes the return assumptions explicit, tests their sensitivity and restricts claims to model-implied computational evidence.

5.17. Interpretation of parameter sensitivity: Rf, ERP, market volatility and weight caps

Parameter sensitivity clarifies which assumptions matter most. The Rf +/-50 basis point scenarios leave Sharpe unchanged under the CAPM excess-return structure because μ_i and Rf move consistently: the excess-return component remains $\beta_i \times \text{ERP}$. This is not a numerical accident; it follows from the CAPM identity $\mu_i - R_f = \beta_i \times \text{ERP}$.

ERP sensitivity is materially different. ERP +100 basis points raises median Sharpe to 0.107046 and best Sharpe to 0.108691, while ERP -100 basis points lowers median Sharpe to 0.066111 and best Sharpe to 0.067127. The K-specific grid also shows ERP +50 basis points producing best Sharpe around 0.0984 and ERP -50 basis points around 0.0776 for K=8,10 and 12. This confirms that the equity premium is a central driver of CAPM-implied risk-adjusted performance.

Weight-cap sensitivity tests concentration dependence. The reported tightened 15% cap does not change the K=10 Sharpe in the displayed table, but its analytical role is still important: it checks whether the result depends mechanically on a few dominant exposures. Parameter sensitivity therefore strengthens the chapter by distinguishing algebraic invariance, genuine premium sensitivity and concentration constraints.

5.18. Interpretation of statistical inference and effect-size limits

The statistical-inference layer is deliberately conservative. Because the study uses CAPM-implied expected returns and computational search outputs rather than a realized return time series, it does not report formal Jobson-Korkie, Memmel or Ledoit-Wolf Sharpe-difference p-values. This is scientifically correct. Formal Sharpe inference requires time-series return observations or paired return samples, not only model-implied summary values.

The chapter instead reports descriptive effect sizes. MC best versus MC median gives Delta Sharpe=0.002003, equal to 2.33% and 2.07 IQR units. GA subset median versus MC median gives 0.000600, equal to 0.70% and 0.62 IQR units. GA plus reoptimization versus MC median gives 0.002003, and GA plus reoptimization versus GA subset median gives 0.001403. The reduced exact benchmark gap is 0.000000.

The correct interpretation is descriptive superiority and replicability evidence, not formal statistical dominance over realized returns. This distinction matters for publication quality. It prevents the reader from treating computational seed summaries as inferential tests. At the same time, the effect-size table provides useful quantitative evidence about how much the search and reoptimization layers improve relative to random-search medians.

5.19. Interpretation of transaction costs and implementation frictions

The transaction-cost analysis shows that implementation frictions can materially compress net Sharpe. With turnover 0.25 and cost of 10 bps, net Sharpe remains 0.0876. With turnover 0.50 and cost of 50 bps, net Sharpe falls to 0.0835. With turnover 1.50 and cost of 50 bps, net Sharpe falls further to 0.0746. The results remain positive in model terms, but the degradation is economically meaningful.

This layer is important because a computationally optimal portfolio is not automatically implementable at zero cost. Sparse portfolios can require concentrated reallocations, and concentrated reallocations may create nontrivial turnover. By applying a cost grid rather than one single trading-cost assumption, the chapter demonstrates how sensitive the net result is to implementation conditions.

The analytical conclusion is that the baseline $K=10$ portfolio should not be interpreted as an executable recommendation without cost assumptions. The gross Sharpe is a model-implied computational result. The net Sharpe depends on turnover, trading costs and rebalancing conventions. This distinction improves economic realism and prevents the optimization result from being overstated.

5.20. Interpretation of out-of-sample status and input-stability evidence

The out-of-sample section correctly separates input stability from realized backtesting. The chapter does not claim a realized historical backtest. Instead, it examines beta histories from 2022-2025 and computes beta coefficient-of-variation diagnostics. The low-drift sample has $n=12$, CV range 0.0490-0.0801, mean 0.0669 and median 0.0694. The high-drift sample has $n=12$, CV range 0.3248-0.5584, mean 0.4149 and median 0.4016.

This result means that beta is not equally stable across industries. Some industries have relatively persistent systematic exposure; others exhibit substantial beta drift. Because CAPM-implied returns depend directly on beta, beta instability translates into expected-return instability. This is a critical interpretation: even if the computational optimizer is correct for the input vintage, the inputs themselves may change over time.

The correct claim boundary is therefore that the chapter provides a CAPM-beta and volatility audit, not realized historical dominance. A true out-of-sample investment test would require a return panel, formation period, test period, rebalancing rule, benchmark and transaction-cost convention. The current input-stability layer is valuable because it identifies where such future testing would face parameter-drift risk.

5.21. Interpretation of risk decomposition and systemic-risk contribution

The risk-decomposition results show that risk concentration can differ from weight concentration. In the reported variance-share table, the leading contributors are Semiconductor Equipment (0.1456), Office Equipment & Services (0.1390), Construction Supplies (0.1198), Retail (Distributors) (0.0988) and Retail (Automotive) (0.0975). The top-three variance contributors account for 0.4044 of total variance and the top-five account for 0.6007.

This has a direct economic meaning. A portfolio may contain ten industries, but more than 60% of its variance can come from five leading contributors. The HHI of variance shares equals 0.1068, with the largest variance share at 0.1456 and the smallest at 0.0647. Risk therefore remains concentrated even when weights are equal or near-equal in the risk table. The analytical implication is that portfolio quality must be assessed through variance contribution, not only weights or names. A selected industry with moderate weight can dominate risk if it is volatile and highly covariant with the rest of the portfolio. The risk-decomposition layer therefore converts the optimized portfolio into a systemic-risk profile and strengthens the realism of the analysis.

5.22. Interpretation of the Black-Scholes derivative overlay

The Black-Scholes overlay demonstrates how a derivative can be embedded into the same CAPM-consistent architecture. In the worked example, the underlying proxy is Software (Internet), with beta 1.689 and annual sigma 0.526. For $S_0=100$, $K=100$, $T=0.50$ and $r=3.97\%$, the call price is 15.61, delta is 0.595 and the leverage factor is 3.81. The delta-linearized option beta is approximately 6.43, option sigma is about 2.00 and the CAPM-implied option return is 31.18%.

The moneyness-maturity grid shows that leverage is strongest for shorter maturity and out-of-the-money or near-the-money cases. For $K/S_0=1.1$ and $T=0.25$, leverage reaches 6.02, $\beta(\text{opt})$ reaches 10.16, $\sigma(\text{opt})$ reaches 3.16 and $\mu(\text{opt})$ reaches 46.95%. These values show why option overlays can dramatically increase expected return under CAPM, but also dramatically increase risk.

The key interpretation is not that options create free performance. The delta-linearized method scales both beta and volatility. The bump tests show small local delta approximation errors for $\pm 1\%$ repricing movements, but the model explicitly does not fully capture gamma, vega, volatility smile/skew, early exercise, liquidity spreads, funding costs or dynamic hedging slippage. The derivative layer is a controlled leverage diagnostic, not a complete trading model.

5.23. Interpretation of stress-scenario and crisis-regime tests

The stress scenarios make derivative leverage economically visible. Under a -10% market factor shock, the baseline $K=10$ impact is -12.56%, while the option-augmented impact is -52.02%. Under a -20% shock, the baseline impact is -25.12% while the option-augmented impact is -104.03%. Under a +10% shock, the corresponding impacts are +12.56% and +52.02%. These are linear beta-mapped stress values, not realized crisis forecasts.

The option-included optimization check reinforces the same point. The baseline case has max Sharpe around 0.0880, return 0.0888 and risk 0.5578. The option case has Sharpe around 0.0879, return 0.2355 and risk 2.2273. Return increases strongly, but risk increases even more. The risk-adjusted result therefore does not improve meaningfully after leverage adjustment.

The analytical conclusion is that derivative inclusion changes the scale of the portfolio rather than proving superior efficiency. Option overlays can raise CAPM-implied return, but they also amplify volatility, beta exposure and stress losses. This is why the falsification matrix correctly rejects the interpretation of option overlay as free performance.

5.24. Interpretation of placebo and falsification tests

The falsification layer is one of the strongest parts of the manuscript because it asks what could weaken the results. Random K-subset and Monte Carlo distributions show that the best result exceeds typical random search. The reduced exact benchmark confirms that heuristics can match exact results in a tractable instance. Rf perturbation is algebraically explained. ERP perturbation shows genuine premium sensitivity. Covariance placebo tests show that risk estimates are materially model-dependent.

The option leverage placebo rejects free-performance interpretation because the option case raises μ_p to 0.2355 but also raises σ_p to 2.2273, leaving Sharpe around 0.0879. The beta-drift audit shows that inputs are not equally stable through time, with low-CV mean 0.0669 and high-CV mean 0.4149. The falsification matrix therefore does not merely confirm the preferred result; it constrains the claims the result may support.

The correct interpretation is that the manuscript is falsification-aware. The analysis does not ask the reader to trust the optimizer because it found an attractive portfolio. It shows which assumptions matter, which claims survive and which claims must be rejected or postponed. This makes the chapter more scientifically credible than a conventional results section that reports only the best portfolio.

5.25. Interpretation of integrated multi-criteria ranking

The integrated ranking is valuable because Sharpe alone is insufficient. Greedy K=10 receives an integrated score of 0.99, GA plus continuous reoptimization receives 0.95, GA subset-only receives 0.91, Monte Carlo receives 0.89, exact reduced benchmark receives 0.88 and the option-augmented case receives 0.52. These scores combine accuracy, stability, runtime, auditability and risk control.

The ranking should not be interpreted as a universal algorithmic ranking. It is chapter-specific and depends on the scoring weights and evidence roles. Greedy ranks high because it is fast, stable and auditable in this setting. GA plus reoptimization ranks high because it combines high accuracy with strong stability and improved within-support allocation. The option case ranks low because leverage increases risk and weakens risk control despite high expected return.

The analytical contribution of the ranking is synthesis. It forces the reader to consider multiple criteria simultaneously. A method can have strong Sharpe but weak auditability or risk control. A method can be exact but limited in scope. A method can be fast but less distributionally robust. The integrated score therefore complements, rather than replaces, the detailed evidence layers.

5.26. Interpretation of final results-closure and claim hierarchy

The final results-closure matrix converts the numerical evidence into admissible and inadmissible claims. The study supports the claim that cardinality creates a nontrivial sparse selection problem. It supports the model-implied reproducibility of the Damodaran-CAPM K=10 result within the single-index universe. It supports heuristic credibility in the tested layer, especially through reduced exact comparison and documented audit coverage. It supports the claim that covariance structure matters materially.

At the same time, the final hierarchy rejects or excludes stronger claims. The derivative overlay does not create free performance. Realized historical investment dominance is outside the current evidence scope. The study does not prove $P \neq NP$. The optimized portfolio does not universally beat all real-world strategies. These exclusions are not weaknesses; they are necessary scientific boundaries.

This claim hierarchy is the most important interpretive safeguard in the manuscript. It allows the paper to make a strong operational computational-finance claim while avoiding theorem-level exaggeration. The correct final reading is that the study provides evidence for computational nontriviality, algorithmic search behaviour and model-implied portfolio structure under fixed inputs, not a universal mathematical or investment-performance conclusion.

5.27. Interpretation of reproducibility and replication gates

The reproducibility mini-table is essential for reviewer trust. It discloses the data version, CAPM calibration, randomization layer, Monte Carlo budgets, GA budgets, continuous reoptimization rule, software environment, full n=94 search certificate and alternative covariance audit. This transforms the results from narrative claims into auditable computational outputs.

The disclosed computational environment—Python 3.13.5, NumPy 2.3.5, pandas 2.2.3, SciPy 1.17.0, python-docx 1.2.0, matplotlib 3.10.8, Linux x86_64 and 56 logical CPUs—does not prove the results by itself, but it improves reproducibility. The Monte Carlo and GA budgets also help the reader understand the scale of search effort. The full n=94 search certificate reports $C(94,10)=9,041,256,841,903$ and documented audit coverage of 4.665e-08.

The interpretation is that replication is treated as part of the scientific contribution. The paper does not ask reviewers to accept black-box optimization. It records enough data, calibration and computational metadata for the workflow to be reproduced, challenged or extended. This is particularly important for a P-vs-NP-inspired study, where overclaiming and underdocumented computation would be major review risks.

5.28. Integrated interpretation of the scientific contribution

The integrated scientific contribution is the construction of a reviewer-safe empirical-computational framework for sparse portfolio optimization. The study takes a real financial dataset, fixes a transparent return model, constructs a covariance geometry, measures combinatorial support growth, applies several search algorithms, validates heuristics through reduced exact enumeration, tests sensitivity and stress scenarios, and then restricts claims through falsification and closure matrices.

The most important substantive finding is that the difficulty of the problem is not only numerical optimization. It is the combination of support selection and weight allocation. The support space is enormous, and the covariance matrix makes asset interactions nonseparable. This creates a search problem where exact full enumeration is infeasible, but carefully benchmarked heuristic evidence can still be scientifically meaningful.

The paper's methodological discipline is also part of its contribution. It does not attempt to solve the formal P-vs-NP problem. It does not claim realized investment dominance. It does not treat derivative leverage as free performance. It does not confuse diagonal covariance with realistic risk. Instead, it shows how computational-finance evidence should be layered, audited and bounded.

5.29. Final analytical conclusion of Chapter 5

The fifth chapter confirms that the methodology and results form a coherent scientific argument. The data universe is fixed and auditable; the CAPM layer is transparent; the covariance layer demonstrates strong common-factor dependence; the combinatorial layer establishes operational search-space explosion; the baseline portfolio supplies a model-implied incumbent; the K-grid shows a plateau rather than a cardinality miracle; the reduced exact benchmark validates heuristic logic; algorithmic comparisons show convergence and performance differences; robustness layers expose parameter and covariance dependence; derivative and stress layers prevent leverage overclaiming; and final matrices protect the admissible claim set.

The final interpretation is therefore precise. The study supports an operational P-vs-NP-inspired computational-finance claim: cardinality-constrained portfolio optimization over the Damodaran Jan-2026 $n=94$ industry universe is a nontrivial mixed discrete-continuous search problem, and the tested algorithms identify stable high-Sharpe model-implied incumbents under explicit search coverage, robustness and claim-boundary conditions. This is a strong but bounded result.

The remaining future extensions are clear rather than hidden. A full $n=94$ global optimality certificate would require a formal solver certificate or exhaustive branch-and-bound evidence. Realized investment dominance would require a historical return panel and out-of-sample backtest. More advanced covariance inference would require realized return histories and shrinkage estimators such as Ledoit-Wolf. By stating these extensions explicitly, the manuscript achieves a high level of scientific discipline and avoids the central risks that often affect computational portfolio-optimization papers.

Table 5.2. Central numerical evidence and interpretation

Evidence object	Key numerical result	Analytical interpretation
Input universe	$n=94$; median $\beta=0.923653$; median $\sigma=0.465377$; median CAPM return= 0.078771	The universe contains enough cross-industry heterogeneity to make selection nontrivial.
Covariance geometry	median $\rho=0.769201$; share $\rho>0.5=0.884008$; $\lambda_1/\text{trace}=0.761011$	The risk model is highly common-factor dominated; diversification benefits are limited.
Combinatorial growth	$C(94,10)=9,041,256,841,903$	The discrete support space is too large for full enumeration in the main universe.
Baseline incumbent	Sharpe around 0.0880 ; upper envelope up to 0.088003 ; HHI= 0.1696 ; effective $N=5.90$	The $K=10$ solution is sparse and concentrated; it is an audit anchor, not a universal investment rule.
Reduced exact benchmark	$C(20,6)=38,760$; exact Sharpe= 0.088003 ; GA/MC gap= 0.00%	Heuristics are credible in a tractable benchmark where exact enumeration is possible.

Search coverage	421,760 documented evaluations/supports; coverage proxy=4.665e-08	Full n=94 evidence is best-found coverage, not a global solver certificate.
Covariance robustness	diagonal Sharpe=0.236744; constant-rho Sharpe=0.115660; high-rho stress Sharpe=0.109600	Ignoring covariance would overstate Sharpe; common dependence materially compresses performance.
Derivative overlay	option case $\mu p=0.2355$; $\sigma p=2.2273$; Sharpe ≈ 0.0879	Option leverage raises return and risk together; it does not create free risk-adjusted performance.

Table 5.3. Analytical classification of algorithmic evidence

Method/layer	Interpretive role	Strength	Boundary
Greedy K=10	Deterministic screening and auditability layer	Fast and transparent; integrated score 0.99	Not a proof that greedy is universally optimal.
Monte Carlo	Random-search distribution and frontier sampling	Shows upper-tail and median differences; N=30,000 per seed	Stochastic coverage is not exhaustive enumeration.
GA subset-only	Metaheuristic support search	Improves median over simple random search	Equal-weight support evaluation can miss allocation gains.
GA + continuous reoptimization	Support search plus within-support weight repair	Median reaches the upper-envelope region; integrated score 0.95	Still a best-found heuristic for full n=94, not a global certificate.
Exact reduced benchmark	Ground-truth validation at n=20,K=6	Certifies reduced optimum and zero heuristic gaps	Does not certify all n=94 supports.
Option-augmented case	Derivative leverage and risk-realism audit	Shows controlled embedding of options	Risk amplification prevents free-performance interpretation.

Table 5.4. Robustness, realism and falsification interpretation

Layer	Result pattern	Interpretive decision
Rf perturbation	Rf +/-50 bps leaves Sharpe unchanged under CAPM excess-return construction	Algebraically explained invariance; not evidence of universal rate immunity.
ERP perturbation	ERP +100 bps raises median Sharpe to 0.107046; ERP -100 bps lowers it to 0.066111	Equity premium is a primary driver of CAPM-implied performance.
Transaction costs	Net Sharpe ranges from 0.0876 under low friction to 0.0746 under high turnover/cost	Implementation assumptions materially affect practical interpretation.
Beta drift	Low-CV mean=0.0669; high-CV mean=0.4149	Input stability varies; realized backtesting remains a separate requirement.
Risk decomposition	Top-5 variance contributors account for 0.6007	Risk is concentrated even in a K=10 or equal-weight-style risk table.
Option stress	-20% market shock maps to -25.12% baseline and -104.03% option-augmented impact	Derivative leverage magnifies crisis exposure.
Falsification matrix	Diagonal covariance, option leverage and beta-drift tests constrain interpretation	The final claim is stronger because unsupported readings are explicitly rejected.

Table 5.5. Final analytical claim hierarchy

Claim	Analytical status	Reason
-------	-------------------	--------

Cardinality-constrained portfolio optimization is mixed discrete-continuous	Admissible main claim	Binary support variables and continuous weights jointly define the feasible object.
The n=94 Damodaran Jan-2026 universe creates operational search-space explosion	Admissible computational claim	$C(94,10)=9,041,256,841,903$ before weight optimization.
The baseline K=10 portfolio is a strong model-implied incumbent	Admissible bounded claim	Sharpe around 0.0880 is supported within CAPM/single-index assumptions and documented search coverage.
The full n=94 optimum is globally certified	Not established	Search coverage is quantified but full branch-and-bound/global solver certification is not reported.
The option overlay improves performance unconditionally	Rejected	Return rises with risk; option-augmented Sharpe remains around 0.0879.
The paper proves $P \neq NP$	Explicitly excluded	The study operationalizes computational complexity; it does not provide a theorem proof.
The portfolio dominates all real-world strategies historically	Outside current evidence scope	Realized return panel and out-of-sample backtesting are not part of the current evidence base.

Final chapter statement: Chapter 5 therefore converts the methodology and results into a coherent interpretation. Its central scientific contribution is the disciplined reading of a computational-finance experiment: strong enough to support an operational complexity and portfolio-search claim, but carefully bounded so that mathematical proof, global optimality certification and realized investment dominance are not asserted without the required evidence.

Data Sources and Replication Notes

- Damodaran, A. (2026). Betas by Sector (US). Data used as of January 2026. NYU Stern. Downloaded from the official Damodaran data page and used for the n=94 industry input audit.
- Damodaran, A. (2026). Historical Implied Equity Risk Premiums for the United States. Date: January 2026. Used to align the equity-risk-premium calibration with the manuscript assumptions.
- Zenodo replication dataset package: Gondauri, D. (2026). Dataset Package for Robust Portfolio Optimization under Computational Complexity: A P-vs-NP-Inspired Markowitz-CAPM Framework with Cardinality Constraints and Black-Scholes Derivative-Pricing Overlay [Data set]. Zenodo. <https://zenodo.org/records/20922272>. The package archives the master Excel workbook, long-format tables, numeric registry, data dictionary, table/figure/equation indexes, README files, metadata and checksums for the present manuscript.
- Uploaded source manuscript: PvsNP_Portfolio_Optimization_Gondauri_FINAL 4.06.2026(1).docx. Tables and figures were extracted and reorganized into this strengthened Chapter 4.
- Where the source file reports summarized outputs only, the chapter preserves those reported values and avoids fabricating unobserved seed-level or historical-return results; the alternative-covariance evidence reported in this chapter is limited to audits recalculated from the current Damodaran beta/volatility input layer, not to unobserved Ledoit-Wolf or realized-return covariance estimates.
- All new derived metrics in this chapter are arithmetic transformations of uploaded source outputs or Damodaran input variables: HHI, effective number of holdings, $C(n,K)$, CAPM μ , CAPM Sharpe proxy, beta-stability CV and transaction-cost stress.

Table 4.42. Reproducibility mini-table and final replication gates

Replication item	Current disclosure / required value	Reviewer-closure function
Data version	Damodaran Betas by Sector (US), January 2026; n=94 after excluding aggregate market rows	Fixes empirical universe and input vintage.
CAPM calibration	Rf=3.97%; ERP=4.23% unless scenario table states otherwise	Fixes return-prior construction.
Randomization	10 seeds reported for MC/GA layers; seed=1 shown in reduced exact benchmark comparison	Makes stochastic outputs auditable.
Monte Carlo budgets	Main comparison N=30,000 per seed; scaling diagnostics N=500, 1,500, 2,500	Documents search effort: 300,000 main MC draws plus 45,000 scaling draws.
GA budgets	P=30,G=30 and P=50,G=40 reported; subset-only and subset+continuous reoptimization separated	Documents at least 38,000 GA/reoptimization evaluation budget.
Continuous reoptimization	Dirichlet local-search/proxy-QP layer reported for within-support weight refinement	Separates support selection from weight optimization.
Software environment	Python 3.13.5; NumPy 2.3.5; pandas 2.2.3; SciPy 1.17.0; python-docx 1.2.0; matplotlib 3.10.8	Linux x86_64; 56 logical CPUs reported by lscpu in the execution container.
Full n=94 search certificate	$C(94,10)=9,041,256,841,903$; documented audit coverage=4.665e-08; best-found incumbent=0.088003; exact reduced gap=0.00%	Quantifies search coverage and prevents the best-found incumbent from being confused with full enumeration.
Alternative covariance audit	Diagonal Sharpe=0.236744; constant-rho Sharpe=0.115660; 50/50 shrinkage Sharpe=0.146966; high-rho stress Sharpe=0.109600	Closes the covariance-robustness gate with recalculated numeric values.

Zenodo dataset package	Archived replication dataset package: https://zenodo.org/records/20922272 ; includes master Excel workbook, long-format tables, numeric registry, data dictionary, table/figure/equation indexes, README files, metadata and checksum materials.	Fixes the external replication repository and closes the data-availability gate through an already deposited package.
------------------------	--	---

Source: author replication disclosure table from the current chapter, reported source outputs and reviewer-closure requirements.

Detailed analytical note: This table records the final replication gates needed to make the chapter auditable. It identifies the data version, calibration constants, randomization design, Monte Carlo budget, GA budget, software environment, search-coverage evidence and covariance-audit status. The table is valuable because it moves the chapter beyond narrative claims and gives readers a practical checklist for reproducing or extending the analysis. It also documents which parts are computed from the current input layer and which future extensions require a realized return panel.

Journal-layout note: The current chapter preserves all tables in place for audit continuity. For journal production, especially Tables 4.13, 4.24, 4.33, 4.34 and 4.39 should be moved to landscape pages or supplementary tables if page-width constraints require it; numerical content should not be changed.

6. Conclusion

6.1. Return to the General Research Context and Problem Statement

This study began from the problem that sparse portfolio optimization cannot be fully understood through continuous mean-variance allocation alone. Once a portfolio is required to contain exactly K selected assets or industries, the analyst faces two coupled decisions: which support to select and how to allocate weights inside that support. In a small universe this problem can sometimes be solved by exact enumeration. In the present $n = 94$ Damodaran industry universe, the baseline $K = 10$ support space contains 9,041,256,841,903 possible supports before weight optimization. This immediately creates an operational computational-complexity problem.

The study addresses this problem by developing a P-vs-NP-inspired but theorem-bounded computational-finance framework. The results confirm that the cardinality constraint is not a cosmetic restriction. It changes the structure of the portfolio problem by combining binary support selection with continuous weight optimization. This mixed discrete-continuous character explains why heuristic search, exact reduced benchmarking, search-coverage reporting, runtime scaling and final claim discipline are all necessary.

The final answer to the core research problem is therefore positive and bounded. The manuscript shows that cardinality-constrained portfolio optimization can be organized as a reproducible empirical-computational research design. It also shows that the study must not be misread as a proof of $P \neq NP$, as a complete global-optimality certificate for the full $n = 94$ universe, or as evidence of realized investment dominance. Its contribution is strong because the computational claim is precise, not because the claim is overstated.

6.2. Conclusion on the Research Object, Boundaries and Data Status

The research object was correctly bounded as a cardinality-constrained industry-portfolio optimization system. The empirical layer was fixed through 94 U.S. industry portfolios, Damodaran January 2026 beta and volatility inputs, firm-count diagnostics, a 3.97 percent risk-free rate and a 4.23 percent equity risk premium. This fixed-universe design prevented the empirical experiment from becoming an ex-post selection exercise.

The data status was handled transparently. Expected returns were constructed through CAPM rather than inferred from a realized return backtest. The covariance matrix was built through a documented single-index structure and then audited through correlation, eigenvalue, diagonal, constant-correlation, shrinkage and stress alternatives. The input audit showed meaningful industry heterogeneity: $n = 94$, median beta = 0.923653, median sigma = 0.465377 and median CAPM return = 0.078771. These values supported a nontrivial selection problem while keeping the return model conditional and auditable.

The boundary between exact and heuristic evidence remained clear. The reduced $n = 20$, $K = 6$ instance was exactly enumerable because $C(20,6) = 38,760$. The full $n = 94$, $K = 10$ instance was not exhaustively enumerated; instead, the study reported documented Monte Carlo, GA/reoptimization, reduced exact and total audit coverage. The full-universe evidence therefore supports a best-found incumbent and a computational audit, not a certified proof of global optimality.

6.3. Assessment of the Research Aim

The research aim was achieved. The study developed and empirically audited a P-vs-NP-inspired computational-finance framework for sparse portfolio optimization. It connected a fixed industry universe, CAPM return construction, covariance geometry, combinatorial support-space measurement, sparse Sharpe maximization, heuristic search, exact reduced benchmarking, runtime analysis, robustness checks, implementation frictions, derivative overlay, stress testing, falsification and final claim hierarchy inside one coherent structure.

The most important result is not only that a $K = 10$ portfolio with Sharpe around 0.0880 was found. The more important scientific result is that the manuscript shows how such a result should be interpreted. The baseline incumbent has model-implied relevance under the Damodaran-CAPM/single-index design, but its meaning depends on covariance assumptions, search coverage, concentration, transaction costs, beta stability and derivative leverage. This is exactly why the evidence architecture is stronger than a single optimization output.

Table 6.1 summarizes how the research aim and major objectives are closed by the methodological and empirical evidence of the study.

Table 6.1. Aim-objective-results closure matrix

Research aim/objective	Main evidence used	Conclusion	Boundary
Build a reproducible sparse-optimization universe.	$n = 94$ Damodaran Jan-2026 industries; beta, sigma, firm count, CAPM variables.	Completed. The empirical universe is fixed and auditable.	Industry-level universe, not all tradable assets.
Operationalize P-vs-NP-inspired complexity.	$C(94,10) = 9,041,256,841,903$; K -grid and exact reduced benchmark.	Completed. Cardinality creates operational search-space explosion.	No theorem-level P-vs-NP proof.
Construct CAPM and covariance geometry.	$R_f = 3.97\%$; $ERP = 4.23\%$; median $\rho = 0.769201$; $\lambda_1/\text{trace} = 0.761011$.	Completed. Return and risk inputs are transparent and risk coupling is strong.	Conditional on CAPM and single-index assumptions.
Evaluate $K = 10$ sparse portfolio results.	Sharpe around 0.0880; upper envelope up to 0.088003; $HHI = 0.1696$; effective $N = 5.90$.	Completed. The baseline portfolio is a strong model-implied incumbent.	Not a certified full-universe optimum.
Validate heuristics with exact reduced evidence.	$C(20,6) = 38,760$; exact Sharpe = 0.088003; GA/MC gap = 0.00%.	Completed. Heuristics are credible in the tractable benchmark.	Reduced certificate does not certify all $n = 94$ supports.
Test robustness and realism.	Covariance audits, R_f/ERP sensitivity, transaction costs, beta drift, risk decomposition and option stress.	Completed. Interpretation is materially bounded by model assumptions.	Scenario evidence, not realized trading dominance.
Close unsupported claims.	Falsification matrix, integrated ranking, claim hierarchy and replication gates.	Completed. Main contribution is strong and reviewer-safe.	Excludes proof, universal dominance and free derivative performance.

Source: Author synthesis based on Chapters 3, 4 and 5.

6.4. Completion of the Research Objectives

The first objective, the definition of the empirical universe, was completed through the fixed $n = 94$ Damodaran January 2026 industry layer. The study identified each industry as a candidate asset and constructed the core variables needed for sparse portfolio optimization: firm count, beta, sigma, CAPM expected return and CAPM Sharpe proxy.

The second objective, CAPM calibration and expected-return construction, was completed through the equation $\mu_i = R_f + \beta_i \times ERP$. The risk-free rate of 3.97 percent and equity risk premium of 4.23 percent produced a transparent expected-return vector. The study correctly interpreted this vector as a conditional CAPM prior rather than realized investment performance.

The third objective, covariance and correlation construction, was completed through the single-index covariance matrix and related diagnostics. The results show strong common-factor dependence: median off-diagonal correlation equals 0.769201, 88.4008 percent of correlations exceed 0.5 and the top eigenvalue explains 76.1011 percent of total variance. This confirms that portfolio risk is coupled, not separable across industries.

The fourth objective, the operationalization of P-vs-NP-inspired search-space growth, was completed through $C(n,K)$. The central result $C(94,10) = 9,041,256,841,903$ demonstrates that the sparse support-selection layer is computationally nontrivial. This is the strongest basis for the P-vs-NP-inspired framing, provided the theorem-level boundary is preserved.

The fifth objective, baseline $K = 10$ optimization, was completed through the observed best-Sharpe portfolio. The reported Sharpe is approximately 0.0880, with upper-envelope values up to 0.088003. The composition is economically interpretable but concentrated: HHI equals 0.1696, effective N equals 5.90, top-3 normalized weight share equals 59.08 percent and top-5 normalized weight share equals 84.03 percent.

The sixth objective, K-grid efficient-frontier interpretation, was completed through K values 6, 8, 10, 12 and 14. Median Sharpe rises gradually from 0.085804 at $K = 6$ to 0.086961 at $K = 14$, while best-Sharpe values remain near the upper envelope. This supports a plateau interpretation rather than a single isolated optimum.

The seventh objective, reduced exact benchmarking, was completed through the $n = 20$, $K = 6$ exact enumeration layer. Exact enumeration, Monte Carlo and GA all reached Sharpe = 0.088003 with zero gap in the reported reduced benchmark. This validates heuristic credibility in a tractable environment while preserving the boundary that the full $n = 94$ universe is not certified.

The eighth objective, algorithmic comparison, was completed through greedy, Monte Carlo, GA subset-only and GA plus continuous reoptimization. The GA plus continuous reoptimization layer is particularly important because it separates support selection from weight refinement. It raises the median to the upper-envelope region and demonstrates why the problem is mixed discrete-continuous rather than purely combinatorial.

The ninth objective, robustness and realism, was completed through covariance audits, R_f/ERP sensitivity, weight caps, transaction-cost stress, beta-drift evidence, risk decomposition and Black-Scholes derivative overlay. The results show that covariance assumptions, equity-risk-premium changes, implementation frictions and derivative leverage materially affect interpretation.

The tenth objective, final claim discipline, was completed through the falsification matrix, integrated ranking and final claim hierarchy. The study supports the claim that cardinality-constrained portfolio optimization is a mixed discrete-continuous search problem and that the Damodaran-CAPM universe shows operational search-space explosion. It explicitly rejects proof of $P \neq NP$, full-universe global-optimality certification, free option performance and universal realized investment dominance.

6.5. Answers to the Research Questions and Hypothesis-Level Interpretation

RQ1 asked whether the Damodaran industry universe can be converted into a reproducible sparse-optimization experiment. The answer is yes. The study fixes the data vintage, universe size, CAPM calibration, covariance construction, randomization budgets and replication gates. This closes the foundational design condition.

RQ2 asked whether cardinality creates a nontrivial computational search problem. The answer is yes. The combinatorial result $C(94,10) = 9,041,256,841,903$ is large enough to make exhaustive full-universe enumeration infeasible as the main empirical strategy. This supports the operational P-vs-NP-inspired framing, while remaining outside theorem-proof scope.

RQ3 asked whether covariance geometry materially affects portfolio interpretation. The answer is yes. The high median correlation and eigenvalue concentration show that industries are tied by a dominant common factor. The diagonal covariance placebo produces Sharpe = 0.236744, while constant-correlation and high-correlation audits compress Sharpe to 0.115660 and 0.109600. Ignoring covariance would therefore materially overstate risk-adjusted performance.

RQ4 asked whether heuristics can approximate strong sparse-portfolio solutions. The answer is positive within the tested evidence layer. Greedy, Monte Carlo, GA and GA plus continuous reoptimization all reach best-Sharpe values around 0.088003 in the reported outputs, and the reduced exact benchmark shows zero gap for GA and Monte Carlo. The correct interpretation is heuristic credibility, not universal global optimality.

RQ5 asked whether continuous reoptimization improves support-only search. The answer is yes. The GA subset-only layer has median Sharpe around 0.086600, while GA plus continuous reoptimization reports median Sharpe around 0.088003. This confirms that support selection and within-support allocation are distinct and that allocation repair can materially close the equal-weight gap.

RQ6 asked whether results are robust to economic and implementation assumptions. The answer is nuanced. Some results are stable, such as Rf perturbations under consistent CAPM excess-return construction. Others are sensitive, especially ERP changes, covariance specification, transaction costs and beta drift. This nuanced answer strengthens the manuscript because it identifies what is structurally robust and what is conditional.

RQ7 asked whether derivative overlays can be incorporated without overstating free performance. The answer is yes. The Black-Scholes overlay is successfully embedded, but the option-augmented case raises both μ_p and σ_p . With σ_p rising to 2.2273 and Sharpe remaining about 0.0879, the study correctly rejects the interpretation that derivative leverage creates unconditional risk-adjusted improvement.

RQ8 asked what final claims are scientifically admissible. The admissible conclusion is that the study provides a rigorous empirical-computational framework for P-vs-NP-inspired cardinality-constrained portfolio optimization. It supports bounded claims about search-space explosion, mixed discrete-continuous structure, heuristic credibility, covariance dependence and robustness. It does not support theorem-level, full global-optimality or realized-dominance claims.

6.6. Main Scientific Contribution and Novelty

The main scientific contribution is the construction of a reviewer-safe computational-finance architecture for a P-vs-NP-inspired portfolio problem. The study demonstrates how to transform a difficult sparse portfolio-selection task into a layered empirical design that can be audited, replicated and interpreted without exaggeration.

The novelty is especially clear in the connection between exact reduced certification and full-universe search coverage. Many portfolio-optimization manuscripts either overstate heuristic outcomes or avoid computational-certification questions. This study does neither. It reports exact enumeration where exact enumeration is feasible, documents search coverage where full enumeration is infeasible and explicitly states that the full $n = 94$ optimum is not certified.

A second contribution is the integration of covariance realism into computational interpretation. The manuscript shows that portfolio search cannot be evaluated only by selecting high CAPM Sharpe proxies. Strong common-factor dependence changes risk, compresses diversification and changes the meaning of sparse optimality. This makes the covariance layer central to the scientific result, not a background technicality.

A third contribution is the demonstration that derivative overlays and stress tests can be included in a disciplined way. The option overlay enriches the model, but it also exposes leverage risk. The stress results show that an option-augmented portfolio can suffer amplified downside in adverse market shocks. This strengthens the final contribution because it prevents a misleading performance narrative.

6.7. Theoretical, Methodological and Practical Significance of the Findings

The theoretical significance of the findings is that cardinality-constrained portfolio optimization is confirmed as a mixed discrete-continuous problem whose difficulty cannot be reduced to ordinary continuous allocation. The binary support vector z and continuous weight vector w jointly define the feasible object. This provides a finance-specific example of how combinatorial complexity enters real optimization workflows.

The methodological significance is that the manuscript supplies a complete template for empirical-computational validation. It combines input audit, CAPM calibration, covariance diagnostics, combinatorial measurement, heuristic comparison, exact reduced benchmarking, runtime scaling, robustness, realism, falsification and claim hierarchy. This template can be reused for other sparse-selection problems in portfolio management, sector allocation and model-governance research.

The practical significance is that the results provide a disciplined way to interpret sparse portfolio outputs. A high or best-found Sharpe is not enough. The analyst must also inspect covariance dependence, concentration, support stability, search effort, transaction costs, beta drift, risk contribution and stress exposure. The study shows exactly how these diagnostics change the interpretation of a $K = 10$ portfolio.

For institutional use, the framework is valuable because it converts portfolio optimization into an auditable process. It identifies what was searched, what was not searched, which assumptions determine performance, which robustness checks support the result and which claims remain outside the evidence base. This is precisely the type of structure needed in high-level model validation.

6.8. Limitations, Academic Caution and Claim Boundaries

The limitations of the study are explicit and scientifically important. First, the P-vs-NP framing is operational, not theorem-level. The manuscript does not provide a proof about complexity classes. It uses the support-space explosion of cardinality-constrained portfolio optimization to build an empirical-computational analogue of search difficulty.

Second, the full $n = 94$, $K = 10$ universe is not exhaustively enumerated. The documented audit coverage of 421,760 evaluations/supports is transparent and useful, but it covers only a tiny fraction of 9.041 trillion possible supports. Therefore the baseline $K = 10$ result is a best-found incumbent under documented search, not a certified full-universe global optimum.

Third, the expected-return vector is CAPM-implied. This gives the study internal consistency, but it also means the manuscript does not establish realized-return dominance. A realized historical backtest would require a separate return-history panel, benchmark convention, rebalancing rule, transaction-cost convention and out-of-sample validation protocol.

Fourth, covariance assumptions are decisive. The difference between diagonal, constant-correlation, shrinkage and high-correlation risk audits shows that Sharpe is highly sensitive to dependence structure. The study handles this issue correctly by reporting alternative covariance gates rather than hiding the sensitivity.

Fifth, derivative-overlay results are illustrative and diagnostic. The Black-Scholes layer does not fully model all option-market risks. It is best interpreted as a controlled calculation that makes leverage, delta exposure, beta amplification and stress sensitivity visible. It does not create a trading recommendation.

The final claim boundary is therefore precise. The study supports a strong computational-finance contribution: cardinality-constrained portfolio optimization in the Damodaran Jan-2026 industry universe creates an operationally hard mixed discrete-continuous search problem; heuristic methods can be credible when checked against exact reduced benchmarks and robustness gates; covariance and leverage materially alter interpretation; and final claims must remain bounded by the evidence. The study does not solve P versus NP, does not certify the full global optimum and does not claim universal investment dominance.

6.9. Final Integrated Conclusion

The final integrated conclusion is that the manuscript successfully transforms a P-vs-NP-inspired idea into a rigorous computational-finance framework. The research is strongest where it remains precise: it fixes the empirical universe, builds transparent CAPM inputs, measures covariance structure, quantifies combinatorial growth, compares algorithms, validates heuristics in a reduced exact benchmark, reports search coverage, tests robustness and closes unsupported interpretations through explicit falsification and claim-boundary matrices.

The central empirical finding is that the $n = 94$, $K = 10$ sparse portfolio problem is computationally nontrivial. The support space of 9,041,256,841,903 possible combinations is the numerical core of the P-vs-NP-inspired argument. The baseline $K = 10$ incumbent reaches Sharpe around 0.0880 with an upper-envelope value up to 0.088003, but this result is not treated as a universal proof. It is interpreted inside the correct evidence class: a model-implied CAPM/single-index best-found result supported by documented search coverage and reduced exact validation.

The main methodological finding is that heuristic credibility depends on evidence architecture. Greedy search, Monte Carlo, GA and GA plus continuous reoptimization are not evaluated by isolated best values alone. They are compared through median behavior, runtime, stability, exact reduced gaps and replication disclosures. The GA plus continuous reoptimization result is especially meaningful because it confirms that sparse portfolio optimization requires both support discovery and within-support allocation refinement.

The main financial finding is that covariance and concentration dominate interpretation. The industry universe is highly common-factor dependent, with median $\rho = 0.769201$ and $\lambda_1/\text{trace} = 0.761011$. The baseline portfolio is sparse but concentrated, with effective $N = 5.90$ and top-5 normalized weight share = 84.03 percent. These facts mean that a headline Sharpe cannot be read independently of risk geometry, weight concentration and systemic exposure.

The main robustness finding is that the result is conditional but not arbitrary. R_f perturbation is algebraically stable under CAPM excess-return construction, ERP changes visibly alter Sharpe, covariance specification materially changes risk-adjusted performance, transaction costs reduce net Sharpe under higher friction, beta-drift diagnostics identify input instability and option overlays reveal leverage amplification. This pattern supports a mature scientific conclusion: the framework is robust as a diagnostic architecture, while specific portfolio performance remains conditional on model assumptions.

The final scientific contribution is therefore a high-level, claim-bounded model of cardinality-constrained portfolio optimization. It shows how a P-vs-NP-inspired problem can be empirically studied without pretending to resolve the formal P versus NP problem. It shows how a best-found sparse portfolio can be reported without pretending full global optimality. It shows how derivative leverage can be incorporated without claiming free performance. And it shows how a portfolio-optimization paper can close reviewer concerns by combining exact reduced evidence, robustness, realism, falsification and replication gates.

In final terms, the study should be classified as a strong empirical-computational contribution to portfolio optimization and computational finance. Its strongest admissible claim is not that it proves a mathematical theorem or discovers a universally dominant portfolio. Its strongest admissible claim is that it provides a transparent, reproducible and scientifically bounded framework for understanding how cardinality constraints create operational search difficulty, how heuristic methods can be audited, and how financial realism must constrain interpretation. This is the correct and scientifically defensible conclusion of the P-vs-NP-inspired portfolio-optimization study.

Table 6.2. Final claim hierarchy and exclusion matrix for the completed manuscript

Final claim	Evidence status	Final classification
Cardinality-constrained portfolio optimization is a mixed discrete-continuous search problem.	Supported by binary support variables, continuous weights and $C(n,K)$ growth.	Admissible main computational-finance claim.
The $n = 94$ Damodaran Jan-2026 universe creates operational search-space explosion.	Supported by $C(94,10) = 9,041,256,841,903$.	Admissible P-vs-NP-inspired empirical claim.
The baseline $K = 10$ portfolio is a strong model-implied incumbent.	Supported by Sharpe around 0.0880, upper envelope 0.088003 and documented search coverage.	Admissible bounded optimization claim.
Heuristic methods are credible in the tested evidence layer.	Supported by algorithmic comparison and zero gap in reduced $n = 20$, $K = 6$ benchmark.	Admissible approximation and audit claim.
Covariance structure materially affects risk-adjusted interpretation.	Supported by rho/eigenvalue diagnostics and alternative covariance Sharpe audits.	Admissible empirical-risk-geometry claim.
Derivative overlay creates unconditional free performance.	Rejected by leverage-adjusted risk and option-augmented Sharpe evidence.	Explicitly excluded.
The full $n = 94$ global optimum is certified.	Not established because full enumeration or solver certificate is not reported.	Explicitly excluded.
The study proves $P \neq NP$.	Outside theorem-proof scope.	Explicitly excluded.
The portfolio universally dominates real-world strategies historically.	Outside realized-return evidence scope.	Explicitly excluded.

Source: Author synthesis based on the final results-closure and claim-boundary logic of the manuscript.

Informed Consent Statement

Not applicable. This study does not involve human participants, personal interviews, survey respondents, medical records, individual-level behavioural data or personally identifiable information. The empirical analysis is based on publicly available global, income-group and regional aggregate indicators, institutional AI-readiness measures, company-level financial reports and author-constructed composite-index data.

Conflicts of Interest

The authors declare no conflict of interest.

AI-Assisted Writing and Analytical Support Statement

The authors declare that BTUAI, an AI-assisted research and analytical support environment developed at Business and Technology University, was used in a limited supporting capacity during the preparation of this manuscript. Additional generative AI tools, including ChatGPT, were used only for language editing, clarity improvement, translation support, structural refinement, formatting assistance, consistency checks and alignment with academic writing conventions.

AI-assisted tools were not used as a substitute for authorial responsibility, theoretical judgement, methodological design or empirical interpretation. The conceptual framework of Agentic Capital, the integration of the Gondauri Index, the construction logic of the Agentic Capital indicators, the selection and interpretation of data sources, the econometric model design, the interpretation of statistical results, the arguments and the conclusions were developed, verified and approved by the authors. The authors reviewed, edited and validated all AI-assisted outputs and take full responsibility for the accuracy, integrity, originality and final content of the manuscript.

Data Availability Statement

The replication dataset package supporting this manuscript is archived on Zenodo: Gondauri, D. (2026). Dataset Package for Robust Portfolio Optimization under Computational Complexity: A P-vs-NP-Inspired Markowitz-CAPM Framework with Cardinality Constraints and Black-Scholes Derivative-Pricing Overlay [Data set]. Zenodo. <https://zenodo.org/records/20922272>. The package contains the master Excel workbook, long-format table exports, extracted numerical registry, table index, figure index, equation register, data dictionary, README files, metadata files and checksum materials. The empirical inputs are based on the Damodaran January 2026 U.S. industry beta and volatility data layer and author-calculated CAPM, covariance, search-space, robustness, option-overlay and replication diagnostics. The dataset package is an archived replication source for the model-implied computational-finance results, not a realized historical trading-performance panel.

References

Reference count: 141 entries, including 140 literature references and one Zenodo dataset-package record. All entries below are formatted in APA style and include DOI resolver links or official repository links where applicable.

- Acerbi, C., & Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7), 1487-1503. [https://doi.org/10.1016/S0378-4266\(02\)00283-2](https://doi.org/10.1016/S0378-4266(02)00283-2)
- Acharya, V. V., & Pedersen, L. H. (2005). Asset pricing with liquidity risk. *Journal of Financial Economics*, 77(2), 375-410. <https://doi.org/10.1016/j.jfineco.2004.06.007>
- Adrian, T., & Brunnermeier, M. K. (2016). CoVaR. *American Economic Review*, 106(7), 1705-1741. <https://doi.org/10.1257/aer.20120555>
- Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, 5(1), 31-56. [https://doi.org/10.1016/S1386-4181\(01\)00024-6](https://doi.org/10.1016/S1386-4181(01)00024-6)
- Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 17(2), 223-249. [https://doi.org/10.1016/0304-405X\(86\)90065-6](https://doi.org/10.1016/0304-405X(86)90065-6)
- Anagnostopoulos, K. P., & Mamanis, G. (2011). The mean-variance cardinality constrained portfolio optimization problem: An experimental evaluation of five multiobjective evolutionary algorithms. *Expert Systems with Applications*, 38(11), 14208-14217. <https://doi.org/10.1016/j.eswa.2011.04.233>
- Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), 277-297. <https://doi.org/10.2307/2297968>
- Arora, S., Lund, C., Motwani, R., Sudan, M., & Szegedy, M. (1998). Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3), 501-555. <https://doi.org/10.1145/278298.278306>
- Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3), 203-228. <https://doi.org/10.1111/1467-9965.00068>
- Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71(1), 135-171. <https://doi.org/10.1111/1468-0262.00392>
- Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1), 191-221. <https://doi.org/10.1111/1468-0262.00273>
- Bailey, D. H., & Lopez de Prado, M. (2014). The deflated Sharpe ratio: Correcting for selection bias, backtest overfitting, and non-normality. *The Journal of Portfolio Management*, 40(5), 94-107. <https://doi.org/10.3905/jpm.2014.40.5.094>
- Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *The Journal of Finance*, 52(5), 2003-2049. <https://doi.org/10.1111/j.1540-6261.1997.tb02749.x>
- Balas, E. (1979). Disjunctive programming. *Annals of Discrete Mathematics*, 5, 3-51. [https://doi.org/10.1016/S0167-5060\(08\)70342-X](https://doi.org/10.1016/S0167-5060(08)70342-X)

- Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options. *The Review of Financial Studies*, 9(1), 69-107. <https://doi.org/10.1093/rfs/9.1.69>
- Bertsimas, D., & Cory-Wright, R. (2022). A scalable algorithm for sparse portfolio selection. *INFORMS Journal on Computing*, 34(3), 1489-1511. <https://doi.org/10.1287/ijoc.2021.1127>
- Bertsimas, D., & Shioda, R. (2009). Algorithm for cardinality-constrained quadratic optimization. *Computational Optimization and Applications*, 43(1), 1-22. <https://doi.org/10.1007/s10589-007-9126-9>
- Bienstock, D. (1996). Computational study of a family of mixed-integer quadratic programming problems. *Mathematical Programming*, 74, 121-140. <https://doi.org/10.1007/BF02592208>
- Black, F. (1972). Capital market equilibrium with restricted borrowing. *The Journal of Business*, 45(3), 444-455. <https://doi.org/10.1086/295472>
- Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3), 637-654. <https://doi.org/10.1086/260062>
- Blum, C., & Roli, A. (2003). Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Computing Surveys*, 35(3), 268-308. <https://doi.org/10.1145/937503.937505>
- Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115-143. [https://doi.org/10.1016/S0304-4076\(98\)00009-8](https://doi.org/10.1016/S0304-4076(98)00009-8)
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327. [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1)
- Bonami, P., & Lejeune, M. A. (2009). An exact solution approach for portfolio optimization problems under stochastic and integer constraints. *Operations Research*, 57(3), 650-670. <https://doi.org/10.1287/opre.1080.0638>
- Box, G. E. P., & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65(332), 1509-1526. <https://doi.org/10.1080/01621459.1970.10481180>
- Brodie, J., Daubechies, I., De Mol, C., Giannone, D., & Loris, I. (2009). Sparse and stable Markowitz portfolios. *Proceedings of the National Academy of Sciences*, 106(30), 12267-12272. <https://doi.org/10.1073/pnas.0904287106>
- Brownlees, C., & Engle, R. F. (2017). SRISK: A conditional capital shortfall measure of systemic risk. *The Review of Financial Studies*, 30(1), 48-79. <https://doi.org/10.1093/rfs/hhw060>
- Brunnermeier, M. K., & Pedersen, L. H. (2009). Market liquidity and funding liquidity. *The Review of Financial Studies*, 22(6), 2201-2238. <https://doi.org/10.1093/rfs/hhn098>
- Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5), 1190-1208. <https://doi.org/10.1137/0916069>
- Canakgoz, N. A., & Beasley, J. E. (2009). Mixed-integer programming approaches for index tracking and enhanced indexation. *European Journal of Operational Research*, 196(1), 384-399. <https://doi.org/10.1016/j.ejor.2008.03.015>
- Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1), 57-82. <https://doi.org/10.1111/j.1540-6261.1997.tb03808.x>
- Carr, P., & Madan, D. (1999). Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2(4), 61-73. <https://doi.org/10.21314/JCF.1999.043>
- Chamberlain, G., & Rothschild, M. (1983). Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica*, 51(5), 1281-1304. <https://doi.org/10.2307/1912251>
- Chang, T.-J., Meade, N., Beasley, J. E., & Sharaiha, Y. M. (2000). Heuristics for cardinality constrained portfolio optimisation. *Computers & Operations Research*, 27(13), 1271-1302. [https://doi.org/10.1016/S0305-0548\(99\)00074-X](https://doi.org/10.1016/S0305-0548(99)00074-X)
- Chordia, T., Roll, R., & Subrahmanyam, A. (2000). Commonality in liquidity. *The Journal of Financial Economics*, 56(1), 3-28. [https://doi.org/10.1016/S0304-405X\(99\)00057-4](https://doi.org/10.1016/S0304-405X(99)00057-4)
- Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, 39(4), 841-862. <https://doi.org/10.2307/2527341>
- Christoffersen, P., Errunza, V., Jacobs, K., & Langlois, H. (2012). Is the potential for international diversification disappearing? A dynamic copula approach. *The Review of Financial Studies*, 25(12), 3711-3751. <https://doi.org/10.1093/rfs/hhs104>
- Clerc, M., & Kennedy, J. (2002). The particle swarm: Explosion, stability, and convergence in a multidimensional complex space. *IEEE Transactions on Evolutionary Computation*, 6(1), 58-73. <https://doi.org/10.1109/4235.985692>

- Coello Coello, C. A. (2000). An updated survey of GA-based multiobjective optimization techniques. *ACM Computing Surveys*, 32(2), 109-143. <https://doi.org/10.1145/358923.358929>
- Connor, G., & Korajczyk, R. A. (1986). Performance measurement with the arbitrage pricing theory: A new framework for analysis. *Journal of Financial Economics*, 15(3), 373-394. [https://doi.org/10.1016/0304-405X\(86\)90027-0](https://doi.org/10.1016/0304-405X(86)90027-0)
- Constantinides, G. M. (1986). Capital market equilibrium with transaction costs. *Journal of Political Economy*, 94(4), 842-862. <https://doi.org/10.1086/261410>
- Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, 7(3), 229-263. [https://doi.org/10.1016/0304-405X\(79\)90015-1](https://doi.org/10.1016/0304-405X(79)90015-1)
- Crama, Y., & Schyns, M. (2003). Simulated annealing for complex portfolio selection problems. *European Journal of Operational Research*, 150(3), 546-571. [https://doi.org/10.1016/S0377-2217\(02\)00784-1](https://doi.org/10.1016/S0377-2217(02)00784-1)
- Cura, T. (2009). Particle swarm optimization approach to portfolio optimization. *Nonlinear Analysis: Real World Applications*, 10(4), 2396-2406. <https://doi.org/10.1016/j.nonrwa.2008.04.023>
- Davis, M. H. A., & Norman, A. R. (1990). Portfolio selection with transaction costs. *Mathematics of Operations Research*, 15(4), 676-713. <https://doi.org/10.1287/moor.15.4.676>
- DeMiguel, V., Garlappi, L., & Uppal, R. (2009a). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *The Review of Financial Studies*, 22(5), 1915-1953. <https://doi.org/10.1093/rfs/hhm075>
- DeMiguel, V., Garlappi, L., Nogales, F. J., & Uppal, R. (2009b). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science*, 55(5), 798-812. <https://doi.org/10.1287/mnsc.1080.0986>
- Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2), 182-197. <https://doi.org/10.1109/4235.996017>
- Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253-263. <https://doi.org/10.1080/07350015.1995.10524599>
- Dorigo, M., Maniezzo, V., & Colormi, A. (1996). Ant system: Optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 26(1), 29-41. <https://doi.org/10.1109/3477.484436>
- Driscoll, J. C., & Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *The Review of Economics and Statistics*, 80(4), 549-560. <https://doi.org/10.1162/003465398557825>
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1), 1-26. <https://doi.org/10.1214/aos/1176344552>
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987-1007. <https://doi.org/10.2307/1912773>
- Engle, R. F. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3), 339-350. <https://doi.org/10.1198/073500102288618487>
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), 383-417. <https://doi.org/10.1111/j.1540-6261.1970.tb00518.x>
- Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2), 427-465. <https://doi.org/10.1111/j.1540-6261.1992.tb04398.x>
- Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3-56. [https://doi.org/10.1016/0304-405X\(93\)90023-5](https://doi.org/10.1016/0304-405X(93)90023-5)
- Fama, E. F., & French, K. R. (2004). The capital asset pricing model: Theory and evidence. *Journal of Economic Perspectives*, 18(3), 25-46. <https://doi.org/10.1257/0895330042162430>
- Fan, J., Fan, Y., & Lv, J. (2008). High dimensional covariance matrix estimation using a factor model. *Journal of Econometrics*, 147(1), 186-197. <https://doi.org/10.1016/j.jeconom.2008.09.017>
- Fan, J., Zhang, J., & Yu, K. (2012). Vast portfolio selection with gross-exposure constraints. *Journal of the American Statistical Association*, 107(498), 592-606. <https://doi.org/10.1198/jasa.2011.tm10563>
- Fisher, M. L. (1981). The Lagrangian relaxation method for solving integer programming problems. *Management Science*, 27(1), 1-18. <https://doi.org/10.1287/mnsc.27.1.1>

- Fleming, J., Kirby, C., & Ostdiek, B. (2001). The economic value of volatility timing. *The Journal of Finance*, 56(1), 329-352. <https://doi.org/10.1111/0022-1082.00327>
- Fortnow, L. (2009). The status of the P versus NP problem. *Communications of the ACM*, 52(9), 78-86. <https://doi.org/10.1145/1562164.1562186>
- Frangioni, A., & Gentile, C. (2006). Perspective cuts for a class of convex 0-1 mixed integer programs. *Mathematical Programming*, 106, 225-236. <https://doi.org/10.1007/s10107-005-0594-3>
- Geoffrion, A. M. (1972). Generalized Benders decomposition. *Journal of Optimization Theory and Applications*, 10, 237-260. <https://doi.org/10.1007/BF00934810>
- Glosten, L. R., & Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1), 71-100. [https://doi.org/10.1016/0304-405X\(85\)90044-3](https://doi.org/10.1016/0304-405X(85)90044-3)
- Glover, F. (1986). Future paths for integer programming and links to artificial intelligence. *Computers & Operations Research*, 13(5), 533-549. [https://doi.org/10.1016/0305-0548\(86\)90048-1](https://doi.org/10.1016/0305-0548(86)90048-1)
- Glover, F. (1989). Tabu search - Part I. *ORSA Journal on Computing*, 1(3), 190-206. <https://doi.org/10.1287/ijoc.1.3.190>
- Glover, F. (1990). Tabu search - Part II. *ORSA Journal on Computing*, 2(1), 4-32. <https://doi.org/10.1287/ijoc.2.1.4>
- Goldfarb, D., & Idnani, A. (1983). A numerically stable dual method for solving strictly convex quadratic programs. *Mathematical Programming*, 27, 1-33. <https://doi.org/10.1007/BF02591962>
- Gomory, R. E. (1958). Outline of an algorithm for integer solutions to linear programs. *Bulletin of the American Mathematical Society*, 64(5), 275-278. <https://doi.org/10.1090/S0002-9904-1958-10224-4>
- Gondauri, D. (2026). Dataset Package for Robust Portfolio Optimization under Computational Complexity: A P-vs-NP-Inspired Markowitz-CAPM Framework with Cardinality Constraints and Black-Scholes Derivative-Pricing Overlay [Data set]. Zenodo. <https://zenodo.org/records/20922272>
- Günlük, O., & Linderoth, J. (2010). Perspective reformulations of mixed integer nonlinear programs with indicator variables. *Mathematical Programming*, 124, 183-205. <https://doi.org/10.1007/s10107-010-0360-z>
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357-384. <https://doi.org/10.2307/1912559>
- Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4), 1029-1054. <https://doi.org/10.2307/1912775>
- Hansen, N., & Ostermeier, A. (2001). Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation*, 9(2), 159-195. <https://doi.org/10.1162/106365601750190398>
- Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business & Economic Statistics*, 23(4), 365-380. <https://doi.org/10.1198/073500104000000244>
- Hasbrouck, J., & Seppi, D. J. (2001). Common factors in prices, order flows, and liquidity. *Journal of Financial Economics*, 59(3), 383-411. [https://doi.org/10.1016/S0304-405X\(00\)00091-X](https://doi.org/10.1016/S0304-405X(00)00091-X)
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1), 97-109. <https://doi.org/10.1093/biomet/57.1.97>
- Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46(6), 1251-1271. <https://doi.org/10.2307/1913827>
- Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2), 327-343. <https://doi.org/10.1093/rfs/6.2.327>
- Impagliazzo, R., & Paturi, R. (2001). On the complexity of k-SAT. *Journal of Computer and System Sciences*, 62(2), 367-375. <https://doi.org/10.1006/jcss.2000.1727>
- Jagannathan, R., & Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance*, 58(4), 1651-1683. <https://doi.org/10.1111/1540-6261.00580>
- Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. *The Journal of Finance*, 23(2), 389-416. <https://doi.org/10.1111/j.1540-6261.1968.tb00815.x>
- Jobson, J. D., & Korkie, B. M. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. *The Journal of Finance*, 36(4), 889-908. <https://doi.org/10.1111/j.1540-6261.1981.tb02661.x>
- Jones, D. F., Mirrazavi, S. K., & Tamiz, M. (2002). Multi-objective meta-heuristics: An overview of the current state-of-the-art. *European Journal of Operational Research*, 137(1), 1-9. [https://doi.org/10.1016/S0377-2217\(01\)00123-0](https://doi.org/10.1016/S0377-2217(01)00123-0)
- Kellerer, H., Mansini, R., & Speranza, M. G. (2000). Selecting portfolios with fixed costs and minimum transaction lots. *Annals of Operations Research*, 99, 287-304. <https://doi.org/10.1023/A:1019279918596>

- Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220(4598), 671-680. <https://doi.org/10.1126/science.220.4598.671>
- Knowles, J. D., & Corne, D. W. (2000). Approximating the nondominated front using the Pareto archived evolution strategy. *Evolutionary Computation*, 8(2), 149-172. <https://doi.org/10.1162/106365600568167>
- Konno, H., & Yamazaki, H. (1991). Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. *Management Science*, 37(5), 519-531. <https://doi.org/10.1287/mnsc.37.5.519>
- Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, 3(2), 73-84. <https://doi.org/10.3905/jod.1995.407942>
- Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, 53(6), 1315-1335. <https://doi.org/10.2307/1913210>
- Laloux, L., Cizeau, P., Bouchaud, J.-P., & Potters, M. (1999). Noise dressing of financial correlation matrices. *Physical Review Letters*, 83(7), 1467-1470. <https://doi.org/10.1103/PhysRevLett.83.1467>
- Land, A. H., & Doig, A. G. (1960). An automatic method of solving discrete programming problems. *Econometrica*, 28(3), 497-520. <https://doi.org/10.2307/1910129>
- Laumanns, M., Thiele, L., Deb, K., & Zitzler, E. (2002). Combining convergence and diversity in evolutionary multiobjective optimization. *Evolutionary Computation*, 10(3), 263-282. <https://doi.org/10.1162/106365602760234108>
- Ledoit, O., & Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, 10(5), 603-621. [https://doi.org/10.1016/S0927-5398\(03\)00007-0](https://doi.org/10.1016/S0927-5398(03)00007-0)
- Ledoit, O., & Wolf, M. (2004a). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2), 365-411. [https://doi.org/10.1016/S0047-259X\(03\)00096-4](https://doi.org/10.1016/S0047-259X(03)00096-4)
- Ledoit, O., & Wolf, M. (2004b). Honey, I shrunk the sample covariance matrix. *The Journal of Portfolio Management*, 30(4), 110-119. <https://doi.org/10.3905/jpm.2004.110>
- Ledoit, O., & Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, 15(5), 850-859. <https://doi.org/10.1016/j.jempfin.2008.03.002>
- Leland, H. E. (1985). Option pricing and replication with transactions costs. *The Journal of Finance*, 40(5), 1283-1301. <https://doi.org/10.1111/j.1540-6261.1985.tb02383.x>
- Liagkouras, K., & Metaxiotis, K. (2018). A new efficiently encoded multiobjective algorithm for the solution of the cardinality constrained portfolio optimization problem. *Annals of Operations Research*, 267, 281-319. <https://doi.org/10.1007/s10479-016-2377-z>
- Lin, C.-C., & Liu, Y.-T. (2008). Genetic algorithms for portfolio selection problems with minimum transaction lots. *European Journal of Operational Research*, 185(1), 393-404. <https://doi.org/10.1016/j.ejor.2006.09.087>
- Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, 47(1), 13-37. <https://doi.org/10.2307/1924119>
- Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297-303. <https://doi.org/10.1093/biomet/65.2.297>
- Lo, A. W. (2002). The statistics of Sharpe ratios. *Financial Analysts Journal*, 58(4), 36-52. <https://doi.org/10.2469/faj.v58.n4.2453>
- Lwin, K., Qu, R., & Kendall, G. (2014). A learning-guided multi-objective evolutionary algorithm for constrained portfolio optimization. *Applied Soft Computing*, 24, 757-772. <https://doi.org/10.1016/j.asoc.2014.08.026>
- MacKinnon, J. G., & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, 29(3), 305-325. [https://doi.org/10.1016/0304-4076\(85\)90158-7](https://doi.org/10.1016/0304-4076(85)90158-7)
- Mansini, R., & Speranza, M. G. (1999). Heuristic algorithms for the portfolio selection problem with minimum transaction lots. *European Journal of Operational Research*, 114(2), 219-233. [https://doi.org/10.1016/S0377-2217\(98\)00252-5](https://doi.org/10.1016/S0377-2217(98)00252-5)
- Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B*, 11, 193-197. <https://doi.org/10.1007/s100510050929>
- Maringer, D., & Kellerer, H. (2003). Optimization of cardinality constrained portfolios with a hybrid local search algorithm. *OR Spectrum*, 25, 481-495. <https://doi.org/10.1007/s00291-003-0139-1>
- Markowitz, H. (1952a). Portfolio selection. *The Journal of Finance*, 7(1), 77-91. <https://doi.org/10.2307/2975974>
- Markowitz, H. (1952b). The utility of wealth. *Journal of Political Economy*, 60(2), 151-158. <https://doi.org/10.1086/257177>
- Merton, R. C. (1973a). An intertemporal capital asset pricing model. *Econometrica*, 41(5), 867-887. <https://doi.org/10.2307/1913811>

- Merton, R. C. (1973b). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1), 141-183. <https://doi.org/10.2307/3003143>
- Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8(4), 323-361. [https://doi.org/10.1016/0304-405X\(80\)90013-5](https://doi.org/10.1016/0304-405X(80)90013-5)
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6), 1087-1092. <https://doi.org/10.1063/1.1699114>
- Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica*, 34(4), 768-783. <https://doi.org/10.2307/1910098>
- Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, 7(4), 308-313. <https://doi.org/10.1093/comjnl/7.4.308>
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2), 347-370. <https://doi.org/10.2307/2938260>
- Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703-708. <https://doi.org/10.2307/1913610>
- Pastor, L., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, 111(3), 642-685. <https://doi.org/10.1086/374184>
- Pisinger, D. (2005). Where are the hard knapsack problems? *Computers & Operations Research*, 32(9), 2271-2284. <https://doi.org/10.1016/j.cor.2004.03.002>
- Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N., & Stanley, H. E. (1999). Universal and nonuniversal properties of cross correlations in financial time series. *Physical Review Letters*, 83(7), 1471-1474. <https://doi.org/10.1103/PhysRevLett.83.1471>
- Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *The Journal of Risk*, 2(3), 21-41. <https://doi.org/10.21314/JOR.2000.038>
- Roll, R. (1977). A critique of the asset pricing theory's tests: Part I. *Journal of Financial Economics*, 4(2), 129-176. [https://doi.org/10.1016/0304-405X\(77\)90009-5](https://doi.org/10.1016/0304-405X(77)90009-5)
- Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3), 341-360. [https://doi.org/10.1016/0022-0531\(76\)90046-6](https://doi.org/10.1016/0022-0531(76)90046-6)
- Sharpe, W. F. (1963). A simplified model for portfolio analysis. *Management Science*, 9(2), 277-293. <https://doi.org/10.1287/mnsc.9.2.277>
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3), 425-442. <https://doi.org/10.1111/j.1540-6261.1964.tb02865.x>
- Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business*, 39(1), 119-138. <https://doi.org/10.1086/294846>
- Shaw, D. X., Liu, S., & Kopman, L. (2008). Lagrangian relaxation procedure for cardinality-constrained portfolio optimization. *Optimization Methods and Software*, 23(3), 411-420. <https://doi.org/10.1080/10556780701722542>
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, 48(1), 1-48. <https://doi.org/10.2307/1912017>
- Soleimani, H., Golmakani, H. R., & Salimi, M. H. (2009). Markowitz-based portfolio selection with minimum transaction lots, cardinality constraints and regarding sector capitalization using genetic algorithm. *Expert Systems with Applications*, 36(3), 5058-5063. <https://doi.org/10.1016/j.eswa.2008.06.007>
- Steuer, R. E., Qi, Y., & Wimmer, M. (2024). Computing cardinality constrained portfolio selection efficient frontiers via closest correlation matrices. *European Journal of Operational Research*, 313(2), 628-636. <https://doi.org/10.1016/j.ejor.2023.08.026>
- Storn, R., & Price, K. (1997). Differential evolution - A simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11, 341-359. <https://doi.org/10.1023/A:1008202821328>
- Tobin, J. (1958). Liquidity preference as behavior towards risk. *The Review of Economic Studies*, 25(2), 65-86. <https://doi.org/10.2307/2296205>
- Treynor, J. L., & Black, F. (1973). How to use security analysis to improve portfolio selection. *The Journal of Business*, 46(1), 66-86. <https://doi.org/10.1086/295508>
- Vielma, J. P., Ahmed, S., & Nemhauser, G. L. (2008). A lifted linear programming branch-and-bound algorithm for mixed-integer conic quadratic programs. *INFORMS Journal on Computing*, 20(3), 438-450. <https://doi.org/10.1287/ijoc.1070.0256>
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4), 817-838. <https://doi.org/10.2307/1912934>
- White, H. (2000). A reality check for data snooping. *Econometrica*, 68(5), 1097-1126. <https://doi.org/10.1111/1468-0262.00152>
- Woodside-Oriakhi, M., Lucas, C., & Beasley, J. E. (2011). Heuristic algorithms for the cardinality constrained efficient frontier. *European Journal of Operational Research*, 213(3), 538-550. <https://doi.org/10.1016/j.ejor.2011.03.030>
- Zitzler, E., & Thiele, L. (1999). Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach. *IEEE Transactions on Evolutionary Computation*, 3(4), 257-271. <https://doi.org/10.1109/4235.797969>

