Predictive Physical Tail Width and the Cross-Section of Returns

Nima Taheri Hosseinkhani
Rutgers Business School
nt612@rutgers.edu
ORCID: 0009-0007-5564-7839

Abstract

Firms differ in the width of their ex ante physical return distributions. I measure it with Physical Tail Uncertainty (PTU), a fixed-rule rank of out-of-sample 5-to-95 percentile predictive return width. PTU forecasts next-month realized kurtosis and high-kurtosis states beyond volatility, historical tails, lottery measures, expected skewness, illiquidity, and firm controls. It contains residual physical-tail information beyond backward-looking benchmarks and remains informative relative to broad option-surface proxies. Option tests separate broad option-surface coverage from the stricter exact raw-chain BKM benchmark. PTU earns positive alphas that survive investability screens and are strongest where tail exposure is harder to absorb.

Keywords: physical tail uncertainty; predictive tail width; realized kurtosis; option-implied moments; cross-section of returns; limits to arbitrage.
JEL classification: G12, G13, G17, C45, C58.

Disclosure Statement

Nima Taheri Hosseinkhani: I have no conflicts of interest to disclose. I received no external funding for this project.

Introduction

Investors care about expected returns and volatility, and also about the shape of the physical return distribution they will face next month. Two firms can have similar expected returns and similar volatilities while exposing investors to very different realized tail states. These physical tails matter for risk management, delegated portfolio constraints, intermediary balance sheets, hedging demand, and the willingness of investors to hold concentrated positions.

Existing measures capture important but different objects. Realized higher moments measure tails after uncertainty resolves. Lottery variables, idiosyncratic volatility, MAX, and expected skewness capture specific features of past or predicted payoff asymmetry [Ang et al., 2006, Boyer et al., 2010, Bali et al., 2011]. Option-implied moments summarize risk-neutral distributions for option-covered firms [Bakshi et al., 2003, Conrad et al., 2013]. The missing object is a broad-universe, ex ante measure of physical predictive tail width: before a month begins, which firms are predicted to have unusually wide realized-return tails?

This paper constructs that object. I refer to the fixed-rule rank of predictive physical tail width as Physical Tail Uncertainty, or PTU. Four quantile-prediction methods estimate out-of-sample return quantiles. For each estimator, W90 is the predicted 95th percentile minus the predicted 5th percentile. Each W90 signal is adjusted within month and rank-normalized. PTU is the equal-weight average of the four adjusted W90 ranks, requires at least three available inputs, and is rank-normalized again within month. A high PTU value therefore means that a stock's predicted physical 5-to-95 interval is wide relative to other stocks in the same month after the fixed adjustment procedure.

A two-sided predictive width is economically relevant because many frictions respond to the width of the physical payoff distribution, not only to one signed tail. Wide physical tails raise margin needs, hedging demand, rebalancing risk, delegated portfolio concerns, and intermediary inventory risk. When these exposures are difficult to hedge or intermediate, investors require compensation for holding stocks whose future physical return distribution is unusually wide. W90 supplies the ex ante physical-width measurement; the pricing sign comes from the stochastic-discount-factor and friction channel.

The construction is deliberately not an optimized ensemble. The equal-weight rule makes PTU a fixed characteristic rather than a fitted trading rule, and no return, realized-kurtosis, or portfolio-performance outcome enters the aggregation weights. The four estimators are measurement devices. If linear, penalized, tree-based, and transformer-based quantile methods point to the same relative tail-width ranking, the economic object is the common physical tail-width phenomenon, not any one model category.

The first result validates PTU on the physical distribution before studying returns. Table 3 shows that PTU forecasts next-month realized physical kurtosis. The PTU coefficient is 0.367 with a t-statistic of 7.89 after size, book-to-market, momentum, realized volatility, and idiosyncratic volatility enter, and it remains precisely estimated in expanded non-option specifications that include historical physical-tail measures, BMV expected skewness, MAX-style lottery controls, and Amihud illiquidity. Tables 4 and 5 translate the continuous result into high-kurtosis states. The evidence therefore identifies forward-looking physical-tail information rather than merely high-volatility firms.

The second result addresses whether PTU is a persistent firm type or a repackaging of backward-looking tail measures. Table 6 shows that firm-demeaned PTU predicts future realized kurtosis and that changes in PTU predict changes in future realized kurtosis. Table 7 shows that residual PTU remains informative after transparent physical-tail forecasts are removed. These tests matter because historical kurtosis, volatility, and MAX are strong benchmarks in their own right. PTU adds information alongside them; it does not replace them.

The option evidence disciplines the interpretation through benchmark separation. In broad option-surface proxy samples, PTU remains informative after option-implied proxy controls enter. Exact raw-chain BKM provides the stricter option-chain benchmark in a much smaller liquid-option universe. This division of labor is central: PTU measures ex ante physical tail width across the broad equity universe, while exact BKM summarizes risk-neutral option information where liquid option-chain data are available.

The return evidence asks whether the validated physical-tail characteristic earns compensation. Table 12 shows that the direct PTU high-minus-low sort earns an FF6 alpha of 3.07 percent per month with a t-statistic of 9.75. The characteristic-managed implementation and the long-only high-μ, high-PTU implementation also earn positive factor-adjusted returns, but they answer implementation questions rather than serving as direct PTU tests. Table 13 adds the necessary discipline: value weighting and price, size, and liquidity screens reduce the alpha and identify where the economics is strongest, while the premium remains positive under stricter implementation screens.

The mechanism evidence locates where the premium is largest. Table 14 groups the pricing relation by information production, institutional risk-bearing, liquidity, monitoring, and hedging-market availability. The largest premia appear among firms with low analyst coverage, low institutional ownership, high Amihud illiquidity, low short interest, no option coverage, and no IV30 availability. The short-interest result is interpreted cautiously because short interest also reflects disagreement, bad news, overvaluation, hedging demand, and short-sale constraints. The broader pattern is consistent with compensation for costly-

to-bear physical tail exposure: PTU earns larger premia where information, liquidity, monitoring, and hedging-market access are thinner.

The contribution is a finance measurement-and-pricing result. PTU is a broad-universe, fixed-rule rank of predicted physical return width. It forecasts future realized physical tail states beyond volatility and historical-tail benchmarks, remains informative relative to broad option-surface proxies, and is benchmarked against exact raw-chain BKM in the selected liquid-option universe. High-PTU stocks earn positive premia, including under investability screens, and the premia are largest where information, liquidity, monitoring, institutional risk-bearing, and hedging-market capacity are thinner. Flexible prediction enters as measurement technology; the economic claim is that predictive physical tail width is an economically meaningful cross-sectional characteristic.

I Literature and Contribution

The first related literature measures higher moments from option prices. Bakshi et al. [2003] provide model-free option-implied variance, skewness, and kurtosis measures and show that option prices contain rich information about the risk-neutral distribution. Subsequent work studies variance premia, higher-moment risks, option-implied skewness, and option-implied kurtosis in asset pricing [Bollerslev et al., 2009, Conrad et al., 2013, Chang et al., 2013]. This paper shares that economic subject matter and measures a different object. PTU is a physical predictive tail-width characteristic. It is estimated from out-of-sample conditional quantiles of realized stock returns and evaluated against future realized physical kurtosis. Option-implied variables enter the analysis as benchmarks and spanning variables. The empirical distinction matters because an option-implied moment is simultaneously a forecast object, a risk-price object, and an option-market coverage object; PTU is designed to isolate the physical predictive component before asking whether it is priced.

The physical-versus-risk-neutral distinction is central. A risk-neutral moment is a state-price-weighted object. If tail payoffs covary with the stochastic discount factor, the option-implied moment differs from the physical moment through the pricing-kernel covariance term. That difference is not a market inefficiency; it is the standard distinction between physical and risk-neutral probabilities. The empirical issue is whether a forward-looking physical tail-width signal adds information about future realized physical tails relative to option-implied benchmarks in the samples where option-implied moments are observed. The evidence separates the two option comparisons: PTU remains positive in broad option-surface proxy spanning tests, while exact raw-chain BKM identifies the tighter benchmark inside a narrower liquid-option overlap. This sample discipline lets the paper use option-

implied moments as serious competitors while keeping the option-overlap exercise distinct from the full-universe tests.

This positioning gives PTU a distinct role from an option-implied factor. PTU complements option-implied moments by measuring a forward-looking physical characteristic that risk-neutral variables do not mechanically subsume in every sample. Many empirical asset-pricing applications require a stock-month signal for the broad CRSP/CZ universe, while option-implied moments are naturally concentrated in the option-covered universe. PTU therefore provides a physical predictive object that can be tested in the full universe and then benchmarked against options where options are observed.

The second related literature studies physical higher moments, expected skewness, lottery demand, and realized distributional characteristics. Boyer et al. [2010] estimate expected idiosyncratic skewness and connect it to returns. Bali et al. [2011] show that maximum daily returns capture lottery-like demand, while Ang et al. [2006] document low average returns for high idiosyncratic-volatility stocks. These papers show that the cross-section prices dimensions of the return distribution beyond beta and volatility. PTU differs from these characteristics in timing and content. It is forward-looking rather than realized, interval-width based rather than maximum-return based, and targeted to future realized physical tail shape rather than to expected returns alone. For this reason, the empirical design includes realized volatility, idiosyncratic volatility, historical skewness and kurtosis, BMV expected skewness, and MAX-style variables as benchmarks. The paper also uses Amihud illiquidity, analyst coverage, institutional ownership, and option availability as mechanism variables because they help describe environments in which information production, trading capacity, and hedging capacity plausibly differ [Amihud, 2002, Diether et al., 2002].

The third related literature uses flexible prediction methods in empirical asset pricing. Gu et al. [2020] show that machine-learning methods improve return prediction, Freyberger et al. [2020] estimate characteristic-return relations nonparametrically, Kelly et al. [2019] connect characteristics to conditional risk exposures, and Giglio et al. [2022] review machine learning in asset pricing. PTU uses flexible prediction in a focused measurement role. The models estimate conditional return quantiles, and the paper extracts a transparent feature of the predictive distribution: the 5-to-95 percent tail width. The quantile-regression foundation follows the logic of Koenker and Bassett [1978]. The tree and transformer implementations draw on modern tabular prediction methods [Ke et al., 2017, Gorishniy et al., 2021], and the penalized quantile-regression implementation follows the high-dimensional quantile-regression logic of Belloni and Chernozhukov [2011]. The multi-estimator implementation provides a way to check whether the signal is a general feature of the predictive distribution across distinct estimation approaches.

The distinction between prediction and measurement is important for contribution discipline. A return-forecasting paper would naturally be judged by out-of-sample return prediction and trading performance. This paper first evaluates the predicted distribution against the future realized physical distribution. Return premia enter after the physical-tail validation. That sequence makes the machine-learning component interpretable: the models are used to extract a tail-width characteristic, and the characteristic is then subjected to the same kinds of cross-sectional regressions, spanning tests, and portfolio tests used elsewhere in empirical asset pricing.

This design shapes the contribution. PTU avoids a one-algorithm model-selection design by using the equal-weight common signal across four distinct estimators. The estimator-specific evidence is also economically important: IA Table IA.7 shows that each estimator's W90 signal earns a positive high-minus-low alpha, with FF6 alpha t-statistics ranging from 8.64 to 9.72. The estimator-specific tables preserve variation across modeling approaches and support the interpretation that PTU captures a stable physical tail-width component. The use of the Chen–Zimmermann signal library also connects the implementation to the recent open-source asset-pricing data infrastructure [Chen and Zimmermann, 2022]; the paper treats the consensus and disagreement variables derived from that library as auxiliary characteristics separate from the main contribution.

The fourth related literature studies limits to arbitrage, costly risk bearing, and intermediary constraints. Shleifer and Vishny [1997] and Gromb and Vayanos [2010] emphasize that arbitrage is limited when capital, horizon, and risk-bearing constraints bind. He and Krishnamurthy [2013] and Adrian et al. [2014] link intermediary conditions to asset prices. The mechanism tests in this paper are motivated by that logic. If PTU captures adverse physical tail uncertainty that is costly to hedge or intermediate, the premium should be larger where information production, liquidity, monitoring, or hedging-market access is thinner. The mechanism evidence supports the predicted direction in all six split dimensions, with the strongest return groups in low analyst coverage, low institutional ownership, high illiquidity, low short interest, no option coverage, and no IV30 availability. The short-interest split is interpreted through monitoring: low short interest is consistent with lower short-seller attention, and the broader short-interest literature also links shorting to disagreement, bad news, overvaluation, hedging demand, and short-sale constraints. These mechanism tests complement the main prediction and pricing evidence by locating the premium where the proposed economic friction has the most natural bite.

The contribution can therefore be stated precisely. Prior work measures risk-neutral higher moments, realized and expected physical higher moments, lottery-like payoffs, and flexible expected-return predictors. This paper measures a forward-looking physical tail-

width characteristic, validates it against future realized physical kurtosis, shows that it contains incremental information beyond historical physical-tail measures, evaluates option-implied spanning in labeled option-overlap samples, and documents return premia consistent with compensation for difficult-to-bear physical tail uncertainty. The novelty is the disciplined construction of a common physical tail-width object and the sequence of tests linking that object to realized tail shape, benchmark spanning, pricing, and mechanism heterogeneity; machine learning supplies the measurement technology while the contribution is the physical tail-width object and its empirical validation.

II Data and Measure Construction

A Sample, Timing, and Outcomes

The empirical unit is a stock-month. Accounting variables, analyst variables, institutional ownership, short interest, option-implied variables, and market variables are aligned using information available no later than the signal month, with standard reporting lags applied where applicable. The analysis panel contains 3,216,583 stock-month observations across 744 calendar months, from January 1963 through December 2024, with no duplicate firm-month keys. The out-of-sample predictive tests begin in 1985, after the training burn-in and subject to the availability of the required predictors, return outcomes, and controls. The core outcome is next-month realized physical kurtosis. Because the outcome is realized after the signal date, the timing is predictive: characteristics and quantile forecasts are measured at or before month t, while realized tail shape is measured in month $t+1$.

Realized physical moments are constructed from CRSP daily returns. For each firm-month, the sample keeps ordinary common stocks on NYSE, Amex, and Nasdaq, converts daily returns to log returns $x_{i,d} = \log(1 + r_{i,d})$, drops returns at or below -100 percent, and requires at least 15 daily observations in the calendar month. Let $\bar{x}_{i,t}$ be the monthly daily-log-return mean and let

$$m_{k,i,t} = \frac{1}{N_{i,t}} \sum_{d \in t} (x_{i,d} - \bar{x}_{i,t})^k. \tag{1}$$

The realized physical skewness and kurtosis are

$$\text{Skew}_{i,t} = \frac{m_{3,i,t}}{m_{2,i,t}^{3/2}}, \qquad \text{Kurt}_{i,t} = \frac{m_{4,i,t}}{m_{2,i,t}^2}. \tag{2}$$

Realized kurtosis is winsorized cross-sectionally within each month at the 1st and 99th

percentiles. The dependent variable is then the one-month lead of this winsorized realized-kurtosis measure within firm. This construction is why the regression outcome is a future physical moment rather than an option-implied moment or a return forecast.

The paper uses returns in two distinct ways. Daily returns define realized physical tail outcomes and physical benchmarks, including realized skewness, realized kurtosis, realized volatility, and CDG-style historical moments. Monthly returns define portfolio returns and factor alphas. This separation matters for interpretation. The first empirical claim is about the mapping from an ex ante physical predictive tail-width signal to future realized physical tail shape. The return tests then evaluate whether that same physical-tail characteristic is priced.

B Predictive Quantiles and Estimator-Specific W90

Each estimator predicts conditional quantiles of the next-month stock return distribution using information available at month t. The four estimators are gradient-boosted trees, a tabular transformer, linear quantile regression, and penalized linear quantile regression. The nonlinear estimators allow interactions and nonlinearities among predictors. The linear and penalized quantile regressions provide simpler distributional benchmarks with a different functional form. A signal that appears across these models is less likely to be a peculiarity of one estimation technology.

The information set is the processed Chen–Zimmermann stock-characteristic panel [Chen and Zimmermann, 2022]. Each raw characteristic is rank-normalized within month; missingness is preserved for consensus construction, while the ranked characteristics are median-filled for model training. The models are trained on rolling expanding windows and produce out-of-sample quantile predictions for the monthly return distribution. The prediction target for the quantile models is the next-month stock return, while the main empirical outcome in the paper is the realized physical kurtosis calculated from daily returns in the next month. This distinction is deliberate: the quantile models estimate a forward-looking return distribution, and W90 extracts a distributional width from those predictions. The model category is not the empirical phenomenon; the phenomenon is that predicted physical tail width varies across firms in a way that is informative about future realized tail shape.

For estimator a, define

$$W90^{a}_{i,t} = \widehat{q}^{a}_{0.95,i,t} - \widehat{q}^{a}_{0.05,i,t}, \tag{3}$$

where $\widehat{q}^{a}_{\tau,i,t}$ is the estimator-specific prediction of the τ conditional return quantile for firm i at month t. W90 is a predicted 90 percent interval width rather than a conditional mean or realized-volatility measure. A high W90 stock is one whose predicted physical return

distribution has a wide 5-to-95 percent interval relative to other firms in the same month. Figure 1 illustrates this object. The same realized return is drawn against a narrow and a wide predictive interval to separate the ex ante width measured by W90 from the realized return observed after the month ends.

Each estimator-specific interval width is adjusted within month by predicted-median-return decile and then ranked cross-sectionally. For width-like measures such as W90, W50, and tail-to-center width ratios, the adjustment is the log ratio of a stock's width to the mean width of firms in the same month and predicted-median decile. For signed asymmetry, the adjustment is additive because the measure can be negative. This adjustment removes mechanical scale differences associated with the predicted median while preserving cross-sectional variation in predictive tail width.

This construction has a simple economic purpose. W90 should measure relative predictive tail width rather than the mechanical scale associated with a high or low predicted median. Demeaning within q50 deciles keeps the signal focused on the conditional width of the predicted physical distribution. Monthly rank-normalization then makes coefficients and portfolio sorts cross-sectional: a high value means high predicted physical tail width relative to other firms in that month.

The demeaning step matters because a common within-month shift would not change rank-normalized widths. Predicted-median-decile adjustment changes the relative width ranking only when stocks are unusually wide or narrow compared with firms that have similar predicted median returns in the same month.

C Physical Tail Uncertainty

PTU aggregates the estimator-specific W90 signals into a common physical tail-width characteristic. For each stock-month, the construction takes the equal-weight average of the four demeaned and normalized W90 signals, requiring at least three nonmissing estimator inputs, and then rank-normalizes the average within month to produce PTU. Formally, if $\mathcal{A}_{i,t}$ is the set of available estimator signals and $|\mathcal{A}_{i,t}| \geq 3$,

$$\widetilde{PTU}_{i,t} = \frac{1}{|\mathcal{A}_{i,t}|} \sum_{a \in \mathcal{A}_{i,t}} W90_{i,t}^{a,\text{adj}}, \qquad PTU_{i,t} = \text{Rank}_t(\widetilde{PTU}_{i,t}). \tag{4}$$

The equal-weight rule is fixed ex ante. It prevents the paper from selecting the best estimator after observing the final empirical performance. The resulting PTU sample contains 2,262,381 nonmissing stock-months across 479 months, covering 70.33 percent of the analysis panel. In the reported sample, each estimator-specific W90 input has the same nonmissing

count. Table 1 reports the construction anatomy, coverage, and estimator correlations, and Figure 2 visualizes the corresponding common-component structure.

IA Table IA.27 reports unconditional quantile calibration, and IA Table IA.13 reports calibration by PTU quintile. Numerically, unconditional quantile coverage is close to nominal, with maximum absolute gaps of about 1.6 percentage points across estimators. Conditional on the top PTU quintile, however, the nominal 90 percent interval covers only about 85–87 percent of realized returns across the four estimators. Thus high PTU does not merely label wider predicted intervals; it identifies stock-months in which realized returns breach even these widened ex ante physical intervals more often.

The rank transformation gives PTU a clear interpretation. Moving from low to high PTU means moving across the monthly cross-section from firms with relatively narrow predicted physical return tails to firms with relatively wide predicted physical return tails. The main regressions therefore estimate whether this cross-sectional ranking predicts next-month realized physical tail shape and whether investors require compensation for holding high-PTU stocks.

The aggregation also separates the common economic signal from estimator-specific variation. The common PTU result is reported alongside each estimator-specific W90 signal in the appendix. If the result appeared only for one estimator, the natural interpretation would be model-specific. In the estimator-specific tables, all four estimator-specific W90 signals have positive high-minus-low alphas and strong physical-tail prediction under the reported non-option control blocks. PTU therefore summarizes a common component while preserving estimator-level evidence.

D Auxiliary Consensus and Disagreement Measures

Some portfolio implementations combine PTU with an auxiliary expected-return consensus formed from the Chen–Zimmermann characteristic panel. The consensus is the average active monthly rank across available characteristics for a firm-month. A corresponding dispersion measure summarizes cross-characteristic disagreement in the active ranked signals. The Internet Appendix reports alternative consensus and disagreement definitions, including theme-weighted versions. These auxiliary variables organize portfolio implementations; they are not main controls in the physical-tail prediction regressions.

E Physical and Option-Implied Benchmarks

Historical physical-tail benchmarks include realized skewness, realized kurtosis, CDG-style 60-day historical skewness and kurtosis, realized volatility, and idiosyncratic volatility.

Current-month realized skewness and kurtosis use the same daily-log-return central-moment formulas described above. Month-t values are used as predictors for the month-$t+1$ outcome.

The CDG-style historical measures use daily log returns over the trailing 60 trading days ending at month-end, require at least 30 daily observations, and compute central standardized skewness and kurtosis from raw first through fourth moments. These variables are historical physical-tail benchmarks, not option-implied quantities.

The tests also include the official Boyer–Mitton–Vorkink expected idiosyncratic skewness series where available and Bali–Cakici–Whitelaw MAX-style lottery controls. The official BMV series is a monthly dataset containing month, CRSP firm identifier, expected idiosyncratic skewness, and expected total skewness. Months are converted to month-ends, duplicate firm-month keys are checked, and the retained variables are expected idiosyncratic skewness and expected total skewness. The official BMV series contains 2,437,392 firm-month observations from January 1978 through December 2023 with no duplicate keys. The BCW controls use the maximum daily return over the prior 21 trading days and the average of the five largest daily returns over that same window, requiring at least 15 daily observations. These variables are natural comparisons because they capture past tail shape, expected skewness, lottery-like payoffs, and volatility.

Option-implied benchmarks include IV30 and BKM-style variance, skewness, and kurtosis. IV30 is built from the OptionMetrics IvyDB 30-day volatility surface. For each OptionMetrics security-month, it averages the call implied volatility at delta 0.50 and the put implied volatility at delta -0.50, then links OptionMetrics securities to CRSP firms using the OptionMetrics–CRSP link table. BKM moments are based on the model-free option-moment formulas of Bakshi et al. [2003]. The broad option-surface proxy uses OptionMetrics volatility-surface rows with 30 days to expiration, deltas from 10 to 90 for calls and -90 to -10 for puts, positive implied volatility and strike, and a Black–Scholes reconstruction of out-of-the-money option prices before applying the BKM quadratic, cubic, and quartic contract integrals. The broad option-surface proxy series contains 1,209,609 observations from January 1996 through December 2024 and provides a high-coverage option-implied benchmark.

The exact raw-chain BKM series is built from raw OptionMetrics option-chain quotes. The construction uses the last OptionMetrics trading date in each calendar month, standard contracts only, positive bid and ask quotes, out-of-the-money calls with $K > S$, out-of-the-money puts with $K < S$, same-day underlying prices, and zero-coupon rates. It computes BKM V/W/X contract values by maturity from option midquotes, then standardizes the contracts to 30 calendar days by exact 30-day maturity, interpolation between bracketing maturities, or a nearest-maturity fallback within the documented tolerance before converting

V/W/X into risk-neutral variance, skewness, and kurtosis. The final exact raw-chain series contains 181,279 observations from January 1996 through December 2024 with no duplicate firm-month keys. The exact series is closer to the canonical option-chain BKM construction, while the surface-proxy series provides broader coverage. The empirical tables therefore label the two sources separately. Table 2 reports the post-1996 coverage denominators for exact raw-chain BKM, broad/proxy BKM, and IV30.

This distinction between benchmark sources is part of the measurement contribution. BKM-style moments are valuable because they summarize the risk-neutral distribution implied by option prices. At the same time, they require observed option prices, usable strikes, maturities around the target horizon, and a mapping from options to the underlying stock. The raw-chain implementation therefore gives the cleanest direct option-chain benchmark but a smaller overlap sample. The broad surface proxy gives a wider option-implied comparison but with a different measurement interpretation. The empirical design uses both because they answer different benchmark questions.

F Control Sets and Sample Scopes

The non-option controls are conventional and benchmark-oriented. The baseline firm-characteristic controls are size, book-to-market, and 12–2 momentum. Because the dependent variable is future realized tail shape, the main predictive specification also includes realized volatility and idiosyncratic volatility. Expanded published non-option controls add historical physical-tail measures, BMV expected skewness, MAX-style lottery controls, and Amihud illiquidity. Macro controls, when used, enter as an additional layer. The PTU and estimator-specific tables report the realized-kurtosis relation under these control blocks.

The empirical design reports three sample scopes. Full-universe non-option tests use stock-months with the required physical outcomes and non-option controls. Broad option-proxy overlap tests use the wider option-implied proxy sample. Exact raw-chain option-overlap tests use the stricter BKM sample based on raw option chains. Table 9 is organized by these scopes because the scope is part of the estimand. This structure keeps the full-universe physical prediction evidence, broad option-proxy spanning evidence, and exact raw-chain option benchmark evidence distinct and interpretable.

III Economic Framework and Testable Predictions

A Physical Tail Width and Future Realized Tails

The framework isolates the sufficient economic restrictions behind the empirical design. First, PTU must contain information about a latent future physical tail state. Second, option-implied moments and physical predictive moments can differ because option moments are risk-neutral and observed in selected option-market samples. Third, if wide physical tails proxy for exposure that is costly to bear, hedge, or intermediate, the PTU premium should be largest where those costs are highest. Table 16 maps these restrictions to the empirical tables.

Let $\theta_{i,t+1}$ denote a latent future physical tail-thickness state for firm i. Let $K_{i,t+1}$ denote next-month realized physical kurtosis, measured from daily returns in month $t+1$. Suppose

$$K_{i,t+1} = k_0 + k_1\theta_{i,t+1} + u_{i,t+1}, \qquad k_1 > 0, \tag{5}$$

where $u_{i,t+1}$ is measurement noise and transitory realized-tail variation. PTU, denoted $P_{i,t}$, is an observable signal available at time t:

$$\theta_{i,t+1} = a_t + bP_{i,t} + c'X_{i,t} + \eta_{i,t+1}, \qquad b > 0, \tag{6}$$

where $X_{i,t}$ contains historical physical-tail measures and other controls. Combining the equations implies that PTU should predict future realized physical kurtosis in the population projection. This is the first empirical prediction and motivates the realized-kurtosis tests before any return interpretation.

B Incremental Information Beyond Historical Tails

Historical tail measures are informative but backward-looking. Let $H_{i,t}$ denote historical physical-tail measures, including lagged realized kurtosis, CDG-style historical kurtosis, realized skewness, and volatility. If $H_{i,t}$ is a backward-looking proxy for the future tail state and PTU contains forward-looking information beyond $H_{i,t}$, then residual PTU should predict residual future realized kurtosis:

$$K_{i,t+1}^{\perp H,X} = \alpha_t + \beta P_{i,t}^{\perp H,X} + \varepsilon_{i,t+1}, \qquad \beta > 0. \tag{7}$$

This residualized equation is the innovation restriction tested in the realized-kurtosis innovation evidence. It asks whether the component of PTU not explained by historical physical-tail

information predicts the component of future realized kurtosis not explained by the same information.

C Physical and Risk-Neutral Tail Objects

Option-implied moments are risk-neutral benchmarks. PTU is a physical predictive-width rank. The two objects need not coincide because risk-neutral moments combine physical expectations with pricing-kernel covariance terms. Let M_{t+1} be the stochastic discount factor and let $Y_{i,t+1}$ be a payoff-like tail object. Then

$$E_t^Q[Y_{i,t+1}] = \frac{E_t^P[M_{t+1}Y_{i,t+1}]}{E_t^P[M_{t+1}]} = E_t^P[Y_{i,t+1}] + \frac{\mathrm{Cov}_t^P(M_{t+1}, Y_{i,t+1})}{E_t^P[M_{t+1}]}. \tag{8}$$

This decomposition is compatible with informative option prices. It means that a risk-neutral moment and a physical predictive moment are different economic objects.

Option tests are therefore benchmark-discipline tests with explicit sample scope. Let $Q_{i,t}$ denote an observed option-implied moment:

$$Q_{i,t} = q_0 + q_1 E_t^P[\theta_{i,t+1}] + \pi_{i,t} + \nu_{i,t}, \tag{9}$$

where $\pi_{i,t}$ collects risk-neutral-to-physical pricing-kernel terms and $\nu_{i,t}$ captures measurement and market-coverage noise. Broad option-surface BKM-style proxies and exact raw-chain BKM moments are observed on different samples and with different measurement properties. Exact raw-chain BKM is the strict benchmark in the selected liquid-option universe; broad option-surface proxies provide higher-coverage option-related comparisons.

D Pricing Interpretation

PTU is a distribution-width object by construction. W90 is symmetric by design because it measures the amount of physical tail exposure investors must finance, hedge, and monitor. The pricing model does not require every wide-tail stock to have negative skewness; it requires the priced component of physical width to load on adverse aggregate tail states, with larger compensation where absorbing the same exposure is costlier.

To state this prediction, let

$$M_{t+1} = m_t + \lambda_t Z_{t+1}, \tag{10}$$

where Z_{t+1} is an adverse aggregate tail-uncertainty shock and $\lambda_t > 0$ is its price. Suppose

firm excess returns load on this shock as

$$R^e_{i,t+1} = \mu_{i,t} - \beta_{i,t} Z_{t+1} + \xi_{i,t+1}. \tag{11}$$

The sign convention means that firms with larger $\beta_{i,t}$ lose more in adverse tail-uncertainty states. The pricing condition $E_t[M_{t+1} R^e_{i,t+1}] = 0$ implies that expected excess returns increase with $\beta_{i,t}$ when the exposure is priced. I model the relevant exposure as

$$\beta_{i,t} = \beta_0 + \beta_P P_{i,t} + \beta_\phi \phi_{i,t} + \beta_{P\phi} P_{i,t} \phi_{i,t} + \nu_{i,t}, \qquad \beta_P > 0, \quad \beta_{P\phi} > 0, \tag{12}$$

where $\phi_{i,t}$ is a high-friction environment. The cross-partial restriction $\beta_{P\phi} > 0$ is the mechanism prediction: PTU should carry a larger premium when the same physical tail-width signal is harder to hedge, monitor, trade, or intermediate. This is why the mechanism tests focus on analyst coverage, institutional ownership, illiquidity, short interest, option coverage, and IV30 availability.

E Empirical Design

The predictive regressions are monthly Fama–MacBeth regressions [Fama and MacBeth, 1973]. In each month, the tests estimate cross-sectional regressions of next-month realized physical kurtosis on PTU and controls. Monthly coefficients are averaged and reported with Newey–West t-statistics using six lags. Full-universe physical-tail specifications use non-option control sets. Option-implied variables enter separately in broad option-proxy and exact raw-chain option-overlap rows.

The return tests use three complementary implementations. The direct high-minus-low test sorts on PTU and reports the highest-minus-lowest PTU quintile alpha. The characteristic-managed strategy asks whether PTU adds value inside an expected-return implementation. The long-only high-μ, high-PTU strategy asks whether the premium remains visible without shorting low-PTU stocks, but it is not a standalone PTU portfolio because it combines PTU with the expected-return consensus. FF6 alphas use the Fama–French five factors plus momentum, following Fama and French [2015] and Carhart [1997].

The empirical language is predictive and pricing-oriented. The paper estimates whether PTU forecasts realized tail shape, whether it contains information beyond historical and option-implied benchmarks in the relevant samples, and whether high-PTU stocks earn returns consistent with compensation for costly-to-bear physical tail exposure.

IV Results

The empirical tables follow a common reporting convention. Coefficients come from monthly cross-sectional Fama–MacBeth regressions unless the table is explicitly a portfolio table. Reported t-statistics use Newey–West standard errors with six lags. "Months" is the number of monthly cross-sections after the required variables for that row are present. "Avg. N" is the average number of stocks in the monthly cross-section, while "Rows" is the total complete-case stock-month count. Portfolio returns and alphas are monthly returns. These conventions are repeated in table notes because the sample changes meaningfully when option-implied benchmarks enter.

The results proceed in four steps. First, the paper validates PTU as a predictor of future realized physical tail shape. Second, it uses historical physical-tail measures and option-implied moments to define benchmark scope. Third, it asks whether the validated characteristic earns compensation in returns. Fourth, it locates the return premium across information, liquidity, monitoring, institutional risk-bearing, and hedging-market environments.

A PTU Predicts Future Realized Physical Kurtosis

Table 3 reports the core predictive evidence. The dependent variable is next-month realized physical kurtosis. The first row includes PTU alone and establishes the unconditional relation: high-PTU firms subsequently realize higher physical kurtosis, with a coefficient of 0.342 and a t-statistic of 8.01 across 479 months. The average monthly cross-section in this row is 4,685 stocks. Because PTU is rank-normalized within month, the coefficient is a cross-sectional slope across the monthly PTU distribution. The result says that the predicted physical tail-width ranking formed at month t maps to realized physical tail shape in month $t + 1$.

The next rows add controls that are natural for a physical-tail prediction exercise. Adding size, book-to-market, momentum, realized volatility, and idiosyncratic volatility produces a coefficient of 0.367 with a t-statistic of 7.89 across 479 months. The average monthly cross-section is 3,707 stocks and the average R^2 is 0.042. Adding the available macro control leaves a coefficient of 0.352 and a t-statistic of 7.15. The expanded published non-option block adds historical realized skewness and kurtosis, CDG-style historical skewness and kurtosis, BMV expected skewness, BCW MAX and MAX5, and Amihud illiquidity. In that block, the PTU coefficient is 0.399 with a t-statistic of 10.07 across 468 months. Adding the macro control produces a coefficient of 0.401 with a t-statistic of 10.19.

These estimates indicate that PTU adds information beyond volatility, historical-tail,

lottery, and illiquidity controls. PTU captures more than broad volatility sorting. The coefficient remains large when realized volatility and idiosyncratic volatility are included, and it remains large when historical physical-tail measures, expected skewness, lottery-like payoffs, and illiquidity are included. The predictive relation is therefore tied to a forward-looking tail-width feature of the conditional physical distribution. A high PTU stock is a stock for which several quantile-prediction methods agree that the future physical return distribution is unusually wide in the tails, and Table 3 shows that this ex ante ranking predicts ex post realized kurtosis.

Table 4 translates the continuous realized-kurtosis result into high-tail-state outcomes, and Table 5 reports the corresponding high-minus-low PTU quintile spreads. Moving from the lowest to the highest PTU quintile raises the probability of a top-decile future realized-kurtosis state by 3.06 percentage points and a top-quintile state by 4.90 percentage points. Figure 3 shows the same pattern visually: high-PTU firms are more likely to enter high realized-tail states next month, and the event rate rises monotonically across PTU quintiles.

The progression across rows clarifies the economics. The unconditional row establishes raw predictive content. The volatility row is the central benchmark because volatility is the most immediate alternative explanation for wider predictive tails. The expanded row is the most demanding published-control specification in the main table. PTU remains positive and precisely estimated after historical physical-tail measures, expected skewness, lottery controls, volatility, and illiquidity enter, so the estimates point to incremental physical-tail information beyond these benchmark variables.

The estimator-specific counterpart in IA Table IA.3 provides an additional estimator-level check. Under the size, book-to-market, momentum, realized-volatility, and idiosyncratic-volatility specification, the estimator-specific W90 t-statistics range from 4.71 for penalized quantile regression to 8.68 for gradient-boosted trees. Under the expanded non-option benchmarks, the t-statistics range from 6.29 to 10.45. The physical-tail relation is present across the estimator-specific W90 signals. The commonality across estimators is part of the result: the width itself carries information about future realized physical tail shape. PTU aggregates strong estimator-level information without selecting the best-performing model, and the estimator-level evidence shows that each model family contains similar physical-tail information.

B Incremental Information Relative to Historical Physical Tails

Table 6 first separates persistent firm exposure from time-varying physical-tail information. The first row repeats the cross-sectional level relation. The second row removes each

firm's time-series mean from PTU and future realized kurtosis, so identification comes from within-firm deviations rather than permanent firm differences. Firm-demeaned PTU remains positive, with a coefficient of 0.255 and a t-statistic of 6.95. The third row uses monthly changes: changes in PTU predict changes in future realized kurtosis with a coefficient of 0.378 and a t-statistic of 12.74. IA Table IA.22 shows that estimator-specific W90 has moderate AR(1) persistence, around 0.66–0.71 across estimators, comparable to realized-volatility and trailing-tail benchmarks but far below near-fixed idiosyncratic-volatility persistence. PTU is therefore persistent enough to measure a tail-width state and time-varying enough for the within-firm and change tests in Table 6 to be economically meaningful.

Table 7 then evaluates whether transparent physical-tail forecasts explain PTU's predictive content. PTU remains positive after lagged realized volatility, realized kurtosis, realized skewness, CDG-style historical kurtosis, MAX-style lottery measures, and related transparent benchmarks enter. More directly, the component of PTU residualized against these simple tail forecasts still predicts future realized kurtosis, with a coefficient of 0.205 and a t-statistic of 7.73. The benchmark evidence therefore changes the interpretation from high-PTU firms tend to be tail-prone to PTU contains time-varying forward-looking information about future physical tail states beyond transparent historical-tail forecasts. IA Table IA.24 shows that related distributional transforms also contain pricing content, while W90 remains the primary object because it is the transparent fixed 5-to-95 physical predictive width and is first validated against future realized physical tail shape.

Table 8 reports the residual innovation version of the same logic. Future realized kurtosis is residualized within month on historical physical-tail measures and controls. PTU is residualized on the same information set. The regression then estimates whether the component of PTU orthogonal to historical physical-tail information predicts the component of future realized kurtosis orthogonal to that same information. The coefficient is 0.198 with a t-statistic of 6.66 across 1,736,062 observations. This result moves the interpretation from correlation with known tail variables to incremental predictive content.

The first row of Table 9 provides the full-universe bridge into the option section. It uses the same future realized-kurtosis outcome as Table 3, includes the non-option historical physical-tail and published benchmark controls described in Section II, and reports the coefficient on PTU. PTU remains positive after controlling for realized kurtosis, CDG-style historical kurtosis, realized skewness, CDG-style historical skewness, BMV expected skewness, MAX-style lottery controls, realized volatility, idiosyncratic volatility, and firm characteristics. The PTU coefficient is 0.359 with a t-statistic of 6.46 across 468 months and 1,736,062 observations. Historical physical-tail variables are informative in the same regression, and PTU remains informative alongside them.

The other coefficients in the full-universe benchmark specification reinforce the benchmark's difficulty. CDG-style historical kurtosis has a coefficient of 0.697 and a t-statistic of 13.05, realized kurtosis has a coefficient of 0.214 and a t-statistic of 4.40, and idiosyncratic volatility has a coefficient of 0.425 and a t-statistic of 14.95. These variables are strong benchmarks in their own right. PTU's positive coefficient in the same regression is therefore incremental information alongside strong historical-tail and volatility benchmarks. Historical tails and volatility forecast future tail shape, and PTU adds forward-looking distributional information on top of them.

C Option-Implied Benchmarks and Coverage Discipline

Option-implied moments are the natural benchmark for any tail-risk paper. The key distinction is that option moments are risk-neutral and observed only where option-market infrastructure is sufficient, while PTU is a physical predictive-width characteristic formed across a broader equity universe. I therefore treat option tests as benchmark-discipline tests with explicit sample labels and explicit coverage distinctions.

Figure 4 and Table 2 show the coverage distinction. Exact raw-chain BKM is available for about 7.4 percent of the post-1996 analysis panel, while broad option-surface proxies and IV30 cover roughly 52.5 percent of the panel. Table 11 shows that the exact-BKM universe is economically selected: exact-BKM firms are larger, more followed, more institutionally held, more liquid, and lower in future realized physical kurtosis than firms outside the broad option-proxy sample. These differences define the option-market environment in which exact option-chain variance is expected to be especially informative.

With this coverage discipline in place, Table 9 reports the broad option-proxy comparison. The broad option-surface BKM-style proxy row uses the higher-coverage OptionMetrics volatility-surface controls: IV30 plus BKM-style variance, skewness, and kurtosis reconstructed from the standardized 30-day delta surface. In this option-proxy overlap, PTU remains positive after the option-surface controls enter. The coefficient is 0.574 with a t-statistic of 12.44 across 724,269 observations and 347 months. This result shows that the physical predictive-width ranking is not mechanically absorbed by broad option-surface information.

Table 10 then reports the stricter exact raw-chain benchmark. Inside the selected liquid-option universe, PTU is positive before exact option controls enter, with a coefficient of 0.439 and a t-statistic of 4.66. Once exact BKM variance enters, the full exact-BKM control set leaves a residual PTU coefficient of -0.342 with a t-statistic of -3.91. The economic interpretation is that exact BKM variance becomes the tight option-chain benchmark when

clean liquid option-chain information is available. The negative residual PTU coefficient in the full exact-BKM specification is therefore a local residual relation inside the selected liquid-option sample after exact option variance absorbs the relevant tail-state information; the economic comparison is not a negative broad-universe PTU result, but the benchmark role of exact option variance where clean option-chain data exist.

The division of labor is clear: exact option-chain moments are the liquid-option benchmark, while PTU provides broad physical-tail measurement outside and around that benchmark. In the broad proxy row, IV30 is positive and significant, with a coefficient of 0.518 and a t-statistic of 6.19, while proxy BKM kurtosis is negative with a t-statistic of -2.98. In the full exact raw-chain BKM specification, BKM variance is strongly positive, with a coefficient of 2.569 and a t-statistic of 9.47, and BKM skewness is positive with a t-statistic of 2.03. Option-implied moments contain information in option-covered stocks, and the relevant comparison depends on whether the benchmark is broad option-surface proxy information or exact raw-chain option information.

D Pricing Evidence

The return tests ask whether the physical-tail characteristic validated above is compensated in the cross-section. This sequence matters. PTU is not introduced as a return predictor and then rationalized ex post. It is first validated as an ex ante predictor of realized physical tail shape. The portfolio tests then ask whether investors require compensation for holding stocks with high predicted physical tail width.

Table 12 begins with the direct PTU high-minus-low sort. This is the clean characteristic test: each month, stocks are sorted by PTU, and the table reports the high-PTU minus low-PTU portfolio. The direct sort earns an average monthly return of 2.83 percent with a t-statistic of 7.27 and an FF6 alpha of 3.07 percent with a t-statistic of 9.75. Because the sort is based only on PTU, the row shows that the physical tail-width characteristic earns factor-adjusted returns before any expected-return interaction is introduced.

The characteristic-managed strategy asks a different question. It uses the auxiliary expected-return consensus μ and PTU together, forming a tail-width implementation inside the expected-return structure described in Section II. It earns an average monthly return of 2.72 percent with a t-statistic of 6.85 and an FF6 alpha of 2.41 percent with a t-statistic of 7.27. The row is an implementation test showing that high predicted physical tail width has value inside a broader expected-return screen.

The long-only high-μ, high-PTU portfolio is also an implementation result. It earns an average monthly return of 4.51 percent with a t-statistic of 10.81 and an FF6 alpha of 3.88

percent with a t-statistic of 12.06. This row shows that the premium is visible without shorting low-PTU stocks. Because the portfolio combines PTU with the expected-return consensus, it is interpreted as a long-only implementation of the physical-width signal.

Table 12 also reports the expected-return-by-PTU difference-in-differences row. It earns a positive mean return with a t-statistic of 4.19, while the FF6 alpha is 0.41 percent with a t-statistic of 1.34. The main pricing conclusion rests on the direct high-minus-low, characteristic-managed, and long-only implementations, all of which have positive and precisely estimated factor-adjusted alphas.

The magnitude of the equal-weight spread makes implementation discipline important. Table 13 reports the direct PTU high-minus-low portfolio under value weighting, price screens, NYSE-size screens, and a liquid-half screen. The full equal-weight FF6 alpha is 3.05 percent per month, while the value-weighted alpha is 0.83 percent. The price-greater-than-five value-weighted alpha is 0.46 percent, and the NYSE-median-size value-weighted alpha is 0.55 percent. The premium remains positive under stricter weighting, price, size, and liquidity screens, with magnitudes that identify where the economics is strongest. The magnitudes are economically meaningful: using measured turnover, IA Table IA.15 reports break-even costs of roughly 428–552 basis points per trip for the long-only implementations and 159–287 basis points for the μ-stratified long-short implementations, while the DiD design is more sensitive. These numbers discipline implementation claims without turning the paper into a capacity study. IA Table IA.25 reports value-weighted estimator-specific implementations and preserves the same qualitative pricing interpretation.

E Mechanism Evidence

Table 14 reports all groups within each friction split and asks where the PTU premium is largest. The text focuses on the friction-consistent side because that is the model prediction. The framework predicts larger premia where information production, hedging access, liquidity provision, monitoring, or risk-bearing capacity is weaker.

Table 15 and Figure 5 connect the measurement object to the friction channel. High-PTU firms are smaller, less covered by analysts, less institutionally held, more illiquid, and have high IV30 where option data exist. The positive IV30 spread is conditional on IV30 availability and does not contradict the larger premium among firms without IV30 availability. The anatomy evidence shows that high predicted physical width is located in precisely the firm environments where the pricing mechanism predicts higher compensation.

The information-production channel appears in the analyst-coverage and institutional-ownership splits. The PTU premium is largest among firms with low analyst coverage, where

the FF6 alpha is 3.21 percent with a t-statistic of 8.67. It is also largest among firms with low institutional ownership, where the FF6 alpha is 4.78 percent with a t-statistic of 10.84. These are settings in which public information production and institutional risk-bearing capacity are thinner, so investors may require greater compensation for exposure to wide physical tails.

The liquidity channel appears in the Amihud split. The high-illiquidity group has the largest PTU premium, with an FF6 alpha of 5.71 percent and a t-statistic of 11.51. This is the clearest trading-friction split: when positions are more costly to enter, exit, or rebalance, the compensation for bearing predicted physical tail width is larger.

The monitoring channel is interpreted more cautiously. The short-interest split has the largest PTU premium in the low short-interest group, with an FF6 alpha of 2.92 percent and a t-statistic of 5.51. Low short interest is consistent with lower short-seller monitoring, but short interest also reflects disagreement, bad news, overvaluation, hedging demand, and short-sale constraints. The split is therefore auxiliary evidence on monitoring rather than the central mechanism result.

The hedging-market channel appears in the option-availability splits. Firms without option coverage have a PTU high-minus-low FF6 alpha of 4.15 percent with a t-statistic of 11.56. Firms without IV30 availability have an FF6 alpha of 4.12 percent with a t-statistic of 11.43. These results are consistent with the idea that PTU premia are larger where direct option-market signals and hedging instruments are thinner.

The mechanism evidence concerns where the return premium is largest, not necessarily where PTU predicts realized kurtosis most strongly. Prediction and pricing are distinct. PTU can be informative about future physical tails in more covered stocks, while the return premium is larger where bearing that exposure is more costly. The mechanism table therefore speaks primarily to pricing: the premium is located where the cost of absorbing physical tail exposure is plausibly highest.

F Robustness Across Quantile Estimators

IA Table IA.7 reports estimator-specific W90 pricing evidence. Each estimator-specific W90 signal earns a positive high-minus-low return and a positive FF6 alpha. The GBT W90 alpha is 2.29 percent with a t-statistic of 8.64. The transformer W90 alpha is 2.90 percent with a t-statistic of 9.44. The linear quantile-regression W90 alpha is 2.75 percent with a t-statistic of 9.72. The penalized quantile-regression W90 alpha is 2.50 percent with a t-statistic of 9.31. This estimator-specific evidence is reported without optimized ensemble weights. Together, these rows show that the economic pattern exists across nonlinear and

linear quantile-prediction approaches.

Figure 1 is the schematic measurement object: it shows that PTU ranks ex ante physical predictive width rather than realized returns. Figure 2 is measurement evidence: it shows why aggregation is sensible by reporting the common physical-width component across estimator-specific W90 signals. Figure 3 is tail-state validation: it turns the regression result into event probabilities across PTU quintiles. Figure 4 is coverage discipline: it shows why exact raw-chain BKM is a selected liquid-option benchmark rather than the estimand of the broad-universe PTU contribution. Figure 5 is friction anatomy: it links high PTU to the firm environments in which the pricing mechanism predicts higher compensation. IA Figure IA.1 complements the schematic by visualizing predicted-distribution shape by PTU state.

V Interpretation and Scope

PTU is a physical predictive-width characteristic formed from predicted 5-to-95 percentile physical return width and ranked within month. W90 intentionally measures physical distribution width; the pricing sign comes from the stochastic-discount-factor and friction channel.

The historical-tail evidence defines the first scope condition. Realized kurtosis, CDG-style trailing kurtosis, volatility, expected skewness, and MAX-style variables are informative benchmarks. PTU remains positive alongside them and predicts residual innovations in future realized kurtosis after historical physical-tail information is removed. The empirical interpretation is therefore incremental: backward-looking tail measures matter, and the forward-looking predictive-width rank contains additional information about next-month physical tail shape.

The option evidence defines the second empirical domain. Broad option-surface BKM-style proxies provide high-coverage option-related benchmarks, and PTU remains informative in that option-proxy overlap. Exact raw-chain BKM provides the stricter benchmark in the selected liquid-option universe. The division of labor is clear: exact option-chain moments are the liquid-option benchmark, while PTU provides broad physical-tail measurement outside and around that benchmark.

The return evidence is interpreted as compensation evidence for a measured physical-tail characteristic. Direct PTU high-minus-low portfolios provide the clean characteristic test. Characteristic-managed and long-only high-μ, high-PTU portfolios are implementation tests that combine PTU with expected-return information. Value-weighting and investability screens reduce the magnitudes and identify where the economics is strongest, while the premium remains positive under stricter implementation discipline.

VI Additional Evidence and Economic Interpretation

The Internet Appendix has four roles. First, it deepens measurement validity: the model-training, glossary, construction, calibration, and predicted-distribution exhibits show how the common physical-width component is formed and why high PTU corresponds to stock-months with wider and more frequently breached ex ante physical intervals. Second, it documents estimator and component robustness: estimator-specific prediction and pricing tables, residualized W90 tests, fuller realized-kurtosis specifications, and related distributional transforms show that the evidence is tied to predicted physical width rather than to one estimator or one mechanical transformation.

Third, the appendix provides implementation discipline. Cost, holding-period, microcap, value-weighted, subperiod, industry, and macro-state evidence clarify where the return premium is strongest and how it changes under stricter implementation screens, without turning the paper into a trading-strategy exercise. Fourth, the appendix documents persistence and time variation. PTU has enough persistence to measure a physical tail-width state, enough movement for within-firm and change specifications to be informative, and rolling pricing and realized-kurtosis evidence that connects the cross-sectional result to time-series stability.

The Internet Appendix also provides international external validation in global JKP data. In a 70-country, 10.79-million-stock-month panel, country-month PTU ranks strongly predict future absolute-tail states, left-tail events, and right-tail events. The evidence is used as external validation of the physical tail-width phenomenon rather than as a replacement for the U.S. option-benchmark and factor-alpha tests. Return spreads are positive in liquid, large, and developed-market screens, while the raw full-sample equal-weight global spread is treated cautiously because it is sensitive to extreme small-country/month observations.

VII Conclusion

This paper introduces PTU, a fixed-rule rank of ex ante physical predictive tail width. PTU forecasts future realized physical kurtosis and high-kurtosis states beyond volatility, historical physical tails, lottery measures, expected skewness, illiquidity, and firm controls. International JKP evidence in the Internet Appendix shows that the physical tail-width validation is not confined to the U.S. sample. The benchmark evidence shows that PTU contains forward-looking physical-tail information beyond backward-looking measures and remains informative relative to broad option-surface proxies, while exact raw-chain BKM defines the stricter benchmark in the selected liquid-option universe.

PTU earns positive alphas that survive investability screens, with the largest pre-

mia in settings where information, liquidity, monitoring, institutional risk-bearing, and hedging-market frictions make tail exposure harder to absorb. The contribution is a finance measurement-and-pricing result: predictive physical tail width is an economically meaningful characteristic in the cross-section of returns.

References

Tobias Adrian, Erkko Etula, and Tyler Muir. Financial intermediaries and the cross-section of asset returns. *Journal of Finance*, 69(6):2557–2596, 2014. doi: 10.1111/jofi.12189.

Yakov Amihud. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, 5(1):31–56, 2002. doi: 10.1016/S1386-4181(01)00024-6.

Andrew Ang, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang. The cross-section of volatility and expected returns. *Journal of Finance*, 61(1):259–299, 2006. doi: 10.1111/j.1540-6261.2006.00836.x.

Gurdip Bakshi, Nikunj Kapadia, and Dilip Madan. Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies*, 16 (1):101–143, 2003. doi: 10.1093/rfs/16.1.0101.

Turan G. Bali, Nusret Cakici, and Robert F. Whitelaw. Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, 99(2):427–446, 2011. doi: 10.1016/j.jfineco.2010.08.014.

Alexandre Belloni and Victor Chernozhukov. ℓ_1-penalized quantile regression in high-dimensional sparse models. *Annals of Statistics*, 39(1):82–130, 2011. doi: 10.1214/10-AOS827.

Tim Bollerslev, George Tauchen, and Hao Zhou. Expected stock returns and variance risk premia. *Review of Financial Studies*, 22(11):4463–4492, 2009. doi: 10.1093/rfs/hhp008.

Brian H. Boyer, Todd Mitton, and Keith Vorkink. Expected idiosyncratic skewness. *Review of Financial Studies*, 23(1):169–202, 2010. doi: 10.1093/rfs/hhp041.

Mark M. Carhart. On persistence in mutual fund performance. *Journal of Finance*, 52(1): 57–82, 1997. doi: 10.1111/j.1540-6261.1997.tb03808.x.

Bo-Young Chang, Peter Christoffersen, and Kris Jacobs. Market skewness risk and the cross section of stock returns. *Journal of Financial Economics*, 107(1):46–68, 2013. doi: 10.1016/j.jfineco.2012.07.002.

Andrew Y. Chen and Tom Zimmermann. Open source cross-sectional asset pricing. *Critical Finance Review*, 11(2):207–264, 2022. doi: 10.1561/104.00000112.

Jennifer Conrad, Robert F. Dittmar, and Eric Ghysels. Ex ante skewness and expected stock returns. *Journal of Finance*, 68(1):85–124, 2013. doi: 10.1111/j.1540-6261.2012.01795.x.

Karl B. Diether, Christopher J. Malloy, and Anna Scherbina. Differences of opinion and the cross section of stock returns. *Journal of Finance*, 57(5):2113–2141, 2002. doi: 10.1111/0022-1082.00490.

Eugene F. Fama and Kenneth R. French. A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22, 2015. doi: 10.1016/j.jfineco.2014.10.010.

Eugene F. Fama and James D. MacBeth. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636, 1973. doi: 10.1086/260061.

Joachim Freyberger, Andreas Neuhierl, and Michael Weber. Dissecting characteristics nonparametrically. *Review of Financial Studies*, 33(5):2326–2377, 2020. doi: 10.1093/rfs/hhz123.

Stefano Giglio, Bryan Kelly, and Dacheng Xiu. Factor models, machine learning, and asset pricing. *Annual Review of Financial Economics*, 14:337–368, 2022. doi: 10.1146/annurev-financial-101521-104735.

Yury Gorishniy, Ivan Rubachev, Valentin Khrulkov, and Artem Babenko. Revisiting deep learning models for tabular data. In *Advances in Neural Information Processing Systems*, volume 34, pages 18932–18943. Curran Associates, Inc., 2021. URL https://proceedings.neurips.cc/paper_files/paper/2021/file/9d86d83f925f2149e9edb0ac3b49229c-Paper.pdf.

Denis Gromb and Dimitri Vayanos. Limits of arbitrage. *Annual Review of Financial Economics*, 2:251–275, 2010. doi: 10.1146/annurev-financial-073009-104107.

Shihao Gu, Bryan Kelly, and Dacheng Xiu. Empirical asset pricing via machine learning. *Review of Financial Studies*, 33(5):2223–2273, 2020. doi: 10.1093/rfs/hhaa009.

Zhiguo He and Arvind Krishnamurthy. Intermediary asset pricing. *American Economic Review*, 103(2):732–770, 2013. doi: 10.1257/aer.103.2.732.

Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. LightGBM: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf.

Bryan T. Kelly, Seth Pruitt, and Yinan Su. Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, 134(3):501–524, 2019. doi: 10.1016/j.jfineco.2019.05.001.

Roger Koenker and Gilbert Bassett. Regression quantiles. *Econometrica*, 46(1):33–50, 1978. doi: 10.2307/1913643.

Andrei Shleifer and Robert W. Vishny. The limits of arbitrage. *Journal of Finance*, 52(1): 35–55, 1997. doi: 10.1111/j.1540-6261.1997.tb03807.x.

Main Tables

Table 1: PTU construction across estimator-specific W90 signals.

Estimator-specific W90 signal	Firm-months	Months	Coverage	Corr. with PTU
GBT W90	2,262,381	479	70.33	0.778
Transformer W90	2,262,381	479	70.33	0.906
Linear QR W90	2,262,381	479	70.33	0.942
Penalized QR W90	2,262,381	479	70.33	0.933

Notes: Each estimator-specific input is a q50-decile-demeaned and within-month rank-normalized W90 signal. PTU is the within-month rank of the equal-weight average, requiring at least three available estimator inputs.

Table 2: Coverage of exact and broad option-implied benchmarks after 1996.

Benchmark	Denominator rows	Available rows	Coverage (%)
Exact BKM variance	1,562,078	115,278	7.38
IV30	1,562,078	821,839	52.61
Exact raw-chain BKM	1,562,078	115,278	7.38
Broad/proxy BKM	1,562,078	820,364	52.52
IV30 available	1,562,078	821,839	52.61

Notes: The denominator is the post-1996 analysis-panel firm-month universe. Exact raw-chain BKM requires stricter option-chain availability and filters, while broad/proxy BKM and IV30 cover a much wider option-related sample.

Table 3: PTU and future realized physical kurtosis.

Specification	Coef.	t-stat	Months	Avg. N	Avg. R^2
PTU alone	0.342	8.01	479	4,685	0.005
Firm characteristics and volatility	0.367	7.89	479	3,707	0.042
Firm characteristics, volatility, and macro beta	0.352	7.15	479	2,917	0.042
Expanded non-option benchmarks	0.399	10.07	468	3,490	0.042
Expanded non-option benchmarks and macro beta	0.401	10.19	468	2,763	0.040

Notes: The table reports monthly Fama–MacBeth regressions of next-month realized physical kurtosis on PTU. Firm characteristics include size, book-to-market, and momentum; volatility controls include realized and idiosyncratic volatility. Macro beta is the 60-month rolling beta of firm excess returns with respect to changes in Jurado–Ludvigson–Ng one-month macro uncertainty. Expanded non-option controls add historical physical-tail measures, BMV expected skewness, MAX-based lottery controls, and illiquidity. Option-implied variables are excluded from this full-universe specification. Newey–West t-statistics use six lags.

Table 4: PTU and high future realized-kurtosis states.

Outcome	Specification	Coef.	t-stat	Months	Avg. N	Avg. R^2
Top decile	PTU only	0.037	9.27	479	4,685	0.003
Top decile	+ historical physical tail controls	0.075	22.93	479	4,653	0.027
Top decile	+ historical + non-option friction controls	0.059	15.36	479	3,238	0.016
Top quintile	PTU only	0.059	11.83	479	4,685	0.004
Top quintile	+ historical physical tail controls	0.114	26.22	479	4,653	0.031
Top quintile	+ historical + non-option friction controls	0.089	18.33	479	3,238	0.021

Notes: The dependent variable equals one when a firm-month's next-month realized physical kurtosis is in the top decile or top quintile of the monthly cross-section. Historical physical-tail controls are realized kurtosis, realized volatility, realized skewness, and CDG-style kurtosis. The non-option friction block adds size, analyst coverage, institutional ownership, and Amihud illiquidity.

Table 5: High-minus-low PTU spreads in future tail outcomes.

Outcome	Q5–Q1	SE	t-stat	Months
Top-decile event rate	0.0306	0.0033	9.35	479
Top-quintile event rate	0.0490	0.0042	11.72	479
Future realized kurtosis	0.2870	0.0354	8.12	479

Notes: Each row reports the time-series average of the monthly difference between the highest and lowest PTU quintiles. The event-rate outcomes are indicators for top-decile and top-quintile future realized physical kurtosis.

Table 6: Within-firm PTU and future realized physical kurtosis.

Specification	PTU measure	Coef.	t-stat	Months	Avg. N	Avg. R^2
Cross-sectional levels	PTU	0.326	6.16	468	3,710	0.058
Firm-demeaned levels	Firm-demeaned PTU	0.255	6.95	468	3,710	0.023
Monthly changes	Change in PTU	0.378	12.74	467	3,655	0.233

Notes: Rows report monthly Fama–MacBeth tests of future realized physical kurtosis on PTU. The firm-demeaned row removes each firm's time-series mean from both PTU and the future realized-kurtosis outcome. The monthly-change row uses adjacent-month changes. Newey–West t-statistics use six lags.

Table 7: PTU relative to transparent physical-tail forecasts.

Specification	Measure	Coef.	t-stat	Months	Avg. N	Avg. R^2
PTU alone	PTU	0.342	8.01	479	4,685	0.005
PTU plus simple physical-tail forecasts	PTU	0.360	7.58	479	3,702	0.059
PTU plus simple physical-tail forecasts and BMV	PTU	0.359	7.51	468	3,710	0.060
Residual PTU after simple physical-tail forecasts	Residual PTU	0.205	7.73	479	3,702	0.002

Notes: The simple physical-tail benchmark set includes lagged realized volatility, realized skewness, realized kurtosis, CDG-style historical kurtosis, MAX-style lottery measures, and related non-option tail forecasts. The residual row first removes these simple forecasts from both future realized kurtosis and PTU within month, then tests whether residual PTU predicts residual future kurtosis.

Table 8: PTU and innovations in future realized physical kurtosis.

Specification	Coef.	t-stat	Months	Avg. N	Avg. R^2
Residual future kurtosis on residual PTU	0.198	6.66	468	3,710	0.002
Level future kurtosis on PTU + innovation controls	0.359	6.46	468	3,710	0.062

Notes: The outcome is residual future realized physical kurtosis after removing historical physical-tail information and non-option controls. Innovation controls are the backward-looking tail and non-option benchmark variables used to residualize both future realized kurtosis and PTU, including realized volatility, realized skewness and kurtosis, CDG-style historical moments, BMV expected skewness, MAX-style lottery controls, illiquidity, and firm controls. The table estimates PTU's information beyond backward-looking tail measures.

Table 9: PTU, historical physical-tail benchmarks, and option-implied benchmarks.

Sample	Specification	PTU coef.	t-stat	Rows	Months
Full universe, non-option	PTU + historical physical-tail benchmarks + FULL non-option	0.359	6.46	1,736,062	468
Broad option-proxy overlap	PTU + broad BKM surface/proxy option controls + FULL non-option	0.574	12.44	724,269	347
Exact raw-chain option overlap	PTU + exact raw-chain BKM controls + FULL non-option	-0.316	-4.93	103,191	248

Notes: The full-universe row uses only non-option benchmarks and controls. The broad option row uses the wider option-proxy overlap sample. The exact raw-chain row uses stricter BKM moments within its selected option-overlap sample; in that row, exact option variance provides the tighter benchmark for PTU.

Table 10: PTU and exact raw-chain BKM benchmark-scope tests.

Specification	Reported variable	Coef.	t-stat	Months	Avg. N	Avg. R^2
Full post-1996 universe	PTU	0.521	12.39	336	3,683	0.032
Broad option-proxy overlap	PTU	0.482	11.59	336	2,071	0.038
Exact-BKM sample before exact option controls	PTU	0.439	4.66	261	360	0.124
Exact-BKM sample plus exact BKM variance	PTU	-0.267	-3.26	261	360	0.151
Exact-BKM sample plus all exact BKM controls	PTU	-0.342	-3.91	261	360	0.174
Exact BKM variance in full exact specification	Exact BKM variance	2.569	9.47	261	360	0.174
Exact BKM skewness in full exact specification	Exact BKM skewness	0.140	2.03	261	360	0.174
Exact BKM kurtosis in full exact specification	Exact BKM kurtosis	-0.287	-3.94	261	360	0.174

Notes: The table separates full-universe, broad option-proxy, and exact raw-chain option-overlap tests. The exact-BKM rows show that PTU is positive in the exact-option sample before exact BKM controls enter and that exact BKM variance provides the tighter benchmark once included. This row is the selected liquid-option-universe comparison, distinct from the full-universe tests.

Table 11: Firm characteristics of the exact-BKM option universe.

Variable	Contrast	Mean difference	t-stat	Months
PTU	Exact BKM minus broad-proxy-only	0.017	1.20	347
PTU	Exact BKM minus no broad proxy	0.020	1.42	347
PTU	Broad-proxy-only minus no broad proxy	0.002	0.30	347
Future realized kurtosis	Exact BKM minus broad-proxy-only	-0.042	-1.02	347
Future realized kurtosis	Exact BKM minus no broad proxy	-0.213	-3.88	347
Future realized kurtosis	Broad-proxy-only minus no broad proxy	-0.171	-6.55	347
Log size	Exact BKM minus broad-proxy-only	2.119	36.91	348
Log size	Exact BKM minus no broad proxy	4.606	87.98	348
Log size	Broad-proxy-only minus no broad proxy	2.487	65.13	348
Size rank	Exact BKM minus broad-proxy-only	0.235	24.15	348
Size rank	Exact BKM minus no broad proxy	0.606	110.75	348
Size rank	Broad-proxy-only minus no broad proxy	0.371	61.11	348
Idiosyncratic volatility	Exact BKM minus broad-proxy-only	-0.016	-1.39	348
Idiosyncratic volatility	Exact BKM minus no broad proxy	-0.065	-4.82	348
Idiosyncratic volatility	Broad-proxy-only minus no broad proxy	-0.049	-9.49	348
Lagged realized kurtosis	Exact BKM minus broad-proxy-only	-0.021	-4.46	348
Lagged realized kurtosis	Exact BKM minus no broad proxy	-0.066	-10.44	348
Lagged realized kurtosis	Broad-proxy-only minus no broad proxy	-0.044	-10.75	348
Analyst coverage	Exact BKM minus broad-proxy-only	0.205	26.68	348
Analyst coverage	Exact BKM minus no broad proxy	0.350	46.05	348
Analyst coverage	Broad-proxy-only minus no broad proxy	0.144	25.78	348
Institutional ownership	Exact BKM minus broad-proxy-only	0.065	12.41	348
Institutional ownership	Exact BKM minus no broad proxy	0.401	70.91	348
Institutional ownership	Broad-proxy-only minus no broad proxy	0.336	64.67	348
Amihud illiquidity	Exact BKM minus broad-proxy-only	-0.254	-29.70	348
Amihud illiquidity	Exact BKM minus no broad proxy	-0.648	-130.37	348
Amihud illiquidity	Broad-proxy-only minus no broad proxy	-0.394	-72.23	348

Notes: Rows compare monthly average characteristics across exact-BKM, broad-proxy-only, and no-broad-proxy option groups. The table documents that exact-BKM firms are a selected liquid-option population with distinct coverage and firm-characteristic profiles.

Table 12: PTU portfolio returns and factor-adjusted alphas.

Strategy	Mean (%)	t-mean	FF6 alpha (%)	t-alpha	Sharpe
Direct PTU high-minus-low	2.83	7.27	3.07	9.75	1.52
Expected-return × PTU DiD	1.14	4.19	0.41	1.34	0.71
Characteristic-managed PTU	2.72	6.85	2.41	7.27	1.25
Long-only high-μ, high-PTU	4.51	10.81	3.88	12.06	2.22

Notes: Returns and alphas are monthly percentages. The high-minus-low portfolio sorts directly on PTU. Expected-return × PTU DiD is the high-minus-low expected-return spread among high-PTU firms minus the same spread among low-PTU firms. Characteristic-managed PTU is the μ-stratified long-short implementation that combines the high-PTU and low-PTU expected-return spread legs. The long-only implementation holds the high-μ, high-PTU corner. FF6 alphas use the six-factor benchmark used throughout the return tests.

Table 13: PTU return premia under value-weighting and investability screens.

Implementation	Mean (%)	t-mean	FF6 alpha (%)	t-alpha	Months
All stocks, equal-weighted	2.81	7.29	3.05	9.77	479
All stocks, value-weighted	0.64	1.98	0.83	4.00	479
Price greater than $5, equal-weighted	0.57	1.98	0.51	3.45	479
Price greater than $5, value-weighted	0.36	1.18	0.46	2.78	479
Above NYSE 20th percentile, equal-weighted	0.56	1.84	0.58	3.56	479
Above NYSE 20th percentile, value-weighted	0.39	1.27	0.52	3.07	479
Above NYSE median size, equal-weighted	0.36	1.20	0.42	2.50	479
Above NYSE median size, value-weighted	0.41	1.34	0.55	3.18	479
Liquid half, equal-weighted	0.65	2.00	0.75	3.95	479
Liquid half, value-weighted	0.44	1.39	0.57	3.21	479

Notes: Rows report monthly percentages for the direct PTU high-minus-low portfolio under equal weighting, value weighting, price screens, NYSE-size screens, and a liquid-half screen. The evidence is interpreted as implementation discipline: PTU premia are largest in the broad cross-section and remain positive under value-weighting, price, size, and liquidity screens.

Table 14: PTU prediction and pricing across firm environments.

Split	Group	Metric	Kurtosis coef.	Kurtosis t	Mean (%)	Mean t	Alpha (%)	Alpha t	Months
Analyst coverage	low	Kurtosis prediction	0.384	9.12					479
Analyst coverage	low	PTU high-minus-low return			3.01	6.86	3.21	8.67	479
Analyst coverage	mid	Kurtosis prediction	0.566	12.76					479
Analyst coverage	mid	PTU high-minus-low return			1.64	4.48	1.84	6.45	479
Analyst coverage	high	Kurtosis prediction	0.583	11.82					479
Analyst coverage	high	PTU high-minus-low return			0.79	2.40	0.95	4.50	479
Institutional ownership	low	Kurtosis prediction	0.099	1.42					479
Institutional ownership	low	PTU high-minus-low return			4.57	9.17	4.78	10.84	479
Institutional ownership	mid	Kurtosis prediction	0.373	6.55					479
Institutional ownership	mid	PTU high-minus-low return			2.12	5.33	2.27	7.23	479
Institutional ownership	high	Kurtosis prediction	0.503	11.97					479
Institutional ownership	high	PTU high-minus-low return			0.96	3.44	0.95	5.01	479
Amihud illiquidity	low	Kurtosis prediction	0.500	10.18					479
Amihud illiquidity	low	PTU high-minus-low return			0.48	1.53	0.51	3.07	479
Amihud illiquidity	mid	Kurtosis prediction	0.478	9.18					479
Amihud illiquidity	mid	PTU high-minus-low return			1.52	3.70	1.68	5.22	479
Amihud illiquidity	high	Kurtosis prediction	0.328	9.30					479
Amihud illiquidity	high	PTU high-minus-low return			5.50	10.75	5.71	11.51	479
Short interest	low	Kurtosis prediction	0.526	8.64					221
Short interest	low	PTU high-minus-low return			3.01	4.38	2.92	5.51	221
Short interest	mid	Kurtosis prediction	0.682	10.76					221
Short interest	mid	PTU high-minus-low return			1.93	3.58	1.76	5.34	221
Short interest	high	Kurtosis prediction	0.634	10.05					221
Short interest	high	PTU high-minus-low return			1.56	2.82	1.38	4.26	221
Option coverage	No	Kurtosis prediction	0.272	5.62					479
Option coverage	No	PTU high-minus-low return			3.93	9.11	4.15	11.56	479
Option coverage	Yes	Kurtosis prediction	0.662	13.83					347
Option coverage	Yes	PTU high-minus-low return			1.39	2.94	1.50	4.60	347
IV30 availability	No	Kurtosis prediction	0.275	5.67					479
IV30 availability	No	PTU high-minus-low return			3.89	9.02	4.12	11.43	479
IV30 availability	Yes	Kurtosis prediction	0.662	13.82					347
IV30 availability	Yes	PTU high-minus-low return			1.43	3.00	1.54	4.70	347

Notes: Rows split the cross-section by information, ownership, liquidity, short-interest, and option-coverage variables. The table characterizes mechanism heterogeneity by showing where PTU is most informative and where the return premium is strongest.

Table 15: Firm characteristics of high-minus-low PTU portfolios.

Characteristic	Q5–Q1	t-stat	Months
Size	-0.292	-27.63	479
Analyst coverage	-0.124	-19.62	479
Institutional ownership	-0.128	-10.52	479
Amihud illiquidity	0.256	22.88	479
Short interest	0.025	1.89	221
IV30	0.595	73.17	347

Notes: Each row reports the time-series average of the monthly high-minus-low PTU quintile spread in the listed characteristic. The table describes the firm environments selected by PTU and complements the mechanism heterogeneity evidence.

Table 16: Economic Framework and Empirical Roadmap

Prediction	Economic logic	Empirical object	Main table
PTU predicts future physical tails	PTU signals latent future physical tail thickness before realized returns are observed.	Future realized kurtosis, high-kurtosis states, and PTU-quintile tail-state spreads.	Tables 3–5.
PTU adds information beyond historical tails	Historical moments are informative but backward-looking; PTU is formed from the ex ante predictive distribution.	Within-firm and change specifications, simple physical-tail benchmarks, spanning regressions, and residual innovation tests.	Tables 6–9.
Option moments provide benchmark discipline	Risk-neutral option moments and physical predictive width are distinct objects, and exact option data are observed in a selected liquid-option universe.	Broad option-surface BKM-style proxy comparisons, exact raw-chain BKM comparisons, and option-sample selection evidence.	Tables 2 and 9–11.
PTU earns compensation	Wide physical tails proxy for exposure that can be costly to bear, hedge, or intermediate.	Direct PTU high-minus-low alphas, characteristic-managed implementations, long-only implementations, and investability screens.	Tables 12–13.
Premia are stronger with frictions	Information, liquidity, monitoring, institutional risk-bearing, and hedging-market frictions raise the cost of absorbing tail exposure.	Subsample pricing patterns across analyst coverage, institutional ownership, illiquidity, short interest, option coverage, and IV30 availability.	Table 14.

Main Figures

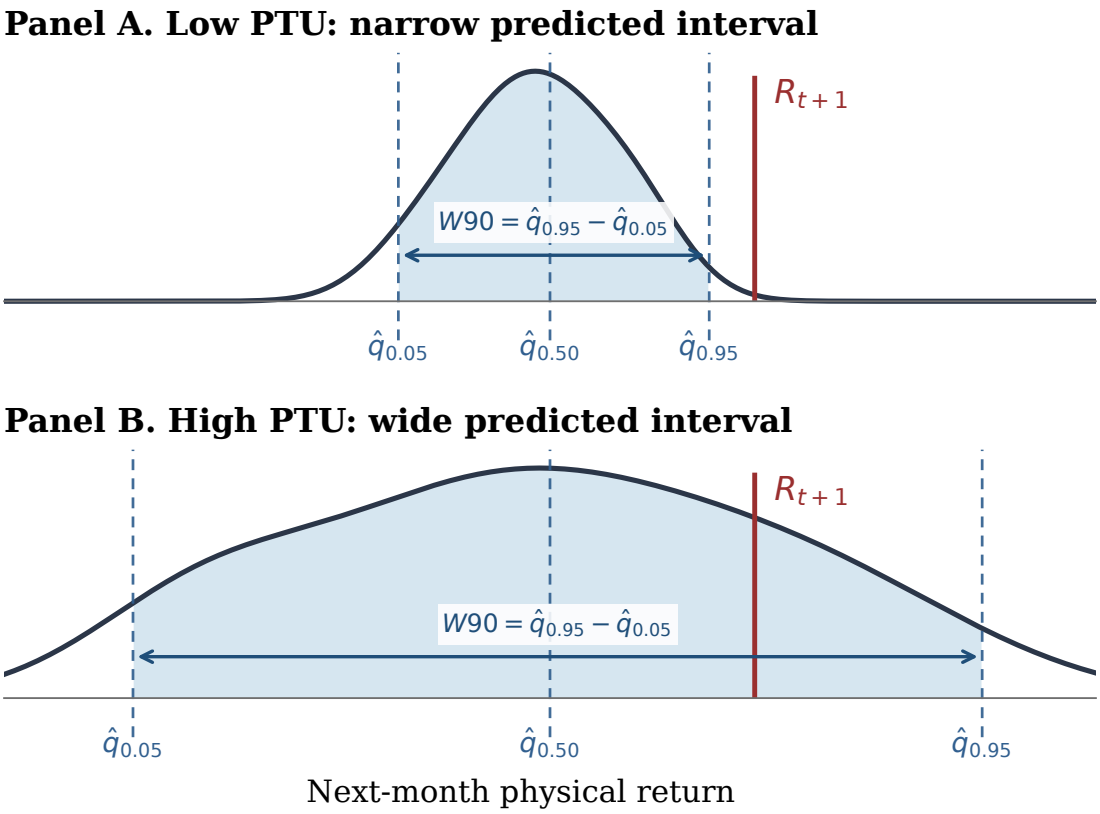


Figure 1: **Schematic illustration of predictive physical tail width.** The shaded region is the predicted 5th-to-95th percentile interval of next-month physical returns, $W90 = \hat{q}_{0.95} - \hat{q}_{0.05}$. Panel A shows a low-PTU stock with a narrow predicted interval; Panel B shows a high-PTU stock with a wider predicted interval. The same realized return R_{t+1} is drawn in both panels to emphasize that PTU ranks ex ante predictive width, not the realized outcome. The curves are generic schematic predictive densities and do not impose normality. The figure is schematic and not drawn to scale.

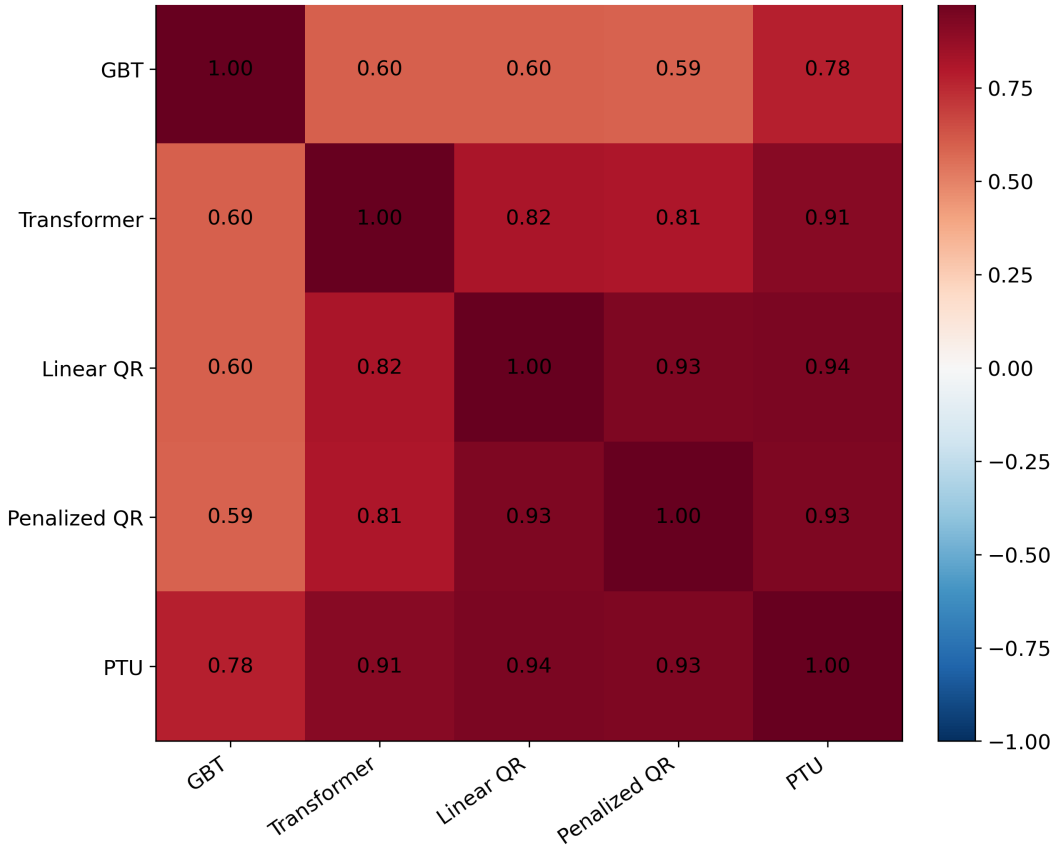


Figure 2: PTU and estimator-specific W90 signals. The heatmap reports correlations among the four demeaned W90 estimator signals and PTU.

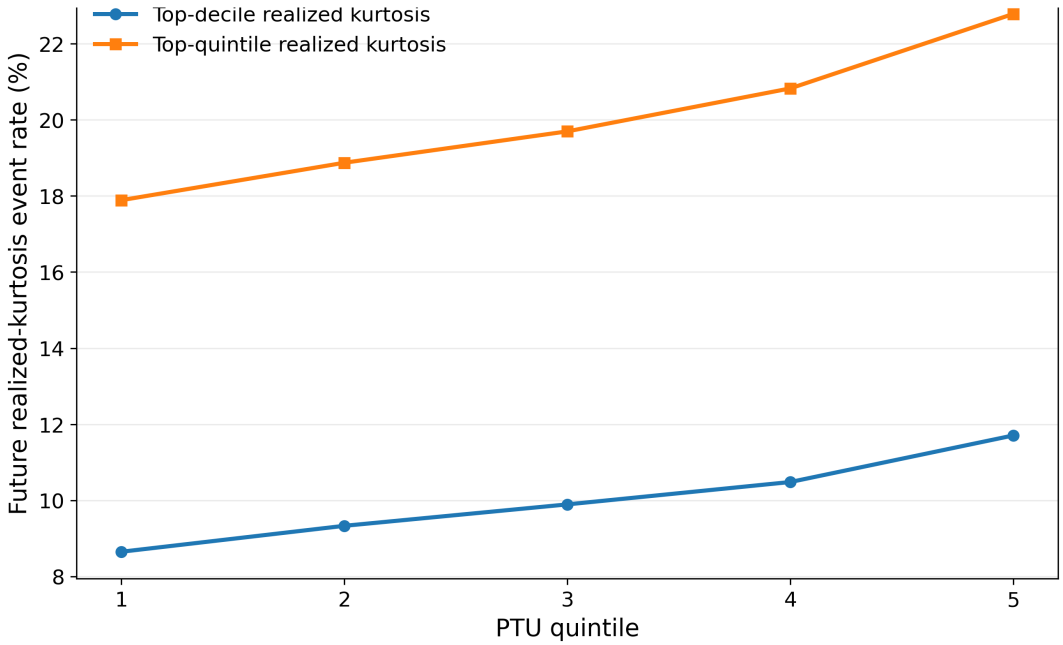


Figure 3: Future high-realized-kurtosis outcomes by PTU quintile. The figure plots the share of firm-months in each PTU quintile that realize top-decile or top-quintile future physical kurtosis.

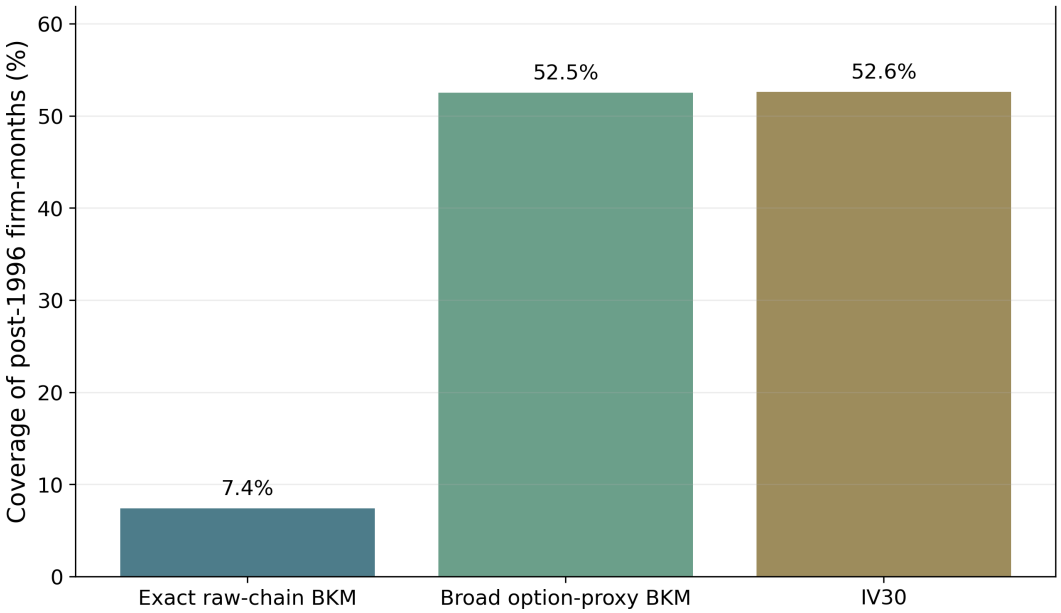


Figure 4: Exact and broad option-benchmark coverage. The figure reports post-1996 coverage for exact raw-chain BKM, broad option-proxy BKM, and IV30.

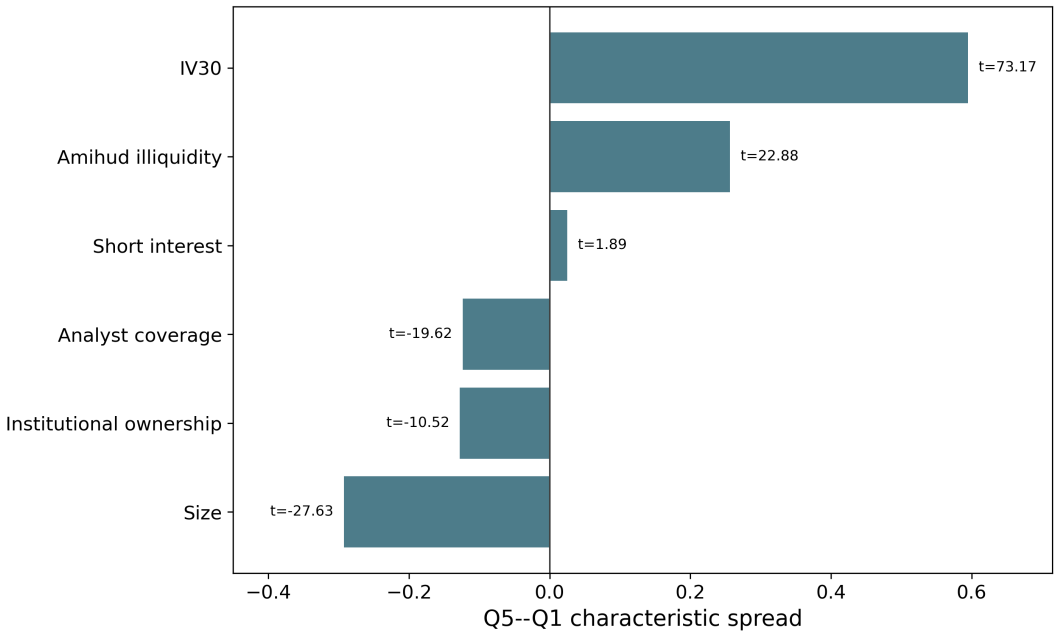


Figure 5: PTU characteristic anatomy. Bars report Q5–Q1 PTU-quintile spreads in firm characteristics. Labels report t-statistics of the monthly spread means.

Internet Appendix

Predictive Physical Tail Width and the Cross-Section of Returns

This Internet Appendix provides model-training details, variable definitions, robustness tables, and supplementary figures. The main manuscript is self-contained; the appendix extends the estimator-specific and implementation evidence behind the aggregate PTU results.

The Internet Appendix deepens the main evidence in four ways. First, it shows that the physical-tail relation appears across the estimator-specific W90 inputs. IA Table IA.1 documents that the four W90 inputs are out-of-sample return-quantile forecasts with realized kurtosis reserved for validation, and IA Table IA.2 separates physical moments, risk-neutral option moments, portfolio variables, and auxiliary consensus variables. The construction and mathematical-definition subsections then state how the benchmark measures, transformations, and portfolio designs are formed. IA Tables IA.3 and IA.4 show that estimator-specific adjusted W90 signals predict future realized physical kurtosis under the relevant firm, volatility, option-overlap, expanded non-option, and macro control blocks.

Second, the appendix refines benchmark scope. IA Table IA.5 evaluates W90 within option-benchmark strata, IA Table IA.6 reports residualized W90 tests after benchmark adjustments, and IA Table IA.7 shows that estimator-specific high-minus-low return evidence is present across nonlinear and linear quantile-prediction approaches. IA Table IA.12 further decomposes the exact raw-chain BKM benchmark relation, clarifying that exact option-chain information is the stricter benchmark in a narrow liquid-option sample.

Third, the appendix sharpens the return interpretation. IA Tables IA.8, IA.10, and IA.11 separate the unconditional high-minus-low, expected-return by W90 difference-in-differences, characteristic-managed, and long-only implementations, so the reader can see which return design carries each pricing claim. IA Table IA.9 summarizes information-environment splits. IA Tables IA.14, IA.15, IA.16, IA.21, IA.25, IA.26, IA.18, and IA.20 treat implementation, cost, horizon, size, weighting, subperiod, industry, and macro-state evidence as scope evidence that sharpens interpretation.

Fourth, the appendix clarifies what the predictive interval-width signal is measuring. IA Tables IA.13 and IA.27 show that unconditional calibration is close to nominal while high-PTU states can under-cover, supporting the interpretation of PTU as a relative physical tail-width ranking. IA Tables IA.17, IA.19, IA.22, IA.23, IA.24, and IA.28 report option and physical benchmark specifications, full realized-kurtosis controls, persistence, alternative physical benchmarks, alternative width families, and alternative consensus variables. IA Table IA.29 adds international external validation by asking whether country-month PTU ranks in global JKP data identify future realized tail states outside the U.S. sample. IA

Figures IA.2, IA.3, IA.4, IA.5, IA.6, IA.7, IA.1, IA.8, IA.9, IA.10, IA.11, IA.12, IA.13, IA.14, and IA.15 visualize investability attenuation, cumulative portfolio evidence, quintile event rates, alpha patterns, rolling estimates, estimator correlations, PTU transitions, estimator-specific characteristic anatomy, realized-kurtosis heat maps, and time variation in alpha and kurtosis prediction. These figures summarize persistence, calibration, implementation, external validity, and scope evidence.

How to Read the Internet Appendix Evidence

IA Table IA.1 verifies that the four W90 inputs are generated from out-of-sample return-quantile forecasts, with realized kurtosis reserved for validation. IA Table IA.2 defines the physical, risk-neutral, portfolio, and auxiliary variables, which keeps the paper's measurement objects distinct. The construction subsections give the measure definitions and construction formulas behind those objects, including realized physical moments, IV30, BKM moments, Amihud illiquidity, the JLN beta, consensus variables, and PTU itself. IA Table IA.3 replaces the common PTU signal with estimator-specific W90 signals and shows that the physical-tail relation appears across estimator-specific W90 signals. IA Table IA.4 repeats that estimator-specific prediction exercise under additional control blocks and option-overlap specifications, which shows where the W90 inputs retain physical-tail content.

IA Table IA.5 evaluates W90 within option-benchmark strata, helping distinguish broad option-proxy overlap from exact raw-chain option coverage. IA Table IA.6 residualizes W90 against benchmark variables and therefore tests the incremental physical-tail information in W90. IA Table IA.7 reports estimator-specific high-minus-low returns; its role is to show that the pricing pattern appears in estimator-specific W90 implementations. IA Table IA.8 decomposes the expected-return-by-width double-sort design, so the stringent DiD evidence can be interpreted separately from the direct PTU high-minus-low evidence. IA Table IA.9 studies information-environment splits and characterizes mechanism heterogeneity. IA Table IA.10 shows how return-regression coefficients evolve as control blocks enter, which helps separate unconditional pricing evidence from controlled return predictability. IA Table IA.11 compares direct, DiD, characteristic-managed, and long-only implementations, clarifying which portfolio design carries each pricing statement. IA Table IA.12 decomposes the exact-BKM benchmark relation and explains why exact option-chain information is treated as the stricter benchmark in a narrow liquid-option sample.

IA Table IA.13 evaluates conditional calibration by PTU quintile; its high-PTU coverage patterns support the rank-based interpretation of PTU as a relative physical tail-width measure. IA Table IA.14 reports coverage and missingness for major control families, which helps

interpret sample changes across specifications. IA Table IA.15 reports break-even trading-cost evidence and serves as implementation discipline rather than a capacity analysis. IA Table IA.16 studies the horizon profile of the return evidence. IA Table IA.17 clarifies the benchmark structure behind option-spanning tests by separating broad option-proxy controls from exact raw-chain BKM controls. IA Table IA.18 evaluates industry concentration, supporting the scope of the cross-sectional interpretation. IA Table IA.19 reports fuller realized-kurtosis specifications and secondary width measures, which shows that the main measurement claim is stable across richer realized-kurtosis specifications. IA Table IA.20 checks macro-state subsamples and therefore evaluates the aggregate-state scope of the PTU relation.

IA Table IA.21 applies a small-stock exclusion and shows that several core results remain while unconditional spreads attenuate, matching the main text's investability language. IA Table IA.22 studies persistence in PTU, separating slow-moving firm exposure from time-varying signal variation. IA Table IA.23 compares PTU and W90 with transparent physical-tail benchmarks, including lagged realized moments, CDG-style moments, BMV expected skewness, and MAX-family lottery measures. IA Table IA.24 evaluates related predictive-width families, clarifying why W90 is the main object. IA Table IA.25 reports value-weighted versions of the portfolio designs and is central for the implementation-discipline interpretation. IA Table IA.26 evaluates time-period stability. IA Table IA.27 reports unconditional quantile calibration, complementing the conditional calibration evidence. IA Table IA.28 shows that portfolio results do not depend on a single expected-return consensus definition.

IA Figure IA.2 visualizes the implementation-screen attenuation of PTU alphas under value-weighting and investability screens. IA Figure IA.3 reports cumulative portfolio evidence as a summary of pricing persistence. IA Figure IA.4 visualizes realized tail-event rates across PTU quintiles, supporting the measurement interpretation. IA Figure IA.5 shows the alpha surface across portfolio designs and helps identify which implementation carries the strongest pricing evidence. IA Figure IA.6 reports rolling portfolio evidence and clarifies time variation in the return relation. IA Figure IA.7 visualizes correlations among estimator-specific W90 signals and PTU, supporting the common-component interpretation. Figure 1 gives the schematic measurement object, and IA Figure IA.1 shows predicted-distribution shape by PTU state. IA Figure IA.8 shows transition behavior in W90 states, which separates persistent state dependence from time-varying tail-width information.

IA Figures IA.9, IA.10, IA.11, and IA.12 report estimator-specific characteristic anatomy, showing how high W90 states load on observable firm environments across models. IA Figure IA.13 visualizes realized-kurtosis evidence across PTU states and supports the core physical-tail validation. IA Figure IA.14 shows rolling alpha evidence and therefore char-

acterizes time-series stability. IA Figure IA.15 shows rolling realized-kurtosis prediction evidence, linking the persistence of return evidence back to the measurement claim.

A Model Training Protocol

This appendix documents how the four estimator-specific W90 measures are trained before they enter PTU. The four estimators are gradient-boosted quantile trees, an FT-Transformer for tabular data, joint linear quantile regression, and elastic-net penalized joint linear quantile regression. They share the same target, feature panel, annual out-of-sample timing, monotone quantile rearrangement, W90 construction, q50-decile demeaning, and monthly rank-normalization.

The key point for interpretation is that the models estimate next-month return quantiles, while realized kurtosis is reserved for validation. Each model is trained to predict the conditional distribution of next-month stock returns. The training panel merges the monthly CRSP return panel with the preprocessed Chen–Zimmermann characteristics. The feature vector contains 213 inputs: 212 rank-normalized characteristic signals plus a missingness-fraction measure. The feature matrix is complete after preprocessing and lies in the unit interval. The training target is the one-month-ahead CRSP return winsorized within month at the first and 99th percentiles. The empirical tests in the paper then evaluate whether the shape of the resulting predictive return distribution forecasts subsequent realized physical tail outcomes.

All four quantile-prediction methods estimate the same five conditional return quantiles,

$$(\hat{q}_{0.05}, \hat{q}_{0.25}, \hat{q}_{0.50}, \hat{q}_{0.75}, \hat{q}_{0.95}),$$

using the same annual out-of-sample timing. For each test year Y, the validation window is the prior 12 calendar years, $Y-12, \ldots, Y-1$, and the test set is year Y. The training window is expanding: it uses the available pre-validation history beginning from the configured training origin, 1965, through $Y - 13$. Thus the first out-of-sample year, 1985, uses a pre-1985 training/validation design, while later test years use progressively more historical data. This protocol fixes the information set before the test-year predictions are formed.

The loss function is the quantile, or pinball, loss. For quantile level τ, prediction $\hat{q}_\tau$, and realized next-month return r, the loss is

$$\ell_\tau(r, \hat{q}_\tau) = \max\{\tau(r - \hat{q}_\tau), (\tau - 1)(r - \hat{q}_\tau)\}.$$

The boosted-tree model fits five separate LightGBM quantile models and tunes hyperpa-

rameters to minimize the sum of validation pinball losses across the five quantiles. The transformer and linear-quantile models fit all five quantiles jointly and minimize the average multi-quantile pinball loss.

The annual fit is two-stage. In phase 1, the model is trained on the expanding training window and evaluated on the 12-year validation window. For the boosted-tree model, phase 1 records the validation-selected number of boosting iterations for each quantile. For the transformer and the linear-quantile models, phase 1 records the validation-selected best epoch. In phase 2, a fresh model is refit on the combined training and validation sample for exactly the selected number of boosting iterations or epochs. Phase 3 predicts the five return quantiles for the test year. This is repeated separately for each out-of-sample year from 1985 through 2024.

After prediction, all estimators use the same post-processing. The five predicted quantiles are monotonically rearranged within observation to remove quantile crossing. The construction then forms $W90 = \hat{q}_{0.95} - \hat{q}_{0.05}$, $W50 = \hat{q}_{0.75} - \hat{q}_{0.25}$, the tail-to-center width ratio $W90/W50$, asymmetry $(\hat{q}_{0.95} - \hat{q}_{0.50}) - (\hat{q}_{0.50} - \hat{q}_{0.05})$, and the center $\hat{q}_{0.50}$. Width-like measures are log-demeaned within month by deciles of the model's own median prediction, signed asymmetry is additively demeaned within the same q50 deciles, and the resulting variables are rank-normalized within month. PTU is then constructed in the main analysis as the equal-weight average of the four demeaned and rank-normalized W90 signals, requiring at least three available estimators, followed by another within-month rank normalization. IA Table IA.1 summarizes the estimator-specific estimation details.

Table IA.1: Training protocol for the four estimator-specific W90 signals used to construct PTU.

Estimator	Model	Loss and target	Training, validation, and tuning	Output used in PTU
Gradient-boosted trees	LightGBM quantile boosting [Ke et al., 2017]. Five independent boosters, one for each quantile.	Monthly target is the winsorized next-month CRSP return, winsorized by month at 1/99 percentiles. Objective is LightGBM quantile loss for each $\tau \in \{.05, .25, .50, .75, .95\}$.	Fixed hyperparameter search, 60 trials, seed 42, on train 1965–1982 and validation 1983–1984. Search space covers learning rate, leaves, minimum leaf size, feature/bagging fractions, bagging frequency, L1/L2 penalties, and tree depth. For each test year, phase 1 trains on the expanding window through $Y - 13$, validates on $Y - 12$ through $Y - 1$, and records validation-selected boosting rounds with 100-round early stopping. Phase 2 refits on train plus validation for the selected rounds.	Predicted quantiles are rearranged, converted to W90, q50-decile demeaned, and rank-normalized. The GBT component of PTU is the GBT W90 demeaned rank.
FT-Transformer	FT-Transformer for tabular data [Gorishniy et al., 2021]. One model with five output heads.	Same winsorized next-month CRSP return target and five quantile levels. Loss is the mean pinball loss over observations and quantile heads.	Fixed hyperparameter search, 60 trials with pruning, seed 42, on train 1965–1982 and validation 1983–1984. Search space covers learning rate, weight decay, dropout, attention dropout, number of layers, token dimension, heads, feed-forward ratio, batch size, warmup fraction, maximum epochs, and patience. Annual phase 1 records the validation-selected best epoch; phase 2 refits a fresh model on train plus validation for exactly that epoch count.	Predicted quantiles are rearranged and processed identically to GBT. The transformer component of PTU is the transformer W90 demeaned rank.

Continued on next page

Estimator	Model	Loss and target	Training, validation, and tuning	Output used in PTU
Linear quantile regression	Joint linear quantile regression following Koenker and Bassett [1978].	Same winsorized next-month CRSP return target. The model outputs all five quantiles jointly and minimizes mean multi-quantile pinball loss.	No automated hyperparameter search. The linear baseline uses fixed hyperparameters: learning rate 10^{-3}, batch size 4096, maximum 150 epochs, early-stop patience 15, no L1 penalty, and a small L2 penalty 10^{-3} for numerical stability. For each test year, phase 1 trains on the expanding window and validates on the prior 12 years; phase 2 refits on train plus validation for the selected best epoch.	Predicted quantiles are rearranged and processed identically. The LQR component of PTU is the LQR W90 demeaned rank.
Penalized linear quantile regression	Elastic-net quantile regression, motivated by high-dimensional penalized quantile regression [Belloni and Chernozhukov, 2011].	Same winsorized next-month CRSP return target and joint multi-quantile pinball loss.	Fixed hyperparameters mirror the LQR implementation except that PLQR adds an L1 penalty of 0.05 and keeps the L2 penalty at 10^{-3}. Optimization uses Adam-style updates, proximal soft-thresholding for L1, validation early stopping, and annual refit on train plus validation for the selected best epoch.	Predicted quantiles are rearranged and processed identically. The PLQR component of PTU is the penalized-QR W90 demeaned rank.

Construction of Benchmark Measures

The benchmark measures used in the paper are constructed under the same timing discipline as PTU. Realized physical skewness and kurtosis are computed for each stock-month from daily CRSP log returns after standard common-share and exchange filters, a minimum daily-observation rule, and cross-sectional monthly winsorization. The main realized-tail outcome is the one-month lead of the winsorized realized physical kurtosis measure. CDG-style historical moments use trailing daily log returns ending at the signal-month end, while MAX and MAX5 use the largest and top-five daily returns over the prior trading-month window. BMV expected skewness is taken from the official predicted-skewness series when available and is cleaned to unique firm-month observations before monthly ranking.

The option-implied measures are separated by construction and sample scope. IV30 is the average of 30-day at-the-money call and put implied volatilities from the OptionMetrics surface, linked to CRSP firms. Broad/proxy BKM moments use 30-day volatility-surface rows, reconstruct option prices from the implied-volatility grid, integrate out-of-the-money call and put curves, and convert the resulting BKM contracts into risk-neutral variance, skewness, and kurtosis. Exact raw-chain BKM moments instead use filtered raw option-chain midpoint quotes on the last available option quote date in the month, keep out-of-the-money calls and puts, standardize BKM contracts to a 30-calendar-day horizon by exact maturity, interpolation, or nearest-maturity fallback, and then map the resulting risk-neutral moments to CRSP firm-months. This stricter construction makes exact BKM the liquid-option benchmark.

W90 and PTU are constructed from out-of-sample predictive quantiles under the same timing discipline. Each estimator predicts next-month return quantiles out of sample. The predicted quantiles are monotonically rearranged within observation, and W90 is the predicted 95th percentile less the predicted 5th percentile. Each W90 signal is adjusted within month by predicted-median-return decile, using log demeaning for positive width measures, and then rank-normalized within month. PTU is the within-month rank of the equal-weight average of the four adjusted W90 ranks, requiring at least three available estimator inputs. The auxiliary expected-return consensus used in portfolio tests is the average active monthly rank across available Chen–Zimmermann characteristics, with a corresponding cross-signal dispersion measure; these variables organize portfolio implementations and remain separate from the PTU construction.

Portfolio returns are formed from signal-month ranks and next-month realized returns. Direct high-minus-low tests sort stocks into monthly PTU quintiles and subtract the lowest quintile from the highest. Managed return tests combine PTU with the expected-return consensus through independent monthly double sorts, characteristic-managed spreads, and

the high-expected-return/high-PTU long-only corner. Equal-weighted tests average stock returns within each portfolio; value-weighted tests use lagged market equity. Factor alphas are estimated from monthly portfolio return series against the six-factor model with HAC/Newey–West inference. Investability screens apply value weighting, a price-above-five-dollar screen, NYSE-size breakpoints, and a liquid-half screen based on Amihud ranks.

Mathematical Definitions of Benchmark Measures

This subsection records the formulas used for the paper's constructed measures. It states the economic and statistical definitions; implementation details are reserved for the replication materials.

For any monthly cross-section t, let $z_{i,t}$ denote a raw measure. The paper's standard winsorized version is

$$z^w_{i,t} = \min\{\max(z_{i,t}, Q_{0.01,t}(z)), Q_{0.99,t}(z)\},$$

where $Q_{p,t}(z)$ is the pth cross-sectional percentile in month t. The rank-normalized version is

$$R_t(z_{i,t}) = \frac{\text{rank}_t(z_{i,t}) - 1}{N^u_t - 1},$$

using dense monthly ranks and the number of distinct nonmissing values N^u_t. Singleton monthly cross-sections are assigned 0.5.

Realized physical moments use daily CRSP log returns $x_{i,d} = \log(1 + r_{i,d})$ within month t, after common-share and exchange filters and a minimum of 15 daily observations. With $\bar{x}_{i,t}$ the monthly mean and

$$m_{k,i,t} = \frac{1}{N_{i,t}} \sum_{d \in t} (x_{i,d} - \bar{x}_{i,t})^k,$$

the realized physical skewness and kurtosis measures are

$$\text{Skew}_{i,t} = \frac{m_{3,i,t}}{m^{3/2}_{2,i,t}}, \qquad \text{Kurt}_{i,t} = \frac{m_{4,i,t}}{m^2_{2,i,t}}.$$

The main realized-tail outcome is $\text{Kurt}^w_{i,t+1}$, the one-month lead of the monthly winsorized realized physical kurtosis measure. Realized volatility is computed from the daily log-return sample variance $s^2_{i,t}$ and scaled to a trading-month horizon as

$$\text{RVol}_{i,t} = \sqrt{21\, s^2_{i,t}}.$$

The CDG-style historical moment controls use the same central standardized skewness and kurtosis formulas on the latest 60 eligible trading days ending at the signal-month end,

requiring at least 30 daily observations. BCW MAX and MAX5 use the latest 21 eligible trading days, requiring at least 15 observations:

$$\mathrm{MAX}_{i,t} = \max_{d \in \mathcal{D}_{21,t}} r_{i,d}, \qquad \mathrm{MAX5}_{i,t} = \frac{1}{5} \sum_{j=1}^{5} r_{i,t}^{(j)},$$

where $r_{i,t}^{(j)}$ is the jth largest daily return in the trailing window. BMV expected skewness is the official Boyer–Mitton–Vorkink predicted-skewness series when available; local fallback estimates are auxiliary; the official series defines the main BMV benchmark.

IV30 is the average of the 30-day OptionMetrics at-the-money call and put implied volatilities:

$$\mathrm{IV30}_{i,t} = \frac{1}{2} \left(\sigma_{i,t}^{C}(\Delta = 0.50) + \sigma_{i,t}^{P}(\Delta = -0.50) \right).$$

For BKM moments, let S be the underlying price, τ the 30-calendar-day maturity, r the zero-coupon rate, and $C(K)$ and $P(K)$ the out-of-the-money call and put prices at strike K. Let $\ell_C(K) = \log(K/S)$ and $\ell_P(K) = \log(S/K)$. Define the call and put components

$$\begin{aligned}
V^C &= \int_S^\infty \frac{2(1 - \ell_C(K))}{K^2} C(K)\, dK, \\
V^P &= \int_0^S \frac{2(1 + \ell_P(K))}{K^2} P(K)\, dK, \\
W^C &= \int_S^\infty \frac{6\ell_C(K) - 3\ell_C(K)^2}{K^2} C(K)\, dK, \\
W^P &= \int_0^S \frac{6\ell_P(K) + 3\ell_P(K)^2}{K^2} P(K)\, dK, \\
X^C &= \int_S^\infty \frac{12\ell_C(K)^2 - 4\ell_C(K)^3}{K^2} C(K)\, dK, \\
X^P &= \int_0^S \frac{12\ell_P(K)^2 + 4\ell_P(K)^3}{K^2} P(K)\, dK.
\end{aligned}$$

The BKM contracts are $V_{i,t} = V^C + V^P$, $W_{i,t} = W^C - W^P$, and $X_{i,t} = X^C + X^P$. With $E = e^{r\tau}$,

$$\mu_{i,t} = E - 1 - \frac{EV_{i,t}}{2} - \frac{EW_{i,t}}{6} - \frac{EX_{i,t}}{24},$$

$$\mathrm{BKMVar}_{i,t} = EV_{i,t} - \mu_{i,t}^2,$$

$$\mathrm{BKMSkew}_{i,t} = \frac{EW_{i,t} - 3\mu_{i,t}EV_{i,t} + 2\mu_{i,t}^3}{\mathrm{BKMVar}_{i,t}^{3/2}},$$

$$\mathrm{BKMKurt}_{i,t} = \frac{EX_{i,t} - 4\mu_{i,t}EW_{i,t} + 6\mu_{i,t}^2 EV_{i,t} - 3\mu_{i,t}^4}{\mathrm{BKMVar}_{i,t}^2}.$$

Broad/proxy BKM uses 30-day OptionMetrics volatility-surface rows, reconstructs option prices from the implied-volatility grid, and applies these integrals to a wider option-related sample. Exact raw-chain BKM uses filtered raw option-chain midpoint quotes on the last available option quote date of the month, keeps out-of-the-money calls and puts, and standardizes the BKM contracts to 30 calendar days by exact maturity, interpolation, or nearest-maturity fallback.

The idiosyncratic-volatility control is the monthly WRDS Fama–French three-factor beta-suite residual volatility formed over a 60-month window. In formula notation, it is the residual scale from

$$r_{i,m} - r_{f,m} = \alpha_i + \beta_{M,i}\,\mathrm{MKT}_m + \beta_{S,i}\,\mathrm{SMB}_m + \beta_{H,i}\,\mathrm{HML}_m + \varepsilon_{i,m},$$

computed by the WRDS beta-suite procedure. Local daily residual-volatility calculations are auxiliary fallback construction; the WRDS beta-suite series defines the main idiosyncratic-volatility control.

Amihud illiquidity is the monthly average of daily absolute return per dollar of trading volume,

$$\mathrm{Amihud}_{i,t} = \frac{1}{N_{i,t}} \sum_{d \in t} \frac{|r_{i,d}|}{|P_{i,d}|\,\mathrm{Vol}_{i,d}},$$

with daily ratios cleaned within month before monthly aggregation. Analyst coverage is the average number of one-year-ahead EPS analysts in the month. Analyst dispersion is $\mathrm{sd}(\widehat{EPS})/|\,\mathrm{mean}(\widehat{EPS})|$ where at least two forecasts and a nontrivial mean forecast are available. Short interest is

$$\mathrm{ShortRatio}_{i,t} = \frac{\mathrm{ShortInterest}_{i,t}}{1000 \times \mathrm{SharesOutstanding}_{i,t}},$$

and institutional ownership is institutional shares divided by CRSP shares outstanding after link-based monthly alignment. The JLN macro-uncertainty beta is the 60-month rolling covariance of firm excess returns with changes in one-month JLN uncertainty divided by the rolling variance of those uncertainty changes.

The auxiliary expected-return consensus uses active monthly ranks from the Chen–Zimmermann characteristic panel. Let $\tilde{c}_{i,t}$ denote the month-t rank-normalized value of

characteristic c, and let $\mathcal{C}_{i,t}$ be the set of available characteristics for firm i in month t. Then

$$\mu_{i,t} = \frac{1}{|\mathcal{C}_{i,t}|} \sum_{c \in \mathcal{C}_{i,t}} \tilde{c}_{i,t}, \qquad \sigma_{i,t} = \left(\frac{1}{|\mathcal{C}_{i,t}|} \sum_{c \in \mathcal{C}_{i,t}} (\tilde{c}_{i,t} - \mu_{i,t})^2 \right)^{1/2}.$$

These variables organize portfolio tests and remain separate from the PTU aggregation rule.

For estimator a, the out-of-sample predictive quantiles are monotonically rearranged and

$$W90_{i,t}^{(a)} = \widehat{Q}_{0.95,i,t+1|t}^{(a)} - \widehat{Q}_{0.05,i,t+1|t}^{(a)}.$$

The quantiles are forecasts of next-month returns conditional on information available at month t. Let $\mathcal{D}_{i,t}^{(a)}$ be the set of firms in firm i's predicted-median-return decile for estimator a in month t. The adjusted width is

$$\widetilde{W}90_{i,t}^{(a)} = \log \left(\frac{W90_{i,t}^{(a)}}{|\mathcal{D}_{i,t}^{(a)}|^{-1} \sum_{j \in \mathcal{D}_{i,t}^{(a)}} W90_{j,t}^{(a)}} \right),$$

and the estimator-specific signal is $R_t(\widetilde{W}90_{i,t}^{(a)})$. With $\mathcal{A}_{i,t}$ the available estimator set and $|\mathcal{A}_{i,t}| \geq 3$,

$$\mathrm{PTU}_{i,t} = R_t \left(\frac{1}{|\mathcal{A}_{i,t}|} \sum_{a \in \mathcal{A}_{i,t}} R_t(\widetilde{W}90_{i,t}^{(a)}) \right).$$

Thus PTU is a fixed equal-weight common physical tail-width rank rather than a performance-weighted forecasting ensemble.

B IA Tables

Table IA.2: Measures and benchmark glossary.

Measure	Construction	Role
PTU	Equal-weight common physical tail-width characteristic formed from four estimator-specific W90 signals; at least three inputs must be available, and the average is rank-normalized within month.	Main prediction, spanning, mechanism, and pricing tests.
W90	Forecasted 90 percent return-quantile width, equal to the predicted 95th percentile less the predicted 5th percentile of next-month returns. Estimator-specific W90 signals are adjusted within month by predicted-median-return deciles before ranking.	Estimator-specific robustness.
Future realized physical kurtosis	Pearson kurtosis of firm daily log returns in the following calendar month, computed as the fourth central moment divided by the squared second central moment.	Primary physical-tail outcome.
Historical physical-tail controls	Lagged realized volatility, realized skewness, realized kurtosis, and 60-trading-day CDG-style historical skewness and kurtosis measures.	Backward-looking benchmark controls.
BMV expected skewness	Official Boyer–Mitton–Vorkink expected idiosyncratic skewness series, aligned to firm-months and rank-normalized for regressions.	Physical skewness benchmark.
BCW MAX and MAX5	Largest daily return and average of the five largest daily returns over the prior month.	Lottery-demand benchmarks.
Exact raw-chain BKM	Risk-neutral variance, skewness, and kurtosis measures constructed from out-of-the-money option prices, standardized to 30 days.	Strict option-overlap benchmark.
Broad/proxy BKM	Option-surface-based BKM-style moments available for a wider option-related sample.	Option-overlap robustness benchmark.
IV30	Thirty-day at-the-money implied volatility from the option surface.	Option-implied volatility benchmark.
Firm and volatility controls	Size, book-to-market, momentum, realized volatility, and idiosyncratic volatility.	Main non-option control block.
Expanded non-option controls	Firm characteristics, volatility controls, historical physical-tail measures, BMV expected skewness, MAX/MAX5, and illiquidity.	Full-universe spanning block without option variables.
Expected-return and scale signals	The portfolio appendix uses an expected-return score and a forecasted return-scale measure to form characteristic-managed and long-only implementations.	Return implementation only.

Notes: The glossary summarizes the economic objects used in the empirical tests. Variables used in option-overlap tests are interpreted within their option-overlap samples.

Table IA.3: Estimator-specific W90 control-block robustness

Signal	Control block	Coef.	t-stat	Months	Avg. N	Avg. R^2
GBT	No controls	0.290	6.45	479	4,685	0.004
GBT	Firm characteristics and volatility	0.327	8.68	479	3,707	0.042
GBT	Firm characteristics, volatility, and macro beta	0.323	8.20	479	2,917	0.042
GBT	Expanded non-option benchmarks	0.296	10.45	468	3,490	0.042
GBT	Expanded non-option benchmarks and macro beta	0.297	9.88	468	2,763	0.040
Transformer	No controls	0.317	8.09	479	4,685	0.004
Transformer	Firm characteristics and volatility	0.238	6.08	479	3,707	0.042
Transformer	Firm characteristics, volatility, and macro beta	0.224	5.27	479	2,917	0.041
Transformer	Expanded non-option benchmarks	0.282	9.76	468	3,490	0.042
Transformer	Expanded non-option benchmarks and macro beta	0.281	9.59	468	2,763	0.040
Linear QR	No controls	0.345	8.21	479	4,685	0.004
Linear QR	Firm characteristics and volatility	0.333	7.85	479	3,707	0.042
Linear QR	Firm characteristics, volatility, and macro beta	0.324	7.26	479	2,917	0.042
Linear QR	Expanded non-option benchmarks	0.331	8.71	468	3,490	0.042
Linear QR	Expanded non-option benchmarks and macro beta	0.334	8.78	468	2,763	0.040
Penalized QR	No controls	0.272	7.26	479	4,685	0.003
Penalized QR	Firm characteristics and volatility	0.185	4.71	479	3,707	0.042
Penalized QR	Firm characteristics, volatility, and macro beta	0.182	4.70	479	2,917	0.041
Penalized QR	Expanded non-option benchmarks	0.204	6.29	468	3,490	0.042
Penalized QR	Expanded non-option benchmarks and macro beta	0.211	7.01	468	2,763	0.040

Notes: Fama-MacBeth regressions of next-month realized physical kurtosis on PTU or estimator-specific W90. Size, value, and momentum includes size, log book-to-market, and 12-2 momentum. Volatility controls are realized volatility and idiosyncratic volatility. Expanded published non-option controls add available historical physical-tail, BMV expected skewness, BCW MAX/MAX5, and Amihud illiquidity controls. The displayed control blocks use firm characteristics, volatility, macro, and published non-option benchmarks. The table reports the positive control-block rows with t-statistics of at least 2.

Table IA.4: **Estimator-specific W90 prediction of future realized physical kurtosis.** Fama–MacBeth regressions of next-month realized physical kurtosis on each estimator-specific adjusted W90 signal. The control blocks are reported as no controls, firm-characteristic controls, firm controls plus option benchmarks in the option-overlap sample, expanded non-option controls, and expanded non-option controls plus macro beta. Each cell is "coef (t)". The key non-option full-control and full+macro specifications are positive and statistically strong across estimators; the firm-controls-plus-option-benchmarks rows make the option-overlap sample and control set explicit. The table reports W90 as the main width measure; secondary widths and quantiles appear in IA Table IA.19. Newey-West t with 6 lags. Out-of-sample prediction window is 1985–2024; N_{months} reports the exact monthly count after required controls. BKM variables are observed for firm-months with usable OptionMetrics option chains; their coverage is naturally lower than the full CRSP/CZ panel, and all BKM rows report their own complete-case month counts. Generic expanded non-option controls specifications exclude option variables; option moments enter in rows explicitly labeled as option-overlap tests. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Estimator	Measure	Spec	Coef (t)	$\bar{R}^2$	N_{months}
GBT	`W90 (GBT)`	No controls	+0.290 (+6.45)***	0.004	479
GBT	`W90 (GBT)`	Firm-characteristic controls	-0.013 (-0.13)	0.030	479
GBT	`W90 (GBT)`	Firm controls plus option benchmarks	+0.127 (+1.60)	0.076	248
GBT	`W90 (GBT)`	Expanded non-option controls	+0.296 (+7.31)***	0.046	479
GBT	`W90 (GBT)`	Expanded non-option controls plus macro beta	+0.299 (+7.06)***	0.046	479
Transformer	`W90 (Transformer)`	No controls	+0.317 (+8.09)***	0.004	479
Transformer	`W90 (Transformer)`	Firm-characteristic controls	-0.060 (-0.60)	0.030	479
Transformer	`W90 (Transformer)`	Firm controls plus option benchmarks	-0.045 (-0.66)	0.076	248
Transformer	`W90 (Transformer)`	Expanded non-option controls	+0.222 (+5.08)***	0.046	479
Transformer	`W90 (Transformer)`	Expanded non-option controls plus macro beta	+0.217 (+4.72)***	0.045	479
Linear QR	`W90 (Linear QR)`	No controls	+0.345 (+8.21)***	0.004	479
Linear QR	`W90 (Linear QR)`	Firm-characteristic controls	-0.079 (-0.70)	0.032	479
Linear QR	`W90 (Linear QR)`	Firm controls plus option benchmarks	-0.068 (-0.97)	0.076	248
Linear QR	`W90 (Linear QR)`	Expanded non-option controls	+0.349 (+7.04)***	0.047	479
Linear QR	`W90 (Linear QR)`	Expanded non-option controls plus macro beta	+0.350 (+7.00)***	0.046	479
Penalized QR	`W90 (Penalized QR)`	No controls	+0.272 (+7.26)***	0.003	479
Penalized QR	`W90 (Penalized QR)`	Firm-characteristic controls	-0.150 (-1.45)	0.031	479
Penalized QR	`W90 (Penalized QR)`	Firm controls plus option benchmarks	-0.535 (-7.74)***	0.077	248
Penalized QR	`W90 (Penalized QR)`	Expanded non-option controls	+0.212 (+4.81)***	0.046	479
Penalized QR	`W90 (Penalized QR)`	Expanded non-option controls plus macro beta	+0.219 (+5.24)***	0.045	479

Table IA.5: Within-BKM-quintile W90 high-minus-low realized-kurtosis spreads.

Estimator	BKM source	W90 measure	BKM kurt Q	Mean kurt H-L	t	N_{months}
GBT	Broad option-proxy overlap	W90 (GBT)	Q1	+0.374	$+8.87^{***}$	347
GBT	Broad option-proxy overlap	W90 (GBT)	Q2	+0.390	$+10.93^{***}$	347
GBT	Broad option-proxy overlap	W90 (GBT)	Q3	+0.392	$+12.04^{***}$	347
GBT	Broad option-proxy overlap	W90 (GBT)	Q4	+0.437	$+15.24^{***}$	347
GBT	Broad option-proxy overlap	W90 (GBT)	Q5	+0.507	$+15.68^{***}$	347
Transformer	Broad option-proxy overlap	W90 (Transformer)	Q1	+0.379	$+10.54^{***}$	347
Transformer	Broad option-proxy overlap	W90 (Transformer)	Q2	+0.384	$+10.47^{***}$	347
Transformer	Broad option-proxy overlap	W90 (Transformer)	Q3	+0.399	$+12.05^{***}$	347
Transformer	Broad option-proxy overlap	W90 (Transformer)	Q4	+0.491	$+15.05^{***}$	347
Transformer	Broad option-proxy overlap	W90 (Transformer)	Q5	+0.531	$+14.15^{***}$	347
Linear QR	Broad option-proxy overlap	W90 (Linear QR)	Q1	+0.424	$+11.09^{***}$	347
Linear QR	Broad option-proxy overlap	W90 (Linear QR)	Q2	+0.410	$+11.12^{***}$	347
Linear QR	Broad option-proxy overlap	W90 (Linear QR)	Q3	+0.439	$+12.51^{***}$	347
Linear QR	Broad option-proxy overlap	W90 (Linear QR)	Q4	+0.510	$+14.31^{***}$	347
Linear QR	Broad option-proxy overlap	W90 (Linear QR)	Q5	+0.552	$+15.81^{***}$	347
Penalized QR	Broad option-proxy overlap	W90 (Penalized QR)	Q1	+0.286	$+7.75^{***}$	347
Penalized QR	Broad option-proxy overlap	W90 (Penalized QR)	Q2	+0.304	$+7.66^{***}$	347
Penalized QR	Broad option-proxy overlap	W90 (Penalized QR)	Q3	+0.349	$+9.26^{***}$	347
Penalized QR	Broad option-proxy overlap	W90 (Penalized QR)	Q4	+0.428	$+12.66^{***}$	347
Penalized QR	Broad option-proxy overlap	W90 (Penalized QR)	Q5	+0.475	$+12.51^{***}$	347
GBT	Exact raw-chain option overlap	W90 (GBT)	Q1	+0.558	$+5.30^{***}$	180
GBT	Exact raw-chain option overlap	W90 (GBT)	Q2	+0.305	$+4.23^{***}$	179
GBT	Exact raw-chain option overlap	W90 (GBT)	Q3	+0.302	$+4.08^{***}$	179
GBT	Exact raw-chain option overlap	W90 (GBT)	Q4	+0.225	$+3.32^{***}$	179
GBT	Exact raw-chain option overlap	W90 (GBT)	Q5	+0.375	$+5.71^{***}$	180
Transformer	Exact raw-chain option overlap	W90 (Transformer)	Q1	+0.279	$+3.02^{***}$	180
Transformer	Exact raw-chain option overlap	W90 (Transformer)	Q2	+0.077	+0.98	179
Transformer	Exact raw-chain option overlap	W90 (Transformer)	Q3	+0.121	+1.56	179
Transformer	Exact raw-chain option overlap	W90 (Transformer)	Q4	+0.176	$+2.67^{***}$	179
Transformer	Exact raw-chain option overlap	W90 (Transformer)	Q5	+0.254	$+3.61^{***}$	180
Linear QR	Exact raw-chain option overlap	W90 (Linear QR)	Q1	+0.403	$+4.92^{***}$	180
Linear QR	Exact raw-chain option overlap	W90 (Linear QR)	Q2	+0.245	$+3.90^{***}$	179
Linear QR	Exact raw-chain option overlap	W90 (Linear QR)	Q3	+0.165	$+1.77^{*}$	179
Linear QR	Exact raw-chain option overlap	W90 (Linear QR)	Q4	+0.159	$+2.59^{***}$	179
Linear QR	Exact raw-chain option overlap	W90 (Linear QR)	Q5	+0.252	$+3.58^{***}$	180
Penalized QR	Exact raw-chain option overlap	W90 (Penalized QR)	Q1	-0.014	-0.19	180
Penalized QR	Exact raw-chain option overlap	W90 (Penalized QR)	Q2	+0.003	+0.04	179
Penalized QR	Exact raw-chain option overlap	W90 (Penalized QR)	Q3	-0.014	-0.18	179
Penalized QR	Exact raw-chain option overlap	W90 (Penalized QR)	Q4	+0.003	+0.04	179
Penalized QR	Exact raw-chain option overlap	W90 (Penalized QR)	Q5	+0.177	$+2.31^{**}$	180

Notes: Rows sort firms within BKM-kurtosis quintiles and report the high-minus-low W90 spread in future realized physical kurtosis. Broad/proxy and exact raw-chain option-overlap samples are kept separate. Newey–West t-statistics use six lags where applicable.

Table IA.6: Residual information-content tests.

Estimator	Sample	Resid. measure	Spec	Coef (t)	$\bar{R}^2$	N_{months}
GBT	Full universe, non-option	W90 residual on historical kurtosis (GBT)	Residual W90 with expanded non-option controls	+0.148 (+3.53)***	0.045	479
GBT	Full universe, non-option	Historical kurtosis residual on W90 (GBT)	Residual historical kurtosis with expanded non-option controls	+0.780 (+9.96)***	0.055	479
Transformer	Full universe, non-option	W90 residual on historical kurtosis (Transformer)	Residual W90 with expanded non-option controls	+0.068 (+1.54)	0.045	479
Transformer	Full universe, non-option	Historical kurtosis residual on W90 (Transformer)	Residual historical kurtosis with expanded non-option controls	+0.781 (+9.91)***	0.056	479
Linear QR	Full universe, non-option	W90 residual on historical kurtosis (Linear QR)	Residual W90 with expanded non-option controls	+0.166 (+3.34)***	0.045	479
Linear QR	Full universe, non-option	Historical kurtosis residual on W90 (Linear QR)	Residual historical kurtosis with expanded non-option controls	+0.773 (+9.75)***	0.055	479
Penalized QR	Full universe, non-option	W90 residual on historical kurtosis (Penalized QR)	Residual W90 with expanded non-option controls	+0.059 (+1.35)	0.045	479
Penalized QR	Full universe, non-option	Historical kurtosis residual on W90 (Penalized QR)	Residual historical kurtosis with expanded non-option controls	+0.781 (+9.90)***	0.056	479
GBT	Broad option-proxy overlap	W90 residual on broad option controls (GBT)	Option-overlap residual, no controls	+0.213 (+7.35)***	0.002	347
GBT	Broad option-proxy overlap	W90 residual on broad option controls (GBT)	Option-overlap residual plus firm controls	+0.234 (+7.91)***	0.020	347
GBT	Broad option-proxy overlap	W90 residual on broad option controls (GBT)	Option-overlap residual plus expanded non-option controls	+0.192 (+7.81)***	0.028	347
Transformer	Broad option-proxy overlap	W90 residual on broad option controls (Transformer)	Option-overlap residual, no controls	+0.215 (+9.02)***	0.002	347
Transformer	Broad option-proxy overlap	W90 residual on broad option controls (Transformer)	Option-overlap residual plus firm controls	+0.235 (+10.20)***	0.020	347
Transformer	Broad option-proxy overlap	W90 residual on broad option controls (Transformer)	Option-overlap residual plus expanded non-option controls	+0.182 (+8.52)***	0.028	347
Linear QR	Broad option-proxy overlap	W90 residual on broad option controls (Linear QR)	Option-overlap residual, no controls	+0.239 (+9.13)***	0.002	347
Linear QR	Broad option-proxy overlap	W90 residual on broad option controls (Linear QR)	Option-overlap residual plus firm controls	+0.299 (+10.71)***	0.020	347
Linear QR	Broad option-proxy overlap	W90 residual on broad option controls (Linear QR)	Option-overlap residual plus expanded non-option controls	+0.278 (+9.93)***	0.028	347
Penalized QR	Broad option-proxy overlap	W90 residual on broad option controls (Penalized QR)	Option-overlap residual, no controls	+0.100 (+3.76)***	0.002	347
Penalized QR	Broad option-proxy overlap	W90 residual on broad option controls (Penalized QR)	Option-overlap residual plus firm controls	+0.161 (+5.79)***	0.020	347
Penalized QR	Broad option-proxy overlap	W90 residual on broad option controls (Penalized QR)	Option-overlap residual plus expanded non-option controls	+0.165 (+5.80)***	0.028	347
GBT	Exact raw-chain option overlap	W90 residual on exact option controls (GBT)	Option-overlap residual, no controls	+0.093 (+1.70)*	0.006	252
GBT	Exact raw-chain option overlap	W90 residual on exact option controls (GBT)	Option-overlap residual plus firm controls	+0.107 (+1.91)*	0.054	248
GBT	Exact raw-chain option overlap	W90 residual on exact option controls (GBT)	Option-overlap residual plus expanded non-option controls	+0.014 (+0.30)	0.075	248
Transformer	Exact raw-chain option overlap	W90 residual on exact option controls (Transformer)	Option-overlap residual, no controls	-0.036 (-0.65)	0.006	252
Transformer	Exact raw-chain option overlap	W90 residual on exact option controls (Transformer)	Option-overlap residual plus firm controls	-0.036 (-0.84)	0.054	248
Transformer	Exact raw-chain option overlap	W90 residual on exact option controls (Transformer)	Option-overlap residual plus expanded non-option controls	-0.066 (-1.66)*	0.075	248
Linear QR	Exact raw-chain option overlap	W90 residual on exact option controls (Linear QR)	Option-overlap residual, no controls	-0.074 (-1.45)	0.006	252
Linear QR	Exact raw-chain option overlap	W90 residual on exact option controls (Linear QR)	Option-overlap residual plus firm controls	-0.047 (-1.00)	0.054	248
Linear QR	Exact raw-chain option overlap	W90 residual on exact option controls (Linear QR)	Option-overlap residual plus expanded non-option controls	-0.036 (-0.74)	0.075	248
Penalized QR	Exact raw-chain option overlap	W90 residual on exact option controls (Penalized QR)	Option-overlap residual, no controls	-0.366 (-9.01)***	0.008	252
Penalized QR	Exact raw-chain option overlap	W90 residual on exact option controls (Penalized QR)	Option-overlap residual plus firm controls	-0.358 (-7.94)***	0.055	248
Penalized QR	Exact raw-chain option overlap	W90 residual on exact option controls (Penalized QR)	Option-overlap residual plus expanded non-option controls	-0.282 (-6.20)***	0.075	248

Notes: Rows report residualized realized-kurtosis tests estimating whether W90 contains information beyond historical physical-tail benchmarks or option-implied benchmarks in the relevant overlap sample. Newey–West t-statistics use six lags where applicable.

Table IA.7: **Estimator independence.** Panel A: estimator-specific EW long-short portfolio on adjusted W90 signal. All four estimators (LightGBM, FT-Transformer, Koenker-Bassett 1978 linear QR, Belloni-Chernozhukov 2011 elastic-net penalized linear QR) yield comparable Sharpe ratios. Panel B: pairwise correlation of monthly EW H-L returns across all six estimator pairs. High cross-estimator return correlations indicate that portfolio returns are similar across estimators. Mean and α in monthly percent. Out-of-sample prediction window is 1985–2024. Newey-West t with 6 lags. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Panel A: Estimator-specific H-L performance

Estimator	H-L mean	t	FF6 α	$t(\alpha)$	Sharpe	N_{months}
GBT	+2.03%	+5.93***	+2.27%	+8.64***	+1.27	479
Transformer	+2.67%	+7.32***	+2.87%	+9.44***	+1.58	479
Linear QR	+2.51%	+7.30***	+2.72%	+9.72***	+1.44	479
Penalized QR	+2.36%	+7.31***	+2.48%	+9.31***	+1.41	479

Panel B: Cross-estimator correlation of monthly H-L returns

	GBT	Transformer	Linear QR	Penalized QR
GBT	+1.000	+0.949	+0.952	+0.935
Transformer	+0.949	+1.000	+0.977	+0.964
Linear QR	+0.952	+0.977	+1.000	+0.986
Penalized QR	+0.935	+0.964	+0.986	+1.000

Table IA.8: **Stambaugh-Yu-Yuan-style DiD ($\mu \times W$).** Panel A: 5×5 double-sort DiD portfolio in the spirit of Stambaugh-Yu-Yuan (2015). Each month, firms are independently sorted into 5 μ-quintiles (rows ordered from low to high expected-return consensus) and 5 W-quintiles (columns ordered from low to high W90). The DiD compares the high-μ minus low-μ spread among high-W firms with the same spread among low-W firms. The main table reports the main expected-return consensus measure μ under both EW and VW; alternative consensus definitions are reported in IA Table IA.28. Panel B: within-μ-quintile estimator-specific high-minus-low portfolios for the low- and high-μ quintiles, decomposing the DiD into its boundary spreads. All in monthly percent. Out-of-sample prediction window is 1985–2024; N_{months} reports exact monthly coverage. Newey-West t with 6 lags. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Panel A: Expected-return-by-W90 difference-in-differences portfolio returns

			DiD return		FF6 α			
Estimator	W90 measure	μ var	Mean	t	α	$t(\alpha)$	Sharpe	N_{months}
GBT	`W90 (GBT)`	μ (EW)	+1.99%	+7.83***	+1.52%	+5.35***	+1.44	479
	`W90 (GBT)`	μ (VW)	+0.79%	+2.96***	+0.70%	+2.45**	+0.46	479
Transformer	`W90 (Transformer)`	μ (EW)	+0.57%	+2.18**	-0.04%	-0.12	+0.40	479
	`W90 (Transformer)`	μ (VW)	+0.17%	+0.57	-0.17%	-0.58	+0.09	479
Linear QR	`W90 (Linear QR)`	μ (EW)	+1.19%	+4.50***	+0.45%	+1.46	+0.77	479
	`W90 (Linear QR)`	μ (VW)	+1.16%	+3.58***	+0.78%	+2.20**	+0.63	479
Penalized QR	`W90 (Penalized QR)`	μ (EW)	+1.18%	+4.53***	+0.51%	+1.69*	+0.78	479
	`W90 (Penalized QR)`	μ (VW)	+1.05%	+3.29***	+0.64%	+1.96**	+0.53	479

Panel B: Estimator-specific high-minus-low within μ-quintile

Estimator	W90 measure	μ Q	Mean H-L	t	N_{months}
GBT	`W90 (GBT)`	Q1	+1.90%	+5.70***	479
GBT	`W90 (GBT)`	Q5	+2.83%	+8.40***	479
Transformer	`W90 (Transformer)`	Q1	+2.65%	+7.54***	479
Transformer	`W90 (Transformer)`	Q5	+3.14%	+8.40***	479
Linear QR	`W90 (Linear QR)`	Q1	+1.83%	+5.28***	479
Linear QR	`W90 (Linear QR)`	Q5	+3.18%	+9.39***	479
Penalized QR	`W90 (Penalized QR)`	Q1	+1.71%	+5.20***	479
Penalized QR	`W90 (Penalized QR)`	Q5	+2.96%	+9.09***	479

Table IA.9: **W90 high-minus-low by information environment.** Within each tercile of an information-environment proxy, firms are sorted into 5 quintiles on adjusted W90 signal and the EW Q5–Q1 return is averaged across months. The table reports the low and high terciles for clarity; the mid tercile lies between them. The H-L spread is largest in the low-information-environment tercile (small firms, low analyst coverage, low institutional ownership) and remains significant in every cell, consistent with limits-to-arbitrage variation while retaining a broad W90 pricing relation. Mean H-L in monthly percent. Sample is 1985–2024 subject to information-environment proxy availability. Newey-West t with 6 lags. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Estimator	W90 measure	Split var	Tercile	Mean H-L	t	N_{months}
GBT	`W90 (GBT)`	`Analyst coverage`	low	+2.02%	+4.99***	479
GBT	`W90 (GBT)`	`Analyst coverage`	high	+0.33%	+1.18	479
Transformer	`W90 (Transformer)`	`Analyst coverage`	low	+2.95%	+6.96***	479
Transformer	`W90 (Transformer)`	`Analyst coverage`	high	+0.74%	+2.34**	479
Linear QR	`W90 (Linear QR)`	`Analyst coverage`	low	+2.73%	+6.92***	479
Linear QR	`W90 (Linear QR)`	`Analyst coverage`	high	+0.93%	+3.06***	479
Penalized QR	`W90 (Penalized QR)`	`Analyst coverage`	low	+2.49%	+6.84***	479
Penalized QR	`W90 (Penalized QR)`	`Analyst coverage`	high	+0.87%	+2.96***	479
GBT	`W90 (GBT)`	`Size`	low	+4.98%	+9.77***	479
GBT	`W90 (GBT)`	`Size`	high	+0.00%	+0.02	479
Transformer	`W90 (Transformer)`	`Size`	low	+5.21%	+9.64***	479
Transformer	`W90 (Transformer)`	`Size`	high	+0.37%	+1.26	479
Linear QR	`W90 (Linear QR)`	`Size`	low	+4.48%	+10.18***	479
Linear QR	`W90 (Linear QR)`	`Size`	high	+0.62%	+2.22**	479
Penalized QR	`W90 (Penalized QR)`	`Size`	low	+4.14%	+10.02***	479
Penalized QR	`W90 (Penalized QR)`	`Size`	high	+0.61%	+2.26**	479

Table IA.10: **FM return regressions on W90 with progressive controls.** Fama-MacBeth cross-sectional regressions of forward 1-month stock returns on each estimator-specific adjusted W90 signal with progressively-added controls. The regression is *unconditional* on μ and estimates W90 return predictability averaged across the entire cross-section (compare Table IA.8, which conditions on μ). Univariate = W90 alone; firm-characteristic controls adds {size, log B/M, mom 12-2, μ, σ}; +OPTION adds {IV30, BKM kurt, BKM skew}; expanded non-option controls adds {rvol, ivol}; +MACRO adds {JLN macro β, analyst dispersion, analyst coverage}. Each cell "coef (t)". Returns are monthly decimals in the regression; sample is 1985–2024 subject to control availability. Newey-West t with 6 lags. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Estimator	W90 measure	Spec	Coef (t)	$\bar{R}^2$	N_{months}
GBT	`W90 (GBT)`	No controls	+0.011 (+2.73)***	0.014	479
GBT	`W90 (GBT)`	Firm-characteristic controls	+0.017 (+5.87)***	0.041	479
GBT	`W90 (GBT)`	Firm controls plus option benchmarks	-0.000 (-0.08)	0.137	248
GBT	`W90 (GBT)`	Expanded non-option controls	+0.017 (+7.90)***	0.048	479
GBT	`W90 (GBT)`	Expanded non-option controls plus macro beta	+0.015 (+7.58)***	0.052	479
Transformer	`W90 (Transformer)`	No controls	+0.020 (+4.72)***	0.016	479
Transformer	`W90 (Transformer)`	Firm-characteristic controls	+0.027 (+7.84)***	0.045	479
Transformer	`W90 (Transformer)`	Firm controls plus option benchmarks	+0.011 (+2.15)**	0.138	248
Transformer	`W90 (Transformer)`	Expanded non-option controls	+0.033 (+10.41)***	0.051	479
Transformer	`W90 (Transformer)`	Expanded non-option controls plus macro beta	+0.030 (+10.20)***	0.054	479
Linear QR	`W90 (Linear QR)`	No controls	+0.020 (+4.83)***	0.016	479
Linear QR	`W90 (Linear QR)`	Firm-characteristic controls	+0.025 (+7.60)***	0.046	479
Linear QR	`W90 (Linear QR)`	Firm controls plus option benchmarks	+0.012 (+2.16)**	0.139	248
Linear QR	`W90 (Linear QR)`	Expanded non-option controls	+0.033 (+11.32)***	0.051	479
Linear QR	`W90 (Linear QR)`	Expanded non-option controls plus macro beta	+0.032 (+11.33)***	0.055	479
Penalized QR	`W90 (Penalized QR)`	No controls	+0.019 (+4.89)***	0.016	479
Penalized QR	`W90 (Penalized QR)`	Firm-characteristic controls	+0.023 (+7.20)***	0.045	479
Penalized QR	`W90 (Penalized QR)`	Firm controls plus option benchmarks	+0.011 (+1.92)*	0.141	248
Penalized QR	`W90 (Penalized QR)`	Expanded non-option controls	+0.029 (+10.91)***	0.051	479
Penalized QR	`W90 (Penalized QR)`	Expanded non-option controls plus macro beta	+0.029 (+11.13)***	0.055	479

Table IA.11: **Strategy comparison: long-only vs DiD vs μ-stratified.** EW long-short and long-only strategies on $\mu \times W$ double-sorts, all four estimators. (a) W **Q5–Q1 (unconditional)**: long high-W, short low-W, without conditioning on μ. (b) **DiD long-leg**: μ Q5 – μ Q1 within high-W slice; this is the upper half of the $\mu \times W$ DiD. (c) **DiD short-leg**: μ Q5 – μ Q1 within low-W slice. (d) **μ-stratified L/S**: sum of (b) and (c), not the average — a tail-symmetric portfolio. (e) **Long-only (μ_5, W_5)**: long the top-right corner of the $\mu \times W$ grid only — the long-only implementation. The table decomposes return implementations used in the portfolio analysis. The long-only corner portfolio has the largest raw return, while the DiD construction is the most tightly hedged implementation. Mean and α in monthly percent. EW-only (VW reported in IA Table IA.25). Newey-West t with 6 lags. Out-of-sample prediction window is 1985–2024. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

		Mean return		FF6 α			
Estimator	Strategy	Mean	t	α	$t(\alpha)$	Sharpe	N
GBT	W Q5–Q1 (unconditional)	+2.05%	+5.92***	+2.29%	+8.66***	+1.28	479
	DiD long-leg (high-W μ spread)	+3.01%	+9.05***	+2.68%	+8.92***	+1.90	479
	DiD short-leg (low-W μ spread)	+1.02%	+5.75***	+1.16%	+7.95***	+1.16	479
	μ-stratified L/S (long + short legs)	+4.03%	+8.59***	+3.84%	+10.18***	+1.87	479
	Long-only (μ_5, W_5)	+4.71%	+10.95***	+4.12%	+11.81***	+2.39	479
Transformer	W Q5–Q1 (unconditional)	+2.71%	+7.32***	+2.90%	+9.39***	+1.59	479
	DiD long-leg (high-W μ spread)	+1.62%	+6.03***	+1.16%	+4.32***	+1.05	479
	DiD short-leg (low-W μ spread)	+1.05%	+6.36***	+1.20%	+8.84***	+1.22	479
	μ-stratified L/S (long + short legs)	+2.67%	+7.37***	+2.37%	+7.99***	+1.30	479
	Long-only (μ_5, W_5)	+4.33%	+10.71***	+3.69%	+11.99***	+2.19	479
Linear QR	W Q5–Q1 (unconditional)	+2.54%	+7.29***	+2.75%	+9.71***	+1.45	479
	DiD long-leg (high-W μ spread)	+1.95%	+7.10***	+1.42%	+5.40***	+1.19	479
	DiD short-leg (low-W μ spread)	+0.76%	+4.46***	+0.97%	+6.94***	+0.83	479
	μ-stratified L/S (long + short legs)	+2.71%	+7.26***	+2.39%	+8.35***	+1.25	479
	Long-only (μ_5, W_5)	+4.09%	+10.88***	+3.42%	+12.73***	+2.11	479
Penalized QR	W Q5–Q1 (unconditional)	+2.39%	+7.28***	+2.50%	+9.26***	+1.42	479
	DiD long-leg (high-W μ spread)	+1.94%	+6.89***	+1.42%	+5.29***	+1.17	479
	DiD short-leg (low-W μ spread)	+0.76%	+4.59***	+0.91%	+7.16***	+0.85	479
	μ-stratified L/S (long + short legs)	+2.70%	+7.08***	+2.34%	+7.93***	+1.22	479
	Long-only (μ_5, W_5)	+3.89%	+10.61***	+3.19%	+12.30***	+2.05	479

Table IA.12: Exact-BKM benchmark scope by estimator, period, industry, and size.

Split	Measure	Coef.	t-stat	Months	Avg. N
Estimator: GBT	GBT W90	-0.034	-0.40	282	335
Estimator: Transformer	Transformer W90	-0.097	-0.98	282	335
Estimator: Linear QR	Linear QR W90	-0.112	-1.31	282	335
Estimator: Penalized QR	Penalized QR W90	-0.396	-4.00	282	335
Period: 2007 2015	PTU	-0.499	-3.31	108	278
Period: 2016 2024	PTU	-0.432	-4.06	96	617
Size: above nyse50	PTU	-0.225	-1.98	223	180
Size: below nyse50	PTU	-0.472	-4.19	248	214
Industry: nonfinancials	PTU	-0.330	-2.92	275	306

Notes: Rows repeat the exact raw-chain BKM overlap specification across estimator-specific W90 measures and selected subsamples. The table documents where the exact-option benchmark relation is concentrated.

Table IA.13: Conditional quantile calibration by PTU quintile.

Estimator	PTU quintile	N	$P(r \leq q_{05})$	$P(r \leq q_{95})$	90% interval	Error
GBT	1	452,663	5.13	95.73	90.60	0.60
GBT	2	452,372	5.62	95.20	89.57	-0.43
GBT	3	452,394	6.44	94.73	88.29	-1.71
GBT	4	452,372	7.29	94.04	86.75	-3.25
GBT	5	452,580	7.78	92.60	84.82	-5.18
Transformer	1	452,663	5.68	95.54	89.86	-0.14
Transformer	2	452,372	5.43	95.66	90.23	0.23
Transformer	3	452,394	5.77	95.46	89.70	-0.30
Transformer	4	452,372	6.19	95.09	88.90	-1.10
Transformer	5	452,580	6.68	94.04	87.37	-2.63
Linear QR	1	452,663	4.86	95.55	90.69	0.69
Linear QR	2	452,372	4.90	95.94	91.03	1.03
Linear QR	3	452,394	5.83	95.45	89.62	-0.38
Linear QR	4	452,372	6.70	94.71	88.01	-1.99
Linear QR	5	452,580	7.14	92.89	85.75	-4.25
Penalized QR	1	452,663	5.22	96.06	90.84	0.84
Penalized QR	2	452,372	5.27	96.20	90.93	0.93
Penalized QR	3	452,394	6.24	95.58	89.34	-0.66
Penalized QR	4	452,372	7.10	94.69	87.59	-2.41
Penalized QR	5	452,580	7.45	92.66	85.20	-4.80

Notes: The table compares realized next-month return coverage to the predicted 5th and 95th percentiles by PTU quintile. High-PTU undercoverage means realized returns fall outside nominal intervals more often than the unconditional calibration would imply.

Table IA.14: **Controls coverage.** Per-control overall coverage (fraction of firm-months with non-null observation), first year with ≥5% coverage, and distributional summary. Option-implied controls are measured on the OptionMetrics overlap and enter as option-overlap variables. `log_bm` is computed only when book-to-market is strictly positive. Sample is the analysis panel through 2024, before table-specific complete-case filters.

Control	Coverage	First yr (≥5%)	Mean	SD	p1	p99
`ivol_n`	84.0%	1963	+0.499	0.288	+0.010	+0.990
`mom_12_2`	91.5%	1963	+0.167	0.761	-0.753	+2.614
`jln_beta_w_n`	61.9%	1968	+0.500	0.294	+0.000	+1.000
`Size`	99.2%	1963	+0.499	0.289	+0.010	+0.990
σ	95.6%	1963	+0.500	0.288	+0.010	+0.990
`log_bm`	81.4%	1963	-0.639	1.024	-3.812	+1.435
`rvol_n`	99.4%	1963	+0.496	0.290	+0.003	+0.990
μ	95.6%	1963	+0.500	0.288	+0.010	+0.990

Table IA.15: **Break-even transaction costs (measured turnover).** Break-even round-trip transaction cost (in bp) for each of the three portfolio implementations on each of the four estimators, using *measured* monthly turnover from holdings rebalancing rather than an assumed turnover rate. Monthly turnover = average fraction of long-or-short holdings that change month-to-month (symmetric difference / union, averaged across rebalances). Long-only (μ_5, W_5) has 1 leg; long-short strategies have 2 legs. Break-even cost equals α divided by turnover times the number of legs. Break-even costs are large for the long-only and characteristic-managed implementations. The DiD implementation is more design-sensitive and is reported as implementation evidence rather than as a capacity analysis. Mean and α in monthly percent. Sample is 1985–2024; N reports monthly return observations.

Estimator	Strategy	Monthly α	Annual α	Monthly turnover	Break-even (bp/trip)	N
GBT	DiD	+1.52%	18.3%	70.2%	108.6	479
GBT	Long-only (μ_5, W_5)	+4.12%	49.4%	74.6%	552.2	479
GBT	μ-stratified L/S	+3.84%	46.1%	66.9%	287.3	479
Transformer	DiD	-0.04%	-0.5%	71.3%	-2.7	479
Transformer	Long-only (μ_5, W_5)	+3.69%	44.3%	73.0%	505.6	479
Transformer	μ-stratified L/S	+2.37%	28.4%	71.2%	166.1	479
Linear QR	DiD	+0.45%	5.4%	71.5%	31.6	479
Linear QR	Long-only (μ_5, W_5)	+3.42%	41.0%	71.6%	477.8	479
Linear QR	μ-stratified L/S	+2.39%	28.6%	71.6%	166.8	479
Penalized QR	DiD	+0.51%	6.1%	73.8%	34.5	479
Penalized QR	Long-only (μ_5, W_5)	+3.19%	38.3%	74.4%	428.4	479
Penalized QR	μ-stratified L/S	+2.34%	28.1%	73.4%	159.4	479

Table IA.16: **Holding-period robustness.** DiD ($\mu \times W$) returns at holding horizons $h \in \{1, 3, 6, 12\}$ months for all four estimators. Mean returns remain positive across horizons, with factor-adjusted alphas becoming smaller for several estimators at longer horizons. Mean and α are cumulative horizon returns in percent, with FF6 factors summed over the same forward horizon. Sample is the 1985–2024 out-of-sample ML window subject to horizon and factor availability. Newey-West t with 6 lags. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Estimator	Horizon	Mean	t	FF6 α	$t(\alpha)$	N_{months}
GBT	1m	+1.99%	+7.83***	+1.52%	+5.35***	479
GBT	3m	+5.46%	+9.38***	+4.17%	+6.26***	477
GBT	6m	+8.61%	+9.10***	+6.24%	+4.66***	474
GBT	12m	+13.25%	+8.81***	+8.06%	+3.19***	468
Transformer	1m	+0.57%	+2.18**	-0.04%	-0.12	479
Transformer	3m	+2.28%	+3.79***	+0.56%	+0.76	477
Transformer	6m	+3.94%	+3.87***	+0.76%	+0.59	474
Transformer	12m	+6.61%	+4.06***	-0.59%	-0.25	468
Linear QR	1m	+1.19%	+4.50***	+0.45%	+1.46	479
Linear QR	3m	+3.55%	+5.55***	+1.40%	+1.75*	477
Linear QR	6m	+5.77%	+5.47***	+2.06%	+1.42	474
Linear QR	12m	+8.93%	+5.62***	+2.93%	+1.18	468
Penalized QR	1m	+1.18%	+4.53***	+0.51%	+1.69*	479
Penalized QR	3m	+3.52%	+5.30***	+1.63%	+2.12**	477
Penalized QR	6m	+5.57%	+5.15***	+2.54%	+1.82*	474
Penalized QR	12m	+8.09%	+4.75***	+3.64%	+1.41	468

Table IA.17: Kurtosis benchmark specification against option-implied moments.

Estimator	Sample	Spec	Term	Coef (t)	$\bar{R}^2$	N_{months}
GBT	Broad option-proxy overlap	W + option moments	W90 (GBT)	+0.263 (+7.43)***	0.015	347
GBT	Broad option-proxy overlap	W + option moments	Broad BKM variance	+0.276 (+2.84)***	0.015	347
GBT	Broad option-proxy overlap	W + option moments	Broad BKM skewness	+0.049 (+1.73)*	0.015	347
GBT	Broad option-proxy overlap	W + option moments	Broad BKM kurtosis	+0.013 (+0.61)	0.015	347
GBT	Broad option-proxy overlap	W + option moments	IV30	+0.151 (+1.64)	0.015	347
GBT	Broad option-proxy overlap	W + option moments + full controls	W90 (GBT)	+0.376 (+10.38)***	0.036	347
GBT	Broad option-proxy overlap	W + option moments + full controls	Broad BKM variance	-0.065 (-1.04)	0.036	347
GBT	Broad option-proxy overlap	W + option moments + full controls	Broad BKM skewness	+0.018 (+0.83)	0.036	347
GBT	Broad option-proxy overlap	W + option moments + full controls	Broad BKM kurtosis	-0.035 (-1.95)*	0.036	347
GBT	Broad option-proxy overlap	W + option moments + full controls	IV30	+0.583 (+6.81)***	0.036	347
GBT	Exact raw-chain option overlap	W + option moments	W90 (GBT)	-0.133 (-1.87)*	0.052	252
GBT	Exact raw-chain option overlap	W + option moments	Exact BKM variance	+2.274 (+8.26)***	0.052	252
GBT	Exact raw-chain option overlap	W + option moments	Exact BKM skewness	+0.055 (+0.84)	0.052	252
GBT	Exact raw-chain option overlap	W + option moments	Exact BKM kurtosis	-0.381 (-4.60)***	0.052	252
GBT	Exact raw-chain option overlap	W + option moments	IV30	-1.634 (-6.28)***	0.052	252
GBT	Exact raw-chain option overlap	W + option moments + full controls	W90 (GBT)	-0.177 (-3.26)***	0.125	231
GBT	Exact raw-chain option overlap	W + option moments + full controls	Exact BKM variance	+2.805 (+10.82)***	0.125	231
GBT	Exact raw-chain option overlap	W + option moments + full controls	Exact BKM skewness	+0.150 (+2.54)**	0.125	231
GBT	Exact raw-chain option overlap	W + option moments + full controls	Exact BKM kurtosis	-0.259 (-3.90)***	0.125	231
GBT	Exact raw-chain option overlap	W + option moments + full controls	IV30	-0.433 (-2.23)**	0.125	231
Transformer	Broad option-proxy overlap	W + option moments	W90 (Transformer)	+0.291 (+8.37)***	0.015	347
Transformer	Broad option-proxy overlap	W + option moments	Broad BKM variance	+0.208 (+2.18)**	0.015	347
Transformer	Broad option-proxy overlap	W + option moments	Broad BKM skewness	+0.058 (+2.20)**	0.015	347
Transformer	Broad option-proxy overlap	W + option moments	Broad BKM kurtosis	+0.012 (+0.59)	0.015	347
Transformer	Broad option-proxy overlap	W + option moments	IV30	+0.156 (+1.62)	0.015	347
Transformer	Broad option-proxy overlap	W + option moments + full controls	W90 (Transformer)	+0.393 (+10.03)***	0.036	347
Transformer	Broad option-proxy overlap	W + option moments + full controls	Broad BKM variance	-0.135 (-2.16)**	0.036	347
Transformer	Broad option-proxy overlap	W + option moments + full controls	Broad BKM skewness	+0.025 (+1.24)	0.036	347
Transformer	Broad option-proxy overlap	W + option moments + full controls	Broad BKM kurtosis	-0.038 (-2.09)**	0.036	347
Transformer	Broad option-proxy overlap	W + option moments + full controls	IV30	+0.603 (+6.88)***	0.036	347
Transformer	Exact raw-chain option overlap	W + option moments	W90 (Transformer)	-0.309 (-3.57)***	0.053	252
Transformer	Exact raw-chain option overlap	W + option moments	Exact BKM variance	+2.361 (+8.16)***	0.053	252
Transformer	Exact raw-chain option overlap	W + option moments	Exact BKM skewness	+0.066 (+1.03)	0.053	252
Transformer	Exact raw-chain option overlap	W + option moments	Exact BKM kurtosis	-0.388 (-4.67)***	0.053	252
Transformer	Exact raw-chain option overlap	W + option moments	IV30	-1.653 (-6.29)***	0.053	252
Transformer	Exact raw-chain option overlap	W + option moments + full controls	W90 (Transformer)	-0.221 (-3.40)***	0.126	231
Transformer	Exact raw-chain option overlap	W + option moments + full controls	Exact BKM variance	+2.819 (+10.72)***	0.126	231
Transformer	Exact raw-chain option overlap	W + option moments + full controls	Exact BKM skewness	+0.142 (+2.32)**	0.126	231
Transformer	Exact raw-chain option overlap	W + option moments + full controls	Exact BKM kurtosis	-0.263 (-3.94)***	0.126	231
Transformer	Exact raw-chain option overlap	W + option moments + full controls	IV30	-0.445 (-2.23)**	0.126	231
Linear QR	Broad option-proxy overlap	W + option moments	W90 (Linear QR)	+0.317 (+8.83)***	0.015	347
Linear QR	Broad option-proxy overlap	W + option moments	Broad BKM variance	+0.199 (+2.13)**	0.015	347
Linear QR	Broad option-proxy overlap	W + option moments	Broad BKM skewness	+0.052 (+1.97)**	0.015	347
Linear QR	Broad option-proxy overlap	W + option moments	Broad BKM kurtosis	+0.006 (+0.29)	0.015	347
Linear QR	Broad option-proxy overlap	W + option moments	IV30	+0.136 (+1.47)	0.015	347
Linear QR	Broad option-proxy overlap	W + option moments + full controls	W90 (Linear QR)	+0.543 (+11.67)***	0.036	347
Linear QR	Broad option-proxy overlap	W + option moments + full controls	Broad BKM variance	-0.166 (-2.67)***	0.036	347

Linear QR	Broad option-proxy overlap	W + option moments + full controls	Broad BKM skewness	+0.010 (+0.51)	0.036	347
Linear QR	Broad option-proxy overlap	W + option moments + full controls	Broad BKM kurtosis	-0.046 (-2.53)**	0.036	347
Linear QR	Broad option-proxy overlap	W + option moments + full controls	IV30	+0.560 (+6.52)***	0.036	347
Linear QR	Exact raw-chain option overlap	W + option moments	W90 (Linear QR)	-0.360 (-4.46)***	0.053	252
Linear QR	Exact raw-chain option overlap	W + option moments	Exact BKM variance	+2.339 (+8.18)***	0.053	252
Linear QR	Exact raw-chain option overlap	W + option moments	Exact BKM skewness	+0.087 (+1.32)	0.053	252
Linear QR	Exact raw-chain option overlap	W + option moments	Exact BKM kurtosis	-0.365 (-4.46)***	0.053	252
Linear QR	Exact raw-chain option overlap	W + option moments	IV30	-1.561 (-6.13)***	0.053	252
Linear QR	Exact raw-chain option overlap	W + option moments + full controls	W90 (Linear QR)	-0.139 (-1.98)**	0.126	231
Linear QR	Exact raw-chain option overlap	W + option moments + full controls	Exact BKM variance	+2.807 (+10.81)***	0.126	231
Linear QR	Exact raw-chain option overlap	W + option moments + full controls	Exact BKM skewness	+0.145 (+2.44)**	0.126	231
Linear QR	Exact raw-chain option overlap	W + option moments + full controls	Exact BKM kurtosis	-0.255 (-3.84)***	0.126	231
Linear QR	Exact raw-chain option overlap	W + option moments + full controls	IV30	-0.450 (-2.31)**	0.126	231
Penalized QR	Broad option-proxy overlap	W + option moments	W90 (Penalized QR)	+0.110 (+3.03)***	0.015	347
Penalized QR	Broad option-proxy overlap	W + option moments	Broad BKM variance	+0.224 (+2.40)**	0.015	347
Penalized QR	Broad option-proxy overlap	W + option moments	Broad BKM skewness	+0.062 (+2.38)**	0.015	347
Penalized QR	Broad option-proxy overlap	W + option moments	Broad BKM kurtosis	+0.014 (+0.67)	0.015	347
Penalized QR	Broad option-proxy overlap	W + option moments	IV30	+0.268 (+2.73)***	0.015	347
Penalized QR	Broad option-proxy overlap	W + option moments + full controls	W90 (Penalized QR)	+0.347 (+8.40)***	0.036	347
Penalized QR	Broad option-proxy overlap	W + option moments + full controls	Broad BKM variance	-0.137 (-2.21)**	0.036	347
Penalized QR	Broad option-proxy overlap	W + option moments + full controls	Broad BKM skewness	+0.023 (+1.18)	0.036	347
Penalized QR	Broad option-proxy overlap	W + option moments + full controls	Broad BKM kurtosis	-0.039 (-2.17)**	0.036	347
Penalized QR	Broad option-proxy overlap	W + option moments + full controls	IV30	+0.629 (+7.09)***	0.036	347
Penalized QR	Exact raw-chain option overlap	W + option moments	W90 (Penalized QR)	-0.726 (-10.20)***	0.056	252
Penalized QR	Exact raw-chain option overlap	W + option moments	Exact BKM variance	+2.418 (+8.43)***	0.056	252
Penalized QR	Exact raw-chain option overlap	W + option moments	Exact BKM skewness	+0.139 (+2.21)**	0.056	252
Penalized QR	Exact raw-chain option overlap	W + option moments	Exact BKM kurtosis	-0.350 (-4.31)***	0.056	252
Penalized QR	Exact raw-chain option overlap	W + option moments	IV30	-1.378 (-5.28)***	0.056	252
Penalized QR	Exact raw-chain option overlap	W + option moments + full controls	W90 (Penalized QR)	-0.394 (-5.77)***	0.126	231
Penalized QR	Exact raw-chain option overlap	W + option moments + full controls	Exact BKM variance	+2.827 (+10.80)***	0.126	231
Penalized QR	Exact raw-chain option overlap	W + option moments + full controls	Exact BKM skewness	+0.177 (+3.03)***	0.126	231
Penalized QR	Exact raw-chain option overlap	W + option moments + full controls	Exact BKM kurtosis	-0.246 (-3.68)***	0.126	231
Penalized QR	Exact raw-chain option overlap	W + option moments + full controls	IV30	-0.384 (-1.97)**	0.126	231

Notes: The table compares W90 with option-implied moments. Broad/proxy and exact raw-chain BKM rows have different option-overlap denominators and answer distinct benchmark questions. Newey–West t-statistics use six lags where applicable.

Table IA.18: Industry robustness.

		μ-strat return		Kurt FM (FULL)		
FF12	Estimator	Mean	t	Coef	t	N rows
BusEq	GBT	+4.79%	+8.29***	+0.291	+6.10***	436,221
BusEq	Transformer	+2.94%	+5.27***	+0.188	+4.48***	436,221
BusEq	Linear QR	+3.66%	+6.17***	+0.352	+7.94***	436,221
BusEq	Penalized QR	+3.42%	+5.61***	+0.202	+4.47***	436,221
Chems	GBT	+2.00%	+2.84***	+0.295	+4.00***	75,912
Chems	Transformer	+0.46%	+0.51	+0.128	+1.46	75,912
Chems	Linear QR	+1.63%	+1.87*	+0.308	+3.49***	75,912
Chems	Penalized QR	+1.24%	+1.63	+0.200	+2.22**	75,912
Durbl	GBT	+3.00%	+3.89***	+0.240	+3.71***	94,321
Durbl	Transformer	+2.39%	+2.63***	+0.088	+1.01	94,321
Durbl	Linear QR	+2.96%	+3.26***	+0.288	+3.17***	94,321
Durbl	Penalized QR	+2.06%	+2.04**	+0.135	+1.72*	94,321
Enrgy	GBT	+3.80%	+3.95***	+0.158	+2.56**	122,975
Enrgy	Transformer	+2.06%	+2.26**	-0.148	-2.24**	122,975
Enrgy	Linear QR	+1.49%	+1.81*	-0.202	-3.01***	122,975
Enrgy	Penalized QR	+2.46%	+3.37***	-0.159	-2.36**	122,975
Hlth	GBT	+4.99%	+6.61***	+0.369	+8.08***	225,552
Hlth	Transformer	+1.54%	+1.85*	+0.216	+3.65***	225,552
Hlth	Linear QR	+2.78%	+4.05***	+0.307	+4.82***	225,552
Hlth	Penalized QR	+2.88%	+4.10***	+0.192	+2.84***	225,552
Manuf	GBT	+3.34%	+6.41***	+0.218	+4.96***	395,108
Manuf	Transformer	+2.71%	+6.12***	+0.048	+1.04	395,108
Manuf	Linear QR	+2.32%	+4.42***	+0.187	+4.05***	395,108
Manuf	Penalized QR	+2.52%	+5.47***	+0.091	+2.01**	395,108
Money	GBT	+2.68%	+7.72***	+0.149	+3.72***	573,990
Money	Transformer	+2.13%	+6.36***	+0.043	+0.89	573,990
Money	Linear QR	+1.98%	+5.38***	+0.156	+3.14***	573,990
Money	Penalized QR	+2.08%	+5.59***	+0.020	+0.45	573,990
NoDur	GBT	+3.23%	+6.30***	+0.249	+3.17***	214,209
NoDur	Transformer	+2.46%	+5.00***	+0.152	+1.64	214,209
NoDur	Linear QR	+2.88%	+5.42***	+0.313	+3.32***	214,209
NoDur	Penalized QR	+2.64%	+4.97***	+0.173	+2.07**	214,209
Other	GBT	+4.49%	+7.62***	+0.226	+4.60***	578,812
Other	Transformer	+2.15%	+4.41***	+0.050	+0.88	578,812
Other	Linear QR	+2.36%	+4.50***	+0.207	+3.37***	578,812
Other	Penalized QR	+2.14%	+4.01***	+0.090	+1.70*	578,812
Shops	GBT	+3.26%	+5.09***	+0.174	+2.33**	327,352
Shops	Transformer	+2.47%	+4.06***	+0.068	+0.77	327,352
Shops	Linear QR	+2.48%	+4.13***	+0.234	+2.94***	327,352
Shops	Penalized QR	+2.17%	+3.41***	+0.083	+1.12	327,352

Telcm	GBT	+4.22%	+2.22**	+0.101	+0.91	66,606
Telcm	Transformer	+4.05%	+2.62***	+0.225	+2.45**	66,606
Telcm	Linear QR	+3.81%	+2.86***	+0.260	+2.10**	66,606
Telcm	Penalized QR	+2.96%	+3.00***	+0.127	+1.04	66,606
Utils	GBT	+1.35%	+2.40**	+0.068	+0.86	105,525
Utils	Transformer	+2.02%	+4.18***	+0.414	+4.11***	105,525
Utils	Linear QR	+2.07%	+5.50***	+0.676	+6.11***	105,525
Utils	Penalized QR	+1.87%	+4.65***	+0.390	+3.94***	105,525

Notes: Rows repeat return and physical-kurtosis tests within Fama–French 12 industries to show whether the evidence is concentrated in one sector. Newey–West t-statistics use six lags where applicable.

Table IA.19: Secondary predictive-width families.

Estimator	Measure	Spec	Coef (t)	$\bar{R}^2$	N_{months}
GBT	W50_gbt_n	Expanded non-option controls	+0.188 (+1.50)	0.047	479
GBT	W50_gbt_n	Expanded non-option controls plus macro beta	+0.177 (+1.34)	0.047	479
GBT	Tail ratio_gbt_n	Expanded non-option controls	+0.335 (+9.84)***	0.046	479
GBT	Tail ratio_gbt_n	Expanded non-option controls plus macro beta	+0.330 (+10.13)***	0.045	479
GBT	Asymmetry width_gbt_n	Expanded non-option controls	+0.045 (+1.09)	0.046	479
GBT	Asymmetry width_gbt_n	Expanded non-option controls plus macro beta	+0.033 (+0.78)	0.045	479
GBT	W50 (GBT)	Expanded non-option controls	+0.136 (+2.68)***	0.046	479
GBT	W50 (GBT)	Expanded non-option controls plus macro beta	+0.153 (+3.05)***	0.045	479
GBT	Tail ratio (GBT)	Expanded non-option controls	+0.267 (+10.68)***	0.046	479
GBT	Tail ratio (GBT)	Expanded non-option controls plus macro beta	+0.265 (+11.15)***	0.045	479
GBT	Asymmetry width (GBT)	Expanded non-option controls	+0.058 (+2.04)**	0.046	479
GBT	Asymmetry width (GBT)	Expanded non-option controls plus macro beta	+0.054 (+1.88)*	0.045	479
Transformer	W50_trans_n	Expanded non-option controls	+0.416 (+2.91)***	0.048	479
Transformer	W50_trans_n	Expanded non-option controls plus macro beta	+0.419 (+2.76)***	0.047	479
Transformer	Tail ratio_trans_n	Expanded non-option controls	-0.139 (-3.19)***	0.046	479
Transformer	Tail ratio_trans_n	Expanded non-option controls plus macro beta	-0.159 (-3.90)***	0.045	479
Transformer	Asymmetry width_trans_n	Expanded non-option controls	+0.210 (+4.94)***	0.045	479
Transformer	Asymmetry width_trans_n	Expanded non-option controls plus macro beta	+0.133 (+2.42)**	0.045	479
Transformer	W50 (Transformer)	Expanded non-option controls	+0.207 (+3.93)***	0.046	479
Transformer	W50 (Transformer)	Expanded non-option controls plus macro beta	+0.211 (+3.96)***	0.046	479
Transformer	Tail ratio (Transformer)	Expanded non-option controls	-0.106 (-2.60)***	0.046	479
Transformer	Tail ratio (Transformer)	Expanded non-option controls plus macro beta	-0.135 (-3.60)***	0.045	479
Transformer	Asymmetry width (Transformer)	Expanded non-option controls	+0.004 (+0.18)	0.045	479
Transformer	Asymmetry width (Transformer)	Expanded non-option controls plus macro beta	-0.031 (-1.27)	0.044	479
Linear QR	W50_lqr_n	Expanded non-option controls	+0.469 (+4.43)***	0.048	479
Linear QR	W50_lqr_n	Expanded non-option controls plus macro beta	+0.473 (+4.10)***	0.047	479
Linear QR	Tail ratio_lqr_n	Expanded non-option controls	-0.283 (-5.55)***	0.046	479
Linear QR	Tail ratio_lqr_n	Expanded non-option controls plus macro beta	-0.308 (-5.98)***	0.046	479
Linear QR	Asymmetry width_lqr_n	Expanded non-option controls	+0.277 (+6.16)***	0.046	479
Linear QR	Asymmetry width_lqr_n	Expanded non-option controls plus macro beta	+0.237 (+5.79)***	0.045	479
Linear QR	W50 (Linear QR)	Expanded non-option controls	+0.364 (+6.43)***	0.047	479
Linear QR	W50 (Linear QR)	Expanded non-option controls plus macro beta	+0.376 (+6.73)***	0.046	479
Linear QR	Tail ratio (Linear QR)	Expanded non-option controls	-0.249 (-5.54)***	0.046	479
Linear QR	Tail ratio (Linear QR)	Expanded non-option controls plus macro beta	-0.269 (-6.11)***	0.046	479
Linear QR	Asymmetry width (Linear QR)	Expanded non-option controls	+0.216 (+8.76)***	0.045	479
Linear QR	Asymmetry width (Linear QR)	Expanded non-option controls plus macro beta	+0.196 (+8.79)***	0.045	479
Penalized QR	W50_plqr_n	Expanded non-option controls	+0.182 (+1.43)	0.047	479
Penalized QR	W50_plqr_n	Expanded non-option controls plus macro beta	+0.187 (+1.41)	0.046	479
Penalized QR	Tail ratio_plqr_n	Expanded non-option controls	-0.138 (-3.14)***	0.046	479
Penalized QR	Tail ratio_plqr_n	Expanded non-option controls plus macro beta	-0.183 (-3.84)***	0.045	479
Penalized QR	Asymmetry width_plqr_n	Expanded non-option controls	+0.181 (+3.71)***	0.045	479
Penalized QR	Asymmetry width_plqr_n	Expanded non-option controls plus macro beta	+0.102 (+2.18)**	0.044	479
Penalized QR	W50 (Penalized QR)	Expanded non-option controls	+0.222 (+4.56)***	0.046	479
Penalized QR	W50 (Penalized QR)	Expanded non-option controls plus macro beta	+0.243 (+5.40)***	0.045	479

Penalized QR	`Tail ratio (Penalized QR)`	Expanded non-option controls	-0.130 (-3.69)***	0.046	479
Penalized QR	`Tail ratio (Penalized QR)`	Expanded non-option controls plus macro beta	-0.167 (-4.36)***	0.045	479
Penalized QR	`Asymmetry width (Penalized QR)`	Expanded non-option controls	+0.158 (+7.03)***	0.045	479
Penalized QR	`Asymmetry width (Penalized QR)`	Expanded non-option controls plus macro beta	+0.135 (+6.30)***	0.044	479

Notes: Rows compare W50, tail-ratio, and asymmetry width families with the W90 evidence. The table checks whether the main result is specific to one definition of predictive width. Newey–West t-statistics use six lags where applicable.

Table IA.20: **NBER recession state robustness.** All-estimator H-L return, long-only (μ_5, W_5) return, and FM kurt coefficient computed in NBER expansion vs recession months. The signal is positive in both states; return spreads are positive in both expansions and recessions and are larger in the recession subsample, while realized-kurtosis coefficients are more state-dependent across estimators in recession months (smaller cross-section and noisier signals). H-L and long-only mean in monthly percent. N is months in the state used for the kurt FM regression. Newey-West t with 6 lags. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

State	Estimator	H-L mean	t	Long (μ_5, W_5) mean	t	Kurt coef.	t	N months
expansion	GBT	+1.75%	+5.57***	+4.42%	+10.60***	+0.302	+7.12***	443
expansion	Transformer	+2.33%	+7.05***	+4.06%	+10.63***	+0.233	+5.11***	443
expansion	Linear QR	+2.18%	+6.95***	+3.87%	+10.61***	+0.356	+6.77***	443
expansion	Penalized QR	+2.08%	+6.87***	+3.69%	+10.31***	+0.219	+4.73***	443
recession	GBT	+5.54%	+2.89***	+8.28%	+4.09***	+0.224	+2.94***	36
recession	Transformer	+6.90%	+3.52***	+7.74%	+3.34***	+0.093	+0.74	36
recession	Linear QR	+6.51%	+3.48***	+6.75%	+3.13***	+0.265	+2.39**	36
recession	Penalized QR	+5.87%	+3.50***	+6.30%	+3.15***	+0.125	+1.09	36

Table IA.21: **Microcap-excluded robustness.** Repeats five analyses after excluding firms below the NYSE 20$^{\text{th}}$ percentile of `log_me_lag1` each month: (i) unconditional H-L, (ii) DiD $\mu \times W$, (iii) μ-stratified L/S, (iv) long-only (μ_5, W_5), and (v) FM kurt with FULL controls. All four estimators and the BKM kurt benchmark are reported. N is months used in the estimate. The microcap-excluded tests show that PTU's physical-tail prediction and several return implementations remain outside the smallest stocks, with expected attenuation in unconditional high-minus-low spreads. Sample is 1985–2024 subject to required controls. Newey-West t with 6 lags. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Estimator	Block	Measure	Mean / Coef	t	N months
GBT	unconditional_HL_EW	`W90 (GBT)`	+0.001	+0.27	479
GBT	did_mu_W	`W90 (GBT)`	+0.006	+3.11***	479
GBT	mu_strat_EW	`W90 (GBT)`	+0.024	+5.18***	479
GBT	long_q5w5_EW	`W90 (GBT)`	+0.019	+5.75***	479
GBT	kurt_FM_FULL	`W90 (GBT)`	+0.337	+8.63***	479
Transformer	unconditional_HL_EW	`W90 (Transformer)`	+0.005	+1.69*	479
Transformer	did_mu_W	`W90 (Transformer)`	+0.004	+1.76*	479
Transformer	mu_strat_EW	`W90 (Transformer)`	+0.022	+5.32***	479
Transformer	long_q5w5_EW	`W90 (Transformer)`	+0.021	+6.40***	479
Transformer	kurt_FM_FULL	`W90 (Transformer)`	+0.400	+9.68***	479
Linear QR	unconditional_HL_EW	`W90 (Linear QR)`	+0.007	+2.50**	479
Linear QR	did_mu_W	`W90 (Linear QR)`	+0.009	+4.28***	479
Linear QR	mu_strat_EW	`W90 (Linear QR)`	+0.021	+5.18***	479
Linear QR	long_q5w5_EW	`W90 (Linear QR)`	+0.022	+6.57***	479
Linear QR	kurt_FM_FULL	`W90 (Linear QR)`	+0.520	+10.97***	479
Penalized QR	unconditional_HL_EW	`W90 (Penalized QR)`	+0.007	+2.58***	479
Penalized QR	did_mu_W	`W90 (Penalized QR)`	+0.009	+3.73***	479
Penalized QR	mu_strat_EW	`W90 (Penalized QR)`	+0.021	+4.90***	479
Penalized QR	long_q5w5_EW	`W90 (Penalized QR)`	+0.021	+6.38***	479
Penalized QR	kurt_FM_FULL	`W90 (Penalized QR)`	+0.347	+9.62***	479
—	bkm_kurt_FM_STANDARD	`Exact BKM kurtosis`	-0.459	-7.12***	248

Table IA.22: **Persistence: AR(1) correlations.** Cross-quintile AR(1) Spearman rank correlations for each measure, all four estimators plus option-implied benchmarks, realized-moment benchmarks, consensus μ and σ, and macro β. ML/QR widths show moderate persistence (AR(1) $\approx$ 0.6–0.9), comparable to BKM and rvol. Full transition matrices for W_{90} are shown in Figure IA.8. Sample uses adjacent firm-month pairs with both lagged and current measure observed; N reports the pair count.

Measure	AR(1) corr.	N pairs
GBT W_90	+0.714	2,241,385
Transformer W_90	+0.709	2,241,385
Linear QR W_90	+0.712	2,241,385
Penalized QR W_90	+0.659	2,241,385
BKM kurt	+0.582	111,740
BKM skew	+0.496	111,740
IV30	+0.923	813,475
Realized vol	+0.670	3,171,455
Idio vol	+0.988	2,678,608
Realized kurt	+0.143	3,150,598
CDG-style 60d kurt	+0.657	3,153,966
BCW MAX	+0.487	3,171,455
JLN macro beta	+0.952	1,977,321
mu (consensus)	+0.855	3,049,924
sigma (consensus)	+0.902	3,049,924

Table IA.23: Historical physical-tail benchmark specification.

Predictor	Spec	Coef (t)	$\bar{R}^2$	N_{months}
realized_skew_w_n	No controls	-0.053 (-1.82)*	0.002	743
realized_skew_w_n	expanded non-option controls	+0.096 (+3.96)***	0.062	733
realized_skew_w_n	expanded non-option controls+GBT	+0.077 (+3.82)***	0.043	479
realized_skew_w_n	expanded non-option controls+Transformer	+0.081 (+4.03)***	0.042	479
realized_skew_w_n	expanded non-option controls+Linear QR	+0.103 (+5.43)***	0.043	479
realized_skew_w_n	expanded non-option controls+Penalized QR	+0.096 (+5.14)***	0.042	479
realized_kurt_w_n	No controls	+1.458 (+8.14)***	0.039	743
realized_kurt_w_n	expanded non-option controls	+0.867 (+8.78)***	0.075	733
realized_kurt_w_n	expanded non-option controls+GBT	+0.584 (+6.63)***	0.051	479
realized_kurt_w_n	expanded non-option controls+Transformer	+0.580 (+6.50)***	0.051	479
realized_kurt_w_n	expanded non-option controls+Linear QR	+0.583 (+6.56)***	0.052	479
realized_kurt_w_n	expanded non-option controls+Penalized QR	+0.576 (+6.45)***	0.051	479
cdg_skew_w_n	No controls	+0.206 (+4.48)***	0.003	743
cdg_skew_w_n	expanded non-option controls	+0.150 (+5.55)***	0.068	733
cdg_skew_w_n	expanded non-option controls+GBT	+0.098 (+4.20)***	0.047	479
cdg_skew_w_n	expanded non-option controls+Transformer	+0.096 (+4.22)***	0.046	479
cdg_skew_w_n	expanded non-option controls+Linear QR	+0.104 (+4.57)***	0.047	479
cdg_skew_w_n	expanded non-option controls+Penalized QR	+0.097 (+4.29)***	0.046	479
cdg_kurt_w_n	No controls	+1.931 (+9.71)***	0.051	743
cdg_kurt_w_n	expanded non-option controls	+1.113 (+11.06)***	0.083	733
cdg_kurt_w_n	expanded non-option controls+GBT	+0.849 (+9.31)***	0.059	479
cdg_kurt_w_n	expanded non-option controls+Transformer	+0.851 (+9.31)***	0.058	479
cdg_kurt_w_n	expanded non-option controls+Linear QR	+0.848 (+9.23)***	0.059	479
cdg_kurt_w_n	expanded non-option controls+Penalized QR	+0.849 (+9.27)***	0.058	479
bmv_eskew_w_n	No controls	+0.974 (+8.78)***	0.019	552
bmv_eskew_w_n	expanded non-option controls	+0.238 (+5.09)***	0.050	552
bmv_eskew_w_n	expanded non-option controls+GBT	+0.160 (+5.16)***	0.047	468
bmv_eskew_w_n	expanded non-option controls+Transformer	+0.154 (+4.97)***	0.046	468
bmv_eskew_w_n	expanded non-option controls+Linear QR	+0.143 (+4.63)***	0.047	468
bmv_eskew_w_n	expanded non-option controls+Penalized QR	+0.146 (+4.64)***	0.046	468
bcw_max_n	No controls	-0.655 (-6.35)***	0.010	743
bcw_max_n	expanded non-option controls	+0.246 (+5.93)***	0.069	733
bcw_max_n	expanded non-option controls+GBT	+0.208 (+6.13)***	0.047	479
bcw_max_n	expanded non-option controls+Transformer	+0.197 (+5.87)***	0.047	479
bcw_max_n	expanded non-option controls+Linear QR	+0.223 (+6.79)***	0.047	479
bcw_max_n	expanded non-option controls+Penalized QR	+0.215 (+6.51)***	0.047	479
bcw_max5_n	No controls	-0.948 (-7.03)***	0.019	743
bcw_max5_n	expanded non-option controls	-1.061 (-9.75)***	0.072	733
bcw_max5_n	expanded non-option controls+GBT	-0.759 (-5.85)***	0.049	479
bcw_max5_n	expanded non-option controls+Transformer	-0.836 (-5.40)***	0.050	479
bcw_max5_n	expanded non-option controls+Linear QR	-0.748 (-4.94)***	0.050	479
bcw_max5_n	expanded non-option controls+Penalized QR	-0.809 (-4.98)***	0.050	479

Notes: Rows compare W90 with lagged realized moments, CDG-style historical moments, BMV expected skewness, and MAX-family lottery benchmarks. Newey–West t-statistics use six lags where applicable.

Table IA.24: Secondary predictive-width and quantile families.

		H-L return		H-L α		μ-strat return		μ-strat α	
Family	Measure	Mean	t	α	$t(\alpha)$	Mean	t	α	$t(\alpha)$
W90	`W90 (GBT)`	+2.03%	+5.93***	+2.27%	+8.64***	+4.03%	+8.59***	+3.84%	+10.18***
W90	`W90 (Transformer)`	+2.67%	+7.32***	+2.87%	+9.44***	+2.67%	+7.37***	+2.37%	+7.99***
W90	`W90 (Linear QR)`	+2.51%	+7.30***	+2.72%	+9.72***	+2.71%	+7.26***	+2.39%	+8.35***
W90	`W90 (Penalized QR)`	+2.36%	+7.31***	+2.48%	+9.31***	+2.70%	+7.08***	+2.34%	+7.93***
W50	`W50 (GBT)`	+1.38%	+4.17***	+1.44%	+5.72***	+4.28%	+9.38***	+4.05%	+10.55***
W50	`W50 (Transformer)`	+2.52%	+6.77***	+2.59%	+8.66***	+3.69%	+9.34***	+3.37%	+9.79***
W50	`W50 (Linear QR)`	+2.11%	+6.08***	+2.15%	+7.92***	+3.54%	+9.14***	+3.31%	+10.36***
W50	`W50 (Penalized QR)`	+1.91%	+5.79***	+1.88%	+7.22***	+3.41%	+8.54***	+3.03%	+9.50***
Tail ratio	`Tail ratio (GBT)`	+1.31%	+5.68***	+1.89%	+8.10***	+2.78%	+5.62***	+2.35%	+7.69***
Tail ratio	`Tail ratio (Transformer)`	-0.12%	-0.72	+0.18%	+1.15	+4.77%	+8.58***	+4.14%	+8.29***
Tail ratio	`Tail ratio (Linear QR)`	+0.66%	+4.34***	+1.09%	+7.76***	+1.76%	+3.97***	+1.17%	+3.86***
Tail ratio	`Tail ratio (Penalized QR)`	+0.83%	+4.98***	+1.20%	+8.36***	+1.61%	+3.61***	+0.91%	+2.86***
Asymmetry width	`Asymmetry width (GBT)`	+3.22%	+12.34***	+3.51%	+13.14***	+2.25%	+4.10***	+1.59%	+4.21***
Asymmetry width	`Asymmetry width (Transformer)`	+3.14%	+9.09***	+3.61%	+10.95***	+2.24%	+4.17***	+1.46%	+3.63***
Asymmetry width	`Asymmetry width (Linear QR)`	+3.13%	+10.33***	+3.54%	+12.33***	+1.02%	+2.24**	+0.36%	+1.13
Asymmetry width	`Asymmetry width (Penalized QR)`	+3.17%	+10.92***	+3.49%	+12.68***	-0.05%	-0.11	-0.80%	-1.96*
Predictive center	`Predictive center (GBT)`	+0.19%	+0.93	-0.12%	-0.70	+3.00%	+7.06***	+2.76%	+10.51***
Predictive center	`Predictive center (Transformer)`	+2.79%	+9.20***	+2.20%	+10.14***	+0.69%	+1.54	+0.51%	+1.30
Predictive center	`Predictive center (Linear QR)`	+0.96%	+3.69***	+0.41%	+2.70***	+3.36%	+7.33***	+3.28%	+8.34***
Predictive center	`Predictive center (Penalized QR)`	+0.58%	+2.13**	-0.07%	-0.44	+3.29%	+7.41***	+3.22%	+8.86***
Lower-tail quantile	`Lower-tail quantile (GBT)`	-0.67%	-1.70*	-1.18%	-4.55***	+4.00%	+8.61***	+4.02%	+9.52***
Lower-tail quantile	`Lower-tail quantile (Transformer)`	-0.14%	-0.36	-0.60%	-2.63***	+3.19%	+7.34***	+3.17%	+7.52***
Lower-tail quantile	`Lower-tail quantile (Linear QR)`	-0.90%	-2.20**	-1.34%	-5.35***	+5.04%	+9.73***	+4.84%	+10.27***
Lower-tail quantile	`Lower-tail quantile (Penalized QR)`	-1.04%	-2.51**	-1.50%	-5.79***	+5.21%	+9.45***	+5.25%	+10.08***
Upper-tail quantile	`Upper-tail quantile (GBT)`	+2.21%	+5.47***	+2.74%	+8.98***	+3.79%	+9.61***	+3.57%	+12.28***
Upper-tail quantile	`Upper-tail quantile (Transformer)`	+2.30%	+5.49***	+2.83%	+9.06***	+4.17%	+9.63***	+3.90%	+11.77***
Upper-tail quantile	`Upper-tail quantile (Linear QR)`	+2.29%	+5.50***	+2.83%	+9.07***	+3.68%	+9.53***	+3.39%	+12.15***
Upper-tail quantile	`Upper-tail quantile (Penalized QR)`	+2.21%	+5.36***	+2.75%	+8.81***	+3.60%	+9.30***	+3.28%	+11.34***

Notes: Rows compare W90 with alternative quantile-width, tail-ratio, asymmetry, and location-style signals across estimators. Newey–West t-statistics use six lags where applicable.

Table IA.25: Equal- and value-weighted strategy comparison.

Arch	Wt	Strategy	Mean	t	α	$t(\alpha)$	Sharpe	N
GBT	EW	W Q5–Q1	+2.05%	$+5.92^{***}$	+2.29%	$+8.66^{***}$	+1.28	479
GBT	EW	DiD	+1.99%	$+7.83^{***}$	+1.52%	$+5.35^{***}$	+1.44	479
GBT	EW	DiD long-leg	+3.01%	$+9.05^{***}$	+2.68%	$+8.92^{***}$	+1.90	479
GBT	EW	DiD short-leg	+1.02%	$+5.75^{***}$	+1.16%	$+7.95^{***}$	+1.16	479
GBT	EW	μ-stratified L/S	+4.03%	$+8.59^{***}$	+3.84%	$+10.18^{***}$	+1.87	479
GBT	EW	Long-only (μ_5, W_5)	+4.71%	$+10.95^{***}$	+4.12%	$+11.81^{***}$	+2.39	479
GBT	VW	W Q5–Q1	+0.36%	+1.26	+0.51%	$+2.94^{***}$	+0.20	479
GBT	VW	DiD	+0.79%	$+2.96^{***}$	+0.70%	$+2.45^{**}$	+0.46	479
GBT	VW	DiD long-leg	+1.61%	$+4.67^{***}$	+1.33%	$+5.09^{***}$	+0.94	479
GBT	VW	DiD short-leg	+0.82%	$+3.46^{***}$	+0.63%	$+3.43^{***}$	+0.66	479
GBT	VW	μ-stratified L/S	+2.43%	$+4.61^{***}$	+1.97%	$+5.60^{***}$	+0.99	479
GBT	VW	Long-only (μ_5, W_5)	+2.44%	$+6.77^{***}$	+1.58%	$+6.31^{***}$	+1.17	479
Transformer	EW	W Q5–Q1	+2.71%	$+7.32^{***}$	+2.90%	$+9.39^{***}$	+1.59	479
Transformer	EW	DiD	+0.57%	$+2.18^{**}$	-0.04%	-0.12	+0.40	479
Transformer	EW	DiD long-leg	+1.62%	$+6.03^{***}$	+1.16%	$+4.32^{***}$	+1.05	479
Transformer	EW	DiD short-leg	+1.05%	$+6.36^{***}$	+1.20%	$+8.84^{***}$	+1.22	479
Transformer	EW	μ-stratified L/S	+2.67%	$+7.37^{***}$	+2.37%	$+7.99^{***}$	+1.30	479
Transformer	EW	Long-only (μ_5, W_5)	+4.33%	$+10.71^{***}$	+3.69%	$+11.99^{***}$	+2.19	479
Transformer	VW	W Q5–Q1	+0.81%	$+2.45^{**}$	+0.97%	$+4.65^{***}$	+0.41	479
Transformer	VW	DiD	+0.17%	+0.57	-0.17%	-0.58	+0.09	479
Transformer	VW	DiD long-leg	+1.21%	$+3.22^{***}$	+0.75%	$+2.83^{***}$	+0.60	479
Transformer	VW	DiD short-leg	+1.04%	$+5.31^{***}$	+0.93%	$+5.44^{***}$	+1.01	479
Transformer	VW	μ-stratified L/S	+2.25%	$+4.37^{***}$	+1.68%	$+5.13^{***}$	+0.87	479
Transformer	VW	Long-only (μ_5, W_5)	+2.79%	$+8.04^{***}$	+1.87%	$+7.86^{***}$	+1.26	479
Linear QR	EW	W Q5–Q1	+2.54%	$+7.29^{***}$	+2.75%	$+9.71^{***}$	+1.45	479
Linear QR	EW	DiD	+1.19%	$+4.50^{***}$	+0.45%	+1.46	+0.77	479
Linear QR	EW	DiD long-leg	+1.95%	$+7.10^{***}$	+1.42%	$+5.40^{***}$	+1.19	479
Linear QR	EW	DiD short-leg	+0.76%	$+4.46^{***}$	+0.97%	$+6.94^{***}$	+0.83	479
Linear QR	EW	μ-stratified L/S	+2.71%	$+7.26^{***}$	+2.39%	$+8.35^{***}$	+1.25	479
Linear QR	EW	Long-only (μ_5, W_5)	+4.09%	$+10.88^{***}$	+3.42%	$+12.73^{***}$	+2.11	479
Linear QR	VW	W Q5–Q1	+0.78%	$+2.56^{**}$	+0.94%	$+5.16^{***}$	+0.39	479
Linear QR	VW	DiD	+1.16%	$+3.58^{***}$	+0.78%	$+2.20^{**}$	+0.63	479
Linear QR	VW	DiD long-leg	+1.79%	$+4.68^{***}$	+1.30%	$+3.83^{***}$	+0.90	479
Linear QR	VW	DiD short-leg	+0.63%	$+3.18^{***}$	+0.52%	$+3.36^{***}$	+0.59	479
Linear QR	VW	μ-stratified L/S	+2.42%	$+4.70^{***}$	+1.82%	$+4.64^{***}$	+0.93	479
Linear QR	VW	Long-only (μ_5, W_5)	+2.80%	$+7.64^{***}$	+1.85%	$+7.76^{***}$	+1.27	479
Penalized QR	EW	W Q5–Q1	+2.39%	$+7.28^{***}$	+2.50%	$+9.26^{***}$	+1.42	479
Penalized QR	EW	DiD	+1.18%	$+4.53^{***}$	+0.51%	$+1.69^{*}$	+0.78	479
Penalized QR	EW	DiD long-leg	+1.94%	$+6.89^{***}$	+1.42%	$+5.29^{***}$	+1.17	479
Penalized QR	EW	DiD short-leg	+0.76%	$+4.59^{***}$	+0.91%	$+7.16^{***}$	+0.85	479
Penalized QR	EW	μ-stratified L/S	+2.70%	$+7.08^{***}$	+2.34%	$+7.93^{***}$	+1.22	479
Penalized QR	EW	Long-only (μ_5, W_5)	+3.89%	$+10.61^{***}$	+3.19%	$+12.30^{***}$	+2.05	479
Penalized QR	VW	W Q5–Q1	+0.75%	$+2.60^{***}$	+0.86%	$+4.86^{***}$	+0.39	479
Penalized QR	VW	DiD	+1.05%	$+3.29^{***}$	+0.64%	$+1.96^{**}$	+0.53	479
Penalized QR	VW	DiD long-leg	+1.61%	$+4.37^{***}$	+1.01%	$+3.42^{***}$	+0.78	479
Penalized QR	VW	DiD short-leg	+0.56%	$+3.01^{***}$	+0.37%	$+2.55^{**}$	+0.49	479
Penalized QR	VW	μ-stratified L/S	+2.17%	$+4.45^{***}$	+1.38%	$+4.16^{***}$	+0.82	479

Penalized QR VW Long-only (μ_5, W_5)	+2.60%	+7.12***	+1.62%	+7.36***	+1.22	479

Notes: Rows compare PTU/W90 strategy implementations under equal and value weighting. Newey–West t-statistics use six lags where applicable.

Table IA.26: Subperiod robustness.

Period	Estimator	Block	Mean / Coef	t	N months
1985_to_1999	GBT	unconditional_HL_EW	+0.023	$+4.78^{***}$	180
1985_to_1999	GBT	mu_strat_EW	+0.053	$+10.52^{***}$	180
1985_to_1999	GBT	kurt_FM_FULL	+0.112	+1.64	180
1985_to_1999	Transformer	unconditional_HL_EW	+0.034	$+6.56^{***}$	180
1985_to_1999	Transformer	mu_strat_EW	+0.032	$+7.92^{***}$	180
1985_to_1999	Transformer	kurt_FM_FULL	-0.058	-0.86	180
1985_to_1999	Linear QR	unconditional_HL_EW	+0.031	$+6.17^{***}$	180
1985_to_1999	Linear QR	mu_strat_EW	+0.036	$+8.99^{***}$	180
1985_to_1999	Linear QR	kurt_FM_FULL	+0.031	+0.45	180
1985_to_1999	Penalized QR	unconditional_HL_EW	+0.031	$+6.64^{***}$	180
1985_to_1999	Penalized QR	mu_strat_EW	+0.034	$+9.19^{***}$	180
1985_to_1999	Penalized QR	kurt_FM_FULL	-0.058	-0.82	180
2000_to_2007	GBT	unconditional_HL_EW	+0.017	$+2.31^{**}$	96
2000_to_2007	GBT	mu_strat_EW	+0.055	$+4.60^{***}$	96
2000_to_2007	GBT	kurt_FM_FULL	+0.253	$+4.65^{***}$	96
2000_to_2007	Transformer	unconditional_HL_EW	+0.023	$+3.03^{***}$	96
2000_to_2007	Transformer	mu_strat_EW	+0.041	$+4.37^{***}$	96
2000_to_2007	Transformer	kurt_FM_FULL	+0.322	$+5.57^{***}$	96
2000_to_2007	Linear QR	unconditional_HL_EW	+0.024	$+3.41^{***}$	96
2000_to_2007	Linear QR	mu_strat_EW	+0.039	$+3.88^{***}$	96
2000_to_2007	Linear QR	kurt_FM_FULL	+0.456	$+8.87^{***}$	96
2000_to_2007	Penalized QR	unconditional_HL_EW	+0.022	$+3.30^{***}$	96
2000_to_2007	Penalized QR	mu_strat_EW	+0.042	$+4.16^{***}$	96
2000_to_2007	Penalized QR	kurt_FM_FULL	+0.376	$+7.27^{***}$	96
2008_to_2019	GBT	unconditional_HL_EW	+0.018	$+3.05^{***}$	144
2008_to_2019	GBT	mu_strat_EW	+0.021	$+3.46^{***}$	144
2008_to_2019	GBT	kurt_FM_FULL	+0.564	$+9.30^{***}$	144
2008_to_2019	Transformer	unconditional_HL_EW	+0.018	$+2.93^{***}$	144
2008_to_2019	Transformer	mu_strat_EW	+0.012	$+2.76^{***}$	144
2008_to_2019	Transformer	kurt_FM_FULL	+0.493	$+8.50^{***}$	144
2008_to_2019	Linear QR	unconditional_HL_EW	+0.018	$+3.01^{***}$	144
2008_to_2019	Linear QR	mu_strat_EW	+0.012	$+2.49^{**}$	144
2008_to_2019	Linear QR	kurt_FM_FULL	+0.751	$+12.86^{***}$	144
2008_to_2019	Penalized QR	unconditional_HL_EW	+0.016	$+2.89^{***}$	144
2008_to_2019	Penalized QR	mu_strat_EW	+0.013	$+2.58^{***}$	144
2008_to_2019	Penalized QR	kurt_FM_FULL	+0.492	$+10.66^{***}$	144
2020_COVID	GBT	unconditional_HL_EW	+0.037	+1.32	24
2020_COVID	GBT	mu_strat_EW	+0.049	$+1.66^{*}$	24
2020_COVID	GBT	kurt_FM_FULL	+0.169	$+2.42^{**}$	24
2020_COVID	Transformer	unconditional_HL_EW	+0.041	+1.44	24
2020_COVID	Transformer	mu_strat_EW	+0.047	$+1.75^{*}$	24
2020_COVID	Transformer	kurt_FM_FULL	+0.223	$+4.55^{***}$	24
2020_COVID	Linear QR	unconditional_HL_EW	+0.035	+1.32	24
2020_COVID	Linear QR	mu_strat_EW	+0.033	+1.29	24
2020_COVID	Linear QR	kurt_FM_FULL	+0.096	+1.06	24
2020_COVID	Penalized QR	unconditional_HL_EW	+0.039	+1.55	24
2020_COVID	Penalized QR	mu_strat_EW	+0.021	+0.65	24

2020_COVID	Penalized QR	kurt_FM_FULL	+0.041	+0.31	24
2022_to_2024	GBT	unconditional_HL_EW	+0.016	+2.37**	35
2022_to_2024	GBT	mu_strat_EW	+0.007	+0.44	35
2022_to_2024	GBT	kurt_FM_FULL	+0.342	+8.98***	35
2022_to_2024	Transformer	unconditional_HL_EW	+0.026	+4.47***	35
2022_to_2024	Transformer	mu_strat_EW	+0.005	+0.47	35
2022_to_2024	Transformer	kurt_FM_FULL	+0.279	+2.70***	35
2022_to_2024	Linear QR	unconditional_HL_EW	+0.020	+3.91***	35
2022_to_2024	Linear QR	mu_strat_EW	+0.008	+0.64	35
2022_to_2024	Linear QR	kurt_FM_FULL	+0.219	+3.19***	35
2022_to_2024	Penalized QR	unconditional_HL_EW	+0.016	+2.82***	35
2022_to_2024	Penalized QR	mu_strat_EW	+0.013	+1.12	35
2022_to_2024	Penalized QR	kurt_FM_FULL	+0.110	+1.16	35
post_2008	GBT	unconditional_HL_EW	+0.020	+3.44***	203
post_2008	GBT	mu_strat_EW	+0.022	+3.13***	203
post_2008	GBT	kurt_FM_FULL	+0.479	+9.37***	203
post_2008	Transformer	unconditional_HL_EW	+0.022	+3.76***	203
post_2008	Transformer	mu_strat_EW	+0.015	+2.70***	203
post_2008	Transformer	kurt_FM_FULL	+0.424	+8.60***	203
post_2008	Linear QR	unconditional_HL_EW	+0.020	+3.72***	203
post_2008	Linear QR	mu_strat_EW	+0.014	+2.43**	203
post_2008	Linear QR	kurt_FM_FULL	+0.582	+8.79***	203
post_2008	Penalized QR	unconditional_HL_EW	+0.019	+3.55***	203
post_2008	Penalized QR	mu_strat_EW	+0.014	+2.23**	203
post_2008	Penalized QR	kurt_FM_FULL	+0.373	+7.05***	203

Notes: Rows repeat the return and realized-kurtosis evidence across calendar subperiods, including crisis and post-crisis windows. Newey–West t-statistics use six lags where applicable.

Table IA.27: **Predictive calibration of all four estimators.** Empirical coverage of nominal quantiles 5%, 25%, 50%, 75%, 95% for each estimator. The table reports unconditional quantile calibration. Empirical coverages are close to nominal on average, with maximum absolute gaps of about 1.6 percentage points. Conditional calibration by PTU quintile is reported separately and shows undercoverage in high-PTU states, which supports the rank-based interpretation of PTU as a relative physical tail-width measure. Sample is the out-of-sample out-of-sample prediction window, 1985–2024; N reports firm-month predictions.

Estimator	Quantile	Empirical coverage	Expected	Gap	N
GBT	0.05	0.065	0.050	+0.015	2,262,381
GBT	0.25	0.254	0.250	+0.004	2,262,381
GBT	0.50	0.493	0.500	-0.007	2,262,381
GBT	0.75	0.752	0.750	+0.002	2,262,381
GBT	0.95	0.945	0.950	-0.005	2,262,381
Transformer	0.05	0.060	0.050	+0.010	2,262,381
Transformer	0.25	0.254	0.250	+0.004	2,262,381
Transformer	0.50	0.503	0.500	+0.003	2,262,381
Transformer	0.75	0.762	0.750	+0.012	2,262,381
Transformer	0.95	0.952	0.950	+0.002	2,262,381
Linear QR	0.05	0.059	0.050	+0.009	2,262,381
Linear QR	0.25	0.252	0.250	+0.002	2,262,381
Linear QR	0.50	0.501	0.500	+0.001	2,262,381
Linear QR	0.75	0.750	0.750	+0.000	2,262,381
Linear QR	0.95	0.949	0.950	-0.001	2,262,381
Penalized QR	0.05	0.063	0.050	+0.013	2,262,381
Penalized QR	0.25	0.266	0.250	+0.016	2,262,381
Penalized QR	0.50	0.505	0.500	+0.005	2,262,381
Penalized QR	0.75	0.761	0.750	+0.011	2,262,381
Penalized QR	0.95	0.950	0.950	+0.000	2,262,381

Table IA.28: **Alternative consensus μ and σ formulations.** The table replaces the expected-return consensus and forecast-scale variables with alternative consensus constructions. Baseline μ is the average active monthly rank across available Chen–Zimmermann characteristics. Alternative μ uses the same active-rank logic with the alternative available characteristic set, and theme-weighted μ averages characteristic ranks within themes before averaging across themes. Baseline σ is the cross-characteristic dispersion of active ranked signals, and theme-weighted σ applies the same theme aggregation before measuring dispersion. The reported coefficient is on the estimator-specific W90 measure in return specifications that use the corresponding consensus variable. Coefficients are stable across consensus formulations. Each cell "coef (t)". Sample is 1985–2024 subject to FULL-control availability. Newey-West t with 6 lags. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Estimator	Replaced var.	Used var.	Coef (t)	$\bar{R}^2$	N_{months}
GBT	replaces μ	μ	+0.017 (+7.90)***	0.048	479
GBT	replaces μ	`alternative` μ	+0.017 (+7.88)***	0.048	479
GBT	replaces μ	`theme-weighted` μ	+0.017 (+7.80)***	0.048	479
GBT	replaces σ	σ	+0.017 (+7.90)***	0.048	479
GBT	replaces σ	`theme-weighted` σ	+0.017 (+7.74)***	0.048	479
Transformer	replaces μ	μ	+0.033 (+10.41)***	0.051	479
Transformer	replaces μ	`alternative` μ	+0.033 (+10.38)***	0.051	479
Transformer	replaces μ	`theme-weighted` μ	+0.032 (+10.18)***	0.050	479
Transformer	replaces σ	σ	+0.033 (+10.41)***	0.051	479
Transformer	replaces σ	`theme-weighted` σ	+0.034 (+10.16)***	0.050	479
Linear QR	replaces μ	μ	+0.033 (+11.32)***	0.051	479
Linear QR	replaces μ	`alternative` μ	+0.033 (+11.36)***	0.051	479
Linear QR	replaces μ	`theme-weighted` μ	+0.031 (+10.73)***	0.051	479
Linear QR	replaces σ	σ	+0.033 (+11.32)***	0.051	479
Linear QR	replaces σ	`theme-weighted` σ	+0.033 (+11.05)***	0.050	479
Penalized QR	replaces μ	μ	+0.029 (+10.91)***	0.051	479
Penalized QR	replaces μ	`alternative` μ	+0.029 (+10.99)***	0.051	479
Penalized QR	replaces μ	`theme-weighted` μ	+0.028 (+10.51)***	0.050	479
Penalized QR	replaces σ	σ	+0.029 (+10.91)***	0.051	479
Penalized QR	replaces σ	`theme-weighted` σ	+0.029 (+10.76)***	0.050	479

C International External Validation in Global JKP Data

This section uses the global JKP panel as international external validation for the physical tail-width interpretation. The analysis is deliberately separated from the U.S. option-benchmark tests. The U.S. evidence remains the main setting for the option-spanning, exact-BKM, and factor-alpha tests; the global JKP evidence asks whether a PTU-style physical predictive tail-width rank identifies future tail realizations outside the U.S. sample.

The global construction follows the same economic object as the U.S. PTU measure. Four quantile-prediction methods estimate next-month excess-return quantiles, W90 is the predicted 95th percentile less the predicted 5th percentile, and PTU is the equal-weight common component of adjusted W90 ranks. All ranks, quintile sorts, and controls are formed within country-month. The return target is the JKP-prepared next-month excess return winsorized within country-month at the 1st and 99th percentiles. The resulting panel contains 10,791,341 stock-months, 70 countries, and 312 months from January 2000 through December 2025.

The validation result is strong. In country-month PTU sorts, high-PTU stocks are 14.27 percentage points more likely than low-PTU stocks to realize a top-decile absolute-return month, with a Newey–West t-statistic of 58.25. The relation appears on both sides of the distribution: left-tail event probabilities rise by 9.85 percentage points and right-tail event probabilities rise by 10.08 percentage points. In country-month Fama–MacBeth regressions with size, volatility, lottery, skewness, illiquidity, momentum, book-to-market, and missingness controls, the PTU coefficient for top-decile absolute-return states remains 11.05 percentage points with a t-statistic of 39.64. Controlled left-tail and right-tail coefficients are also positive, at 5.57 and 9.07 percentage points, respectively. The international evidence therefore supports the core measurement claim: predicted physical tail width identifies future realized tail states.

The global return evidence is interpreted as supporting scope evidence rather than as the paper's main pricing claim. Raw full-sample equal-weight global return spreads are sensitive to extreme small-country/month observations, so Table IA.29 emphasizes liquid, large, and developed-market screens. In those screens, PTU high-minus-low return spreads are positive: 2.35 percent per month in the liquid half on an equal-weight basis, 0.53 percent on a value-weight basis, 0.67 percent in the large half, 1.00 percent in developed markets, 1.52 percent in Asia-Pacific, and 1.62 percent in Europe. Robust aggregation checks lead to the same qualitative conclusion: clipping country-month high-minus-low spreads at the 2.5/97.5 percentiles gives a 1.23 percent equal-weight spread with a t-statistic of 6.77, requiring at

least 20 stocks in both the low- and high-PTU quintiles gives a 1.22 percent spread with a t-statistic of 6.24, and the median country-level equal-weight spread is 1.17 percent with 67 of 70 countries positive. These results broaden the external validity of the physical tail-width evidence without replacing the U.S. benchmark and option-market analysis.

Table IA.29: **International external validation in global JKP data.** This table summarizes the global JKP PTU external-validation exercise. All ranks and sorts are formed within country-month.

Panel A. Sample and construction	
Stock-months	10,791,341
Countries	70
Months	312
Date range	January 2000–December 2025
Ranking unit	Country-month
Return target	JKP next-month excess return, country-month 1/99 winsorized
Samples	Full, large-half, liquid-half, large-and-liquid
Panel B. Tail validation	
Top absolute-return decile H–L	14.27 pp ($t = 58.25$)
Left-tail event H–L	9.85 pp ($t = 35.57$)
Right-tail event H–L	10.08 pp ($t = 55.61$)
Controlled Fama–MacBeth, top absolute-return decile	11.05 pp ($t = 39.64$)
Controlled Fama–MacBeth, left-tail event	5.57 pp ($t = 16.24$)
Controlled Fama–MacBeth, right-tail event	9.07 pp ($t = 38.90$)
Controlled Fama–MacBeth, absolute return	11.10 percent ($t = 2.05$)
Panel C. Return spreads under implementation screens	
Liquid-half equal-weight H–L	2.35 percent ($t = 2.63$)
Liquid-half value-weight H–L	0.53 percent ($t = 2.83$)
Large-half equal-weight H–L	0.67 percent ($t = 2.46$)
Developed markets equal-weight H–L	1.00 percent ($t = 3.70$)
Asia-Pacific equal-weight H–L	1.52 percent ($t = 8.63$)
Europe equal-weight H–L	1.62 percent ($t = 4.31$)
Panel D. Cross-country breadth	
Countries with at least 60 H–L observations	62
Countries with positive equal-weight H–L means	60/62
Countries with equal-weight H–L $t > 1.96$	35/62
Countries with positive absolute-tail H–L	62/62

Notes: The table reports international external-validation evidence from the global JKP panel. PTU quintiles, controls, and portfolio ranks are formed within country-month. Return tests use the JKP-prepared next-month excess return winsorized within country-month at the 1st and 99th percentiles. The date range refers to signal months with available next-month return targets. Portfolio-sort t-statistics are Newey–West t-statistics with six lags computed from country-equal monthly high-minus-low spreads. Fama–MacBeth coefficients are estimated within country-month, averaged across countries within month, and evaluated with Newey–West six-lag t-statistics over the monthly coefficient series. The global return evidence is interpreted through liquid, large, and developed-market screens because raw full-sample equal-weight global averages are sensitive to extreme small-country/month observations. This table is external validation of the physical tail-width measure; it does not replace the U.S. option-benchmark tests.

D IA Figures

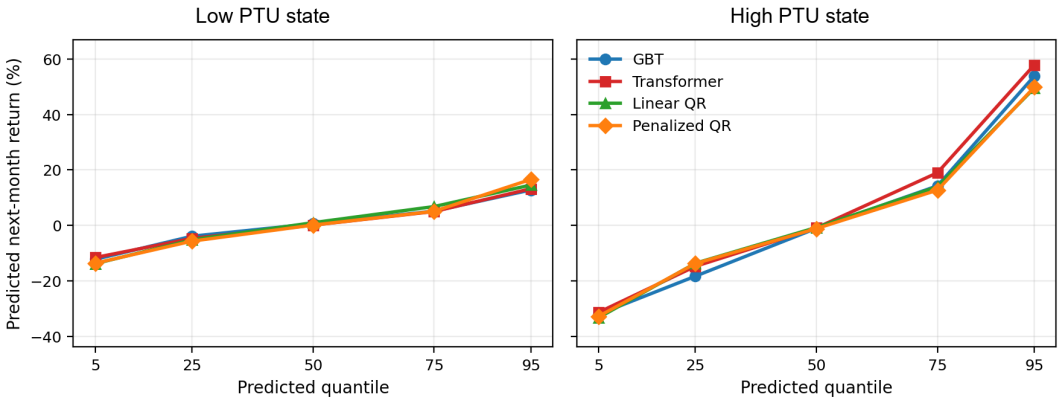


Figure IA.1: Predicted physical return distributions by PTU state. The figure visualizes predicted distribution shape by PTU state and complements the schematic main-text illustration of predictive physical tail width.

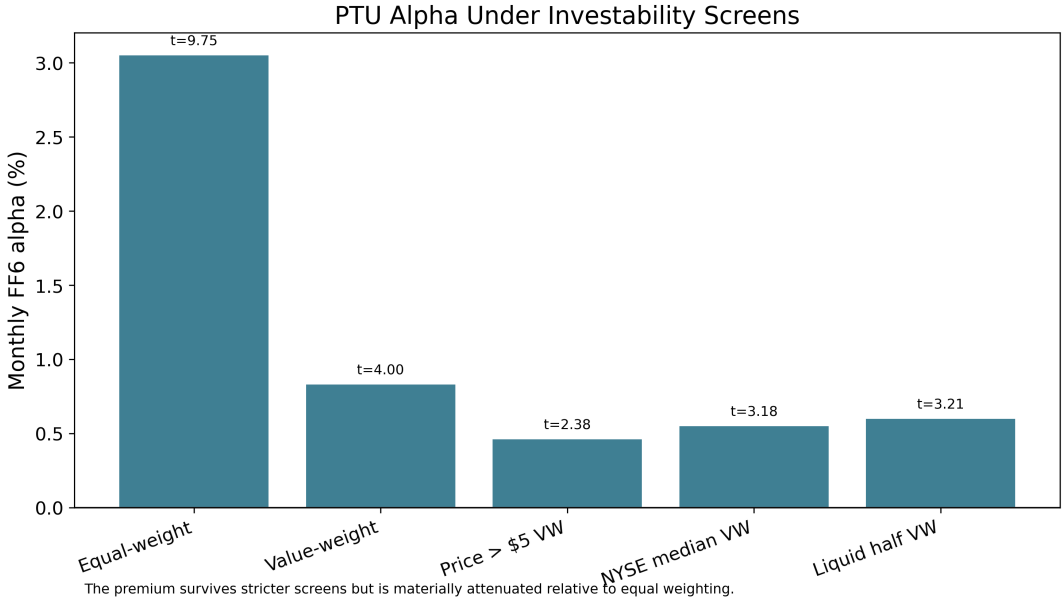


Figure IA.2: PTU alphas under investability screens. The figure reports monthly FF6 alphas for the direct PTU high-minus-low portfolio under equal weighting, value weighting, price screens, NYSE-size screens, and liquid-half screens.

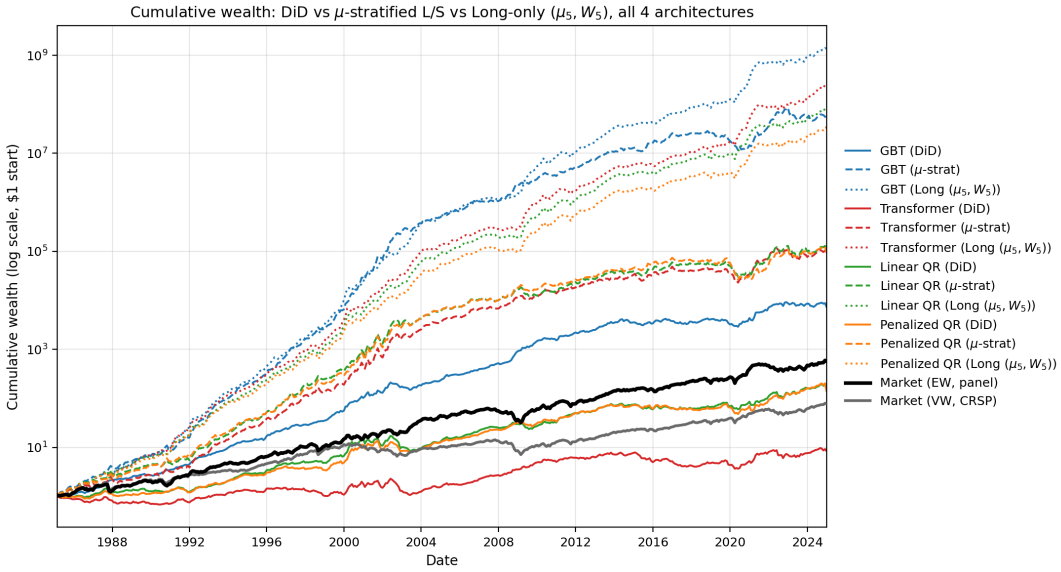


Figure IA.3: Cumulative wealth for PTU-related strategy implementations. The figure plots the time series of portfolio performance corresponding to the pricing tests.

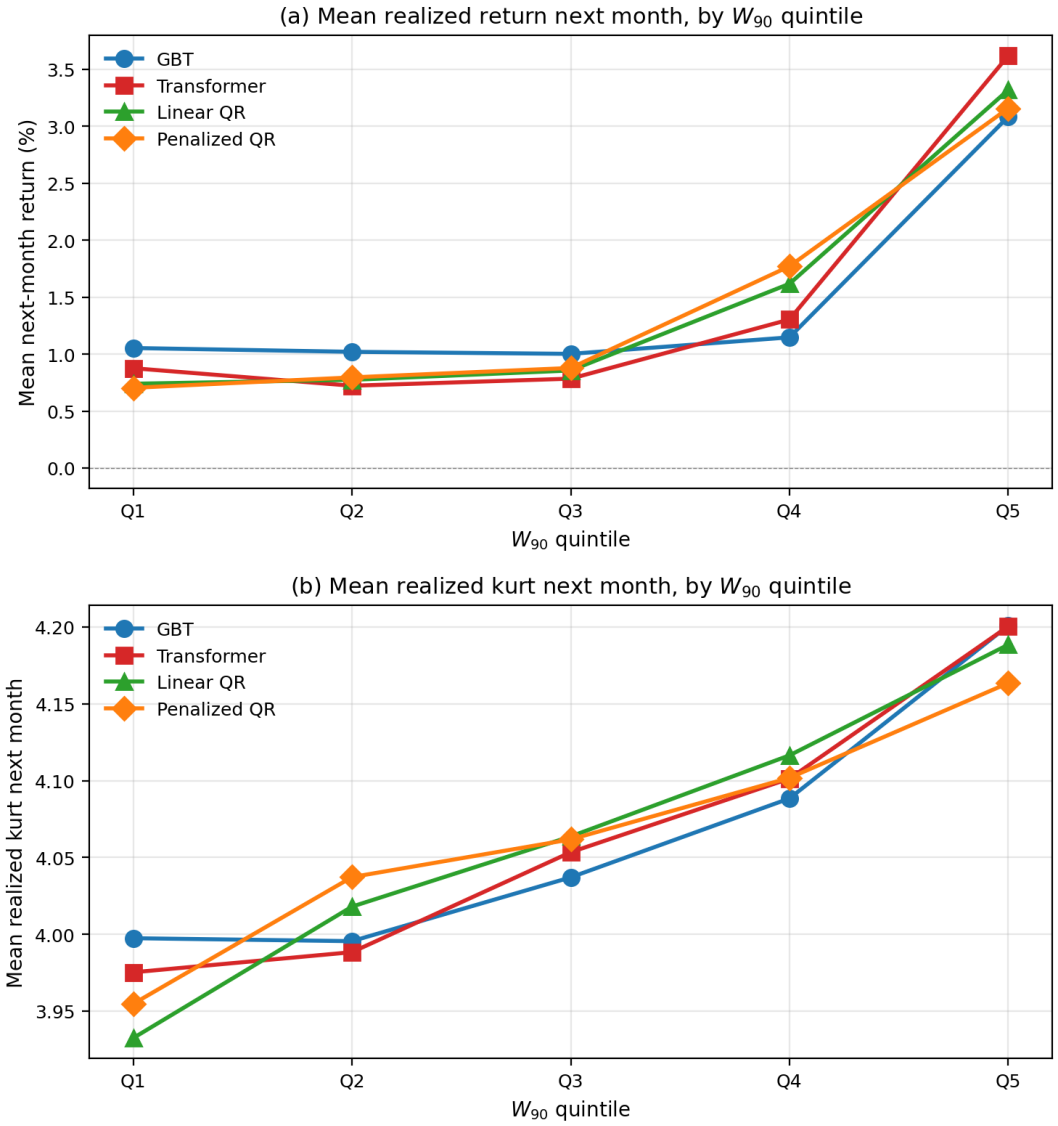


Figure IA.4: Quintile evidence for physical tail-width risk and return patterns.

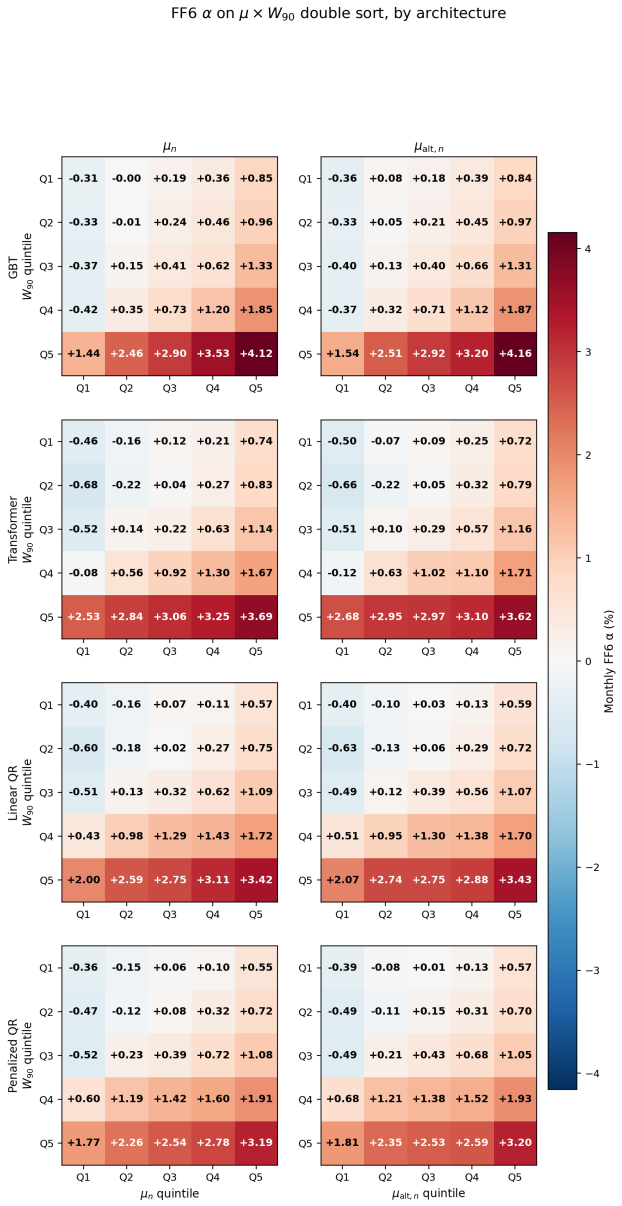


Figure IA.5: Alpha grid across PTU/W90 strategy variants.

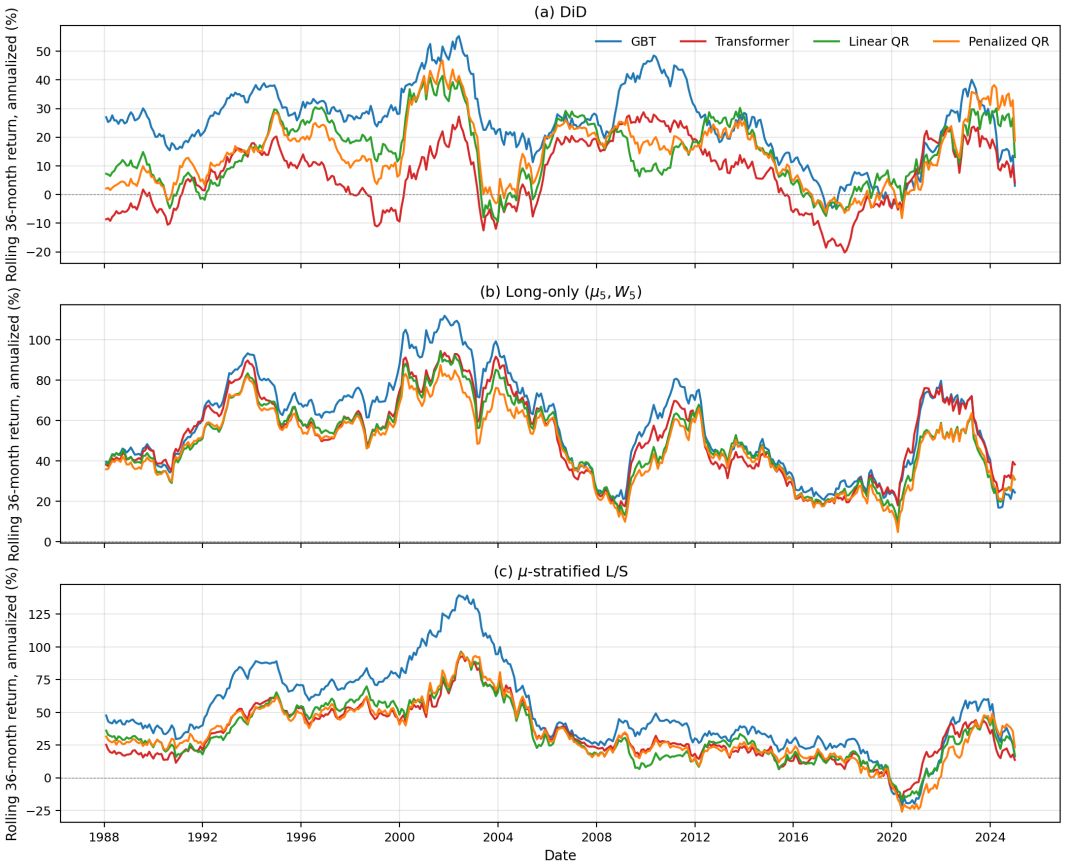


Figure IA.6: Rolling PTU-related strategy performance.

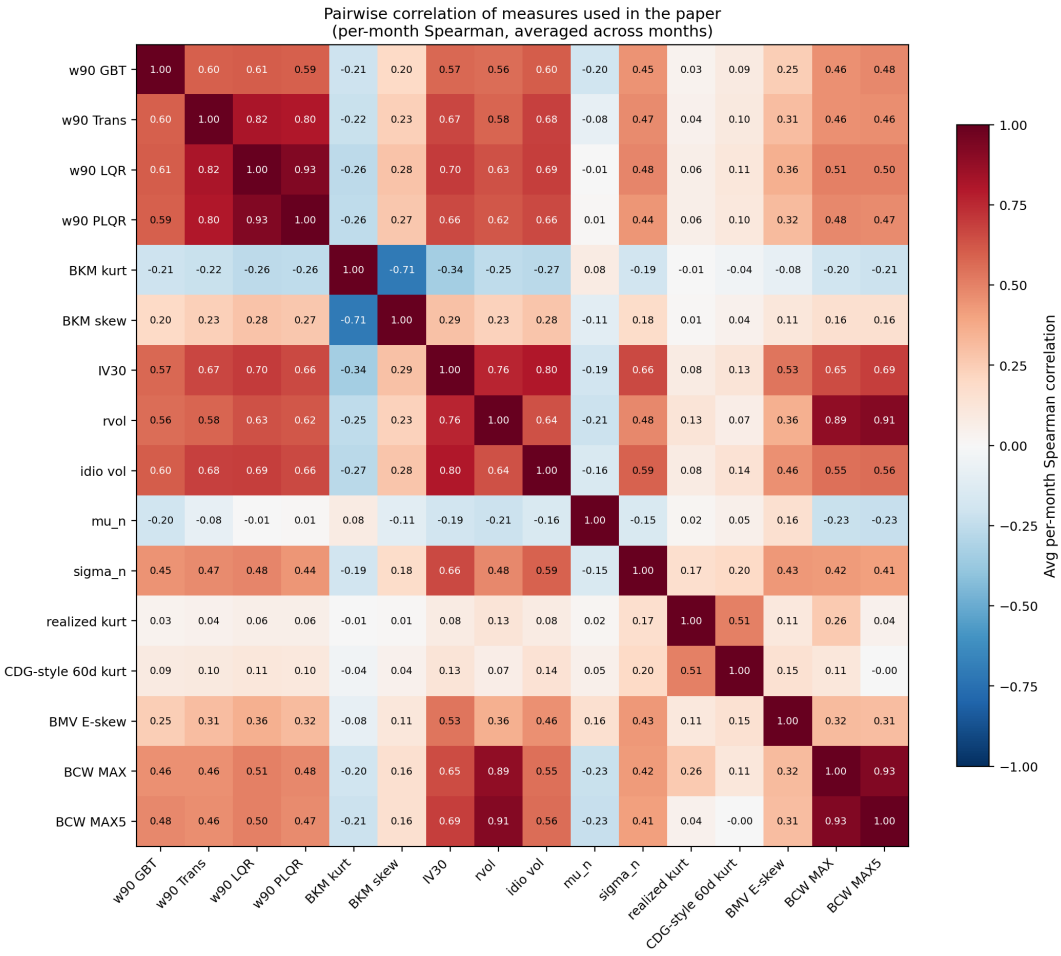


Figure IA.7: Correlation heatmap for W measures and related signals.

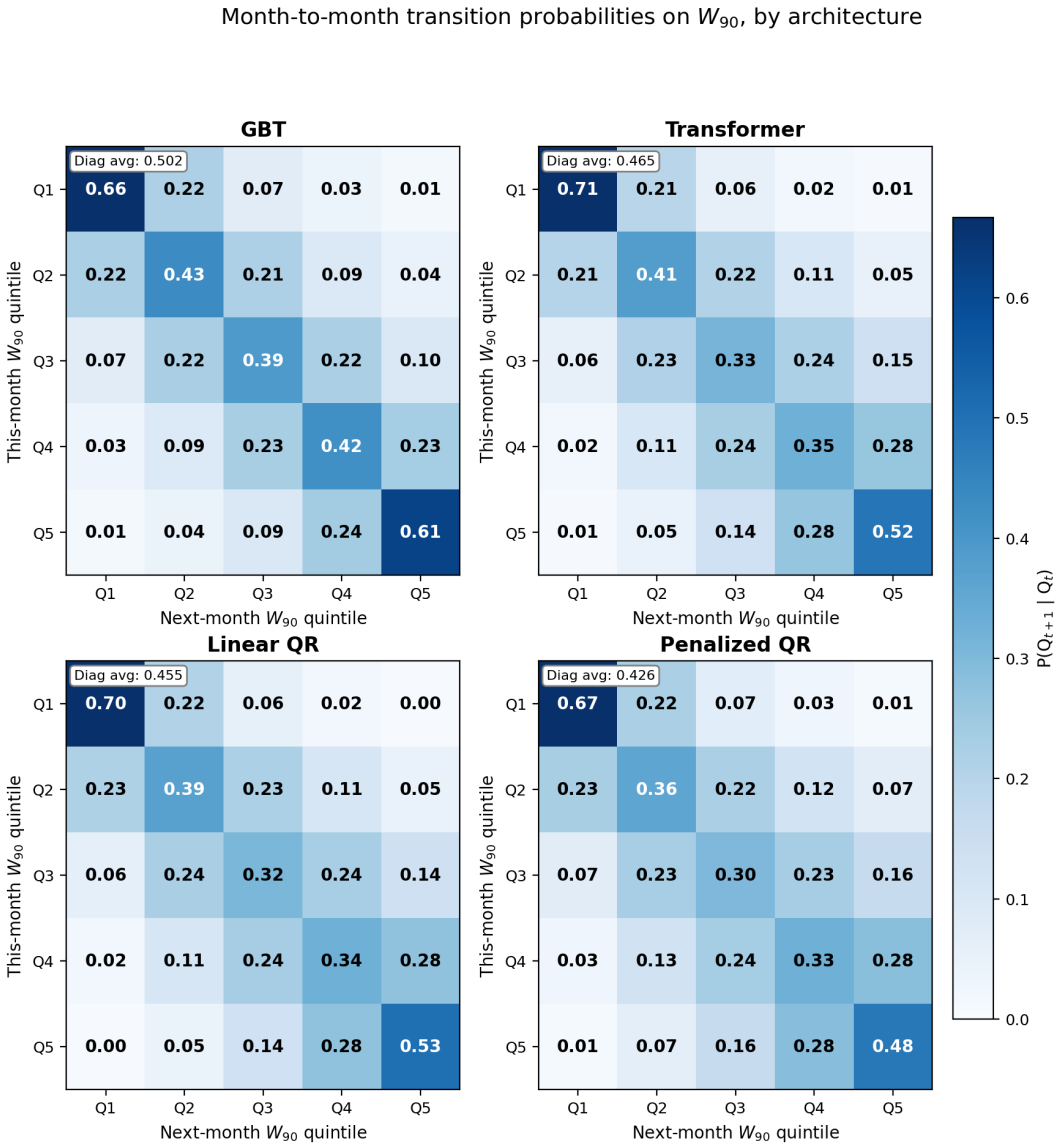


Figure IA.8: Transition behavior for PTU/W90 states.

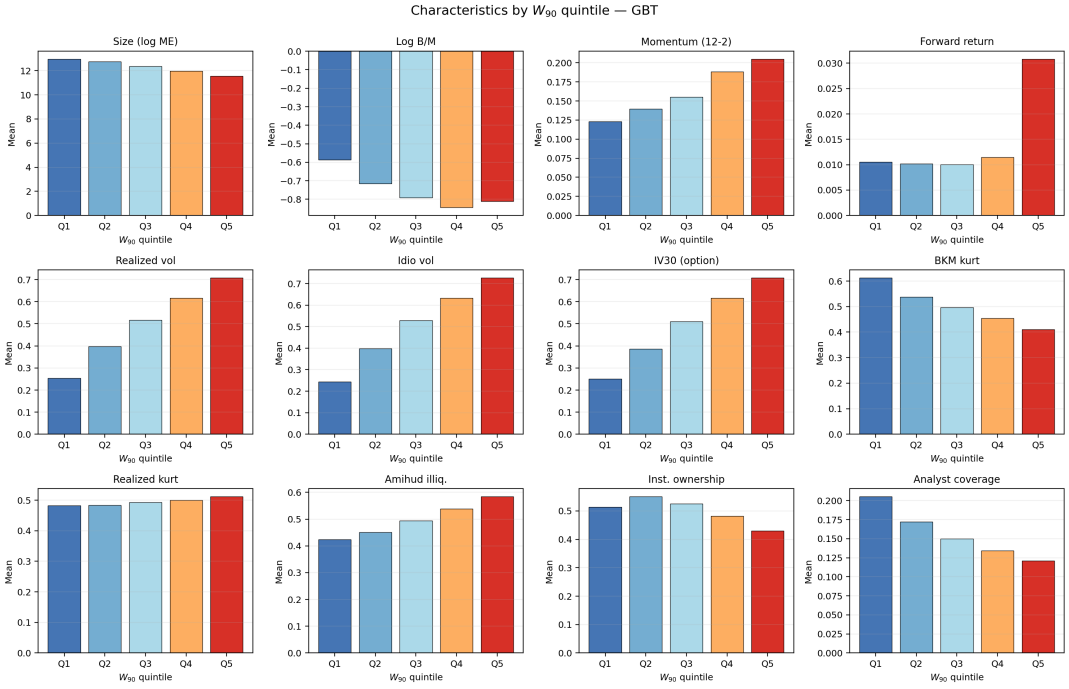


Figure IA.9: Characteristic profiles for the gradient-boosted-tree W90 signal.

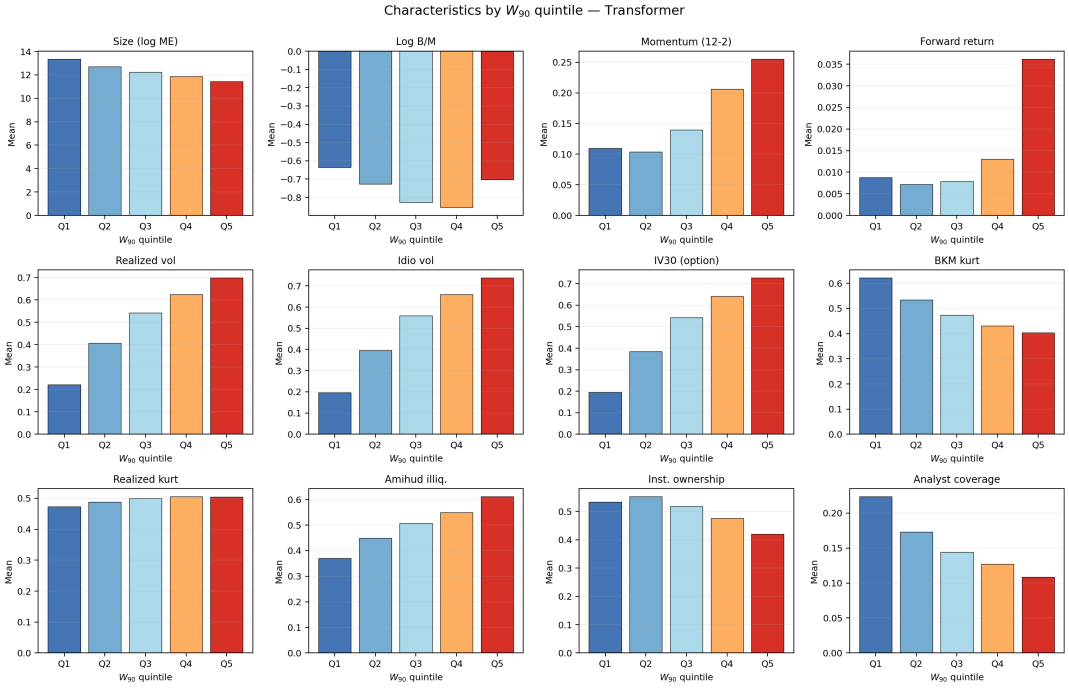


Figure IA.10: Characteristic profiles for the transformer W90 signal.

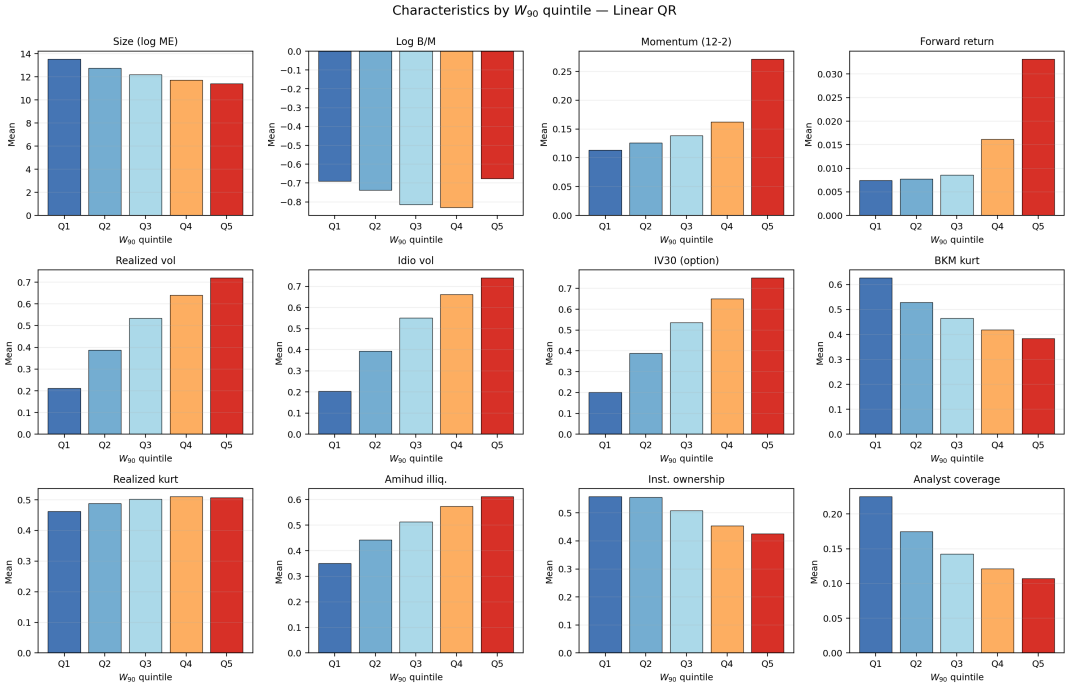


Figure IA.11: Characteristic profiles for the linear quantile-regression W90 signal.

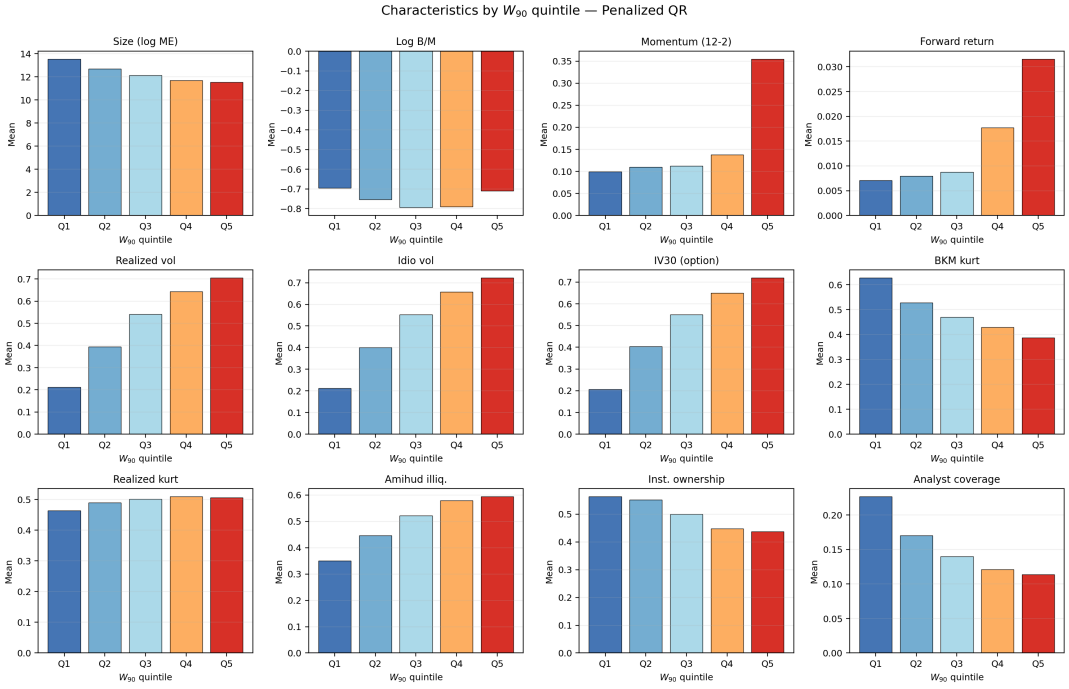


Figure IA.12: Characteristic profiles for the penalized quantile-regression W90 signal.

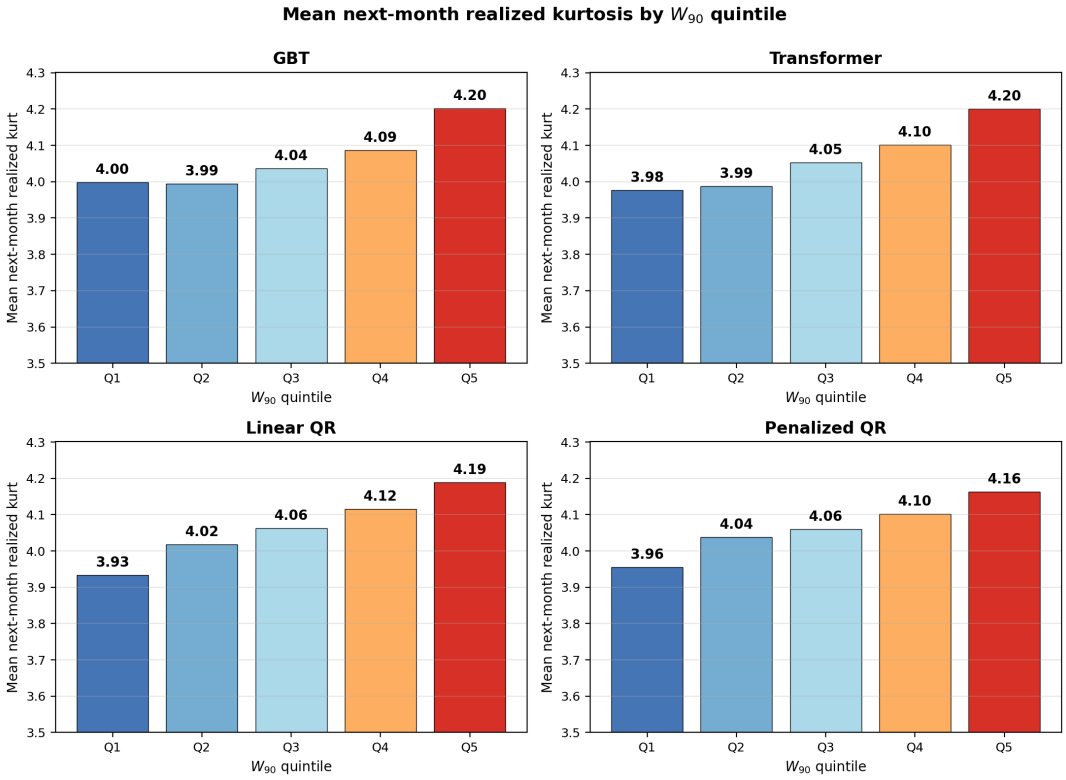


Figure IA.13: Future realized-kurtosis heatmap across sorting cells.



Figure IA.14: Rolling factor-adjusted alpha.

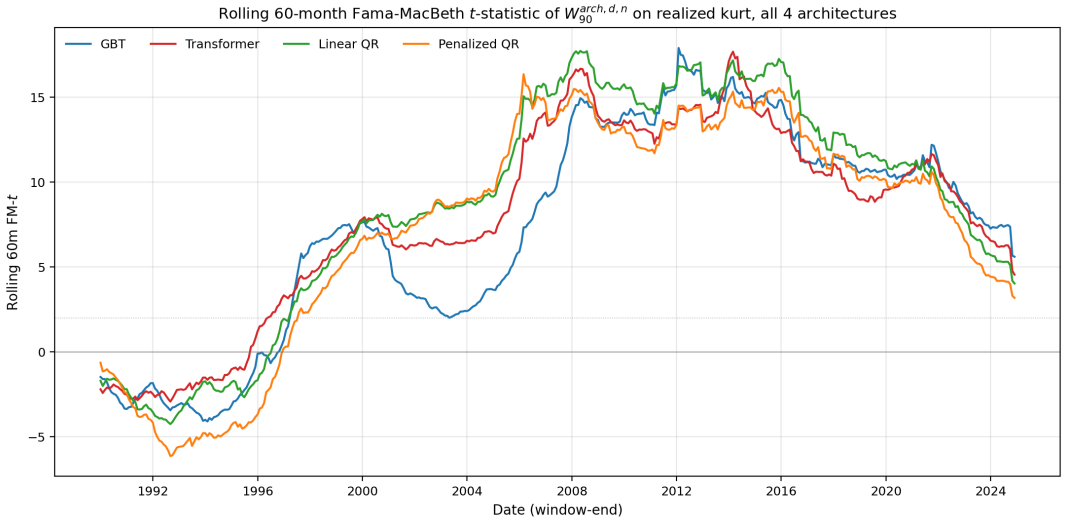


Figure IA.15: Rolling realized-kurtosis prediction t-statistic.