

From Academic Factor Portfolios to Investable Factor Products: Return Differences, Mapping Losses, Implementation Frictions, and a Conditional Framework

Yanbo Lu*

School of Finance, Jiangxi University of Finance and Economics, Nanchang 330100, China

**Corresponding author: Yanbo Lu.*

Abstract

The academic factor premium comes from the “laboratory measurement” of an idealized long-short combination. For investors to actually gain factor exposure, they must first map academic signals into portfolioable data, and then execute the data through the product vehicle. This article organizes its overview along the three-layer filtering logic of “existence–persistence–availability” and summarizes a two-layer explanatory framework of combination mapping loss (short selling constraints, benchmark constraints, factor definition changes) and implementation delivery loss (transaction costs, rates, strategy capacity). The review found that: (1) the shrinkage system exists systematically, and existing major studies generally show a shrinkage ranking of momentum > value > market $\approx$ size (mainly based on US active mutual funds samples); (2) the shrinkage estimates of the same factor vary greatly in different studies, due to differences in three systematic methodologies: sample period, cost model and factor definition; (3) the impact of turnover rate and short selling contribution on the shrinkage magnitude has strong cross-documentary evidence, while the role of execution capability and constraint strength mainly comes from inferences from comparisons between different studies. Based on this, this paper proposes an inductive framework based on factor type $\times$ product carrier conditionalization. The contribution of this paper is not to provide a universally applicable reduction ratio, but to establish a conditional analytical framework that distinguishes between three levels of evidence: “original conclusions, cross-documentary consensus, and inferences from this paper,” in order to explain the discrepancies between studies, and to identify, in particular, the empirical gap in the measurement of factor capture rate in Smart Beta ETFs.

Keywords

factor investing, implementation costs, portfolio mapping, product carrier, transaction costs

1. Introduction

Over the past three decades, factor models have evolved into the core paradigm for asset pricing. The academic community has identified a series of common factors that explain the differences in cross-sectional

returns by constructing long-short portfolios [2]. Meanwhile, global Smart Beta ETFs, multi-factor mutual funds, and style index products collectively manage trillions of dollars in assets.

Academic factors are constructed under idealized conditions—long-short hedging, monthly rebalancing, and zero transaction costs—while investors need to go through two steps to obtain factor exposure: first, map academic signals into executable combinations (losses such as short-selling constraints and benchmark deviation limits), and then execute them through product carriers (additional losses such as transaction costs, rates, and capacity constraints).

Discussing “factor investability” requires a three-layer filtering logic. The first level: Existence – does the factor really exist statistically? Harvey, Liu, and Zhu (2016) [10] proposed a threshold of $t > 3.0$ under the framework of multiple hypothesis testing. Second layer: Persistence – do existing factors persist out of the sample and after publication? McLean and Pontiff (2016) [9] found that factor returns decayed by about 26% out of the sample and then decayed significantly after publication. The third layer: Availability – how much can investors actually receive after deducting implementation frictions? This is the core concern of this article.

To differentiate the strength of evidence, the entire text marks key arguments with three levels of evidence, which are used consistently in both the main text and tables: [Original Conclusion] refers to the direct findings of a single study with a specific sample and method; [Cross-literature Consensus] refers to the consistent support of at least two independent studies; [Inference of This Paper] refers to the author’s judgment based on literature summarization, which is yet to be tested. This distinction separates “a study found X” from “X is a universal law.”

Existing literature is scattered across three areas, factor construction, market microstructure, and fund performance, often leading to seemingly contradictory conclusions due to a lack of distinction between product carriers. Taking momentum as an example, Arnott, Kalesnik and Wu (2017) [1] deduced from the net asset value of US actively managed mutual funds that the momentum capture rate was only about 7% [Original Conclusion]. Patton and Weller (2020) [17] used an extended Fama-MacBeth direct estimate to find that the momentum gap was about 7%/year and the value gap was about 3%/year, while the market and size gaps were not significant [Original Conclusion]. Based on a real transaction record of an institution, Frazzini, Israel and Moskowitz (2012) [15] found that the actual transaction cost was only about one-tenth of the estimate of commonly used models [Original Conclusion]. The fact that estimates of the same factor but with vastly different orders of magnitude demonstrates that the implementation of discounts cannot be discussed in isolation from the factor type and the product carrier. Johansson, Sabbatucci and Tamoni (2025) [16] further decomposed the total gap into two parts: transaction friction and opportunity cost.

Together, these evidences point to a defensible conclusion: implementing discounts depends on factor definitions, portfolio mapping methods, product mechanics, asset size, net turnover and execution capabilities, and should not be reduced to fixed parameters. Among them, turnover rate and short-selling constraints have relatively direct evidence [4,15,17] [Cross-literature Consensus], while the impact of execution capabilities, constraint intensity, and some product mechanisms are more explanations proposed by this article based on differences between different studies [Inference of This Paper]. The contributions of this paper are: integrating scattered evidence into a progressive analytical framework of “existence–persistence–availability”; proposing a two-level inductive framework of combinatorial mapping loss and implementation delivery loss and providing operational definitions to explain the discrepancies among existing studies; and focusing on identifying the empirical gap in the measurement of factor capture rate of Smart Beta ETFs.

This paper includes 20 articles with three selection criteria: (1) Core articles must directly measure implementation losses and have complementary methods—net value back-calculation (Arnott et al. [1]; Patton and Weller [17]), real transaction records (Frazzini et al. [15]), simulated deductions (Novy-Marx and Velikov [4]; Li and Shim [5]), delayed execution (Dickerson et al. [7]) and ETF factor exposure (Glushkov [18]; Johansson et al. [16]), covering the main combination of “measurement method × product carrier”; (2) Supporting articles provide evidence of existence and persistence for the three-layer filtering (Harvey et al. [10]; McLean and Pontiff [9]), or provide evidence of mechanism and institutional background (factor crowding [19], fund flow [20], Active Share [8], factor definition [12,14], etc.); (3) Journal publications are preferred, and working papers are only included when they involve the latest ETF evidence. The sample is mainly from the US market, and the lack of cross-market evidence is a limitation of this review.

2. Analytical Framework: From Academic Signals to Investor Returns

2.1 Mapping Chain and Double-Layer Loss

The academic long-short portfolio is a “laboratory measurement” of factor premium. It does not answer “how much investors can gain,” but rather “whether this factor exists under ideal conditions.”

Step 1: Portfolio Mapping. This involves transforming academic long/short signals into a portfolio of investable targets. Major losses include: discarding short-selling signals (institutional losses for long-only products), benchmark tracking error caps, portfolio concentration and liquidity screening, and changes in factor definitions. Common forms include rule-based long-only constructions (Smart Beta ETFs) and actively managed constructions (where fund managers deviate from stock selection patterns under benchmark constraints).

Step 2: Product delivery. Standardized portfolios can be delivered through ETFs or institutional accounts (low fees, physical subscription and redemption, but not available for short selling); actively managed portfolios are mainly delivered through public funds (flexible stock selection but higher fees and stricter constraints [3]). Losses include bid-ask spreads, market shocks, delayed execution, management fees and cash drag, capacity constraints, and subscription and redemption frictions. Table 1 summarizes the complete mapping chain from academic signals to investor returns.

Table 1: Mapping Chain from Academic Signals to Investor Returns

Stage	Core Question	Main Sources of Loss / Decay
Academic Long-Short Signal	Does the factor exist under ideal conditions?	(Benchmark)
Portfolio Mapping	How is the long-short signal converted into an executable portfolio?	Short-selling constraints, benchmark constraints, position limits, changes in factor definition and weighting
Product Delivery	How is the portfolio executed through specific products?	Bid-ask spread / market impact / latency, fees / cash drag, capacity constraints, subscription/redemption frictions

2.2 Conceptual Matrix and Operational Definition of Two-Layer Loss

“Combined mapping loss” and “implementation and delivery loss” are analytical classifications in this paper, rather than mutually exclusive factual categories—several frictions (liquidity screening, cash holding, benchmark constraints) may act on both layers simultaneously, and mechanical summation will result in double counting. Table 2 provides the operational definitions, measurement indicators, applicable carriers, and representative literature for each of the main frictions.

Table 2: Conceptual Matrix of Two-Layer Losses

Type of Friction	Operational Definition	Measurement Indicator	Main Applicable Vehicles	Main Impact Layer	Representative Literature
Short-selling Constraints	Inability to construct the short leg; long-only signal loss	Contribution of the short leg to returns	ETFs, mutual funds	Portfolio Mapping	[1,17]
Benchmark / Tracking Error Constraints	Upper limit on deviation from benchmark compresses factor exposure	Tracking error, Active Share	Mutual funds	Portfolio Mapping	[6,8]
Changes in Factor Definition and Weighting	Academic construction converted to operational construction	Turnover, correlation with academic factor	ETFs, mutual funds	Portfolio Mapping	[12,14]
Liquidity Screening and Position Concentration	Exclusion of low-liquidity stocks; concentration limits set	Investable stock universe, position limits	ETFs, mutual funds	Related to both layers	[4,17]

Type of Friction	Operational Definition	Measurement Indicator	Main Applicable Vehicles	Main Impact Layer	Representative Literature
Rebalancing Frequency Adjustment	Reducing rebalancing frequency to control costs	Turnover, factor exposure drift	All	Related to both layers	[4,5]
Cash Holdings and Fund Flows	Cash drag, subscription/redemption impact	Cash ratio, fund flows	Mutual funds	Implementation / Delivery	[1,17]
Trading Costs (Spread / Impact / Latency)	Execution frictions	Effective spread, implementation shortfall	All	Implementation / Delivery	[4,5,7,15]
Management Fees and Expense Ratios	Product operating expenses	Expense ratio (%/year)	ETFs, mutual funds	Implementation / Delivery	[3,16]
Strategy Capacity	Break-even scale at which net returns approach zero	Break-even AUM	All	Implementation / Delivery	[4,15]

The “Main Impact Layer” column in Table 2 allows for “related to both levels” precisely to handle overlapping categories. For example, liquidity screening narrows the range of investable assets during the mapping phase and raises the impact cost during the execution phase. For this type of cross-layer friction, it should be classified into the side of direct measurement, and the cross-effect should be declared when interpreting it.

There may be a two-way interaction between the two levels [Inference of This Paper]: tighter tracking error constraints may lead to more frequent rebalancing and higher costs; conversely, investors facing high costs may strategically reduce the frequency of rebalancing, which may further dilute factor exposure. Therefore, the two levels should not be simply summed up. Johansson et al. (2025) [16] provided a direct measurement of this: the total gap was broken down into transaction costs and short-selling fees (58%) and opportunity costs (42%), the latter corresponding to the residual of incomplete combination mapping [Original Conclusion]. This decomposition also reveals asymmetry between factors: for momentum, transaction costs dominate; for value, the two types of losses have more equal weights. This is consistent with the relative structure of the two layers in this paper, but the definition of the two types does not completely overlap with the definition of the two layers in this paper [Inference of This Paper].

3. Classification of Evidence

3.1 Combinatorial Mapping Loss

The institutional loss of short-selling signals. Both Smart Beta ETFs and actively managed mutual funds are subject to short-selling restrictions—the index methodology of ETFs is long-only by definition. Arnott et al. [1] estimated that the short-selling side contributes about 50% of the value factor and about 40% of the total academic returns [Original Conclusion]. Patton and Weller [17] arrived at a consistent estimate using different methods: about half of the momentum gap comes from short-selling constraints, and removing micro-cap stocks only explains about 1% of the gap—the common “investability correction” is far from enough to eliminate the gap [Cross-literature Consensus]. The short selling side of some factors (profit, investment) cannot be effectively replicated even through ETFs [16].

The size factor aligns with the mechanism. The size factor has minimal dependence on short selling—exposure can be obtained simply by buying small-cap stocks and under-allocating to large-cap stocks. Both Arnott et al. [1] and Patton and Weller [17] found that the implementation gap was almost zero, which is consistent with the expectation that “the combined mapping loss is close to zero when both the short-selling contribution and turnover rate are low” [Cross-literature Consensus].

Benchmark constraints and factor exposure dilution. Actively managed mutual funds face tracking error constraints, portfolio concentration limits, and professional risks—excessive deviation from the benchmark

can lead to redemption pressure. Cremers and Petajisto (2009) [8]’s Active Share is the standard tool for characterizing this loss, but Cline and Gilstrap (2021) [6] revealed its distortion: since the publication of the metric, low-skilled fund managers have strategically increased their active shares to “disguise” active management—high active shares only bring positive excess returns when combined with real stock-picking ability.

Changes in factor definitions. HML is updated annually [2], while many Smart Beta indices use monthly updated HMLm [12]—minor differences in academic definitions correspond to substantially different turnover rates and cost structures. No research has yet systematically quantified this mapping relationship of “definition differences translating into implementation losses” [Inference of This Paper].

3.2 Implementation and Delivery Losses a: Transaction Costs

The empirical criticality of turnover rate. Novy-Marx and Velikov (2015) [4] simulated net income after deduction in 23 anomaly strategies and found that most of the monthly turnover rates below 50% were still significantly positive, while most of the monthly turnover rates above 50% were no longer significant [Original Conclusion]. This is an empirical rule under specific strategy sets and cost assumptions rather than a universal threshold, but the finding that “turnover rate is the strongest single predictor of net income after deductions” is consistent with cross-document results showing that momentum (highest turnover) has the largest difference and scale (lowest turnover) has a near-zero difference [Cross-literature Consensus]. The additional transaction costs of equal weighting usually outweigh the additional benefits—academic equal-weighted backtesting cannot be directly used as expected returns.

Sensitivity of the range of transaction cost assumptions. From bid-ask spreads alone [4] to market shock models [5] to delayed execution [7], the choice of the range of assumptions directly affects the net return estimate. Frazzini et al. (2012) [15] provided a key empirical benchmark: the actual transaction costs of institutions are about one-tenth of those estimated by commonly used price data models. This does not mean that transaction costs are unimportant, but rather that pure price data models may systematically overestimate costs—institutions will strategically avoid high-cost transactions, and the model cannot capture this behavioral choice.

The pervasiveness of delay costs across assets. Dickerson, Nozawa, and Roboti (2024) [7] found that the efficiency factor in the corporate bond market decreased from 51 to almost zero after deducting delay costs. Delay costs are proportional to the profitability of the strategy—the more profitable the signal, the higher the opportunity cost of delaying execution, and this cannot be eliminated even when restricted to the most liquid subset of bonds. This suggests that implementation losses are prevalent in markets where “signal changes outpace trade execution.”

The correct interpretation by Li and Shim (2019) [5]. Their report showed that the cost of momentum trading exceeded the estimated excess returns, but this was a capacity stress test under a specific scale and cost model, rather than proof that “momentum is not feasible.” Their more important finding was that the turnover rate of multi-factor portfolios was significantly lower than the simple average of individual factors: the factor signals were negatively correlated and the trades canceled each other out—echoing the evidence from Frazzini et al. [15] that multi-factor portfolios break even far more than pure momentum strategies [Cross-literature Consensus].

Cost advantage of shorting ETFs. Johansson et al. [16] demonstrated an important mechanism innovation: using ETFs to short the short side of factors is much cheaper than directly shorting individual stocks. After deducting fees, the differences between factors are generally narrowed, but the short side of some factors cannot be effectively replicated even at the ETF level.

3.3 Implementation and Delivery Losses B: Product Mechanisms and Investment Constraints

Methodological sensitivity of the market beta factor. Arnott et al. [1] estimated that the market factor shrinks by about 51%, but Patton and Weller [17] found that the implementation gap of the market factor is not statistically or economically significant—mutual funds can obtain market exposure with compensation close to the paper level. The difference stems from methodology: the former classifies cash positions and low

beta values as “shrinkage,” while the latter separates factor compensation from residuals. More generally, different loss classification boundaries will change the shrinkage value of the same factor, and any shrinkage estimate should state the loss attribution criterion [Inference of This Paper].

Strategy capacity is an endogenous variable. Novy-Marx and Velikov [4] estimated the momentum capacity to be only about \$5 billion, while Frazzini et al. [15] estimated the break-even size using real trading data to be one to two orders of magnitude larger. The difference mainly comes from the fact that the actual trading costs of institutions are much lower than the model estimates; the actual turnover demand of long-short combinations after net settlement is much lower than the gross turnover rate. However, the latter is calibrated with data from a highly efficient institution, and extrapolation should be cautious [Inference of This Paper]. Capacity is a function of construction method and execution capability, rather than an inherent property of the factor.

Fund heterogeneity: Average shrinkage masks important differences. Patton and Weller [17] found that funds with higher factor exposure replication ratios have factor compensation that is closer to academic level but still has gaps; the momentum compensation of small funds is several percentage points/year higher than that of large funds. High-efficiency institutions belong to the high-efficiency end of the distribution, and extrapolating to typical investors will be misleading [Inference of This Paper].

Equilibrium Perspective and Factor Crowding. Patton and Weller [17] found that implementation costs did not show a significant downward trend between 1970 and 2016, despite a significant decrease in bid-ask spreads and commissions. The equilibrium explanation is that improved trading techniques reduce the cost of the first dollar, and the increase in perceived alpha attracts capital inflows, resulting in a synchronous increase in non-proportional costs such as price shocks, while the average dollar cost remains basically unchanged [Inference of This Paper]. The consequences of crowding vary depending on the factor type [Original Conclusion]: Baltas (2019) [19] measured crowding by the contemporaneous pairwise correlation of factor-adjusted residual returns and found that after a period of high crowding, divergent premiums such as momentum significantly underperformed, while convergent premiums such as value outperformed. Factors lacking fundamental anchoring were pushed away from fundamentals by funds, while factors with natural anchoring accelerated valuation convergence with inflows. Akbas et al. (2015) [20] provided complementary evidence of fund flows: retail funds disproportionately flowed into overvalued growth and accrual stocks, while institutional funds reversed course by shorting to correct pricing, but arbitrage capital was limited— anomalies and implementation costs could therefore coexist in equilibrium [Original Conclusion].

3.4 Conditional Shrinkage of Factor Type × Product Carrier

Core principle: Any single “shrinkage ratio” is a point estimate under a specific factor type, product carrier, sample period, and methodology, and should not be cited out of context. Table 3 summarizes the evidence for conditional shrinkage by factor type and product carrier.

Table 3: Conditional Shrinkage Evidence by Factor Type × Product Vehicle (Two measurement approaches presented side by side; not directly comparable horizontally)

Factor	Monthly Turnover	Arnott et al. [1]: Shrinkage Ratio of Capture Rate Inferred from NAV	Patton & Weller [17]: Directly Estimated Annualized Return Gap λ^{Δ}	Dominant Loss Channel	Optimization Potential
Momentum	>100%	~93%	Approx. 7.2–7.3%/year	Mainly trading costs, followed by short-leg losses [Original Conclusion]	High — multi-factor netting, trade optimization, and lower rebalancing frequency can compress losses [Inference of This Paper]
Value	30–40%	~39%	Approx. 2.6–3.8%/year	Short leg and trading costs roughly equally responsible	Medium — optimizing factor definition (e.g., HML vs. monthly-updated

Factor	Monthly Turnover	Arnott et al. [1]: Shrinkage Ratio of Capture Rate Inferred from NAV	Patton & Weller [17]: Directly Estimated Annualized Return Gap λ^{Δ}	Dominant Loss Channel	Optimization Potential
Market β	<5%	~51% (methodological controversy)	Insignificant	[Inference of This Paper] Near zero [17] or investment constraints [1] — depends on measurement method [Original Conclusion]	HMLm) can reduce turnover [Inference of This Paper] Low — losses mainly at the product mechanism level [Inference of This Paper]
Size	15–20%	~0%	Insignificant	Lowest short-leg dependence and lowest turnover [Cross-literature Consensus]	Almost unnecessary [Cross-literature Consensus]
Profitability / Quality	To be estimated	No direct reverse-engineered estimate	No independent estimate	Inferred that unavailability of the short leg is the main obstacle [Inference of This Paper]	To be studied

Table 3 retains the core estimates from two independent methods to highlight cross-document comparisons, but the two columns have different measurement standards and cannot be directly compared horizontally: the Arnott et al. column is “Capture Rate/Reduction Ratio”—the gap between the proportion of factor compensation actually captured by the fund and the academic portfolio's return; the Patton & Weller column is “Annualized Return Gap”—the directly estimated annual return difference (λ^{Δ}) between the fund's factor compensation and the academic benchmark. The numerical meanings, estimation methods, and samples of the two are different; in addition, neither column includes measures of transaction cost studies (such as simulated fees, real transaction records). The columns “Dominant Loss Channel” and “Optimization Space” are marked with three levels of evidence: only the momentum “transaction cost as the main factor” has direct evidence from the original text (e.g., Johansson et al. [16] decomposed the total gap into 58% transaction friction and 42% opportunity cost), but this decomposition is not equivalent to the two levels of “implementation delivery loss” and “combination mapping loss” defined in this paper; the rest are inferences based on inter-study comparisons. Supplementary evidence such as capacity and ETF fee deduction improvement can be found in §3.2–3.3.

This paper derives a conditional relationship from the above evidence—this is an explanatory framework to be tested, rather than a verified functional form: the magnitude of shrinkage may vary with factor turnover rate, short-selling contribution, constraint strength, and execution capability [Inference of This Paper]. Turnover rate is the single best predictor [4] but not sufficient—the contribution of short selling helps explain the phenomenon that the value factor is highly dependent on short selling but shrinks less than momentum (low turnover may compensate for short selling losses), while investment constraints are consistent with the evidence that the market beta turnover rate is extremely low but still shrinks controversially.

4. Consensus, Controversy, and Methodological Evaluation

4.1 Established Consensus

The following judgments have robust support across research and methods and can be regarded as [Cross-literature Consensus]: (1) The implementation of shrinkage system exists [1,4,7,17]; (2) Existing major studies show that the shrinkage ranking is momentum > value > market $\approx$ size $\approx$ 0, and the direction is consistent in the three independent estimates [1,4,17], but this ranking is based on US active public funds as the main sample, and its promotion to other carriers and markets needs to be tested; (3) Turnover rate is the strongest single

predictor of net income after deduction [4,17], and academic equal weight backtest income cannot be directly used as expected income [4]; (4) Existing studies provide support for the equilibrium explanation—Patton and Weller [17] found that the implementation cost did not decrease with the improvement of technology, and Baltas [19] found that the income after crowding differentiated with the factor type. The two are related but not completely the same, and are not enough to establish a specific equilibrium mechanism as a fact [Inference of This Paper]; (5) Unified implementation of multiple factors is better than summing them separately [5,15].

4.2 Current Status and Gaps in the Integration Framework

The three-layer filtering framework has not yet been integrated into a unified evaluation model, and the inter-layer connections are weak links. The connection from existence to persistence has been provided by McLean and Pontiff [9], but the mean decay was not decomposed according to factor type—the decay of the arbitrage factor should be the most significant, a hypothesis that has not yet been systematically tested [Inference of This Paper]. The connection from persistence to availability is the weakest: existing implementation loss studies use “appearance in the literature” as the factor selection criterion, rather than “passing the existence and persistence filters.”

The evidence gap for Smart Beta ETFs. Almost all existing evidence comes from reverse engineering of mutual funds or simulated fee deductions of academic portfolios. ETFs differ from active funds in key dimensions—daily holdings disclosure, low fees, and in-kind creation and redemption—so there is reason to expect their factor capture rate to be better than that of active funds, but this expectation still needs to be systematically tested. Scattered preliminary evidence includes: Glushkov (2016) [18] found that many Smart Beta ETFs do load target factors, but their excess returns relative to the benchmark are not consistent after fee deduction; Johansson et al. [16] used the short end of ETFs as another starting point. What is really missing is a reusable measurement of ETF factor capture rate that connects “factor exposure—fees—turnover—net return.” This is one of the empirical gaps that this paper focuses on identifying. Table 4 compares the sample and methodological boundaries of the main existing studies.

Table 4: Comparison of Samples and Methods in Major Implementation Cost Studies

Research	Method	Data / Sample	Cost Assumptions	Investor / Vehicle	Core Estimates
Arnott et al. [1]	NAV reverse-engineering + Fama-MacBeth	5,323 U.S. funds, 1990–2016	No explicit cost model	Active mutual funds	Momentum capture rate ~7%, Value ~61%
Patton & Weller [17]	Extended Fama-MacBeth	4,267 U.S. funds, 1970–2016	No explicit cost model (direct gap measurement)	Active mutual funds	Momentum $\lambda^{\Delta} \approx 7.2\%$, Value ≈ 2.6 – 3.8%
Frazzini et al. [15]	Real trading records	\$721B trades, 1998–2011, 19 markets	Actual spreads + impact	Large long-short institutions	Costs about 1/10 of TAQ; Momentum break-even \$52B–\$122B
Novy-Marx & Velikov [4]	Simulated cost deduction	23 anomalies, 1963–2013	Bayesian effective spreads	Academic portfolios	Most with turnover >50% become insignificant after costs
Li & Shim [5]	Simulated cost deduction	Multi-factor Smart Beta	Market impact model	Institutional portfolios	Momentum cost 2.77%; multi-factor turnover 55.7%
Dickerson et al. [7]	Delayed execution	Corporate bond factors	Delay costs	Academic portfolios	Effective factors drop from 51 to near 0
Johansson et al. [16]	ETF + MF tradeable factors	Institutional + retail	Actual expense ratios + shorting costs	ETFs, mutual funds	Trading frictions explain 58% of the total gap
Glushkov [18]	Holdings-based factor exposures	Smart Beta ETF	Actual expense ratios	ETFs	Relative performance after fees is inconsistent

4.3 Three Systemic Sources of Discrepancies in Literature Conclusions

Most seemingly contradictory conclusions can be traced back to methodological differences rather than contradictions in the facts themselves. (1) Sample period selection: Factor returns are significantly time-varying—momentum shrinkage is concentrated after 2002 [1], which coincides with the time when momentum strategies were widely adopted. Any single estimate is embedded with the cyclical conditions unique to the sample period. (2) Assumption range of transaction cost model: Effective spread [4], market shock model [5], delayed execution [7] and real record [15] correspond to different investor types and execution assumptions, rather than right or wrong. (3) Factor definition standard: The same factor label can point to different constructions—HML vs. HMLm [12], OP vs. RMWCP [14]—small differences in definition have a substantial impact on turnover and cost in implementation. Three sources complement each other: the sample period determines the representativeness of historical mean, the cost model determines the closeness of simulation to reality, and the factor definition determines the “translation loss” between academic benchmarks and portfolios.

5. Conclusions

5.1 Key Conclusions

Based on the reliability of the evidence, the conclusions of this article can be divided into three levels. Each level has a different basis of evidence and should be treated differently when citing them.

Robust conclusions across multiple independent methods: The systematic existence of the implementation of the shrinkage system—net value back-calculation [1,17], simulated deduction [4,5], delayed execution [7] and real transaction records [15]—are supported in a consistent direction from different methods.

The conclusions are consistent in direction but have limited evidence: the shrinkage ranking (momentum > value > market $\approx$ size $\approx$ 0) is consistent in direction among the three independent estimates in [1,4,17], but the sample is mainly US active public funds; turnover rate is the strongest single predictor of net income after deducting fees, supported by the complementary research of the two methods in [4,17]; the evidence for the unified execution of multi-factor is better than the sum of their respective methods is weak, mainly from the two studies in [5,15], and should be regarded as a signal of strong consistency rather than a conclusion.

The judgment still depends on the method setting: any specific shrinkage figure is embedded in the triple setting of sample period, cost model and factor definition—momentum shrinkage is concentrated after 2002 [1], the actual cost of the institution can be an order of magnitude lower than the model estimate [15], and the factor label can point to substantially different constructions [12,14]. Therefore, the two-layer inductive framework proposed in this paper, which is “shrinkage with factor turnover rate, short-selling contribution, constraint strength and execution capability conditionalization,” should be understood as a tool for organizing existing evidence and explaining the discrepancies between studies. Its functional form and the two-layer interaction relationship need to be systematically tested [Inference in This Paper].

5.2 Application Implications and Future Directions

Discounts are expected to be applied based on factor type $\times$ product carrier differentiation—existing evidence is consistent with the expectation that “the expected discount for low-turnover factors is lower than that for high-turnover factors” [4,17]; unified multi-factor implementation is better than separate summation [5,15]; strategy capacity is the core constraint, especially for divergent factors such as momentum [17,19]. As for whether low-fee carriers such as ETFs systematically provide higher net capture rates, existing literature is insufficient to establish a general conclusion, and it can only be regarded as a reasonable expectation that needs to be tested.

The next verifiable steps include: P1 (highest priority)—leveraging the data advantage of publicly available daily holdings and known fees of ETFs, extending the methods of [1] or [17] to Smart Beta ETFs to directly measure their factor capture rate (responding to the preliminary evidence from Glushkov [18]); P2—performing cross-factor matching between the factor-level decay of McLean and Pontiff [9] and the shrinkage estimates of [1] or [17], connecting the persistence estimate and the availability estimate; P3—quantifying the impact of differences in factor definitions (HML vs. HMLm [12], OP vs. RMWCP [14]) on turnover and net real-money returns. Of the three, P1 directly fills the main gap in the existing evidence at the product carrier

level, while P2 and P3 respectively test the interlayer connections and factor definitions in this paper's framework, which are still inferences.

References

- [1] Arnott, R. D., Kalesnik, V., & Wu, L. (2017). The Incredible Shrinking Factor Return. Working Paper, Research Affiliates.
- [2] 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.
- [3] Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance*, 52(1), 57–82.
- [4] Novy-Marx, R., & Velikov, M. (2015). A Taxonomy of Anomalies and Their Trading Costs. *Review of Financial Studies*, 29(1), 104–147.
- [5] Li, F., & Shim, J. (2019). Trade-Off in Multifactor Smart Beta Investing: Factor Premium and Implementation Cost. *Journal of Portfolio Management*, 45(3), 115–124.
- [6] Cline, B. N., & Gilstrap, C. (2021). Active Share: A Blessing and a Curse. *Journal of Financial Research*, 44(2), 365–394.
- [7] Dickerson, A., Nozawa, Y., & Robotti, C. (2024). Factor Investing with Delays. Working Paper.
- [8] Cremers, K. J. M., & Petajisto, A. (2009). How Active Is Your Fund Manager? A New Measure That Predicts Performance. *Review of Financial Studies*, 22(9), 3329–3365.
- [9] McLean, R. D., & Pontiff, J. (2016). Does Academic Research Destroy Stock Return Predictability? *Journal of Finance*, 71(1), 5–32.
- [10] Harvey, C. R., Liu, Y., & Zhu, H. (2016). ... and the Cross-Section of Expected Returns. *Review of Financial Studies*, 29(1), 5–68.
- [11] Asness, C. S., Chandra, S., Iltanen, A., & Israel, R. (2017). Contrarian Factor Timing is Deceptively Difficult. *Journal of Portfolio Management*, 43(5), 72–87.
- [12] Asness, C. S., & Frazzini, A. (2013). The Devil in HML's Details. *Journal of Portfolio Management*, 39(4), 49–68.
- [13] Fama, E. F., & French, K. R. (2015). A Five-Factor Asset Pricing Model. *Journal of Financial Economics*, 116(1), 1–22.
- [14] Fama, E. F., & French, K. R. (2018). Choosing Factors. *Journal of Financial Economics*, 128(2), 234–252.
- [15] Frazzini, A., Israel, R., & Moskowitz, T. J. (2012). Trading Costs of Asset Pricing Anomalies. Working Paper, AQR Capital Management.
- [16] Johansson, T., Sabbatucci, R., & Tamoni, A. (2025). Tradable Risk Factors for Institutional and Retail Investors. *Review of Finance*, forthcoming.
- [17] Patton, A. J., & Weller, B. M. (2020). What You See Is Not What You Get: The Costs of Trading Market Anomalies. *Journal of Financial Economics*, 137(2), 515–549.
- [18] Glushkov, D. (2016). How Smart Are "Smart Beta" ETFs? Analysis of Relative Performance and Factor Exposure. SSRN Working Paper.
- [19] Baltas, N. (2019). The Impact of Crowding in Alternative Risk Premia Investing. *Financial Analysts Journal*, 75(3), 89–104.
- [20] Akbas, F., Armstrong, W. J., Sorescu, S., & Subrahmanyam, A. (2015). Smart Money, Dumb Money, and Capital Market Anomalies. *Journal of Financial Economics*, 118(2), 355–382.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Open Access

This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

