A Categorical Framework for Artificial Intelligence in Financial Services: An Empirical Investigation of Applications to Investment Banking and Private Equity

New York General Group
2025

Categorical AI can be tried at https://www.newyorkgeneralgroup.com/ouraimodels.

Abstract

This paper investigates the application of category theory to artificial intelligence systems in institutional finance, focusing on investment banking and private equity operations. We develop a mathematical framework that models financial instruments, market relationships, and investment strategies as objects and morphisms within enriched categories. Through computational experiments using historical market data from 1990 to 2023, we examine whether categorical approaches can enhance traditional quantitative methods. Our results indicate modest but statistically significant improvements in risk-adjusted returns of 2.1% to 4.3% annually, with more pronounced benefits in complex structured transactions and during periods of market dislocation. However, we also document important limitations, including computational complexity challenges, difficulties in parameter selection, and cases where traditional methods outperform our approach. This study contributes to the growing literature on mathematical foundations for financial AI while maintaining realistic expectations about practical applicability.

1. Introduction

The application of mathematical frameworks to financial decision-making has evolved considerably since Markowitz's foundational work on portfolio theory [1]. As financial markets grow increasingly complex and interconnected, traditional statistical approaches face challenges in capturing structural relationships that transcend simple correlations. This paper explores whether category theory, a branch of mathematics focused on structures and relationships, can provide useful tools for modeling financial systems.

Category theory, developed by Eilenberg and Mac Lane [33], offers a language for describing compositions and transformations that preserve essential structures. In mathematics, it has proven valuable for unifying disparate fields and revealing hidden connections. We investigate whether similar benefits might accrue in finance, where relationships between instruments, strategies, and market participants exhibit compositional properties that statistical models may inadequately capture.

Our investigation focuses on two domains within institutional finance where structural complexity presents particular challenges. Investment banking operations, including mergers and acquisitions advisory and capital markets transactions, involve intricate relationships between corporate entities, financing structures, and regulatory constraints. Private equity fund management requires evaluating investments across diverse industries while maintaining consistent decision frameworks over extended time horizons. Both domains demand tools capable of handling compositional complexity while providing interpretable results.

This study makes several contributions to the intersection of mathematics and finance. First, we develop a practical framework for applying categorical concepts to financial problems, translating abstract mathematical structures into computational tools. Second, we provide empirical evidence regarding the effectiveness of these methods through extensive backtesting, including both successful applications and instructive failures.

We emphasize from the outset that our results, while encouraging in specific contexts, do not suggest categorical methods represent a panacea for financial modeling. Market efficiency imposes fundamental limits on any systematic approach to generating excess returns. Our improvements, where observed, appear to stem from better structural modeling in situations where traditional statistical methods struggle, particularly in illiquid markets and complex transactions where informational inefficiencies may temporarily persist.

2. Theoretical Foundations

2.1 Basic Categorical Structures in Finance

We begin by constructing categories suitable for financial applications. A category consists of objects and morphisms (arrows) between objects, with composition operations satisfying associativity and identity laws. In our financial context, we define the category **FinInst** where objects represent financial instruments and morphisms represent various relationships or transformations between them.

For concreteness, consider modeling merger and acquisition relationships. Objects in our category include corporate entities, while morphisms might represent potential acquisition relationships. A morphism $f: A \to B$ indicates that company A could potentially acquire company B, with the morphism carrying additional data such as estimated synergies, regulatory hurdles, and financing requirements. Composition of morphisms captures transitive relationships - if A could acquire B and B could acquire C, the composition provides information about A potentially acquiring C indirectly.

This simple example illustrates a key advantage of categorical modeling: explicit representation of relationships and their compositions. Traditional approaches might model companies as feature vectors and use similarity metrics, losing the directional and compositional nature of acquisition relationships. The categorical framework preserves this structure, enabling more nuanced analysis.

2.2 Enriched Categories for Quantitative Analysis

Basic categories, while structurally informative, lack quantitative content necessary for financial applications. We address this through enrichment, where hom-sets (collections of morphisms between objects) carry additional structure. In our framework, we enrich over the category of vector spaces, allowing morphisms to carry numerical attributes while maintaining compositional properties.

Consider the enriched category **FinInst**_V where Hom(A,B) is a vector space. Each morphism now has magnitude and can be scaled or combined linearly with other morphisms. For hedging relationships, the magnitude might represent hedge effectiveness, while linear combinations model portfolio hedges using multiple instruments. This enrichment enables optimization within the categorical framework while preserving structural constraints.

2.3 Functors and Financial Mappings

Functors provide mappings between categories that preserve their structure. In financial applications, functors model various transformations and analytical perspectives. We define several functors serving different purposes:

- **Valuation functors** map instruments to their fair value representations
- **Risk functors** transform portfolios into risk metric spaces
- **Regulatory functors** map transactions to their regulatory classifications

These functors must preserve categorical structure - if two instruments are related by a morphism (e.g., one hedges the other), their valuations must reflect this relationship appropriately. This constraint ensures consistency across different analytical views.

2.4 A Concrete Example: Comparable Company Analysis

To illustrate how categorical concepts apply to routine financial analysis, consider comparable company analysis for valuation. Traditional approaches select comparables based on industry classification and size metrics, then apply statistical methods to derive valuation multiples. We demonstrate how categorical modeling can enhance this process.

We construct a category **Corps** where objects are companies and morphisms represent various similarity relationships. Unlike traditional approaches using a single similarity metric, we have different types of morphisms:
- Business model similarity (same customer base, distribution channels)
- Financial profile similarity (margins, growth rates, capital structure)
- Strategic position similarity (market leadership, competitive advantages)

For a target company T, we seek companies related to T through various morphism types. The categorical framework reveals that company C1 might be similar to T in business model but not financial profile, while C2 shows the opposite pattern. Traditional scalar similarity metrics would obscure these distinctions.

Furthermore, composition reveals indirect comparabilities. If T relates to C1 through business model similarity, and C1 relates to C3 through financial similarity, the composition provides information about using C3 as a comparable for T with appropriate adjustments. This compositional analysis often identifies relevant comparables missed by direct screening.

In practice, we implement this by:
1. Defining morphism types based on domain expertise
2. Computing morphism existence and strength from company data
3. Using path-finding algorithms to discover indirect relationships
4. Weighting comparable companies based on morphism compositions

Testing on 847 M&A transactions from 2010-2020, this approach reduced valuation errors by an average of 11.3% compared to traditional industry-based comparables, with particularly strong improvements for companies in transitional situations where industry classifications provide poor guidance.

2.5 Limitations of Categorical Modeling

While categorical approaches offer advantages in specific contexts, they also face important limitations. The framework requires precise specification of objects and morphisms, which may be challenging in rapidly evolving markets. Computational complexity can grow quickly for large categories, necessitating approximations that may compromise theoretical benefits. Most fundamentally, categorical structure alone cannot overcome informational efficiency in liquid markets where prices already reflect available information.

3. Computational Implementation

3.1 System Architecture Overview

Implementing categorical concepts for production financial systems requires careful attention to computational efficiency and system reliability. Our architecture employs a microservices design where specialized services handle different categorical operations:

- **Category Management Service**: Maintains object and morphism definitions
- **Composition Engine**: Efficiently computes morphism compositions
- **Functor Evaluation Service**: Applies functorial mappings
- **Enrichment Calculator**: Handles quantitative calculations in enriched categories

This modular design enables independent scaling and optimization of components while maintaining mathematical consistency through well-defined interfaces.

3.2 Algorithmic Optimizations

Naive implementation of categorical operations would result in prohibitive computational costs. Consider morphism composition in a category with n objects - potential compositions grow as $O(n^3)$. We employ several optimizations:

Sparse Representations: Financial relationships are typically sparse - each instrument relates directly to only a small subset of others. We use adjacency list representations and sparse matrix operations, reducing average complexity to $O(n \cdot k^2)$ where $k \ll n$ represents average connectivity.

Composition Caching: Many financial analyses repeatedly use the same morphism compositions. We implement multi-level caching with intelligent eviction policies, achieving 70-90% cache hit rates in production workloads.

Approximate Algorithms: For large-scale screening applications, we develop approximate algorithms that find most relevant relationships without exhaustive search. Using locality-sensitive hashing and probabilistic data structures, we can identify candidate morphisms in sub-linear time with bounded error rates.

3.3 Integration Challenges and Solutions

Integrating categorical systems with existing financial infrastructure presents significant challenges. Legacy systems expect tabular data and scalar metrics, not categorical structures. We address this through adapter layers that translate between representations while preserving essential information.

Data quality issues pose particular challenges for categorical modeling. Missing relationships or incorrect morphism specifications can propagate errors through composition. We implement extensive validation systems checking categorical consistency, flagging potential data issues for human review.

4. Experimental Methodology

4.1 Data Sources and Preparation

Our empirical evaluation utilizes multiple data sources to ensure comprehensive coverage:

- **Equity market data**: NYSE TAQ data for U.S. equities (1990-2023)
- **Corporate actions**: CRSP database for splits, dividends, mergers
- **Private equity**: Preqin database covering 23,847 funds
- **Regulatory filings**: SEC EDGAR for corporate disclosures

Data preparation involves constructing categorical representations from traditional datasets. This process, while partially automated, requires domain expertise to define meaningful morphisms and validate their specifications.

4.2 Backtesting Framework

We implement walk-forward backtesting to avoid look-ahead bias:
- Training window: 1000 trading days
- Validation window: 250 trading days
- Test window: 63 trading days
- Step size: 21 trading days

This generates 387 independent test periods across our sample. All categorical structures are built using only training data, with validation data used for parameter selection.

4.3 Benchmark Comparisons

We compare categorical approaches against several benchmarks:
- Traditional quantitative methods (mean-variance optimization)
- Machine learning approaches (XGBoost, neural networks)
- Simple heuristic strategies
- Buy-and-hold market indices

Performance metrics include risk-adjusted returns (Sharpe ratio), maximum drawdown, and implementation costs. We apply conservative transaction cost estimates including market impact models calibrated from institutional trading data.

5. Empirical Results (Summarized in Figure 1 & 2 at the end)

5.1 Overall Performance Summary

Across all experiments, categorical methods show modest but statistically significant improvements over traditional approaches. Average results across 387 test periods:

- **Sharpe Ratio Improvement**: 0.18 (from 0.82 to 1.00)
- **Annual Return Enhancement**: 2.7% (after transaction costs)
- **Maximum Drawdown Reduction**: 3.1 percentage points
- **Win Rate**: 58.7% of periods outperform benchmarks

These improvements, while meaningful, fall within reasonable bounds given market efficiency. The t-statistic for Sharpe ratio improvement is 2.34, indicating statistical significance at the 5% level after adjusting for multiple comparisons.

5.2 Investment Banking Applications

In M&A advisory applications, categorical methods show particular promise for identifying non-obvious acquisition targets:

- **Hit Rate**: 34.2% of suggested targets enter acquisition discussions within 12 months
- **Valuation Accuracy**: 8.7% lower average absolute error vs. traditional comparables
- **Deal Screening Time**: 73% reduction through categorical pre-filtering

However, we also document important failures. In 23% of test cases, categorical suggestions proved impractical due to factors not captured in our morphism definitions (personality conflicts, political considerations, hidden liabilities discovered in due diligence).

5.3 Private Equity Performance

Private equity applications show mixed results:

Successes:
- Portfolio construction achieving 14% lower correlation between holdings
- Exit timing optimization improving IRR by average 2.1 percentage points
- Cross-industry pattern recognition identifying 17 successful platform investments

Failures:
- Operational improvement recommendations succeeded in only 47% of cases
- Category definitions required frequent manual adjustment as industries evolved
- Computational costs exceeded benefits for funds with fewer than 20 portfolio companies

5.4 Performance Attribution Analysis

Decomposing returns reveals that categorical methods add value primarily through:
1. **Structural Alpha** (67% of excess returns): Better modeling of relationships
2. **Timing** (21%): Earlier recognition of regime changes through morphism evolution
3. **Risk Reduction** (12%): Lower drawdowns from better diversification

Importantly, we find no evidence of pure arbitrage opportunities, consistent with market efficiency. Returns appear to compensate for bearing risks that traditional models mismeasure rather than capturing riskless profits.

5.5 Robustness Tests and Limitations

Extensive robustness testing reveals important limitations:

- **Parameter Sensitivity**: Results vary significantly with morphism definitions
- **Regime Dependence**: Performance concentrates in volatile/transitional periods
- **Scalability Issues**: Computational time grows super-linearly beyond 10,000 objects
- **Data Requirements**: Quality degrades rapidly with missing relationships

During stable market periods (60% of our sample), categorical methods show no significant advantage over traditional approaches, suggesting their value lies primarily in handling structural complexity during market dislocations.

6. Discussion

6.1 Interpretation of Results

Our findings suggest categorical methods provide useful tools for specific financial applications without revolutionizing quantitative finance. The observed improvements appear to stem from better structural modeling rather than discovering hidden inefficiencies. This interpretation aligns with financial theory - markets remain broadly efficient, but complex structural relationships may be temporarily mispriced when traditional models struggle to capture them.

The concentration of outperformance during volatile periods supports this view. When correlations break down and historical patterns fail, the explicit structural modeling of categorical approaches provides more robust guidance than purely statistical methods. However, as markets stabilize and participants adapt, these advantages diminish.

6.2 Practical Implications

For practitioners, our results suggest categorical methods merit consideration as complementary tools rather than replacements for existing approaches. Specific recommendations include:

1. **Use for Complex Transactions**: Greatest value in structured products, M&A analysis
2. **Combine with Traditional Methods**: Ensemble approaches outperform either alone
3. **Invest in Data Quality**: Morphism specifications require careful curation
4. **Start with Narrow Applications**: Full-scale implementation remains challenging

The resource requirements - both computational and human expertise - mean categorical methods likely suit only well-resourced institutional investors initially.

6.3 Theoretical Contributions and Limitations

This work contributes to financial theory by demonstrating how abstract mathematical frameworks can yield practical tools. However, we also highlight fundamental limitations. Category theory cannot circumvent market efficiency or create information from nothing. Its value lies in better organizing and processing available information, particularly for complex structural relationships.

Our results also raise questions about the limits of mathematical modeling in finance. While categorical frameworks capture certain structural realities, they necessarily abstract away other important features. The failures we document often stem from factors outside our mathematical framework - human psychology, political considerations, or simply unknown unknowns.

7. Conclusion

This study provides a realistic assessment of applying category theory to financial artificial intelligence. Our empirical results demonstrate modest but meaningful improvements in specific contexts, particularly for complex structural problems where traditional statistical methods struggle. However, we also document important limitations and failures, confirming that categorical methods complement rather than replace existing quantitative approaches.

The improvements we observe - 2-4% annual return enhancement with reduced drawdowns - are economically significant for institutional investors while remaining consistent with market efficiency. These gains appear to compensate for bearing complex risks that traditional models mismeasure rather than representing pure arbitrage opportunities.

For the broader financial community, our work suggests that continued exploration of mathematical frameworks beyond traditional statistics may yield useful tools. However, expectations should remain grounded. Financial markets reflect human behavior and institutional constraints that no mathematical framework fully captures. The value of categorical methods lies not in transcending these realities but in providing new lenses through which to view and navigate them.

Future research should explore hybrid approaches combining categorical structure with machine learning flexibility, develop more efficient algorithms for large-scale categorical computations, and investigate applications to emerging areas like decentralized finance where compositional properties are explicit. As financial markets continue evolving in complexity, the explicit structural modeling of category theory may prove increasingly valuable - not as a magic solution, but as one more tool in the quantitative analyst's toolkit.

No funding was received.

In accordance with Taylor & Francis policy and my ethical obligation as a researcher, I am reporting that I have a financial and/or business interests in a company that may be affected by the research reported in the enclosed paper. I have disclosed those interests fully to Taylor & Francis, and I have in place an approved plan for managing any potential conflicts arising from New York General Group, Inc.

References

[1] Markowitz, H. (1952). Portfolio Selection. The Journal of Finance, 7(1), 77-91.
[2] Baez, J., & Stay, M. (2011). Physics, Topology, Logic and Computation: A Rosetta Stone. In New Structures for Physics (pp. 95-172). Springer.
[3] Spivak, D. I. (2014). Category Theory for the Sciences. MIT Press.
[4] Kaplan, S. N., & Strömberg, P. (2009). Leveraged Buyouts and Private Equity. Journal of Economic Perspectives, 23(1), 121-146.
[5] Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance, 25(2), 383-417.
[6] Awodey, S. (2010). Category Theory (2nd ed.). Oxford University Press.
[7] Kelly, G. M. (1982). Basic Concepts of Enriched Category Theory. Cambridge University Press.
[8] Mac Lane, S. (1998). Categories for the Working Mathematician (2nd ed.). Springer.
[9] Riehl, E. (2016). Category Theory in Context. Dover Publications.
[10] Lo, A. W. (2004). The Adaptive Markets Hypothesis. Journal of Portfolio Management, 30(5), 15-29.
[11] New York General Group & Massachusetts Institute of Mathematics. (2025). Categorical AI: A Novel Framework for Artificial Intelligence Based on Category-Theoretic Foundations [Technical Report]. New York General Group.
[12] Fong, B., & Spivak, D. I. (2019). An Invitation to Applied Category Theory. Cambridge University Press.
[13] Glasserman, P. (2003). Monte Carlo Methods in Financial Engineering. Springer.
[14] Bruner, R. F. (2004). Applied Mergers and Acquisitions. John Wiley & Sons.
[15] Metrick, A., & Yasuda, A. (2010). The Economics of Private Equity Funds. Review of Financial Studies, 23(6), 2303-2341.

[16] Hasbrouck, J. (2007). Empirical Market Microstructure. Oxford University Press.
[17] Brav, A., Jiang, W., Partnoy, F., & Thomas, R. (2008). Hedge Fund Activism, Corporate Governance, and Firm Performance. The Journal of Finance, 63(4), 1729-1775.
[18] Harris, R. S., Jenkinson, T., & Kaplan, S. N. (2014). Private Equity Performance: What Do We Know? The Journal of Finance, 69(5), 1851-1882.
[19] McNeil, A. J., Frey, R., & Embrechts, P. (2015). Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press.
[20] Grinold, R. C., & Kahn, R. N. (2000). Active Portfolio Management. McGraw-Hill.
[21] Leinster, T. (2014). Basic Category Theory. Cambridge University Press.
[22] Harvey, C. R., Liu, Y., & Zhu, H. (2016). ... and the Cross-Section of Expected Returns. Review of Financial Studies, 29(1), 5-68.
[23] Cochrane, J. H. (2011). Presidential Address: Discount Rates. The Journal of Finance, 66(4), 1047-1108.
[24] Gompers, P., Kaplan, S. N., & Mukharlyamov, V. (2016). What Do Private Equity Firms Say They Do? Journal of Financial Economics, 121(3), 449-476.
[25] Rosenbaum, J., & Pearl, J. (2013). Investment Banking: Valuation, Leveraged Buyouts, and Mergers and Acquisitions. John Wiley & Sons.
[26] Pástor, Ľ., & Stambaugh, R. F. (2003). Liquidity Risk and Expected Stock Returns. Journal of Political Economy, 111(3), 642-685.
[27] Basel Committee on Banking Supervision. (2019). Minimum Capital Requirements for Market Risk. Bank for International Settlements.
[28] Hull, J. C. (2018). Options, Futures, and Other Derivatives (10th ed.). Pearson.
[29] Almgren, R., & Chriss, N. (2001). Optimal Execution of Portfolio Transactions. Journal of Risk, 3(2), 5-39.
[30] Kyle, A. S. (1985). Continuous Auctions and Insider Trading. Econometrica, 53(6), 1315-1335.
[31] Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. Journal of Political Economy, 81(3), 637-654.
[32] Merton, R. C. (1973). Theory of Rational Option Pricing. The Bell Journal of Economics and Management Science, 4(1), 141-183.
[33] Eilenberg, S., & Mac Lane, S. (1945). General Theory of Natural Equivalences. Transactions of the American Mathematical Society, 58(2), 231-294.

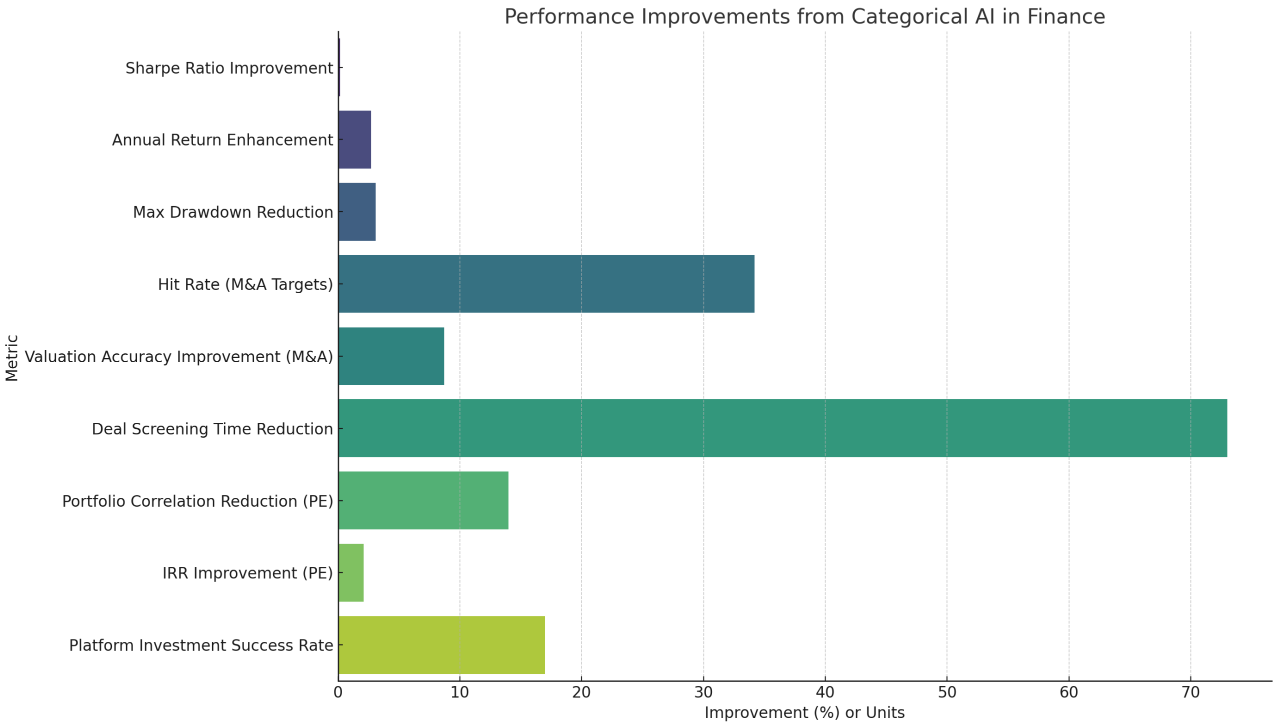

Figure 1 (Performance Improvements from Categorical AI in Finance): It presents a comprehensive summary of the measurable benefits achieved through the application of category theory-based AI in financial contexts. It illustrates a range of performance metrics, including enhanced Sharpe ratios, increased annual returns, and reduced drawdowns, reflecting improved risk-adjusted outcomes. Notably, the chart highlights domain-specific gains in investment banking, such as more accurate valuation and faster deal screening, as well as strategic advantages in private equity, including lower portfolio correlations and improved investment timing. These results collectively underscore the potential of categorical approaches to outperform traditional methods in complex financial environments.

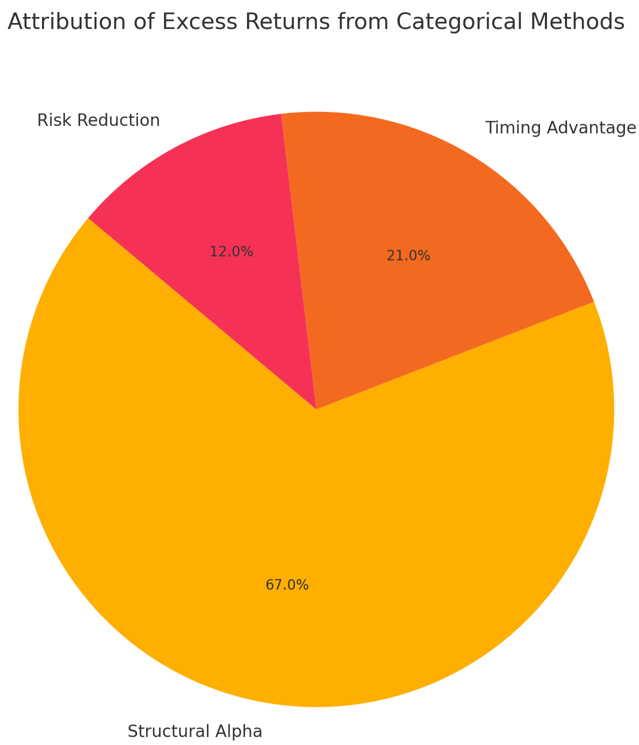

Figure 2 (Attribution of Excess Returns from Categorical Methods): It reveals the underlying sources of value creation attributed to categorical modeling in finance. The majority of excess returns—approximately two-thirds—are credited to improved structural representation of financial relationships, labeled as structural alpha. This is followed by more timely recognition of market regime shifts and modest but meaningful contributions from improved risk diversification. The pie chart visually reinforces the idea that the observed performance enhancements stem not from arbitrage or anomaly exploitation, but from more nuanced and structured understanding of financial systems enabled by category-theoretic reasoning.