The Priced Dimensions of Asymmetric Risk

Matěj Nevrla*

August 11, 2026

Abstract

Do the many measures of asymmetric risk identify one downside factor or several priced dimensions? We study compensation for asymmetric risk in U.S. equities, separating robust premia, distinct return variation, and priced dimensions. Five measure families organize around three recurring priced dimensions rather than a single downside factor. Multiple dimensions receive compensation across periods and portfolio constructions, although their relative pricing strength varies over time. The resulting three-factor benchmark is not jointly spanned by leading factor models. The dimensions also exhibit distinct predictive relations with aggregate conditions. The results distinguish new measures from new priced risks.

Keywords: Asset pricing; Asymmetric risk; Bad-state risk; Factor models

JEL Classification: G11; G12; G14; C58

Data availability: The public data release accompanying this paper is available at https://github.com/matejnevrla/priced-dimensions-asymmetric-risk. It contains the ARM basis returns, ARM PCA portfolio returns, and the underlying long-format ARM implementation returns.

*University of Liverpool Management School and IITA, The Czech Academy of Sciences. Email: matej.nevrla@liverpool.ac.uk.

1 Introduction

A large literature studies how investors price exposure to large losses. Its empirical measures include specifications based on conditional downstate covariance, higher-moment risk, and exposures to tail-based factors, among others. These measures share a broad economic intuition: assets that perform poorly in adverse states should command compensation. Yet they are derived from different models and capture different features of return distributions. Although some existing studies compare measures or include competing controls in cross-sectional regressions, such comparisons speak mainly to whether one measure subsumes another in predicting average returns, not to whether the underlying compensation is the same, distinct, or partly redundant. The literature therefore leaves unresolved what its many measures collectively represent: alternative proxies for one latent downside factor, distinct priced dimensions, or a mixture of priced and unpriced variation.

This issue matters beyond the asymmetric-risk literature. Empirical asset pricing often moves from an economic mechanism to an empirical measure and then to a candidate factor. Yet a new measure need not identify a new source of expected returns. A measure motivated by a different theory may produce a premium that loads on an existing priced dimension, generate orthogonal but unpriced return variation, or reveal a low-variance dimension that nevertheless commands compensation because it covaries with the stochastic discount factor (Cochrane, 2005). The relevant question is therefore not how many asymmetric-risk measures have been proposed, but how many independent priced dimensions they reveal. Establishing that structure requires distinguishing a measure's economic motivation, the distinct return variation it generates, and whether that variation is priced.

We provide such a mapping for five measures proposed by distinct strands of the asymmetric-risk literature: hybrid tail covariance risk (HTCR) (Bali et al., 2014), multivariate crash risk (MCRASH) (Chabi-Yo et al., 2022), tail-risk beta (TR β) (Kelly and Jiang, 2014), lower-tail common idiosyncratic quantile beta (CIQ) (Baruník and Nevrla, 2026), and residual coskewness (COSK) (Harvey and Siddique, 2000). For each measure, 30 portfolio implementations vary sorting rules, weighting schemes, price screens, and breakpoint universes. This common grid separates within-family robustness from cross-family dimensionality and reveals whether the premia associated with differently motivated measures are substitutes, complements, or statistically distinct without being separately priced.

Table 1 illustrates the distinction. All five ARM slopes remain positive and significant in a joint monthly Fama–MacBeth (1973) regression after a Holm adjustment, yet HTCR, MCRASH, and TR β share one broad tail/crash priced dimension. Significant measure-level slopes therefore need not identify independent priced dimensions.

Table 1: **Cross-Sectional ARM Slopes and Priced Return Dimensions**

	Fama–MacBeth regressions		
ARM	Separate	Joint	Main priced dimension
HTCR	0.16	0.14	Broad tail/crash
	(4.64)	(4.45)	
MCRASH	0.08	0.04	Broad tail/crash
	(2.74)	(2.00)	
TR β	0.11	0.09	Broad tail/crash
	(2.71)	(2.46)	
CIQ	0.15	0.14	Lower-tail CIQ
	(3.51)	(3.48)	
COSK	0.09	0.07	Residual coskewness
	(2.99)	(2.58)	

Notes: This table reports monthly Fama–MacBeth slopes for five stock-level asymmetric-risk measures (ARMs) and their main common mappings to the priced ARM dimensions. The Separate column reports one-ARM regressions, the Joint column includes all five ARMs, and the final column reports the main priced dimension assigned by the return-space analysis. The ARMs are hybrid tail covariance risk (HTCR) of Bali et al. (2014), multivariate crash risk (MCRASH) of Chabi-Yo et al. (2022), tail-risk beta (TR β) of Kelly and Jiang (2014), lower-tail common idiosyncratic quantile beta (CIQ) of Baruník and Nevrla (2026), and residual coskewness (COSK) of Harvey and Siddique (2000). Each month, month-t excess stock returns are regressed on ARM estimates formed using information through month $t-1$. Each ARM is standardized cross-sectionally within month, so slopes are monthly percentage points per one-standard-deviation increase in adverse-risk exposure. COSK is sign-reversed so that larger values denote greater adverse exposure. The common sample covers January 1968–December 2024 and includes stocks with a formation-month price of at least \$5. Returns include delisting returns when available. The regressions use $T = 684$ months and an average of $\bar{n} = 1{,}820$ stocks per month. Reported slopes are time-series averages of the monthly cross-sectional estimates. Newey–West t-statistics using six lags are reported in parentheses. The final column records each measure's main common priced mapping. TR β also contains a distinct differential component represented by PC3, which does not satisfy the paper's priced-dimension criteria.

Our primary finding is thus that the priced ARM space is compact but multidimensional. The five ARM families organize around a recurring, economically anchored core of three priced dimensions: broad tail/crash, lower-tail CIQ, and residual coskewness. HTCR, MCRASH, and the common priced content of TR β load on the broad dimension. TR β also contains statistically distinct differential variation, but that component is not robustly priced. Lower-tail CIQ and residual coskewness contribute separate compensated variation. The evidence points to multiple compensated dimensions in both earlier and later periods, although the dimensions receiving the strongest pricing support vary over time. Their common return content also transfers across portfolio constructions.

This three-dimensional core provides a benchmark for evaluating whether additional asymmetric-risk measures contribute incremental compensated variation, while leaving open whether measures outside the five families examined here reveal further priced dimensions. The paper's central contribution is therefore a taxonomy of the priced asymmetric-risk space rather than another isolated premium.

The first implication is conceptual. Economic mechanisms, empirical measures, and priced dimensions are different objects. A new ARM should therefore be evaluated by locating it within the existing map and asking whether its residual return variation adds a priced dimension. For the same reason, controlling for one asymmetric-risk measure does not generally control for the full priced ARM space.

The second implication is practical. The three economically labeled anchor portfolios form a simple PCA-free ARM basis that translates the map into a benchmark for model evaluation. Its ability to carry priced content across portfolio constructions excluded from factor formation shows how the taxonomy can be used beyond the returns that define it. The basis provides a common reference for classifying new ARMs and identifying which dimensions of the ARM space standard factor models explain. It also provides an economically interpretable classification of the independently constructed option-based Bear Beta portfolios of Lu and Murray (2019). These applications call for dimension-specific alphas, joint tests, and residual diagnostics rather than one aggregate measure of model performance.

The third implication is economic. Broad tail/crash and lower-tail CIQ exhibit different predictive relations with aggregate states, whereas residual coskewness is not related to the risk-aversion and uncertainty states examined. This contrast provides supporting economic validation of the statistical taxonomy and is consistent with different forms of adverse-state compensation rather than interchangeable representations of one downside state. Taken together, the results show that asymmetric-risk measures form a compact, interpretable, and multidimensional priced structure. Mapping candidate measures into that structure provides a disciplined foundation for deciding whether they identify new priced risk or re-express compensation already present in the market.

Related literature. First, we contribute to research on compensation for adverse-state exposure. Farago and Tédongap (2018) derive multiple disappointment-related factors, while Massacci et al. (2025) allow the factor structure and risk premia to differ between normal and bear-market states. We take a complementary, measure-centered approach: rather than impose one adverse-state model, we ask how several existing asymmetric-risk measures jointly map into common return variation and distinct priced dimensions.

Second, our organizing objective builds on the economically interpretable factor structure of Fama and French (1993) and the lower-dimensional SDF representations of Kozak et al. (2020) and Lettau and Pelger (2020). Related work documents overlap across competing factor models (Hou et al., 2019; Ahmed et al., 2019) and tests whether candidate factors add pricing information beyond a high-dimensional factor set (Feng et al., 2020). We apply this organizing logic within an economically related ARM universe and explicitly distinguish covariance dimensionality from priced dimensionality. The common implementation grid, sample-separated tests, and multiplicity-adjusted inference also respond to concerns about factor proliferation and specification search (Harvey et al., 2016).

The paper proceeds as follows. Section 2 describes the measures, data, and portfolio construction. Section 3 maps the five measure families into economically interpretable return dimensions and evaluates their recurrence and relative pricing strength over time.

Section 4 uses the resulting map as an asset-pricing benchmark. Section 5 examines external option-based mapping and predictive aggregate-state relations, and Section 6 concludes.

2 Data and Variable Construction

This section defines the five asymmetric-risk measures (ARMs), describes their common data environment, and explains how stock-level estimates are converted into ARM premium returns. The premium construction holds the underlying ARM fixed while varying its portfolio implementation. The resulting return panel allows us to assess both within-family robustness and cross-family dimensionality.

2.1 Asymmetric-Risk Measures

We select five measures that represent different ways in which a stock can contribute to adverse portfolio outcomes. Their distinctions concern both the adverse event being measured and the form of dependence they capture. Some condition directly on extreme return realizations, some estimate exposure to a common tail state, and others capture nonlinear dependence outside a conventional covariance model. The goal is not to include every proposed downside or tail statistic, but to cover prominent, economically distinct constructions from the asymmetric-risk literature.[1]

A second selection criterion is empirical comparability. The measures considered here can be constructed from equity returns, standard firm information, and traded factor returns over a long time span and broad cross-section. This common data environment is important because the paper compares the return space generated by different ARMs rather than isolated estimates from different samples. Measures that require option data are economically valuable, but they typically imply shorter samples, different coverage, and additional data requirements. We therefore reserve option-based evidence for external validation rather than using it to define the baseline ARM map.

Hybrid tail covariance risk (HTCR) is the lower partial cross-moment of a stock return and the market return on days when the stock is below its own 10th percentile (Bali et al., 2014). The market return is centered at its own 10th-percentile threshold, but the market need not itself be in its tail. HTCR therefore measures how much an individual stock contributes to portfolio tail risk when that stock experiences an extreme loss. Its motivation is especially

[1] Appendix A.7 and Table A.1 apply the same implementation environment to related downside and higher-moment measures. They report the measures' composite premia, stability across implementations, and nonredundancy with conventional factors, thereby assessing whether any candidate warrants inclusion in the baseline ARM universe.

natural for underdiversified investors, for whom a stock-specific tail event can materially affect total wealth.

Multivariate crash risk (MCRASH) is the probability that a stock is in its own conditional lower 5% tail given that at least one of seven systematic factors is simultaneously in its conditional lower tail (Chabi-Yo et al., 2022). The conditioning factors are the market, size, value, profitability, investment, momentum, and betting-against-beta factors. Unlike HTCR, which is a tail cross-moment conditioned on a stock event, MCRASH is a conditional crash probability defined by the coincidence of stock and systematic-factor tail events.

Tail-risk beta (TR β) measures a stock's sensitivity to the aggregate tail-risk factor of Kelly and Jiang (2014). That factor summarizes the severity of daily idiosyncratic return realizations below the pooled cross-sectional 5th-percentile threshold. We estimate TR β from the predictive relation between the monthly tail-risk factor and the next month's stock return. The measure therefore captures exposure to aggregate variation in severe idiosyncratic tail realizations rather than the probability of a particular joint crash event.

Lower-tail common idiosyncratic quantile beta (CIQ) measures a stock's exposure to innovations in a quantile factor that captures common variation in the conditional 20th percentile of standardized residual returns (Baruník and Nevrla, 2026). The residualization removes linear Fama and French (1993) three-factor exposure before the quantile factor is estimated following Chen et al. (2021). The time-series beta captures whether a stock covaries with this common lower-tail state, not merely whether the stock has high idiosyncratic volatility. Economically, the measure is motivated by the idea that distress among constrained intermediaries can generate common deterioration in the lower tail of stocks' idiosyncratic returns, making exposure to this state a potential source of priced risk.

Residual coskewness (COSK) is a normalized comoment, estimated using monthly observations, between a stock's market-model residual and the squared demeaned market excess return (Harvey and Siddique, 2000). Negative COSK means that the stock tends to experience adverse idiosyncratic outcomes when absolute market movements are large. It is consequently a higher-moment dependence measure: it can detect an asymmetric relation with the market even when linear market beta has been removed. Because investors are expected to dislike negative COSK, we reverse the conventional high-minus-low portfolio return so that a positive reported premium corresponds to compensation for greater adverse COSK exposure.

These definitions show why the measures are economically related without being interchangeable. HTCR and MCRASH both relate stock-tail losses to market or systematic conditions, but they differ in the conditioning event: HTCR conditions on a stock-tail realization, whereas MCRASH conditions on a factor crash. TR β and CIQ both measure exposure to aggregate tail states constructed from idiosyncratic returns; the former summarizes extreme exceedance severity, whereas the latter captures common movements in a conditional quantile. COSK instead captures nonlinear market dependence without defining a discrete tail event. Appendix A provides the exact estimators, rolling windows, data-frequency choices, and timing conventions used in the analysis.

Although these measures are based on different theoretical principles, they do not necessarily correspond to separate priced dimensions. Two measures motivated by different mechanisms may ultimately capture the same compensated return variation; alternatively, a measure may generate statistically distinct return variation that does not command a distinct premium, or it may remain separately priced after accounting for the others. The central empirical purpose of the paper is to distinguish between these outcomes.

The empirical map developed below should therefore be interpreted as a benchmark for current and future ARMs rather than as an exhaustive list. Its value lies in providing a common return-space comparison of measures that are usually studied separately.

2.2 Data Sources and Sample Construction

The stock-level data come from CRSP. We include common stocks (share codes 10 and 11) traded on NYSE, AMEX, and NASDAQ. We use daily and monthly stock returns, prices, shares outstanding, and exchange identifiers, from which we construct lagged market capitalizations and excess returns. Monthly portfolio returns incorporate CRSP delisting returns when available. The risk-free rate and the Fama–French factors are obtained from Kenneth French's data library.

The asymmetric-risk measures are estimated at the stock-month level using either monthly or daily return data over a rolling window; Appendix A gives the exact estimators, minimum-observation requirements, and data frequencies for each measure. For a portfolio return realized in month t, all stock characteristics used for portfolio formation are based on information available through the end of month $t-1$. Portfolio-level price screens use month-$t-1$ prices, and value-weighted returns use month-$t-1$ market capitalizations, so the sorting variables and portfolio weights are predetermined with respect to the return being measured.

The data begin in January 1963 and end in December 2024. Because the ARM estimators require a return history before a stock can be sorted, we allow a five-year burn-in period. The first ARM portfolio return is therefore January 1968. A stock is eligible for a given portfolio return only when the holding-month CRSP return and the corresponding formation-month price, lagged market capitalization, exchange classification, and relevant ARM estimate are available. Consequently, the eligible cross-section may vary across measures and months, but each ARM family is evaluated using the same portfolio-construction grid described below.

2.3 Premia Construction

For each asymmetric-risk measure j, we construct premium returns $r_{j,s,t}$ under portfolio implementations indexed by s. Across implementations, the ARM definition is held fixed while price screens, breakpoint universes, sorting rules, and weighting methods vary. Common return behavior across implementations within family j indicates that its premium is not generated by one construction convention. Relationships across ARM families then allow us to assess whether their premium variation is empirically redundant.

Throughout the portfolio-construction and return-space analyses, month t denotes the month in which the portfolio return is realized. Portfolio formation uses information available at the end of month $t-1$. The sorting variable $X_{i,t-1}$, price $P_{i,t-1}$, and market capitalization used for portfolio weights are therefore observed before the return month. For each asymmetric-risk measure, we sort the eligible cross-section on $X_{i,t-1}$ and construct monthly high-minus-low portfolio returns. We consider three alternatives based on the absolute formation-month stock price: no minimum-price screen, $P_{i,t-1} \geq 1$, and $P_{i,t-1} \geq 5$. For each price filter, we sort stocks into portfolios based on the respective asymmetric-risk measure values using two alternative breakpoint conventions. In the first case, breakpoints are computed using all NYSE, AMEX, and NASDAQ stocks. In the second case, breakpoints are computed using NYSE stocks only, while all eligible stocks are assigned to portfolios using the resulting NYSE breakpoints.

In the decile specification, the low portfolio contains stocks below the 10th percentile of the sorting variable and the high portfolio contains stocks above the 90th percentile. In the quintile specification, the low and high portfolios contain stocks below the 20th percentile and above the 80th percentile, respectively. Stocks between the two cutoff points are not included in either leg of the high-minus-low return. The long-short portfolio return is therefore defined as

$$r_t^{H-L} = r_t^H - r_t^L, \tag{1}$$

where r_t^H is the return on the high-X portfolio and r_t^L is the return on the low-X portfolio. Portfolio returns are computed using both equal weights and value weights, where value weights are based on month-$t-1$ market capitalization. Each discrete long-short portfolio has zero net exposure and gross exposure of two.[2]

[2]We orient all premia so that positive values correspond to compensation for greater adverse asymmetric-risk exposure. For COSK, this requires multiplying the high-minus-low return by -1, because more negative coskewness corresponds to the less desirable risk exposure.

Finally, we construct a characteristic-managed long-short portfolio following Kozak et al. (2020). Within each formation month $t-1$, stocks are ranked according to the empirical distribution of the sorting variable, where the reference distribution is computed either from all NYSE, AMEX, and NASDAQ stocks or from NYSE stocks only. The rank is centered by subtracting its average across retained eligible stocks, and portfolio weights are normalized by the sum of absolute centered ranks:

$$w_{i,t-1} = \frac{\mathrm{rank}_{i,t-1} - \overline{\mathrm{rank}}_{t-1}}{\sum_{\ell} \left| \mathrm{rank}_{\ell,t-1} - \overline{\mathrm{rank}}_{t-1} \right|}. \tag{2}$$

The centering imposes zero net exposure, while the normalization sets gross exposure to one. The managed portfolio return is then given by

$$r_t^{\mathrm{managed}} = \sum_i w_{i,t-1} r_{i,t}^e. \tag{3}$$

This portfolio is long stocks with high values of the asymmetric-risk measure and short stocks with low values of the measure. Before constructing the managed portfolio, we exclude stocks whose formation-month market capitalization accounts for less than 0.01% of the total market capitalization of the eligible universe in that month.

Overall, for each asymmetric-risk measure the discrete grid combines three price filters, two breakpoint conventions, two weighting schemes, and two sorting definitions—top-minus-bottom 10% and top-minus-bottom 20%—to produce $3 \times 2 \times 2 \times 2 = 24$ implementations. The characteristic-managed portfolio provides a continuous-rank long-high, short-low strategy for each combination of the three price filters and two breakpoint conventions, adding six implementations. This yields 30 implementations for each ARM family and 150 ARM return series in total. The resulting 150-series panel is the input to our dimensionality analysis.

3 From ARM Measures to Priced Dimensions

The previous section constructs a balanced panel of 150 ARM long-short return series: five ARM families observed across a common grid of portfolio implementations. This panel allows us to ask whether the proposed measures represent one compensated bad-state premium, several distinct compensated dimensions, or a mixture of priced and unpriced variation.

Throughout, we use a priced ARM dimension in a reduced-form empirical sense to denote a statistically distinct component of the ARM return space that earns compensation that is robust across our empirical designs. The label identifies a compensated return component, not a structurally identified primitive risk factor or unique economic mechanism; the premium may reflect risk, mispricing, frictions, omitted exposures, or a combination.

3.1 Asymmetric-Risk Premia

We begin the empirical map by examining each proposed measure separately. Let $r_{j,s,t}$ denote the month-t return for ARM j under implementation s. We estimate specification-level premia from

$$r_{j,s,t} = \alpha_{j,s} + \varepsilon_{j,s,t}, \tag{4}$$

annualize the intercept, and use Newey–West standard errors with six lags for inference (Newey and West, 1987).

We also form the pre-specified family composite for each ARM,

$$\bar{r}_{j,t} = \frac{1}{30}\sum_{s=1}^{30} r_{j,s,t}, \tag{5}$$

which represents the return common to the implementations of ARM j. To test whether an ARM family contains incremental compensated variation relative to the other proposed ARM families, we estimate its other-family alpha as

$$\bar{r}_{j,t} = a_j + \sum_{k\neq j} b_{j,k}\bar{r}_{k,t} + u_{j,t}. \tag{6}$$

Table 2 summarizes the resulting ARM premia. It reports the family-composite premium, the other-family alpha a_j, and the cross-implementation distribution of specification-level premia across the 30 construction choices. All five composites earn positive and statistically significant premia, ranging from 2.73% to 4.85% per year. COSK is the most uniform: all 30 specifications are positive and significant at the 5% level. MCRASH, CIQ, and TR β are positive in every implementation, while HTCR is positive in 96.7% of implementations. TR β is the least uniform statistically, with only one-third of its individual implementations significant at the 5% level.[3]

The other-family alphas show why positive premia are not enough to establish five priced dimensions. HTCR and MCRASH retain positive but statistically weak alphas after controlling for the other family composites, suggesting that their premia are largely shared with the broader ARM space. TR β also has an insignificant other-family alpha. CIQ and COSK instead retain annualized alphas of 4.05% and 2.79%, respectively, indicating

[3]Appendix Figure B.1 reports the full distribution of premia within each ARM family, Appendix Figure B.2 disaggregates the premia by individual portfolio specification, and Appendix Table B.1 aggregates premia within the weighting, sorting, minimum-price, and breakpoint subsamples used in the implementation-robustness analyses.

Table 2: **ARM Premia and Incremental Family Alphas**

ARM	Composite premium (% p.a.)	Other-family α (% p.a.)	IQR (% p.a.)	Median Sharpe	Positive premia	Significant at 5%
HTCR	4.85	1.46	[2.58, 6.70]	0.46	96.7%	83.3%
	(3.34)	(1.24)				
CIQ	4.49	4.05	[2.92, 5.88]	0.35	100.0%	80.0%
	(3.01)	(3.12)				
COSK	3.37	2.79	[3.36, 3.95]	0.43	100.0%	100.0%
	(3.53)	(2.43)				
MCRASH	2.91	1.47	[1.61, 3.92]	0.34	100.0%	70.0%
	(2.82)	(1.77)				
TR β	2.73	1.92	[0.92, 3.70]	0.23	100.0%	33.3%
	(2.18)	(1.32)				

Notes: This table reports composite premia, incremental family alphas, and implementation-level summary statistics for the five ARM families. For each family, the composite premium is the equal-weighted monthly average across 30 portfolio implementations that vary weighting scheme, sorting rule, minimum-price screen, and breakpoint universe. Other-family α is the intercept from regressing the family composite on the other four family composites. Composite premia, other-family alphas, and implementation-level premia are annualized and reported in percent. IQR is the interquartile range of annualized implementation-level premia, and Median Sharpe is the median annualized Sharpe ratio across implementations. Positive premia and Significant at 5% are percentages of the 30 implementations. The monthly return sample covers January 1968–December 2024. Newey–West t-statistics using six lags are reported in parentheses below the composite premium and other-family alpha. Implementation-level significance uses conventional two-sided 5% tests with six-lag Newey–West standard errors.

incremental compensated variation.

3.2 Statistical versus Priced Dimensionality

The other-family regressions assess incremental compensation among the five family composites, but they do not reveal how many orthogonal components organize return variation in the full 150-portfolio panel or whether each component carries a premium. We therefore use PCA to recover common orthogonal variation and then test the mean returns of the resulting PC portfolios. A PC portfolio denotes a statistical return extracted by PCA; its premium and stability determine whether it provides evidence of a priced ARM dimension. This two-step analysis separates statistical dimensionality from priced dimensionality.

We estimate principal components from the centered covariance matrix of the 150 ARM returns. The eigenvectors are applied to raw returns, normalized to unit gross exposure, and signed so that average PC portfolio returns are positive. Because each input ARM return is self-financing, every PC portfolio is also a zero-cost portfolio; the unit-gross normalization refers to weights across the 150 ARM portfolios rather than to gross exposure in the underlying stocks. PCA eigenvectors are scale-identified only up to a constant. The unit-gross normalization therefore fixes the economic scale of the reported premia and loading shares; it does not affect the geometric PCA directions, explained-variance shares, PC correlations, or mean-return t-statistics.

Table 3 reports the PCA map for the first five principal components (PC1–PC5). The

Table 3: **PCA Map of ARM Portfolio Returns**

	Variance		PC portfolio			ARM-family loading shares				
PC	Share	Cum.	Mean	*t*-stat.	Sharpe	HTCR	CIQ	COSK	MCRASH	TR β
PC1	24.3	24.3	4.08	3.67	0.52	**47.9**	3.0	1.2	24.2	23.7
						(+47.9)	(-2.7)	(+0.2)	(+24.2)	(+23.7)
PC2	21.6	45.9	3.86	3.10	0.44	8.4	**68.4**	8.6	9.9	4.8
						(+8.4)	(+68.4)	(+8.6)	(-9.9)	(+2.7)
PC3	15.9	61.8	0.41	0.45	0.06	21.6	4.5	1.3	14.4	**58.3**
						(-21.6)	(+0.7)	(-0.6)	(-14.4)	(+58.3)
PC4	9.9	71.7	2.33	2.89	0.38	7.7	6.3	**70.2**	10.1	5.6
						(-6.5)	(-6.3)	(+70.2)	(+10.1)	(+1.7)
PC5	5.6	77.4	0.73	1.23	0.17	37.2	8.0	6.0	**37.3**	11.5
						(-9.8)	(+6.6)	(-0.0)	(+33.5)	(-1.4)

Notes: This table reports the variance, return, and ARM-family loading composition of the first five principal-component (PC) portfolios. Covariance PCA is applied to the centered matrix of 150 monthly ARM portfolio returns, formed from five ARM families and 30 implementations per family. Eigenvectors are applied to the raw returns, normalized so that the sum of absolute portfolio weights equals one, and signed so that reported average PC returns are positive. Share and Cum. are the individual and cumulative percentages of total return variance. Mean is the annualized PC-portfolio return in percent, and Sharpe is annualized. The family columns report gross loading shares, with signed net loading shares in parentheses; gross shares sum to 100 up to rounding. The monthly sample covers January 1968–December 2024. Mean-return *t*-statistics use Newey–West standard errors with six lags. The largest gross family share in each row is bold. TR β denotes tail-risk beta, CIQ denotes lower-tail commonality in quantiles, and COSK denotes residual coskewness. PCA loadings and return statistics are estimated on the same sample.

left two blocks report explained variance and PC portfolio premia; the right block reports ARM-family loading shares, with gross loading shares above signed net shares in parentheses.[4] PC1 and PC2 explain 24.3% and 21.6% of variance and earn annualized premia of 4.08% and 3.86%, respectively. PC4 explains 9.9% of variance and earns 2.33% per year. By contrast, PC3 explains 15.9%—more than PC4—but has a small and insignificant mean. PC5 is also unpriced. The ordering by variance therefore differs from the ordering by expected return: explaining an important part of the covariance structure is not sufficient for membership in the priced return space. PCA maps common return variation, while the subsequent mean-return tests determine which PC portfolios are priced; related approaches that incorporate expected returns directly into factor extraction likewise emphasize that economically important factors need not be the leading covariance PCs (Lettau and Pelger, 2020).

The loading shares give the economic content of the principal components. PC1 combines the three broad tail/crash families: HTCR supplies 47.9% of its gross loading mass, while MCRASH and TR β each supply about 24%. PC2 is dominated by CIQ, PC3 by TR β, and PC4 by COSK. PC5 is a mixed HTCR–MCRASH contrast rather than a separate family-specific dimension. Appendix Table C.4 shows that these PCs are not driven by a small number of individual portfolio implementations.

[4] For PC k, let $w_{j,s}^{(k)}$ denote its loading on implementation s of family j. Because the weights satisfy $\sum_j \sum_s |w_{j,s}^{(k)}| = 1$, the gross and signed net shares are $G_{j,k} = \sum_{s \in \mathcal{S}_j} |w_{j,s}^{(k)}|$ and $N_{j,k} = \sum_{s \in \mathcal{S}_j} w_{j,s}^{(k)}$, respectively. Gross shares sum to one; net shares allow opposing loadings within a family to offset. Both are reported in percent.

Combining the PCA evidence with results from Table 2 separates three concepts. Theoretical distinctiveness concerns the mechanism used to motivate a measure. Statistical distinctiveness concerns whether the measure generates orthogonal return variation. Priced distinctiveness additionally requires that the orthogonal variation command compensation. HTCR and MCRASH are theoretically distinct but largely priced substitutes. TR β has two empirical roles: it contributes materially to the priced broad dimension represented by PC1, while its highest-correlation family-level mapping is the unpriced PC3. Its common variation therefore belongs to broad tail/crash, but its differential variation does not identify another priced dimension. CIQ and COSK are both statistically and priced distinct. This is the economic map that the remaining tests evaluate.

To label and orient the three priced dimensions with observable returns, we define three anchor portfolios:

$$\begin{aligned} A_t^{\mathrm{BROAD}} &= \frac{1}{3}\left(\bar{r}_{\mathrm{HTCR},t} + \bar{r}_{\mathrm{MCRASH},t} + \bar{r}_{\mathrm{TR}\,\beta,t}\right), \\ A_t^{\mathrm{CIQ}} &= \bar{r}_{\mathrm{CIQ},t}, \\ A_t^{\mathrm{COSK}} &= \bar{r}_{\mathrm{COSK},t}. \end{aligned} \tag{7}$$

The broad anchor gives equal family weight to HTCR, MCRASH, and TR β because each contributes to PC1; it does not use PCA-estimated or optimized weights. This construction captures their common broad variation while allowing the contrast specific to TR β in PC3 to remain a separate statistical object. Each anchor portfolio supplies an economic label and sign orientation; we refer to the PC portfolio assigned to it as the matched PC portfolio. Section 4.1 later uses the same observable returns directly as ARM basis factors.

Appendix Table C.1 reports PCs 6–10. PC8–PC10 have conventionally significant in-sample means, reinforcing the point that pricedness is not mechanically ordered by explained variance. We therefore do not infer the stopping point from eigenvalue rank. The main table reports PC1–PC5 because these PCs contain the dominant mappings of the five family composites. The three-dimensional taxonomy captures the recurring, economically anchored core; later low-variance components identify finer and less stable residual structure.[5]

Appendix Table C.2 shows that multiple priced PC portfolios also appear under correlation PCA and when the panel is collapsed to five family-composite returns. Together, these exercises show that multidimensionality is not an artifact of covariance scaling, return-panel aggregation, or the five-component presentation of the main map.

[5] Section 3.3 and Appendix Table C.3 examine the stopping rule using training-only selection over PC1–PC10.

3.3 Stability of the Priced Map

The full-sample analysis yields a multidimensional economic map. We next assess whether its mapped dimensions remain compensated when estimation and evaluation months are separated. The map-informed cross-fit estimates PCA weights, component matching, and signs in training months and applies them to omitted months. It tracks the broad tail/crash, lower-tail CIQ, and residual coskewness dimensions using the anchor portfolios defined in equation (7).

The blocked cross-fit divides the monthly sample into contiguous time blocks. In each iteration, one block is omitted. Using the remaining blocks, we estimate PCA weights for PC1–PC5, correlate the candidate PC portfolio returns with the three anchor portfolios, choose the one-to-one assignment that maximizes total absolute correlation, and orient each matched PC portfolio toward a positive training-sample correlation with its assigned anchor. The resulting signed weights are applied to the omitted block, and stacking the omitted-block returns produces a full-length cross-fitted series for each dimension. We report 4-, 8-, and 12-fold designs to vary the omitted-block length and amount of training data. Inference uses Holm-adjusted two-sided p-values across the three dimensions within each design.[6] We therefore test the portability of the map-informed weights, matching, and signs, conditional on the three-dimension taxonomy.

Table 4 reports the results. All three dimensions earn positive premia and pass the 5% Holm threshold in every design. Across 4, 8, and 12 folds, annualized premia range from 3.00% to 3.79% for broad tail/crash, from 3.29% to 4.18% for lower-tail CIQ, and from 2.26% to 2.40% for residual coskewness. The results show that the map-informed weights, matching, and signs recover positively compensated portfolios in omitted observations across all three block designs.

Appendix Table C.3 asks the complementary selection question. Within every training sample, it re-estimates the PCA weights, maps all five family composites to PC1–PC10, orients the signs, and determines the selected dimension set before evaluating the omitted block. A principal component is selected when its oriented annualized training-sample mean is at least 1% and its two-sided mean-return test has a Holm-adjusted p-value below 0.10 across PC1–PC10. The 1% cutoff imposes minimal economic materiality; the deliberately permissive 10% Holm cutoff retains weaker but potentially recurring dimensions.

The average selected count ranges from 2.25 to 2.83, and the modal count is three in the 8- and 12-fold designs. Lower-tail CIQ and residual coskewness recur among the selected family-

[6] The Holm procedure orders the three conventional p-values and applies a step-down adjustment that controls the family-wise error rate. The table reports the adjusted p-values themselves, which are compared directly with 0.05.

Table 4: **Cross-Fitted Stability of the Priced ARM Map**

	Blocked cross-fits					
	4 folds		8 folds		12 folds	
Dimension	Mean	Holm p	Mean	Holm p	Mean	Holm p
Broad tail/crash	3.25	0.002	3.00	0.003	3.79	< 0.001
	(3.25)		(3.27)		(4.09)	
Lower-tail CIQ	4.18	< 0.001	3.35	0.008	3.29	0.012
	(4.10)		(2.88)		(2.74)	
Residual coskewness	2.26	0.005	2.40	0.008	2.29	0.012
	(2.80)		(2.89)		(2.75)	

Notes: This table reports held-out mean returns for the three matched PC portfolios under 4-, 8-, and 12-fold blocked cross-fits. Each pair of columns gives the annualized held-out mean and its Holm-adjusted p-value. For each design, the monthly sample is divided into contiguous blocks. PCA weights are estimated on the training blocks; PC1–PC5 are matched one-to-one to the broad tail/crash, lower-tail CIQ, and residual-coskewness anchor portfolios using training-sample correlations; and signs are oriented toward positive training-sample correlations with the assigned anchors. The signed weights are applied to the omitted block, and omitted-block returns are stacked across folds. Means are annualized percentages. The blocked designs use monthly ARM returns from January 1968 through December 2024. Newey–West t-statistics using six lags are reported in parentheses below the means. Holm p-values (Holm, 1979) adjust the three two-sided mean-return tests within each cross-fit design.

matched dimensions, while HTCR and MCRASH usually map to the same broad principal component. The PC portfolio matched to the TR β family is never selected, consistent with its differential variation not being separately priced. PC9 is the only unmatched later principal component to enter, doing so in 25% of folds under every partition and with less uniform held-out premia. The two exercises thus have distinct roles: Table 4 establishes the portability of the map-informed dimensions, while the selection-complete analysis shows that a training-only search generally recovers a small space centered on the same recurring priced dimensions.

The chronological diagnostic in Panel C of Appendix Table C.3 adds a temporal qualification. The pre-1996 training sample selects broad tail/crash, residual coskewness, and PC9. After 1995, lower-tail CIQ earns 4.13% per year with a Holm-adjusted p-value of 0.040, residual coskewness remains positive and significant, and broad tail/crash remains positive but is estimated less precisely. Multidimensionality therefore recurs, but the dimensions receiving the strongest pricing support vary over time. We next examine the economic mechanisms consistent with those dimensions.

3.4 Economic Interpretation of the Priced Map

The empirical map is consistent with a pricing kernel in which several adverse-state innovations matter rather than one scalar downside state. A schematic reduced-form representation is

$$m_{t+1} = m_0 + \theta_{\text{BROAD}} Z_{t+1}^{\text{BROAD}} + \theta_{\text{CIQ}} Z_{t+1}^{\text{CIQ}} + \theta_{\text{COSK}} Z_{t+1}^{\text{COSK}} + \eta_{t+1}, \tag{8}$$

where the Z terms denote conceptual adverse-state innovations in the pricing kernel and the coefficients summarize the associated reduced-form prices of exposure. They are distinct from the observable anchor portfolios in equation (7) and from the matched PC portfolios used in the empirical map. Equation (8) provides an organizing interpretation for why ARM families motivated by different adverse states can collapse into fewer priced return dimensions. After orienting the portfolio returns so that positive premia correspond to compensation for adverse exposure, a separately priced dimension is consistent with exposure to a marginal-utility component not represented by the other dimensions. These mappings are economic interpretations rather than structural identifications, and the channels need not be mutually exclusive.

The broad tail/crash dimension is naturally interpreted as exposure to severe joint downside states. HTCR conditions on a stock's own tail event, MCRASH conditions on crashes in systematic factors, and TR β captures aggregate variation in severe idiosyncratic tail realizations. Despite these differences, all three are designed to identify stocks whose losses occur when adverse outcomes become systemically concentrated. Their common contribution to PC1 is therefore consistent with compensation for one broad tail/crash dimension rather than separate compensation for each tail-event definition. This mapping is naturally related to nonlinear downside and disappointment-aversion pricing kernels, in which marginal utility responds disproportionately to sufficiently adverse aggregate outcomes (Farago and Tédongap, 2018).[7] TR β also loads on PC3, which isolates statistically important measure-specific variation without a robust premium. The map thus assigns its common priced content to broad tail/crash while separating its unpriced differential variation.

The lower-tail CIQ dimension has a different economic interpretation. The underlying CIQ measure captures exposure to common movements in the lower conditional quantile of residual stock returns after removing linear Fama–French three-factor exposure. The dimension's premium is therefore consistent with compensation for exposure to common deterioration in idiosyncratic downside outcomes. Intermediary asset-pricing models provide one possible mechanism: when constrained intermediaries lose risk-bearing capacity, assets exposed to them can experience synchronized adverse outcomes beyond conventional market-beta exposure (He and Krishnamurthy, 2013; Brunnermeier and Sannikov, 2014; He et al., 2017). This mechanism provides one reason why CIQ need not be subsumed by aggregate downside models or by a single traded intermediary-capital proxy. Its greater strength in equal-weighted implementations is consistent with greater portfolio weight on smaller or less

[7] Consistent with this interpretation, the factor-model evidence in Section 4.3 shows that downside and generalized-disappointment-aversion benchmarks produce the strongest attenuation of the broad tail/crash premium.

liquid stocks, while the extreme-sort results concentrate the premium among stocks with the largest CIQ exposures (Table 6 and Appendix Table C.5).

Residual coskewness is instead a higher-order dependence dimension. A stock with adverse residual coskewness tends to have poor idiosyncratic realizations when market movements are large in absolute value, even after removing linear market exposure. Compensation for this exposure is consistent with nonlinear marginal utility or higher-order aversion: investors dislike assets that worsen the skewness of wealth in turbulent states (Harvey and Siddique, 2000; Dittmar, 2002). Related evidence in Schneider et al. (2020) shows that coskewness can organize beta- and volatility-sorted low-risk anomalies. This interpretation is consistent with stronger attenuation by polynomial higher-moment controls than by conventional linear models. The empirical attenuation patterns are suggestive rather than formal model tests. Its separation from both broad tail/crash risk and lower-tail CIQ nevertheless indicates that higher-moment dependence is not merely another empirical label for crash exposure or common residual quantile risk.

These mechanisms sharpen the economic labels but do not determine the dimension count. The empirical stopping point instead combines stable anchor representation, distinct return variation, positive compensation, and sample-separated recurrence. The training-only PC1–PC10 exercise typically selects two to three principal components associated with the recurring priced dimensions. The three-dimensional taxonomy therefore captures the recurring, economically anchored core, while intermittent PC9 selection and implementation-level residual alphas identify finer and less stable residual structure. The next section uses the anchor portfolios directly as basis factors and evaluates their usefulness as an asset-pricing benchmark.

4 The ARM Map as an Asset-Pricing Benchmark

The recurrence documented in Section 3 makes the map useful beyond the PCA exercise itself. For asset-pricing applications, however, researchers need a representation that can be applied without re-estimating the full PCA rotation. The three anchor portfolios provide a compact set of controls for a simple economic question: does a new asymmetric-risk measure load on an existing priced ARM dimension, or does it contain a distinct source of compensated risk?

This section develops and evaluates that diagnostic benchmark in three steps. First, we use the anchor portfolios as a PCA-free ARM basis. Second, we assess whether its premia and incremental content persist across portfolio constructions and how well it represents the matched PC portfolios and implementation returns. Third, we examine whether standard factor models price the basis factors and matched PC portfolios and whether the remaining errors contain additional common structure.

4.1 A Three-Factor ARM Basis

We translate the map into a simple PCA-free basis by using the anchor portfolios in equation (7) directly as factor returns. We refer to these directly usable benchmark returns as basis factors. For each dimension $d \in \{\text{BROAD}, \text{CIQ}, \text{COSK}\}$, define

$$\begin{aligned} F_t^{\text{BROAD}} &\equiv A_t^{\text{BROAD}}, \\ F_t^{\text{CIQ}} &\equiv A_t^{\text{CIQ}}, \\ F_t^{\text{COSK}} &\equiv A_t^{\text{COSK}}. \end{aligned} \tag{9}$$

The return construction is unchanged; only its empirical role changes. In Section 3, these returns are anchor portfolios that label and orient priced dimensions. Here they are basis factors used directly in time-series regressions.[8] The broad factor equally weights the HTCR, MCRASH, and TR β family composites, while lower-tail CIQ and residual coskewness retain their respective family composites. Thus, TR β contributes its common broad-dimension variation without being treated as a separate priced dimension; PC3 continues to represent its unpriced differential variation. Each family composite averages its 30 implementations, and the broad factor gives each of its three families one-third weight. The construction therefore represents the broad implementation-space return without using PCA loadings or optimized weights.

Table 5 reports the basis premia and mutual alphas. In the full sample, annualized mean returns range from 3.37% to 4.49%, while mutual alphas range from 2.93% to 4.12%; all are positive and statistically significant. After 1995, all three factors continue to earn positive and significant mean returns. Lower-tail CIQ retains a significant mutual alpha of 4.36%, with a t-statistic of 2.36, whereas the positive mutual alphas for broad tail/crash and residual coskewness are estimated less precisely. The basis therefore preserves the three mapped dimensions without inheriting the in-sample PCA rotation, with lower-tail CIQ providing the most precise post-1995 incremental estimate.

[8] These factors are observable proxy portfolios for the mapped return dimensions, not direct estimates of the conceptual pricing-kernel innovations in equation (8).

Table 5: **Three-Factor ARM Basis**

	Full sample		1996–2024	
Factor	Mean	Mutual α	Mean	Mutual α
Broad tail/crash	3.49	3.51	2.71	2.34
	(4.02)	(3.97)	(1.99)	(1.64)
Lower-tail CIQ	4.49	4.12	5.98	4.36
	(3.01)	(2.84)	(3.12)	(2.36)
Residual coskewness	3.37	2.93	4.18	2.25
	(3.53)	(2.50)	(3.12)	(1.59)

Notes: This table reports mean returns and mutual alphas for the three map-informed ARM basis factors in the full sample and over 1996–2024. The basis factors are the anchor portfolios defined in equation (7) and are used directly as factor returns. The broad tail/crash factor equally weights the HTCR, MCRASH, and TR β family composites; the lower-tail CIQ and residual-coskewness factors are the CIQ and COSK family composites. Each family composite equally weights its 30 portfolio implementations. Mean is the annualized average return in percent. Mutual α is the annualized intercept from regressing one basis factor on the other two. The full monthly sample covers January 1968–December 2024; the second column block covers January 1996–December 2024. Newey–West t-statistics using six lags are reported in parentheses below the estimates.

4.2 Implementation Robustness of the ARM Basis

The full-grid factors in Table 5 average across the broad portfolio-construction space. We next reconstruct all three factors within the eleven weighting, sorting, price-screen, and breakpoint subsamples used in the implementation-specific analysis. This exercise holds the economic definition of each basis factor fixed while varying the portfolio implementations used to construct it. The mutual alpha of each subgroup-specific factor measures whether its premium is absorbed by the other two factors formed from the same subgroup.

Table 6 reports the results. Using Holm-adjusted p-values across the six estimates within each implementation subgroup, all 33 factor–subsample mean returns are positive and significant at the 5% level. Thirty-two of the 33 mutual alphas are also significant at 5%, and the remaining estimate is significant at 10%. The three basis factors therefore retain both their premia and their incremental content across implementation choices.[9]

We next ask whether the observable basis represents the three matched PC portfolios associated with the priced dimensions. We regress these portfolios on increasingly complete versions of the basis: broad tail/crash alone, broad tail/crash plus lower-tail CIQ, and the full three-factor basis. Root mean squared alpha and maximum absolute alpha summarize the remaining pricing errors, while the joint test evaluates whether all three alphas are zero. We conduct the exercise using both the full-sample PC portfolios and PC portfolios whose weights, matching, and signs are estimated through December 1995 and whose returns are evaluated over 1996–2024.

The left block of Table 7 shows why all three factors are useful. In the full sample, broad tail/crash alone leaves a root mean squared annualized alpha of 2.55% across the matched PC

[9] Appendix Table C.5 provides complementary covariance-space evidence by re-estimating and matching PC portfolios within the same subsamples. Every subgroup contains at least two priced matched PC portfolios, eight of the eleven contain all three, and the anchor matches remain strong despite changes in PC ordering.

Table 6: **ARM Basis within Implementation Subsamples**

Implementation restriction	N	Broad tail/crash Mean	Broad tail/crash Mutual α	Lower-tail CIQ Mean	Lower-tail CIQ Mutual α	Residual coskewness Mean	Residual coskewness Mutual α
Panel A. Weighting scheme							
Equal, discrete sorts	60	4.07***	3.88***	5.79***	5.24***	3.88***	2.59*
		(3.70)	(3.29)	(3.92)	(3.69)	(3.57)	(1.83)
Value, discrete sorts	60	4.06***	4.29***	4.61**	4.62**	3.87**	3.74**
		(3.56)	(3.96)	(2.18)	(2.26)	(2.81)	(2.51)
Equal, including managed	90	3.11***	2.97***	4.40***	4.01***	3.04***	2.17**
		(3.86)	(3.40)	(3.73)	(3.54)	(3.61)	(2.01)
Panel B. Sorting rule							
Top–bottom 10%	60	4.74***	4.80***	6.23***	5.85***	4.11***	3.36**
		(4.11)	(4.07)	(3.29)	(3.17)	(3.43)	(2.22)
Top–bottom 20%	60	3.39***	3.43***	4.17**	3.75**	3.65***	3.34**
		(3.68)	(3.70)	(2.68)	(2.47)	(3.56)	(2.75)
Characteristic-managed	30	1.21***	1.18***	1.63**	1.60**	1.36***	1.20**
		(4.02)	(3.96)	(2.40)	(2.33)	(3.17)	(2.53)
Panel C. Minimum-price screen							
No price screen	50	2.96***	2.93***	4.74***	4.29***	3.55***	3.06***
		(3.09)	(3.00)	(3.14)	(2.93)	(3.71)	(2.63)
Price $\geq$ \$1	50	3.89***	3.88***	4.62***	4.23***	3.34***	2.87**
		(4.29)	(4.21)	(3.08)	(2.88)	(3.49)	(2.39)
Price $\geq$ \$5	50	3.64***	3.70***	4.10**	3.85**	3.22***	2.83**
		(4.58)	(4.66)	(2.75)	(2.65)	(3.35)	(2.47)
Panel D. Breakpoint universe							
All exchanges	75	3.97***	3.93***	5.02***	4.70***	3.43***	2.82**
		(4.24)	(4.12)	(3.24)	(3.09)	(3.57)	(2.37)
NYSE	75	3.02***	3.09***	3.95**	3.55**	3.32***	3.03**
		(3.70)	(3.76)	(2.73)	(2.53)	(3.48)	(2.66)

Notes: This table reports mean returns and mutual alphas for ARM basis factors reconstructed within implementation subsamples. Panels vary the weighting scheme, sorting rule, minimum-price screen, and breakpoint universe. Within each row, the broad tail/crash factor equally weights the subgroup-specific HTCR, MCRASH, and TR β composites; the lower-tail CIQ and residual-coskewness factors are the subgroup-specific CIQ and COSK composites. Mean is the annualized average return in percent, and mutual α is the annualized intercept from regressing one subgroup-specific basis factor on the other two. The monthly sample covers January 1968–December 2024. N is the number of underlying ARM portfolio-return series used to construct the row's family composites. Newey–West t-statistics using six lags are reported in parentheses. Within each implementation row, Holm adjustment is applied jointly across the three mean-return and three mutual-alpha tests. The first two weighting rows exclude characteristic-managed portfolios because they have no value-weighted counterpart; the third weighting row includes them in the equal-weighted set. *, **, and *** denote two-sided Holm-adjusted significance at the 10%, 5%, and 1% levels, respectively.

portfolios. Adding lower-tail CIQ reduces it to 1.38%, and the full basis reduces it to 0.15%, at which point the joint alpha null is no longer rejected ($p = 0.196$). For the pre-1996 map evaluated over 1996–2024, the full-basis root mean squared alpha is 0.20% and the joint-test p-value is 0.688. Thus, the basis is not rejected as spanning the matched PC portfolios in either design.

The right block repeats the nested comparison using each of the 150 implementation returns as a test asset. These returns comprise five ARM families formed under the same 30 construction rules, each defined by a weighting scheme, minimum-price screen, sorting rule, and breakpoint universe. For each rule, we remove from factor construction all five corresponding portfolios—one from each ARM family—recompute each family composite from its remaining 29 implementations, and reconstruct the three basis factors. This leave-one-rule-out (LORO) design prevents the held-out rule from entering the basis through any ARM family.

Table 7: **Pricing Diagnostics for the ARM Basis**

Sample	Basis	K	PC regressions RMS α	Max $\lvert\alpha\rvert$	Joint p	LORO implementation regressions Median $\lvert\alpha\rvert$	90th pct. $\lvert\alpha\rvert$	Sig. (%)
Full	Broad	1	2.55	3.81	0.001	1.48	4.81	44.0
Full	Broad + CIQ	2	1.38	2.39	0.032	1.33	3.42	38.7
Full	Broad + CIQ + COSK	3	0.15	0.25	0.196	0.59	2.45	18.7
1996–2024	Broad	1	2.81	3.86	0.009	1.74	6.01	30.7
1996–2024	Broad + CIQ	2	0.87	1.38	0.312	1.82	3.40	15.3
1996–2024	Broad + CIQ + COSK	3	0.20	0.31	0.688	1.10	3.27	14.7

Notes: This table reports pricing diagnostics for sequential one-, two-, and three-factor versions of the ARM basis. The left block evaluates three matched PC portfolios; the right block evaluates the 150 ARM implementations using leave-one-rule-out (LORO) factors. K is the number of basis factors. In the PC block, the test assets are the PC portfolios matched to the broad tail/crash, lower-tail CIQ, and residual-coskewness anchors. RMS α is the root mean squared annualized alpha across the three assets, Max $|\alpha|$ is the largest absolute annualized alpha, and Joint p tests whether the three alphas are jointly zero. In the LORO block, all five portfolios formed under a test asset's weighting, price-screen, sort, and breakpoint rule are excluded, one from each ARM family; family composites and basis factors are reconstructed from the remaining 29 implementations per family. Median $|\alpha|$ and 90th pct. $|\alpha|$ summarize absolute annualized implementation-level alphas, and Sig. is the percentage significant at 5%. Full uses monthly returns from January 1968 through December 2024. For the 1996–2024 PC block, PCA weights, matching, and signs are estimated through December 1995 and applied from January 1996 through December 2024; the LORO block is evaluated over the same latter period. PC Joint p is based on a six-lag HAC Wald test. LORO significance uses two-sided 5% tests with six-lag Newey–West standard errors. All alphas are annualized percentages.

For each target portfolio, we estimate the same three nested time-series regressions used in the left block: broad tail/crash alone, broad tail/crash plus lower-tail CIQ, and the full basis. We repeat the procedure across all 150 targets in the full and 1996–2024 samples and use six-lag Newey–West inference for each annualized alpha. The table summarizes the alpha distribution by its median, 90th percentile, and share significant at the 5% level.

The LORO results also favor the complete basis, although the sequential improvement is less uniform after 1995. Broad tail/crash alone leaves median absolute alphas of 1.48% in the full sample and 1.74% over 1996–2024. The complete basis lowers these medians to 0.59% and 1.10%, respectively. The corresponding 90th percentiles are 2.45% and 3.27%, while 18.7% and 14.7% of the implementation alphas are significant. The reductions remain when each test rule is excluded from basis construction, while the upper tail of residual alphas identifies implementation-specific variation beyond the mapped core.

Appendix Table C.7 extends the separation by withholding entire classes of portfolio-construction rules. Across eleven train–test configurations based on weighting, sorting, breakpoints, and price screens, the full basis lowers root mean squared pricing errors in every case. The reductions range from 45.2% to 70.7% in the full sample and from 40.3% to 60.9% over 1996–2024. The economic map consequently transfers beyond the particular implementations used to construct each set of basis factors.

Taken together, the PC regressions, LORO, and disjoint-implementation results make the basis useful as a diagnostic benchmark for the broad implementation universe. A new measure whose alpha is statistically indistinguishable from zero after conditioning on the basis may offer a new mechanism or implementation of an existing priced dimension; one that retains both a premium and common residual return variation is a candidate addition to the taxonomy.

4.3 Dimension-Specific Factor-Model Evaluation

We ask whether leading asset-pricing models span the three recurring priced ARM dimensions represented by the map. We regress both the directly usable ARM basis factors and PC1–PC5 on benchmarks from six classes: traded-factor, behavioral/mispricing, intermediary-capital, statistical-factor, higher-moment, and downside/disappointment models. Differences in attenuation across the three dimensions provide mechanism-consistent diagnostics for the interpretations in Section 3.4.

The traded benchmarks range from the CAPM of Sharpe (1964) and the three-, five-, and six-factor models of Fama and French (1993, 2015, 2018) to the momentum extension of Carhart (1997), the q5 model of Hou et al. (2021), and the six-factor model of Barillas and Shanken (2018) with the HML Devil value factor of Asness and Frazzini (2013). Additional specifications augment FF6 with the betting-against-beta factor of Frazzini and Pedersen (2014), the short-term reversal factor of Jegadeesh (1990), or the liquidity factor of Pastor and Stambaugh (2003).

The remaining classes draw on the mispricing model of Stambaugh and Yuan (2017), the behavioral factors of Daniel et al. (2020), the intermediary-capital factor of He et al. (2017), the statistical factors of Lettau and Pelger (2020), the higher-moment and skewness specifications of Harvey and Siddique (2000); Dittmar (2002); Langlois (2020), and the downside/disappointment models of Farago and Tédongap (2018). Appendix Table C.8 gives the exact factor vector, notation, and primary reference for every model. “PCA” and “RP-PCA” in this subsection denote external statistical factors, not components extracted from the ARM portfolios.[10]

Table 8 reports annualized alphas for the three ARM basis factors.[11] The results reveal a clear hierarchy: lower-tail CIQ remains unexplained across all benchmark classes, whereas broad tail/crash is particularly attenuated by downside models and residual coskewness by several higher-moment specifications. Broad tail/crash remains positive under every benchmark, with alphas from 0.55% to 3.22%, and is significant at the 5% level in 13 of the 20 specifications. Under Downside2, GDA3, and GDA5, its alpha falls to 0.75%, 1.33%, and 1.61%, respectively, and is statistically insignificant in each case. Because these benchmarks

[10]Benchmark-factor returns not obtained from Kenneth French’s data library are obtained from the authors’ public data libraries or other maintained public sources when available.

[11]Each regression uses the longest sample available for its factor set; the rows therefore assess spanning by each model rather than compare models over a common sample.

Table 8: **Means and Factor-Model Alphas of the ARM Basis**

Model	K	T	Broad tail/crash	Lower-tail CIQ	Residual coskewness
Unadjusted					
Mean return	0	684	3.49*** (4.02)	4.49*** (3.01)	3.37*** (3.53)
Benchmark traded factors					
CAPM	1	684	3.12*** (3.49)	6.22*** (4.59)	3.63*** (3.91)
FF3	3	684	3.22*** (4.28)	4.45*** (3.71)	3.02*** (3.35)
Carhart4	4	684	2.53*** (3.28)	5.43*** (4.29)	3.17*** (3.53)
FF5	5	684	3.08*** (4.11)	4.31*** (3.58)	2.40*** (2.72)
FF6	6	684	2.46*** (3.16)	5.17*** (4.07)	2.53*** (2.79)
FF6+BAB+STR	8	684	3.09*** (3.76)	4.78*** (3.18)	2.34** (2.43)
FF6+LIQ	7	684	2.55*** (3.24)	4.89*** (3.75)	2.40*** (2.59)
q5	5	684	1.69* (1.92)	5.14*** (3.04)	2.82*** (2.74)
BS6	6	684	1.56* (1.79)	4.13*** (2.97)	1.47 (1.49)
Behavioral / mispricing					
SY4	4	588	2.80*** (2.87)	4.42*** (3.23)	2.36** (2.27)
DHS3	3	558	0.55 (0.54)	5.02*** (3.38)	2.83** (2.51)
Intermediary capital					
HKM2	2	587	2.48*** (2.80)	5.63*** (4.21)	3.41*** (3.40)
Statistical factors					
PCA	5	600	2.39*** (2.82)	3.84*** (3.08)	2.27** (2.26)
RP-PCA	5	600	2.24** (2.05)	4.13** (2.51)	1.99 (1.27)
Higher moments / skewness					
Coskewness	2	684	1.86 (1.60)	4.25*** (2.88)	2.52* (1.96)
Cokurtosis	3	684	2.21** (1.99)	4.07*** (2.69)	2.05 (1.55)
FF6+PSS	7	600	2.28*** (2.69)	5.46*** (4.33)	2.24** (2.25)
Downside / disappointment					
Downside2	2	684	0.75 (0.59)	6.38*** (3.89)	3.10** (2.29)
GDA3	3	684	1.33 (1.07)	6.11*** (3.78)	2.78* (1.93)
GDA5	5	684	1.61 (1.31)	5.39*** (3.27)	2.34* (1.69)

Notes: This table reports unadjusted mean returns and factor-model alphas for the three ARM basis factors. The first row contains mean returns; subsequent row blocks contain intercepts from regressions on alternative benchmark models. The broad tail/crash factor equally weights the HTCR, MCRASH, and TR β family composites; the lower-tail CIQ and residual-coskewness factors are formed from the CIQ and COSK composites. Each family composite equally weights 30 portfolio implementations. Returns and alphas are annualized percentages. The underlying ARM returns are monthly from January 1968 through December 2024. Each benchmark regression uses the model's maximum available overlap, and T reports the resulting number of months. Newey–West t-statistics using six lags are reported in parentheses below the estimates. K is the number of benchmark factors. Appendix Table C.8 defines every model and its factor vector. *, **, and *** denote conventional two-sided significance at the 10%, 5%, and 1% levels, respectively.

are designed to price nonlinear exposure to sufficiently adverse aggregate outcomes, their ability to absorb the premium provides mechanism-consistent support for interpreting the broad dimension as compensation for concentrated downside exposure.

Lower-tail CIQ is the most robust basis factor: its alpha ranges from 3.84% to 6.38% and is significant at the 5% level under every benchmark, including downside/GDA models. The intermediary-capital specification leaves an alpha of 5.63%, with a t-statistic of 4.21, so its traded factor does not span the CIQ basis factor. Residual coskewness is positive throughout and significant at the 5% level in 14 specifications. Its alpha is marginal under the quadratic market-return specification and insignificant under the cubic specification, consistent with higher-moment controls absorbing part of this dimension.

Appendix Table C.9 reports the corresponding alphas for PC1–PC5 under the same benchmark models. PC2, the PC portfolio aligned with lower-tail CIQ, has annualized alphas from 3.20% to 5.27% and remains significant under every benchmark. PC1, aligned with broad tail/crash, is significant at the 5% level under 13 of the 20 models but is attenuated by several behavioral, higher-moment, and downside specifications. PC4, aligned with residual coskewness, remains significant under four models. The two representations therefore yield the same broad ordering: lower-tail CIQ is most difficult for the benchmark models to explain, while broad tail/crash and residual coskewness are more sensitive to downside and higher-moment controls.

Appendix Tables C.10 and C.11 report the corresponding six-lag HAC Wald tests of jointly zero alphas. The tests reject for both the three basis factors and PC1–PC5 under every benchmark. Thus, individual dimensions differ in their sensitivity to existing models, but no benchmark jointly prices either the three basis factors or the first five ARM PC portfolios.

4.4 Are the Remaining Errors Systematic?

Section 4.3 shows that no benchmark jointly spans the ARM basis or PC1–PC5. Those alpha tests establish that pricing errors remain, but they do not reveal whether implementation-level errors are isolated intercepts or share common return variation across the 150 ARM portfolios. We therefore use PCA as a residual diagnostic. We compare the common priced structure left after adjustment for each standard benchmark with that left after ARM-basis adjustment, and then supplement this comparison with an ex post portfolio-specific oracle.

For each factor set and ARM implementation, we remove the fitted factor component but retain the intercept. The adjusted return therefore equals the estimated alpha plus residual, and its sample mean equals the corresponding factor-model alpha. We re-estimate PCA from the centered covariance matrix of the adjusted returns separately for every previously considered benchmark model.

Table 9 reports the first five residual principal components. Every standard benchmark leaves at least two PC portfolios priced at the 5% level, 19 of the 20 leave at least three, and five leave four. PC1–PC5 jointly explain between 74.2% and 78.4% of adjusted-return variance across these models. By contrast, ARM-basis adjustment leaves only PC5 priced

Table 9: **Priced Principal Components after Factor Adjustment**

Model	T	Priced PCs 5%	Which PCs	PC1 share	PC1 α	$t(PC1)$	PC1–PC5 share
Benchmark traded factors							
CAPM	684	3	1, 2, 4	25.2	4.25	4.02	77.3
FF3	684	3	1, 3, 4	20.9	3.30	3.89	75.4
Carhart4	684	3	1, 2, 4	21.2	3.29	3.64	75.1
FF5	684	3	1, 2, 4	20.4	3.43	3.57	75.0
FF6	684	3	1, 2, 4	20.9	3.38	3.46	74.6
FF6+BAB+STR	684	4	1, 2, 3, 4	21.3	3.11	2.66	74.2
FF6+LIQ	684	3	1, 2, 4	20.9	3.64	3.91	74.6
q5	684	2	1, 4	21.1	3.64	4.75	75.3
BS6	684	3	1, 2, 5	20.5	2.40	2.47	74.7
Behavioral / mispricing							
SY4	588	4	1, 3, 4, 5	23.1	3.17	3.72	77.3
DHS3	558	3	2, 3, 4	25.6	0.13	0.12	78.1
Intermediary capital							
HKM2	587	4	1, 2, 3, 4	26.2	3.11	2.91	78.4
Statistical factors							
PCA	600	3	1, 2, 4	23.6	2.68	3.15	75.9
RP-PCA	600	3	1, 2, 4	23.3	2.34	2.71	76.4
Higher moments / skewness							
Coskewness	684	4	1, 2, 3, 4	24.9	2.20	2.04	77.2
Cokurtosis	684	3	1, 2, 3	25.0	2.50	2.34	77.2
FF6+PSS	600	4	1, 2, 3, 4	23.3	3.34	3.66	76.5
Downside / disappointment							
Downside2	684	3	2, 3, 4	24.8	1.30	1.23	77.2
GDA3	684	3	2, 3, 4	24.7	2.02	1.93	77.2
GDA5	684	3	1, 2, 3	24.9	2.35	2.32	77.2
ARM basis							
ARM basis	684	1	5	35.7	0.19	0.22	70.4

Notes: This table reports the priced principal components remaining after adjusting the 150 ARM implementation returns for each benchmark model and for the three-factor ARM basis. Each row re-estimates PCA after the indicated adjustment. For each implementation return, fitted factor exposures are subtracted while the regression intercept is retained; the adjusted-return mean therefore equals the factor-model alpha. Covariance PCA is applied to the centered adjusted returns. PC1 share and PC1–PC5 share are percentages of adjusted-return variance, and PC1 α is the annualized mean of the adjusted PC1 portfolio in percent. The inputs are monthly ARM returns beginning in January 1968 and ending in December 2024. Each row uses the benchmark model's maximum available overlap, reported as T. Priced PCs counts PC1–PC5 portfolios with two-sided six-lag Newey–West p-values below 0.05; Which PCs identifies them, and $t(PC1)$ is the corresponding PC1 statistic. PCA is re-estimated separately by row, so PC numbering is row-specific. PC signs are oriented so that sample means are positive.

among the first five. The table therefore uses the same residual diagnostic to show that standard benchmarks leave several common priced components, whereas the ARM basis removes most of that priced structure.

The location of the residual premium varies across benchmarks. PC1 is priced at the 5% level under 17 of the 20 models. Under DHS3, Downside2, and GDA3, however, PC1 is insignificant while PC2–PC4 are priced. A model can therefore attenuate the leading principal component while leaving priced variation in later PCs. This pattern reinforces the value of evaluating the ARM space PC by PC rather than through a single aggregate alpha.

The ARM-basis comparison also separates residual covariance from residual compensation. PC1 remains the leading residual principal component and explains 35.7% of residual variance, but its annualized alpha is only 0.19%, with a t-statistic of 0.22. The priced PC5 explains 4.9% of residual variance and has an annualized alpha of 0.77%, with a t-statistic of 2.63. The basis therefore absorbs most of the common priced structure while leaving a smaller residual component. Appendix Table C.12 reports the detailed residual decomposition.

The final diagnostic summarizes the best fit available from the benchmark set. Appendix Table C.13 lets an ex post oracle select, for every family–implementation cell, the model with the smallest absolute alpha. Because selection is portfolio-specific and may compare different samples, the exercise is a descriptive best-case benchmark. The all-model oracle reduces the median absolute alpha by 80.1% relative to the CAPM, to 0.73% per year. The 90th percentile remains 3.47%, and 12.0% of selected alphas have model-specific $|t| > 1.96$. Portfolio-specific model choice therefore reduces typical pricing errors but leaves a nontrivial upper tail.

Together, the diagnostics sharpen the joint rejections in Section 4.3. Existing benchmarks leave several priced residual PC portfolios, whereas the ARM basis leaves one smaller priced residual component; even portfolio-specific benchmark selection leaves sizeable alphas in the upper tail. Combined with the implementation-separated evidence in Section 4.2, the residual comparison indicates that the basis captures common priced structure that transfers across portfolio-construction rules. The basis consequently provides a compact diagnostic benchmark for the ARM return universe.

5 External Mapping and Economic-State Evidence

We conclude the empirical analysis with two exercises that take the ARM map beyond its baseline return panel and unconditional tests. First, we examine whether the basis organizes option-based Bear Beta portfolios formed from an independent signal. Second, we test whether the mapped dimensions vary predictably with aggregate risk aversion and uncertainty. The first exercise evaluates external classification, while the second provides economic evidence that the dimensions are not interchangeable representations of one generic bad-state premium.

5.1 Option-Based External Mapping: Bear Beta

Bear Beta measures a stock's exposure to the option-based aggregate bad-state factor of Lu and Murray (2019). Neither the Bear Beta sorting signal nor the resulting portfolio returns enter the ARM-basis construction, so these portfolios provide an external mapping exercise.[12]

We use decile-based long–short portfolios for three stock universes indexed by s: all eligible stocks, liquid stocks, and large stocks. The common sample contains 225 months

[12]We downloaded the Bear Beta portfolio returns from Zhongjin Lu's webpage. We thank him for making the data available.

from January 1997 through September 2015. For each portfolio, we estimate

$$r_{s,t}^{BEAR} = \alpha_s + \beta_s^{BROAD} F_t^{BROAD} + \beta_s^{CIQ} F_t^{CIQ} + \beta_s^{COSK} F_t^{COSK} + \varepsilon_{s,t}. \tag{10}$$

We report the raw premium and nested regressions using broad tail/crash, broad tail/crash plus lower-tail CIQ, and the full basis. This sequence shows the incremental mapping of the external portfolios into the three basis dimensions. For the full-basis specification, we test the three portfolio intercepts jointly using a six-lag HAC Wald test.

Table 10 reports the results. The raw Bear Beta spreads earn annualized mean returns from -13.61% to -10.85%. Broad tail/crash alone leaves the intercepts nearly unchanged and produces adjusted R^2s from -0.3% to 0.2%. Adding lower-tail CIQ reduces the absolute intercepts and raises adjusted R^2 to 5.4–8.6%. The full basis reduces the annualized intercepts further, to between -7.41% and -4.29%, and raises adjusted R^2 to 13.2–15.9%. These intercepts have absolute t-statistics from 1.08 to 1.54.

The full-basis loadings on residual coskewness range from -0.96 to -0.80, with t-statistics from -2.21 to -2.01; this is the only basis loading significant across all three stock universes. The negative loading reflects the portfolios' original return orientation. The joint test of the three full-basis intercepts yields a p-value of 0.336.

The evidence supports external classification of Bear Beta within the ARM map. Taken together, the progression in adjusted R^2, the attenuation of the intercepts, and the consistent loading on residual coskewness indicate that the independently constructed option-based portfolios contain variation represented by the basis. At the same time, the full-basis intercepts remain economically sizeable, so their joint non-rejection does not establish exact economic spanning. The exercise extends the map beyond the portfolios used to construct it and illustrates the basis's role as a classification tool. The next subsection studies how compensation for the mapped dimensions varies with aggregate risk aversion and uncertainty.

5.2 Predictive Economic-State Dependence

The second exercise examines an economic implication of the map: next-month premia on dimensions representing different adverse-state exposures may vary differently with lagged states that proxy for the price and quantity of risk. We therefore relate next-month premia to aggregate risk aversion and uncertainty. Distinct predictive signatures provide supporting economic validation of the empirical taxonomy.

We use the time-varying risk-aversion and economic-uncertainty measures of Bekaert et al. (2022). These series are useful because they separate two related adverse-state concepts. Risk aversion captures variation in investors' willingness to bear risk, whereas uncertainty captures

Table 10: **Option-Based External Mapping: Bear Beta Regressions**

	Raw premium	Broad	Broad + CIQ	Full basis
Panel A. All stocks				
Raw premium / alpha	-13.61***	-13.70***	-9.17*	-7.41
	(-2.80)	(-2.70)	(-1.95)	(-1.54)
Broad tail/crash		0.14	0.22	0.12
		(0.20)	(0.32)	(0.19)
Lower-tail CIQ			-0.67	-0.40
			(-1.32)	(-0.94)
Residual coskewness				-0.96**
				(-2.08)
Adj. R^2 (%)		-0.3	5.4	13.2
T	225	225	225	225
Panel B. Liquid stocks				
Raw premium / alpha	-12.98**	-13.19**	-7.78*	-6.04
	(-2.49)	(-2.47)	(-1.77)	(-1.39)
Broad tail/crash		0.32	0.42	0.32
		(0.49)	(0.64)	(0.56)
Lower-tail CIQ			-0.79	-0.53
			(-1.56)	(-1.23)
Residual coskewness				-0.95**
				(-2.21)
Adj. R^2 (%)		0.2	8.5	15.9
T	225	225	225	225
Panel C. Large stocks				
Raw premium / alpha	-10.85**	-11.02**	-5.76	-4.29
	(-2.20)	(-2.17)	(-1.46)	(-1.08)
Broad tail/crash		0.26	0.36	0.27
		(0.42)	(0.58)	(0.50)
Lower-tail CIQ			-0.77	-0.55
			(-1.53)	(-1.27)
Residual coskewness				-0.80**
				(-2.01)
Adj. R^2 (%)		0.0	8.6	14.3
T	225	225	225	225
Joint test that the three full-basis Bear Beta portfolio alphas are zero: $p = 0.336$.				

Notes: This table reports nested time-series regressions of three Bear Beta long–short equity portfolios on the ARM basis. Panel A uses all eligible stocks, Panel B excludes the most illiquid stocks, and Panel C uses stocks above the NYSE median market capitalization. Columns add the broad tail/crash, lower-tail CIQ, and residual-coskewness basis factors sequentially. The Bear Beta portfolios are from Lu and Murray (2019); the underlying stock-level measure is defined relative to the authors' option-based aggregate bad-state factor. Returns retain the authors' Port10-minus-Port1 orientation. Raw premium / alpha is the unadjusted annualized mean in the first column and the annualized regression intercept in the remaining columns. Factor-loading coefficients are reported in the corresponding factor rows, and adjusted R^2 is reported in percent. The monthly sample covers January 1997–September 2015 and contains $T = 225$ observations in each regression. Newey–West t-statistics using six lags are reported in parentheses below estimates. The joint p-value is from a six-lag HAC Wald test that the three full-basis portfolio intercepts are jointly zero. Blank cells denote coefficients not included in a specification. We thank Zhongjin Lu for making the portfolio data available on his webpage. *, **, and *** denote conventional two-sided significance at the 10%, 5%, and 1% levels, respectively.

variation in the perceived quantity of risk. The two monthly series are highly correlated in our sample, with a correlation of 0.77, so entering them directly makes the partial slopes difficult to interpret. We therefore rotate them into a common adverse-state component and a differential risk-aversion-minus-uncertainty component.

Let RA_t denote the risk-aversion measure and UNC_t denote the uncertainty measure observed at the end of month t. We first standardize both variables:

$$\widetilde{RA}_t = \frac{RA_t - \bar{RA}}{\sigma(RA)}, \qquad \widetilde{UNC}_t = \frac{UNC_t - \overline{UNC}}{\sigma(UNC)}. \tag{11}$$

We then rotate and standardize the two series directly:

$$\begin{aligned} \text{COMMON}_t &= \frac{\widetilde{RA}_t + \widetilde{UNC}_t}{\sigma\left(\widetilde{RA} + \widetilde{UNC}\right)}, \\ \text{RA}-\text{UNC}_t &= \frac{\widetilde{RA}_t - \widetilde{UNC}_t}{\sigma\left(\widetilde{RA} - \widetilde{UNC}\right)}. \end{aligned} \tag{12}$$

Because the standardized inputs have equal variance, their sum and difference are orthogonal. The common-state variable, COMMON_t, captures periods in which risk aversion and uncertainty rise together. It represents broadly adverse aggregate conditions in which investors are both less willing to bear risk and perceive more risk in the economy. The $\text{RA}-\text{UNC}_t$ variable captures periods in which risk aversion is high relative to measured uncertainty. This distinction is useful because compensation for bad-state exposure may vary not only with the amount of risk investors perceive, but also with the price they require for bearing that risk. Unit-variance normalization implies that each slope is the annualized change in the next-month ARM premium associated with a one-standard-deviation increase in the corresponding state variable.

For the matched PC portfolios representing the three recurring priced dimensions—PC1, PC2, and PC4 in the full-sample map—we estimate

$$PC_{j,t+1} = \alpha_j + \beta_j^{\text{COMMON}}\text{COMMON}_t + \beta_j^{\text{RA}-\text{UNC}}\text{RA}-\text{UNC}_t + \varepsilon_{j,t+1}. \tag{13}$$

The state variables are observed in month t and matched to PC returns in month $t+1$. The sample contains 462 monthly observations from July 1986 through December 2024. Inference uses six-lag Newey–West standard errors. Because the state-dependence exercise tests six pre-specified slopes, we report Benjamini–Hochberg false-discovery-rate adjustments (Benjamini and Hochberg, 1995) and Holm family-wise-error adjustments.

Table 11 reports distinct predictive signatures across the priced ARM dimensions. Next-month broad tail/crash returns are higher following states in which risk aversion is elevated relative to uncertainty: the RA–UNC coefficient for PC1 is 4.07, with an unadjusted t-statistic of 2.55. Next-month lower-tail CIQ returns rise following states in which risk aversion and uncertainty are jointly elevated: the COMMON coefficient for PC2 is 4.17, with an unadjusted t-statistic of 2.66. Next-month residual coskewness returns are unrelated to either state in this specification.

Both leading slopes survive false-discovery-rate adjustment. The PC2 COMMON coefficient also survives Holm adjustment at the 5% level ($p = 0.048$), while the PC1 RA–UNC

Table 11: **Predictive State Dependence of the Priced ARM Dimensions**

	(1) PC1 Broad tail/crash	(2) PC2 Lower-tail CIQ	(3) PC4 Residual coskewness
α / mean	3.83	4.47	2.32
	(2.60)	(3.63)	(2.45)
COMMON	-0.34	4.17	0.33
	(-0.20)	(2.66)	(0.27)
BH q / Holm p	0.907 / 1.000	0.034 / 0.048	0.907 / 1.000
RA−UNC	4.07	0.18	-0.93
	(2.55)	(0.12)	(-0.89)
BH q / Holm p	0.034 / 0.056	0.907 / 1.000	0.745 / 1.000
R^2	0.021	0.024	0.002
Observations	462	462	462

Notes: This table reports predictive state-dependence regressions for PC1, PC2, and PC4, matched respectively to the broad tail/crash, lower-tail CIQ, and residual-coskewness dimensions. Columns correspond to PC portfolios, and rows report the intercept, two state coefficients, adjusted p-values, R^2, and observations. For each PC portfolio, the dependent variable is its month-$t+1$ return. COMMON is the standardized average of the standardized risk-aversion and uncertainty series of Bekaert et al. (2022); RA−UNC is their standardized difference. Both state variables are observed in month t and have unit variance. Coefficients and the intercept are annualized percentages. The common monthly sample covers July 1986–December 2024 and contains 462 observations. Newey–West t-statistics using six lags are reported in parentheses below the intercept and state-coefficient estimates. BH q-values and Holm p-values adjust jointly across the six state-coefficient tests; the intercept tests are not included in these adjustments. Because the state variables are centered, the reported intercept equals the sample mean of the corresponding PC portfolio.

coefficient is marginal under Holm adjustment ($p = 0.056$). These predictive signatures support the interpretation that broad tail/crash and lower-tail CIQ are not interchangeable rotations of one generic bad-state premium. Next-month broad tail/crash returns are highest following states in which risk aversion is high relative to measured uncertainty, whereas next-month lower-tail CIQ returns are highest following states in which risk aversion and uncertainty are jointly elevated. Residual coskewness remains priced unconditionally but appears less connected to these aggregate risk-aversion and uncertainty states.

The robustness exercises reinforce this interpretation. Appendix Table D.1 reproduces the PC1–PC2 contrast using PCA weights and PC matching estimated through 1995. Appendix Table D.2 shows the same predictive contrast when the state variables are constructed from an alternative RA/UC series: the largest PC1 slope remains on RA–UNC, whereas the largest PC2 slope remains on COMMON. Appendix Table D.3 shows that the pattern also survives additive benchmark-factor controls with fixed factor betas. Together, these checks reinforce the predictive-state contrast.

6 Conclusion

The paper shows that the priced ARM space is multidimensional but more compact than the set of proposed measures. The five ARM families organize around three recurring, economically interpretable priced dimensions: broad tail/crash, lower-tail CIQ, and residual coskewness. Multidimensional compensation recurs over time, but its composition is not fixed, consistent with different adverse-state exposures becoming more salient in different periods. The dimensions' common return content also transfers across portfolio constructions. Later low-variance PCs and implementation-specific residual alphas reveal finer, less stable structure that future research may interpret or extend.

The map also distinguishes empirical measures from priced dimensions. The portfolio returns associated with a new measure may map entirely to an existing dimension, contain statistically distinct but unpriced variation, or combine exposure to an existing dimension with incremental compensated variation.

The three anchor portfolios also form a simple PCA-free benchmark that performs well across portfolio implementations, including constructions not used to build it. The basis offers a practical test: locate a candidate ARM within the map and ask whether its residual return variation is priced. Factor-model comparisons should likewise report dimension-specific alphas, joint tests, and residual dimensionality. The Bear Beta exercise illustrates this external classification role.

Finally, the distinct predictive signatures provide supporting economic validation of the map. Next-month broad tail/crash returns are highest following states in which risk aversion is elevated relative to uncertainty, next-month lower-tail CIQ returns rise following states in which both are elevated, and residual coskewness is comparatively insensitive to these states. These patterns are consistent with several forms of adverse-state compensation rather than one scalar downside channel. The broader lesson is to locate each new ARM in the priced return space before treating it as a new risk.

References

Ahmed, S., Bu, Z., and Tsvetanov, D. (2019). Best of the best: A comparison of factor models. *Journal of Financial and Quantitative Analysis*, 54(4):1713–1758.

Ang, A., Chen, J., and Xing, Y. (2006). Downside risk. *Review of Financial Studies*, 19(4):1191–1239.

Asness, C. S. and Frazzini, A. (2013). The devil in HML's details. *Journal of Portfolio Management*, 39(4):49–68.

Bali, T. G., Cakici, N., and Whitelaw, R. F. (2014). Hybrid Tail Risk and Expected Stock Returns: When Does the Tail Wag the Dog? *Review of Asset Pricing Studies*, 4(2):206–246.

Barillas, F. and Shanken, J. (2018). Comparing asset pricing models. *Journal of Finance*, 73(2):715–754.

Baruník, J. and Nevrla, M. (2026). Common idiosyncratic quantile factors and asset prices. *Journal of Financial and Quantitative Analysis*. Accepted manuscript, published online June 10, 2026.

Bekaert, G., Engstrom, E. C., and Xu, N. R. (2022). The time variation in risk appetite and uncertainty. *Management Science*, 68(6):3975–4004.

Bekaert, G., Hoerova, M., and Lo Duca, M. (2013). Risk, uncertainty and monetary policy. *Journal of Monetary Economics*, 60(7):771–788.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300.

Bollerslev, T., Patton, A. J., and Quaedvlieg, R. (2022). Realized semibetas: Disentangling "good" and "bad" downside risks. *Journal of Financial Economics*, 144(1):227–246.

Brunnermeier, M. K. and Sannikov, Y. (2014). A macroeconomic model with a financial sector. *American Economic Review*, 104(2):379–421.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52(1):57–82.

Chabi-Yo, F., Huggenberger, M., and Weigert, F. (2022). Multivariate crash risk. *Journal of Financial Economics*, 145(1):129–153.

Chen, L., Dolado, J. J., and Gonzalo, J. (2021). Quantile factor models. *Econometrica*, 89(2):875–910.

Cochrane, J. H. (2005). *Asset Pricing*. Princeton University Press.

Daniel, K., Hirshleifer, D., and Sun, L. (2020). Short- and long-horizon behavioral factors. *Review of Financial Studies*, 33(4):1673–1736.

Dittmar, R. F. (2002). Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *Journal of Finance*, 57(1):369–403.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22.

Fama, E. F. and French, K. R. (2018). Choosing factors. *Journal of Financial Economics*, 128(2):234–252.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636.

Farago, A. and Tédongap, R. (2018). Downside risks and the cross-section of asset returns. *Journal of Financial Economics*, 129(1):69–86.

Feng, G., Giglio, S., and Xiu, D. (2020). Taming the factor zoo: A test of new factors. *Journal of Finance*, 75(3):1327–1370.

Frazzini, A. and Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, 111(1):1–25.

Gibbons, M. R., Ross, S. A., and Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica*, 57(5):1121–1152.

Harvey, C. R., Liu, Y., and Zhu, H. (2016). ... and the cross-section of expected returns. *Review of Financial Studies*, 29(1):5–68.

Harvey, C. R. and Siddique, A. (2000). Conditional skewness in asset pricing tests. *Journal of Finance*, 55(3):1263–1295.

He, Z., Kelly, B., and Manela, A. (2017). Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics*, 126(1):1–35.

He, Z. and Krishnamurthy, A. (2013). Intermediary asset pricing. *American Economic Review*, 103(2):732–770.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70.

Hong, Y., Tu, J., and Zhou, G. (2007). Asymmetries in stock returns: Statistical tests and economic evaluation. *Review of Financial Studies*, 20(5):1547–1581.

Hou, K., Mo, H., Xue, C., and Zhang, L. (2019). Which factors? *Review of Finance*, 23(1):1–35.

Hou, K., Mo, H., Xue, C., and Zhang, L. (2021). An augmented q-factor model with expected growth. *Review of Finance*, 25(1):1–41.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *Journal of Finance*, 45(3):881–898.

Jiang, L., Wu, K., and Zhou, G. (2018). Asymmetry in stock comovements: An entropy approach. *Journal of Financial and Quantitative Analysis*, 53(4):1479–1507.

Kelly, B. and Jiang, H. (2014). Tail risk and asset prices. *Review of Financial Studies*, 27(10):2841–2871.

Kozak, S., Nagel, S., and Santosh, S. (2020). Shrinking the cross-section. *Journal of Financial Economics*, 135(2):271–292.

Langlois, H. (2020). Measuring skewness premia. *Journal of Financial Economics*, 135(2):399–424.

Lettau, M. and Pelger, M. (2020). Factors that fit the time series and cross-section of stock returns. *Review of Financial Studies*, 33(5):2274–2325.

Lu, Z. and Murray, S. (2019). Bear beta. *Journal of Financial Economics*, 131(3):736–760.

MacKinlay, A. C. and Richardson, M. P. (1991). Using generalized method of moments to test mean-variance efficiency. *Journal of Finance*, 46(2):511–527.

Massacci, D., Sarno, L., and Trapani, L. (2025). Factor models of asset returns and bear market risk. *Management Science*, 72(8):6806–6825.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.

Pastor, L. and Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, 111(3):642–685.

Schneider, P., Wagner, C., and Zechner, J. (2020). Low-risk anomalies? *Journal of Finance*, 75(5):2673–2718.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442.

Stambaugh, R. F. and Yuan, Y. (2017). Mispricing factors. *Review of Financial Studies*, 30(4):1270–1315.

Appendices

A ARM Measure Definitions and Implementation

A.1 Notation and Timing

This appendix documents the estimators used to construct the five stock-level sorting variables. Let $r_{i,d}$ and $r_{m,d}$ denote the daily returns of stock i and the market, and let $r^e_{i,s}$ and $r^e_{m,s}$ denote monthly excess returns. Consistent with the main text, month t denotes the portfolio-return month and portfolio formation occurs at the end of month $t-1$. In the estimator definitions below, τ denotes the endpoint of a rolling estimation window. Thus, an estimate available for formation at $t-1$ uses information no later than $\tau = t-1$ and cannot contain the return it is intended to predict. The portfolio-level price screens and breakpoint choices described in Section 2.3 are applied after the stock-level measures have been estimated.

A.2 Hybrid Tail Covariance Risk

Following Bali et al. (2014), let $q_{i,\tau}(0.10)$ and $q_{m,\tau}(0.10)$ denote the empirical 10th percentiles of the stock and market daily return distributions in the six-month window $\mathcal{W}_\tau$. We estimate

$$HTCR_{i,\tau} = \sum_{d \in \mathcal{W}_\tau} \mathbb{1}\{r_{i,d} < q_{i,\tau}(0.10)\} \left[r_{i,d} - q_{i,\tau}(0.10)\right] \left[r_{m,d} - q_{m,\tau}(0.10)\right]. \tag{14}$$

At least 100 daily observations are required. The conditioning set contains the stock's own lower-tail days; it does not require the market to be below its threshold. This hybrid conditioning is the defining difference between HTCR and a conventional systematic tail beta.

A.3 Multivariate Crash Risk

MCRASH follows Chabi-Yo et al. (2022). For each stock and each systematic factor, we fit a GARCH(1,1) marginal model with skewed Student-t innovations over a rolling 12-month window and obtain probability integral transforms $U_{i,d}$ and $U_{k,d}$. If the primary fit fails, the implementation uses a Gaussian integrated-GARCH/EWMA fallback. The factor set is

$$\mathcal{K} = \{\mathrm{MKT}, \mathrm{SMB}, \mathrm{HML}, \mathrm{RMW}, \mathrm{CMA}, \mathrm{UMD}, \mathrm{BAB}\}. \tag{15}$$

With $p = 0.05$, define a systematic crash day as one on which at least one factor enters its conditional lower tail,

$$C_d = \mathbb{1}\left\{\min_{k \in \mathcal{K}} U_{k,d} \leq p\right\}. \tag{16}$$

The stock-level measure is the corresponding conditional crash frequency,

$$MCRASH_{i,\tau} = \frac{\sum_{d \in \mathcal{W}_\tau} \mathbb{1}\{U_{i,d} \leq p\}\, C_d}{\sum_{d \in \mathcal{W}_\tau} C_d}. \tag{17}$$

We require at least 200 nonzero daily stock-return observations. The union in C_d is important: MCRASH conditions on at least one systematic factor crashing, not on all seven factors crashing simultaneously.

A.4 Tail-Risk Beta

We construct the aggregate tail-risk factor following Kelly and Jiang (2014). Within each month, daily stock returns are residualized against the Fama–French three factors. Let $e_{k,s}$ index the pooled daily residual observations in month s, let $u_s < 0$ be their empirical 5th percentile, and let K_s be the number of residuals below this threshold. Aggregate tail severity is

$$TR_s = \frac{1}{K_s} \sum_{e_{k,s} < u_s} \log\left(\frac{e_{k,s}}{u_s}\right). \tag{18}$$

Larger values indicate more severe exceedances relative to the contemporaneous tail threshold. For each stock, TR β is estimated from

$$r_{i,s+1} = a_{i,\tau} + \beta_{i,\tau}^{TR} TR_s + \varepsilon_{i,s+1}, \tag{19}$$

using a rolling 120-month window and requiring at least 36 monthly observations. Equation (19) reflects the implementation's timing: the tail factor measured in month s is matched to the stock return in month $s + 1$. The resulting beta is itself lagged before portfolio formation.

A.5 Lower-Tail Common Idiosyncratic Quantile Beta

The CIQ construction follows Baruník and Nevrla (2026). In each 60-month window ending in month τ, we first estimate the Fama–French three-factor model (Fama and French, 1993)

$$r^e_{i,s} = \alpha_{i,\tau} + \beta_{i,\tau}^{MKT}\mathrm{MKT}_s + \beta_{i,\tau}^{SMB}\mathrm{SMB}_s + \beta_{i,\tau}^{HML}\mathrm{HML}_s + e_{i,s,\tau}, \qquad s = \tau - 59, \ldots, \tau, \tag{20}$$

and standardize each stock's residuals by their within-window standard deviation. Stocks used to extract the factor must have a complete 60-month return history and a lagged price above \$1. At quantile q, the standardized residual panel is represented by the one-factor quantile model of Chen et al. (2021),

$$\widetilde{e}_{i,s,\tau} = \gamma_{i,\tau}(q) CIQ_{s,\tau}(q) + u_{i,s,\tau}(q), \qquad \mathbb{P}[u_{i,s,\tau}(q) < 0 \mid CIQ_{s,\tau}(q)] = q. \tag{21}$$

The factor is oriented to be positively correlated with the corresponding realized cross-sectional residual quantile. We use innovations

$$\Delta CIQ_{s,\tau}(q) = CIQ_{s,\tau}(q) - CIQ_{s-1,\tau}(q) \tag{22}$$

and estimate stock-level time-series exposures from

$$r^e_{i,s} = a_{i,\tau} + \beta^{CIQ}_{i,\tau}(q) \Delta CIQ_{s,\tau}(q) + v_{i,s,\tau}. \tag{23}$$

The active measure is $\beta^{CIQ}_{i,\tau}(0.20)$, rather than an average across several downside quantiles. Beta estimation uses the same 60-month window and requires at least 48 stock-month observations. The quantile-model loading $\gamma_{i,\tau}(0.20)$ in Equation (21) describes a stock's contribution to the conditional cross-sectional quantile; the sorting variable in this paper is instead the time-series exposure $\beta^{CIQ}_{i,\tau}(0.20)$ to innovations in that factor.

A.6 Residual Coskewness

Following Harvey and Siddique (2000), we first estimate a market model within each rolling window,

$$r^e_{i,s} = a_{i,\tau} + b_{i,\tau} r^e_{m,s} + \epsilon_{i,s}. \tag{24}$$

COSK is then

$$\text{COSK}_{i,\tau} = \frac{\frac{1}{T} \sum_{s \in \mathcal{W}_\tau} \epsilon_{i,s} \left(r^e_{m,s} - \overline{r}^e_{m,\tau} \right)^2}{\sigma_{\epsilon,i,\tau} \, \text{Var}_\tau(r^e_m)}. \tag{25}$$

The implementation uses monthly observations over a rolling 60-month window and requires at least 48 observations. A negative value indicates that the stock's idiosyncratic return tends to be low when market movements are large in absolute value. Because such exposure is the adverse side of the measure, the COSK portfolio return is reported as low-minus-high, equivalently minus the standard high-minus-low return.

A.7 Additional Downside and Higher-Moment Measures Considered

Table A.1 makes the boundary of the baseline ARM universe transparent by applying the same 30-specification portfolio-construction grid to related measures available in the common data environment. The candidates cover downside dependence, downside beta and semibeta, and alternative coskewness and cokurtosis constructions. For each measure, the table reports both its unadjusted composite premium and its alpha relative to the six-factor model of Fama and French (2018). The table screens the candidates for a robust composite premium, stability across implementations, and nonredundancy with this conventional factor benchmark. The subsequent raw-coskewness comparison additionally examines overlap with the retained ARM basis.

Several candidates lack reliable unadjusted pricedness. The monthly and daily downside betas, daily COSK, and monthly and daily cokurtosis have statistically insignificant composite premia and few significant individual implementations. Downside correlation is marginal in unadjusted returns, but its FF6 alpha is economically and statistically indistinguishable from zero. Downside asymmetry and the positive-stock/negative-market semibeta produce significant FF6 alphas, yet their unadjusted composites are statistically weak and only 8 and 3 of their 30 respective implementations are significant. The negative–negative semibeta likewise lacks robust composite or adjusted evidence. Across the common construction grid, these measures therefore provide useful boundary cases but not stable additional baseline dimensions.

The alternative coskewness measures clarify why the baseline uses monthly COSK. Raw daily coskewness has statistically significant unadjusted premia in 29 of 30 implementations, but its FF6 alpha falls to -0.31% per year with a t-statistic of -0.25, indicating that its premium is absorbed by conventional factor exposures. Raw monthly coskewness remains significant after FF6 adjustment. Its composite, however, has an absolute correlation of 0.72 with the retained residual coskewness dimension, and a separate regression on the three-factor ARM basis reduces its annualized alpha to -0.52% with a t-statistic of -0.76. The daily COSK construction has no reliable composite premium. Taken together, the evidence favors monthly COSK: it removes linear market exposure before portfolio formation, preserves a robust premium, and captures the priced monthly coskewness content without assigning raw coskewness a separate ARM dimension. Table A.1 thereby documents the empirical boundary of the retained ARM universe under a common set of criteria.

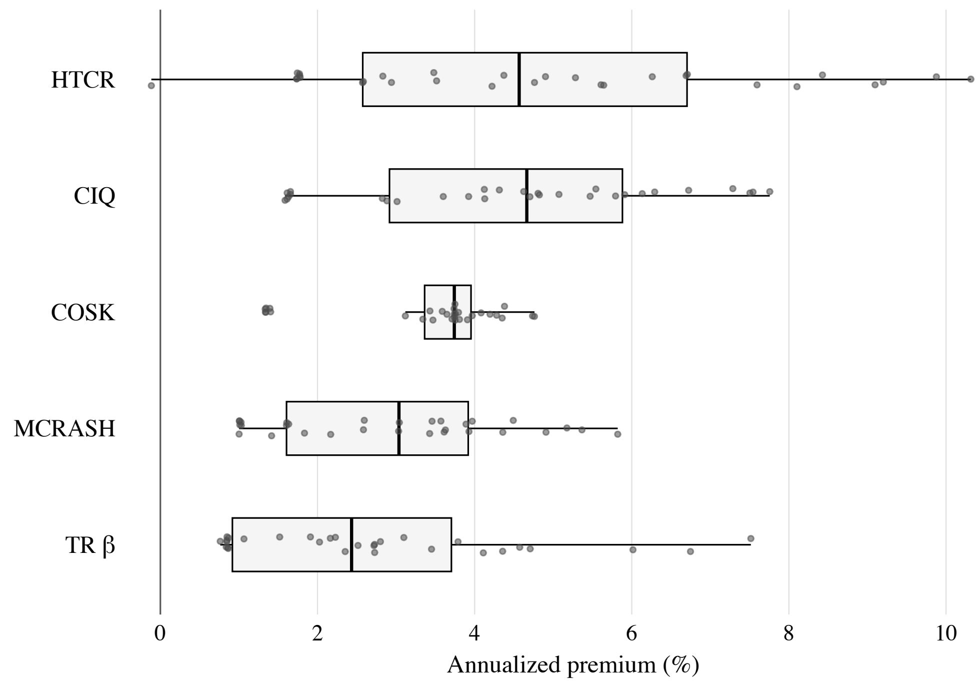
Table A.1: **Additional Downside and Higher-Moment Measures Considered**

		Unadjusted evidence			FF6-adjusted evidence	
Measure	T	Composite	Median [IQR]	Sig. specs	α	Sig. alphas
Downside correlation	673	1.74 (1.89)	1.76 [1.23, 2.04]	12/30	-0.06 (-0.07)	0/30
Downside asymmetry	684	1.70 (1.69)	1.33 [0.64, 2.90]	8/30	2.54 (2.74)	18/30
Downside beta (monthly)	684	1.78 (0.82)	1.71 [1.01, 2.38]	0/30	2.13 (1.24)	0/30
Downside beta (daily)	684	-1.09 (-0.45)	-0.21 [-3.29, 0.77]	1/30	-0.26 (-0.16)	4/30
Negative–negative semibeta	673	-4.54 (-1.45)	-4.02 [-7.18, -1.84]	7/30	-2.56 (-1.48)	10/30
Positive-stock/negative-market semibeta	673	-2.45 (-1.20)	-2.27 [-3.48, -0.74]	3/30	-2.63 (-2.17)	17/30
Raw coskewness (monthly)	684	-3.60 (-3.52)	-3.94 [-4.34, -2.62]	26/30	-3.71 (-3.47)	30/30
Raw coskewness (daily)	684	-3.30 (-3.04)	-3.43 [-4.05, -2.82]	29/30	-0.31 (-0.25)	0/30
COSK (daily)	684	-1.65 (-1.60)	-1.51 [-2.37, -1.08]	0/30	0.22 (0.20)	0/30
Cokurtosis (monthly)	684	0.85 (0.70)	0.36 [-0.19, 1.58]	3/30	0.44 (0.39)	0/30
Cokurtosis (daily)	684	-0.24 (-0.16)	0.16 [-0.21, 0.44]	0/30	-0.11 (-0.07)	3/30

Notes: This table reports unadjusted composite premia and FF6-adjusted alphas for additional downside and higher-moment measures considered when defining the ARM universe. The left block reports the composite, its implementation-level distribution, and the number of significant implementations; the right block reports the FF6 alpha and the corresponding count. For each measure, the composite equally weights 30 monthly portfolio implementations combining top–bottom 10% and 20% sorts and characteristic-managed portfolios with equal or value weighting where applicable, minimum-price screens of zero, $1, and $5, and breakpoints based on all exchanges or NYSE stocks. Discrete-sort and characteristic-managed returns are high-minus-low, and signs are not reoriented according to realized premia. Composite premia, implementation-level means, and FF6 alphas are annualized percentages. Median [IQR] summarizes the 30 annualized implementation means. FF6 contains the market, size, value, profitability, investment, and momentum factors. Portfolio formation, return timing, the stock universe, and delisting-adjusted excess returns follow the baseline analysis. T is the number of monthly composite-return observations. Daily measures require at least 200 daily observations per stock-month, and monthly measures require at least 48 observations in the 60-month estimation window. Composite and alpha t-statistics use Newey–West standard errors with six lags and are reported in parentheses. Sig. specs and Sig. alphas count conventional, unadjusted two-sided 5% tests across the 30 implementations. Downside correlation follows Hong et al. (2007); Jiang et al. (2018); downside asymmetry follows Jiang et al. (2018); downside beta follows Ang et al. (2006); raw and residual coskewness and cokurtosis follow Harvey and Siddique (2000); Dittmar (2002); and the semibeta rows are the negative–negative and positive-stock/negative-market components of Bollerslev et al. (2022).

B ARM Premia Across Portfolio Constructions

Figure B.1: **Distribution of ARM Premia**

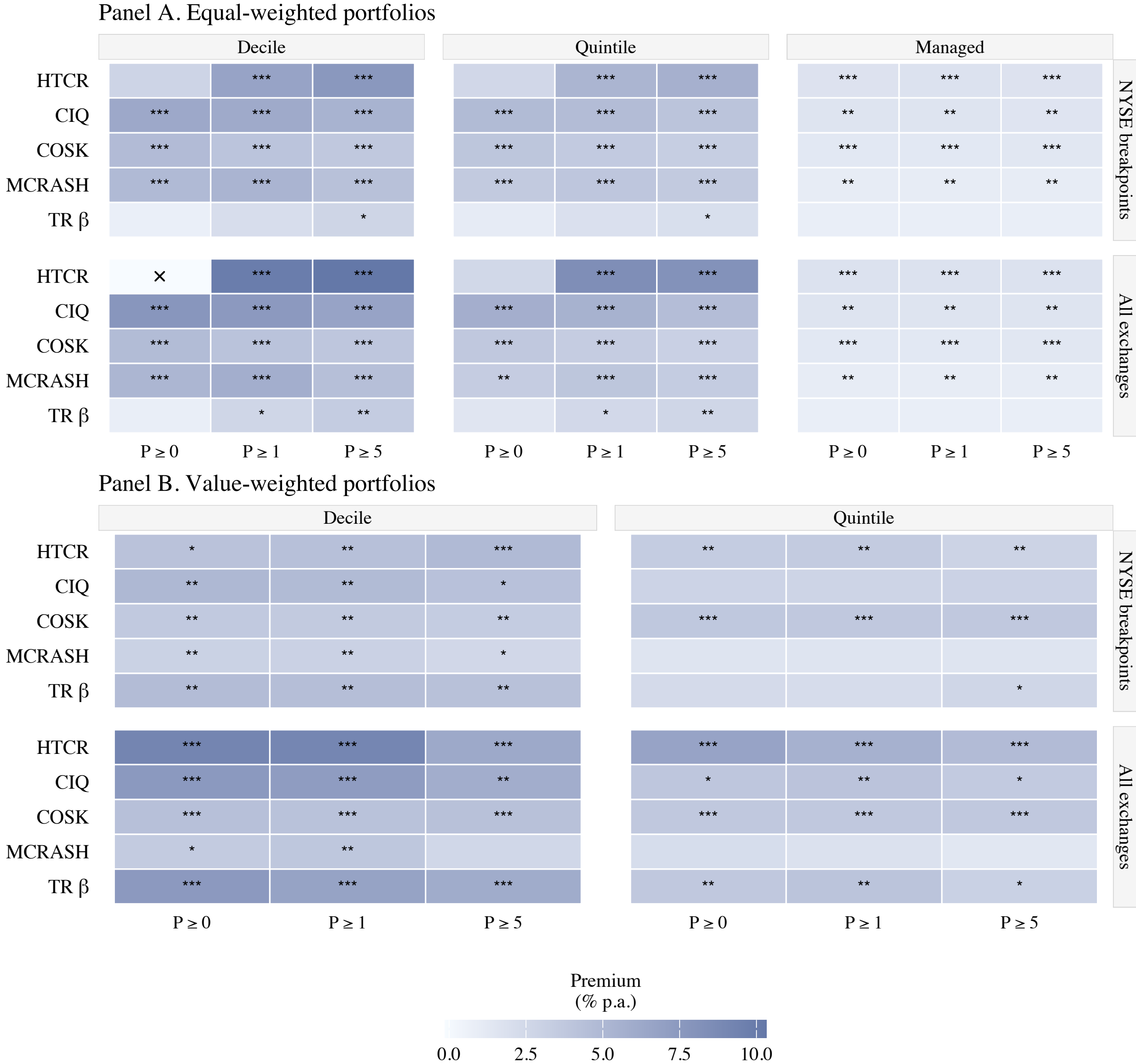


Notes: The figure reports the distribution of annualized ARM premia across all portfolio-construction specifications. Each point is one implementation. Boxes report the interquartile range and median; whiskers extend to the most extreme observation within 1.5 times the interquartile range. The vertical line marks zero.

B.1 ARM Premia Within Implementation Subsamples

Table B.1 reports the raw family-composite premia for the implementation subsamples used in Table 6 and Appendix Table C.5. It shows whether variation in a subgroup-specific basis or matched PC estimate reflects the underlying ARM premium or a change in how that premium enters the common ARM space. Figure B.2 provides the more granular specification-level counterpart.

Figure B.2: **ARM Premia across Specifications**



Notes: This figure reports annualized premia on asymmetric-risk-managed portfolios across portfolio-construction choices. Panel A reports equal-weighted portfolios and Panel B reports value-weighted portfolios. Columns vary by price screen, sorting rule, and breakpoint universe. Stars denote statistical significance based on Newey–West t-statistics, with *, **, and *** corresponding to the 10%, 5%, and 1% levels, respectively. Crosses indicate negative premia.

Table B.1: **ARM Premia by Implementation Subsample**

Implementation restriction	N	HTCR	CIQ	COSK	MCRASH	TR β
Panel A. Weighting scheme						
Equal, discrete sorts	60	5.82***	5.79***	3.88***	4.35***	2.04
		(3.16)	(3.92)	(3.57)	(3.24)	(1.52)
Value, discrete sorts	60	5.42***	4.61**	3.87***	2.41*	4.35**
		(2.87)	(2.18)	(2.81)	(1.76)	(2.44)
Equal, including managed	90	4.46***	4.40***	3.04***	3.23***	1.65
		(3.35)	(3.73)	(3.61)	(3.23)	(1.58)
Panel B. Sorting rule						
Top–bottom 10%	60	6.27***	6.23***	4.11***	4.07***	3.89**
		(3.26)	(3.29)	(3.43)	(2.98)	(2.36)
Top–bottom 20%	60	4.97***	4.17***	3.65***	2.69**	2.50*
		(3.21)	(2.68)	(3.56)	(2.49)	(1.92)
Characteristic-managed	30	1.76***	1.63**	1.36***	1.01**	0.85
		(3.57)	(2.40)	(3.17)	(2.34)	(1.59)
Panel C. Minimum-price screen						
No price screen	50	3.50**	4.74***	3.55***	2.95***	2.43*
		(2.07)	(3.14)	(3.71)	(2.69)	(1.90)
Price $\geq$ \$1	50	5.64***	4.62***	3.34***	3.15***	2.87**
		(3.74)	(3.08)	(3.49)	(2.97)	(2.27)
Price $\geq$ \$5	50	5.40***	4.10***	3.22***	2.62***	2.89**
		(4.19)	(2.75)	(3.35)	(2.70)	(2.32)
Panel D. Breakpoint universe						
All exchanges	75	5.74***	5.02***	3.43***	2.99***	3.18**
		(3.54)	(3.24)	(3.57)	(2.78)	(2.35)
NYSE	75	3.95***	3.95***	3.32***	2.82***	2.28*
		(2.97)	(2.73)	(3.48)	(2.83)	(1.95)

Notes: This table reports raw ARM-family premia within implementation subsamples. Panels vary the weighting scheme, sorting rule, minimum-price screen, and breakpoint universe. Each entry is the annualized mean, in percent, of the equal-weighted monthly return across all implementations in the indicated row and ARM family. The monthly sample covers January 1968–December 2024. N is the total number of ARM return series entering the row, equal to five families times the number of implementations per family. Newey–West t-statistics using six lags are reported in parentheses below the estimates. The first two weighting rows exclude characteristic-managed portfolios because they have no value-weighted counterpart; the third weighting row includes them in the equal-weighted set. *, **, and *** denote conventional two-sided significance at the 10%, 5%, and 1% levels, respectively.

C Further Robustness Checks

C.1 Later PCs and Alternative PCA Specifications

Table C.1 reports PCs 6–10 from the baseline covariance PCA and completes the variance summary. PC8–PC10 have conventionally significant in-sample means, so low variance does not by itself imply an irrelevant premium. The selection-complete analysis below therefore gives every principal component from PC1 through PC10 the same training-sample pricedness and materiality screen. This common treatment separates the recurring, economically anchored core from finer and less stable residual structure without relying on an eigenvalue cutoff.

Table C.1: **Later PCA Map of Asymmetric-Risk-Managed Portfolio Returns**

	Variance		PC portfolio			ARM-family loading shares				
PC	Share	Cum.	Mean	t-stat.	Sharpe	HTCR	CIQ	COSK	MCRASH	TR β
PC6	4.6	82.0	0.17	0.39	0.05	31.6	17.4	9.4	**36.6**	5.0
						(-18.8)	(+6.5)	(-9.4)	(+36.6)	(+4.2)
PC7	2.7	84.7	0.06	0.21	0.03	8.7	20.6	24.5	11.7	**34.5**
						(+5.3)	(-3.0)	(-0.6)	(-5.9)	(-6.1)
PC8	2.2	86.9	0.90	3.06	0.44	21.4	9.9	26.5	**31.9**	10.3
						(+0.6)	(+0.9)	(+5.8)	(+8.8)	(-4.1)
PC9	2.1	88.9	0.61	2.20	0.30	20.8	27.4	12.0	8.2	**31.6**
						(-4.9)	(+8.6)	(+3.2)	(-1.1)	(-8.5)
PC10	1.7	90.6	0.54	2.25	0.29	10.0	27.9	**29.5**	18.3	14.3
						(-0.4)	(+9.8)	(-9.2)	(+6.7)	(-4.0)

Notes: This table reports the variance, return, and ARM-family loading composition of principal-component portfolios 6–10. Covariance PCA is applied to the centered matrix of 150 monthly ARM portfolio returns, formed from five ARM families and 30 implementations per family. Eigenvectors are applied to raw returns, normalized to unit gross exposure, and signed so that reported means are positive. Share and Cum. are the individual and cumulative percentages of total variance. Mean is the annualized PC-portfolio return in percent, and Sharpe is annualized. Family columns report gross loading shares, with signed net shares in parentheses; gross shares sum to 100 up to rounding. The monthly sample covers January 1968–December 2024. Mean-return t-statistics use Newey–West standard errors with six lags. The largest gross family share in each row is bold. PCA loadings and return statistics are estimated on the same sample.

Table C.2 reports two alternatives to the baseline covariance PCA.

Panel A standardizes each of the 150 ARM portfolio returns before estimating the principal components. PC1 explains 22.7% of total standardized variance, and PC1–PC5 jointly explain 76.1%. Positive and statistically significant average returns appear in several PC portfolios rather than only the first, although the ordering and economic interpretation of individual PCs differ from the baseline because standardization changes the PCA rotation.

Panel B first collapses the 30 portfolio-construction specifications within each ARM family into a single family composite and then applies covariance PCA to the resulting five series. PC1 explains 31.8% of total variance, PC1–PC2 explain 60.6%, and PC1–PC4 explain 94.3%. PC1, PC2, and PC4 earn statistically significant average returns. Thus, multiple priced PC portfolios remain when closely related implementations are collapsed before PCA. Family-composite matching and split-sample tests determine which dimensions recur in the benchmark map, and the selection-complete exercise next applies a common stopping rule to the wider PC1–PC10 candidate set.

C.2 Selection-Complete Dimensionality

Table C.3 complements the map-informed cross-fits in Table 4 by estimating the selected taxonomy and its stopping point inside every training sample. The candidate set consists of

Table C.2: **Alternative PCA Maps**

	Variance		PC portfolio			ARM-family loading shares				
PC	Share	Cum.	Mean	t-stat.	Sharpe	HTCR	CIQ	COSK	MCRASH	TR β
Panel A. Correlation PCA of 150 ARM portfolios										
PC1	22.7	22.7	2.13	3.29	0.47	32.1	10.6	2.5	**41.8**	13.0
						(+32.1)	(-10.6)	(+2.3)	(+41.8)	(+13.0)
PC2	19.1	41.7	2.63	4.55	0.63	9.9	**36.0**	23.3	9.8	21.0
						(+9.9)	(+36.0)	(+23.3)	(-9.8)	(+21.0)
PC3	15.2	57.0	1.00	2.03	0.27	1.8	4.6	**55.1**	13.6	24.8
						(-1.7)	(-4.6)	(+55.1)	(+13.6)	(-24.8)
PC4	13.7	70.7	0.19	0.40	0.06	15.2	19.4	23.6	9.0	**32.7**
						(+15.2)	(+19.4)	(-23.6)	(+9.0)	(-32.7)
PC5	5.4	76.1	0.13	0.33	0.05	24.2	10.6	9.9	**47.1**	8.2
						(-24.2)	(+10.6)	(-5.5)	(+46.1)	(+7.3)
PC6	4.0	80.1	0.32	1.01	0.13	30.7	7.7	8.6	**46.9**	6.1
						(-23.2)	(+7.6)	(-8.1)	(+31.7)	(+2.9)
PC7	3.0	83.1	0.16	0.54	0.07	8.3	12.4	**51.8**	10.9	16.6
						(-0.7)	(-2.3)	(+4.5)	(+8.7)	(-2.3)
PC8	2.6	85.7	0.34	1.64	0.22	36.9	6.8	6.6	**39.4**	10.3
						(+21.2)	(-1.1)	(+4.1)	(+14.3)	(-3.0)
PC9	2.0	87.7	0.46	1.92	0.26	18.5	12.0	15.1	18.7	**35.8**
						(-11.7)	(+5.1)	(+5.9)	(-7.5)	(-6.9)
PC10	1.7	89.4	0.16	0.96	0.13	**26.6**	19.1	17.4	25.3	11.6
						(-14.6)	(+2.6)	(-8.2)	(+12.9)	(-3.1)
Panel B. Covariance PCA of five ARM-family averages										
PC1	31.8	31.8	3.85	4.09	0.57	**48.5**	2.3	0.9	24.2	24.1
						(+48.5)	(+2.3)	(+0.9)	(+24.2)	(+24.1)
PC2	28.8	60.6	3.23	2.87	0.41	1.1	**70.2**	9.0	14.5	5.2
						(+1.1)	(+70.2)	(+9.0)	(-14.5)	(+5.2)
PC3	20.7	81.2	0.19	0.24	0.03	21.6	6.5	1.0	13.2	**57.6**
						(+21.6)	(+6.5)	(+1.0)	(+13.2)	(-57.6)
PC4	13.1	94.3	2.19	3.05	0.40	7.1	7.0	**73.9**	11.6	0.5
						(-7.1)	(-7.0)	(+73.9)	(+11.6)	(+0.5)
PC5	5.7	100.0	0.44	1.07	0.16	26.6	11.7	9.3	**49.8**	2.6
						(-26.6)	(+11.7)	(-9.3)	(+49.8)	(+2.6)

Notes: This table reports two alternative PCA constructions. Panel A uses correlation PCA for the 150 ARM implementations; Panel B uses covariance PCA for five ARM-family composite returns. In Panel A, eigenvectors from the correlation matrix of standardized implementation returns are mapped back to raw returns and normalized to unit gross exposure. In Panel B, each family composite equally weights its 30 implementations before covariance PCA is applied. Share and Cum. are the individual and cumulative percentages of total variance. Mean is the annualized PC return in percent, and Sharpe is annualized. Family columns report gross loading shares with signed net shares in parentheses; gross shares sum to 100 up to rounding. Both panels use monthly returns from January 1968 through December 2024. Mean-return t-statistics use Newey–West standard errors with six lags. PC signs are oriented so that sample means are positive, and the largest gross family share in each row is bold. PC numbering is specific to each extraction method.

PC1–PC10. For each training sample, we estimate covariance PCA, map each of the five ARM family composites to the PC with which it has the highest absolute correlation, and orient PC signs using only training observations. A principal component enters the selected system when its annualized training mean is at least 1% and its two-sided mean-return test has a Holm-adjusted p-value below 0.10 across all ten PCs. The 1% cutoff imposes minimal economic materiality; the deliberately permissive 10% Holm cutoff retains weaker but potentially recurring dimensions. The trained weights, matches, signs, and selection decisions are then carried into the omitted block. Table 4 thus evaluates a fixed three-dimension map, whereas this exercise evaluates the dimension count recovered by the training data.

Panel A shows that the selected return space remains small. The average number of selected principal components is 2.25, 2.50, and 2.83 in the 4-, 8-, and 12-fold designs,

Table C.3: **Selection-Complete Dimensionality of the ARM Return Space**

Panel A. Selected dimensionality							
Folds	Mean K	Median K	Range	Family-matched K	Unmatched K	$K = 3$ (%)	HTCR=MCRASH (%)
4	2.25	2	2–3	2.00	0.25	25.0	100.0
8	2.50	3	1–3	2.25	0.25	62.5	75.0
12	2.83	3	1–4	2.58	0.25	41.7	91.7

Panel B. Family-composite matching and conditional held-out premia		4 folds		8 folds		12 folds	
Family match	Modal PC	Selected (%)	Mean	Selected (%)	Mean	Selected (%)	Mean
Broad tail/crash (HTCR)	PC1	50.0	1.15	87.5	3.57	91.7	3.78
			(0.79)		(3.64)		(3.83)
Broad tail/crash (MCRASH)	PC1	50.0	1.15	62.5	2.11	83.3	3.45
			(0.79)		(2.17)		(3.48)
TR β	PC3	0.0		0.0		0.0	
CIQ	PC2	100.0	4.18	75.0	4.45	83.3	3.67
			(4.10)		(3.28)		(2.74)
COSK	PC4	50.0	3.91	75.0	1.96	83.3	1.88
			(3.50)		(2.08)		(2.17)
Unmatched later PC	PC9	25.0	0.47	25.0	1.75	25.0	0.71
			(1.01)		(2.02)		(0.74)

Panel C. Strict chronological selection and holdout				1968–1995 training		1996–2024 holdout	
Training-sample match	PC	Var. (%)	Selected	Mean	Holm p	Mean	Holm p
CIQ	PC1	31.66	No	2.62	1.000	4.13	0.040
				(1.36)		(2.84)	
Broad tail/crash	PC2	24.23	Yes	4.04	0.005	1.82	1.000
				(3.43)		(1.20)	
TR β	PC3	14.48	No	-0.18	1.000	-0.94	1.000
				(-0.18)		(-0.77)	
COSK	PC4	10.00	Yes	2.89	0.006	3.67	0.001
				(3.36)		(3.87)	
Unmatched later PC	PC9	1.41	Yes	1.46	< 0.001	1.28	0.078
				(4.95)		(2.54)	

Notes: This table reports a training-sample selection exercise for the dimensionality of ARM returns. Panel A summarizes the number of selected PCs across blocked cross-fits; Panel B reports family matching and held-out returns conditional on selection; Panel C reports a chronological design trained over 1968–1995 and evaluated over 1996–2024. In each training sample, covariance PCA is applied to the 150 ARM implementation returns and PC1–PC10 are normalized to unit gross exposure. Each of the five ARM-family composites is assigned to the PC with which it has the largest absolute training-sample correlation, and signs are fixed from training observations. A PC is selected when its oriented annualized training mean is at least 1% and its two-sided mean-return test has a Holm-adjusted p-value below 0.10 across PC1–PC10. The 1% condition is the economic-materiality threshold, and the 10% condition is the multiplicity-adjusted statistical threshold. Panel B Mean is the annualized held-out return in percent for folds in which the PC is selected. Panel C Var. is the training-sample percentage of total variance. The blocked cross-fits partition monthly observations from January 1968 through December 2024. The chronological design estimates weights, matches, signs, and selection decisions using January 1968–December 1995 and applies them to January 1996–December 2024. Panel B reports six-lag Newey–West t-statistics in parentheses. Holm p-values in Panel C adjust the ten PC mean-return tests within the training or holdout period. Family-matched K and Unmatched K partition Mean K. HTCR=MCRASH is the percentage of folds in which those families map to the same PC. The unmatched row reports PC9. PC5–PC8 and PC10 are omitted from Panel C because they do not pass the pre-1996 selection rule.

respectively, and the modal count is three in the latter two designs. PCs matched to at least one family composite account for 2.00 to 2.58 selected dimensions on average, compared with 0.25 for unmatched PCs in every design. HTCR and MCRASH map to the same principal component in 75.0% to 100.0% of folds, reinforcing their role as the core of the broad tail/crash dimension.

Panel B identifies which family-matched dimensions recur. Lower-tail CIQ is selected in 75.0% to 100.0% of folds and earns conditional held-out premia between 3.67% and 4.45% per year. Residual coskewness is selected in 50.0% to 83.3% of folds and earns 1.88% to 3.91%. Selection of the PC matched to the HTCR–MCRASH core becomes more frequent as the omitted blocks become shorter, while the TR β family's standalone matched PC is never selected. This distinction is consistent with using TR β as part of the broad anchor while treating its differential component as unpriced. PC9 is the only unmatched later principal component that passes the rule: it is selected in 25.0% of folds in each design, with conditional held-out means of 0.47%, 1.75%, and 0.71%. PC9 therefore appears as a secondary residual candidate rather than a recurring family-matched dimension.

Panel C provides a strict chronological diagnostic. The 1968–1995 training sample selects broad tail/crash, residual coskewness, and PC9. Over 1996–2024, residual coskewness remains positive and significant, broad tail/crash remains positive but is estimated imprecisely, and lower-tail CIQ—which is not selected in the early sample—earns 4.13% per year with a ten-test Holm p-value of 0.040. The early-sample system is therefore compact, but the dimensions receiving the most precise compensation evolve over time.

C.3 Loading Concentration Across Portfolio Constructions

A separate concern is that a principal component could be driven by a small number of unusual portfolio implementations rather than by common variation across construction choices. Table C.4 summarizes the concentration of the absolute PCA loadings. Across PC1–PC5, the largest individual ARM portfolio receives at most 3.9% of absolute loading mass, the five largest portfolios jointly receive at most 17.9%, and the inverse-Herfindahl effective number of portfolios ranges from 49.5 to 73.8. Thus, even the most concentrated PC draws material weight from approximately 50 portfolios, and no five portfolios account for as much as one-fifth of its absolute loading mass.

Table C.4: **Concentration of PC Loadings Across Portfolio Constructions**

PC	Max portfolio	Top five	Effective N	Weighting	Price screen	Sort	Breakpoints
PC1	2.7	12.2	73.8	Equal (51.6)	\$0 (36.6)	Decile (52.5)	All (53.6)
PC2	3.6	17.9	52.6	Value (52.6)	\$0 (33.8)	Decile (50.8)	All (51.8)
PC3	3.4	15.8	61.8	Value (52.7)	\$0 (35.5)	Decile (52.0)	All (54.4)
PC4	3.6	17.9	49.5	Value (54.5)	\$0 (33.8)	Decile (51.5)	All (50.7)
PC5	3.9	16.2	58.4	Value (53.6)	\$0 (36.7)	Decile (52.0)	All (51.7)

Notes: This table reports portfolio- and implementation-level concentration measures for the loadings of PC1–PC5. Loadings come from covariance PCA of the 150 ARM implementation returns and are normalized to unit gross exposure. Max portfolio is the largest absolute loading on one portfolio, Top five is the combined absolute-loading share of the five largest portfolios, and Effective N is the inverse Herfindahl index of absolute portfolio loadings. The final four columns report the implementation category with the largest absolute-loading share; the corresponding share appears in parentheses. All loading shares are percentages. The PCA uses monthly returns from January 1968 through December 2024.

The marginal loading shares are also distributed across implementation choices. Neither equal nor value weighting receives more than 54.5% of loading mass; no price-screen category receives more than 36.7%; and neither breakpoint universe receives more than 54.4%. Decile implementations carry the largest sorting-rule shares, but no PC is dominated by a small number of portfolios. Together with the family-average PCA exercise, these results indicate that the principal components capture broad ARM return variation rather than a narrow portfolio-construction artifact.

C.4 Implementation-Specific PCA

Implementation-specific PCA provides a covariance-space counterpart to the basis evidence in Table 6. We re-estimate covariance PCA while fixing, in turn, the weighting scheme, sorting rule, minimum-price screen, or breakpoint universe. Each subgroup retains all five ARM families. The three anchor portfolios are reconstructed from the subgroup-specific family composites and matched one-to-one to PC1–PC5. Inference uses six-lag Newey–West standard errors and within-row Holm adjustment.

Table C.5 reports the subgroup-specific matched PC premia and pricing counts.

Table C.5: **Priced ARM Dimensions within Implementation Subsamples**

Implementation restriction	N	Broad tail/crash	Lower-tail CIQ	Residual coskewness	Priced	Min. $\lvert\rho\rvert$	PC1–5 var.
Panel A. Weighting scheme							
Equal, discrete sorts	60	4.10***	3.68***	1.56**	3	0.89	92.1
		(3.05)	(4.04)	(2.31)			
Value, discrete sorts	60	4.77***	0.25	2.96**	2	0.81	89.0
		(4.50)	(0.20)	(2.80)			
Equal, including managed	90	3.81***	3.29***	1.51**	3	0.90	90.4
		(3.10)	(3.94)	(2.42)			
Panel B. Sorting rule							
Top–bottom 10%	60	3.92***	5.34***	2.50***	3	0.97	78.7
		(3.21)	(3.94)	(2.76)			
Top–bottom 20%	60	3.85***	2.31**	2.34***	3	0.96	81.1
		(3.90)	(2.01)	(2.97)			
Characteristic-managed	30	1.16***	0.84*	0.81**	2	0.94	99.9
		(3.96)	(1.87)	(2.64)			
Panel C. Minimum-price screen							
No price screen	50	3.39**	3.86***	2.31***	3	0.93	78.3
		(2.43)	(3.02)	(2.91)			
Price $\geq$ \$1	50	5.02***	3.76***	2.41***	3	0.97	79.0
		(4.31)	(2.95)	(2.98)			
Price $\geq$ \$5	50	4.42***	2.10*	2.31***	2	0.96	80.9
		(5.14)	(1.77)	(2.91)			
Panel D. Breakpoint universe							
All exchanges	75	4.74***	4.35***	2.17***	3	0.96	78.0
		(3.79)	(3.32)	(2.70)			
NYSE	75	3.18***	3.40***	2.46***	3	0.97	80.3
		(3.25)	(2.86)	(3.16)			

Notes: This table reports matched PC-portfolio means within implementation subsamples. Panels vary the weighting scheme, sorting rule, minimum-price screen, and breakpoint universe; the final three columns summarize the number of priced matches, the minimum matching correlation, and variance explained. Within each row, covariance PCA is re-estimated from the five ARM families represented in that subsample. Subsample-specific broad tail/crash, lower-tail CIQ, and residual-coskewness anchors are constructed from family composites, and three of PC1–PC5 are matched one-to-one to the anchors using absolute correlations. PC signs are oriented toward positive correlations with the assigned anchors. Matched-PC means are annualized percentages. Priced counts matches with Holm $p < 0.05$, Min. $\lvert\rho\rvert$ is the smallest absolute anchor correlation, and PC1–5 var. is the percentage of subsample variance explained by the first five PCs. The monthly sample covers January 1968–December 2024. N is the number of ARM return series entering each row. Newey–West t-statistics using six lags are reported in parentheses. Holm adjustment is applied across the three matched-PC mean-return tests within each row. The first two weighting rows exclude characteristic-managed portfolios because they have no value-weighted counterpart; the third weighting row includes them in the equal-weighted set. *, **, and *** denote two-sided Holm-adjusted significance at the 10%, 5%, and 1% levels, respectively.

Every one of the eleven subgroup PCAs contains at least two matched PC portfolios with positive premia significant at the 5% adjusted level, and eight contain all three. Broad tail/crash and residual coskewness are priced in every subgroup. The lower-tail CIQ matched PC portfolio is significant at 5% in eight subgroups and at 10% in ten. Thus, the covariance decomposition consistently identifies a multidimensional priced space, although the separate CIQ-aligned component is less uniform than the directly observable CIQ basis premium.

Table C.6 reports the corresponding assignments. Across subgroups, the smallest absolute correlation among the three matched pairs ranges from 0.81 to 0.97. PC numbers change in several cases, demonstrating why economic dimensions must be matched across changing eigenvalue orderings.

Table C.6: **PC Matching within Implementation Subsamples**

Dimension varied	Restriction	Broad tail/crash	Lower-tail CIQ	Residual coskewness
Weighting	Equal, discrete sorts	PC1 (0.94)	PC2 (0.89)	PC4 (0.91)
Weighting	Value, discrete sorts	PC2 (0.81)	PC1 (0.85)	PC4 (0.98)
Weighting	Equal, including managed	PC1 (0.93)	PC2 (0.90)	PC4 (0.92)
Sorting rule	Top–bottom 10%	PC1 (0.97)	PC2 (0.97)	PC4 (0.97)
Sorting rule	Top–bottom 20%	PC1 (0.96)	PC2 (0.98)	PC4 (0.97)
Sorting rule	Characteristic-managed	PC2 (0.99)	PC1 (0.94)	PC4 (0.98)
Minimum price	No price screen	PC1 (0.93)	PC2 (0.99)	PC4 (0.96)
Minimum price	Price $\geq$ \$1	PC1 (0.97)	PC2 (0.99)	PC4 (0.97)
Minimum price	Price $\geq$ \$5	PC2 (0.96)	PC1 (0.97)	PC4 (0.97)
Breakpoint universe	All exchanges	PC1 (0.96)	PC2 (0.99)	PC4 (0.97)
Breakpoint universe	NYSE	PC1 (0.98)	PC2 (0.99)	PC4 (0.97)

Notes: This table reports the PC-to-anchor assignments underlying the implementation-subsample estimates. Rows identify the varied construction dimension and its restriction; columns identify the three priced ARM dimensions. Within each row, covariance PCA is estimated from that subsample's ARM returns. PC1–PC5 are matched one-to-one to the broad tail/crash, lower-tail CIQ, and residual-coskewness anchors by maximizing absolute time-series correlations. Each cell reports the assigned PC followed by the absolute matching correlation in parentheses. The PCA and matching use monthly observations from January 1968 through December 2024. PC signs in Appendix Table C.5 are oriented toward positive correlations with their assigned anchors.

In the three subgroups where the CIQ-matched PC portfolio falls short of the 5% Holm threshold—value-weighted discrete sorts, characteristic-managed portfolios, and the \$5 price-screen sample—the corresponding CIQ family-composite premium remains positive and conventionally significant at 5% (Table B.1). In all three cases, the CIQ anchor maps to PC1 rather than PC2. The weaker matched-PC estimate therefore reflects a change in how the CIQ premium enters the subgroup covariance structure, rather than disappearance of the premium itself. This distinction explains why the basis results in Table 6 are more uniform than the subgroup-PC evidence.

C.5 Disjoint Implementation Validation

The LORO regressions in Table 7 remove one exact portfolio-construction rule at a time. Table C.7 extends this separation to entire rule classes. Each row constructs the broad

Table C.7: **Disjoint-Implementation Validation of the ARM Basis**

Axis	Basis rules → test rules	Rules	N	Median $\lvert\bar{r}\rvert$	Median $\lvert\alpha\rvert$	90th pct. $\lvert\alpha\rvert$	RMS reduction (%)	Sig. (%)
Panel A. Full sample								
Weighting	Equal weighted → Value weighted	18/12	60	3.79	1.10	2.31	67.7	5.0
Weighting	Value weighted → Equal weighted	12/18	90	2.97	1.48	3.24	45.2	34.4
Sort	Decile → Quintile	12/12	60	3.59	0.78	2.38	62.7	18.3
Sort	Quintile → Decile	12/12	60	4.53	1.17	3.09	61.1	25.0
Sort	Discrete → Managed	24/6	30	1.34	0.40	0.58	67.2	0.0
Sort	Managed → Discrete	6/24	120	3.92	0.97	3.36	56.5	20.8
Breakpoints	All stocks → NYSE	15/15	75	3.47	0.63	2.43	61.2	17.3
Breakpoints	NYSE → All stocks	15/15	75	3.76	0.51	3.10	61.1	17.3
Price screen	Other screens → No screen	20/10	50	3.53	0.74	3.14	51.3	22.0
Price screen	Other screens → Price above $1	20/10	50	3.84	0.83	2.32	70.7	18.0
Price screen	Other screens → Price above $5	20/10	50	3.46	0.63	2.30	63.1	24.0
Panel B. 1996–2024								
Weighting	Equal weighted → Value weighted	18/12	60	3.86	1.59	3.65	51.8	0.0
Weighting	Value weighted → Equal weighted	12/18	90	2.18	1.54	5.05	40.3	16.7
Sort	Decile → Quintile	12/12	60	3.45	1.19	3.52	50.4	15.0
Sort	Quintile → Decile	12/12	60	4.62	1.93	3.62	55.4	18.3
Sort	Discrete → Managed	24/6	30	1.31	0.46	1.18	49.2	0.0
Sort	Managed → Discrete	6/24	120	4.26	1.33	4.59	49.6	15.8
Breakpoints	All stocks → NYSE	15/15	75	3.43	1.06	3.55	50.7	18.7
Breakpoints	NYSE → All stocks	15/15	75	4.14	1.10	3.53	56.3	12.0
Price screen	Other screens → No screen	20/10	50	2.73	0.97	3.40	51.8	12.0
Price screen	Other screens → Price above $1	20/10	50	3.63	1.15	3.11	60.9	12.0
Price screen	Other screens → Price above $5	20/10	50	3.54	1.21	3.19	49.3	22.0

Notes: This table reports ARM-basis pricing diagnostics when factor construction and test portfolios use disjoint implementation rules. Rows identify the construction axis and the basis-to-test rule split; Panel A uses the full sample and Panel B uses 1996–2024. For each row, the three anchor portfolios are constructed only from the rules to the left of the arrow and are used directly as basis factors for all five ARM-family portfolios formed under the rules to the right. Rules gives the numbers of basis and test construction rules, and N is the number of test portfolios. Median $|\bar{r}|$ is the median absolute annualized test-portfolio mean. Median $|\alpha|$ and 90th pct. $|\alpha|$ summarize absolute annualized three-factor alphas. RMS reduction is $100[1 - \mathrm{RMS}(\alpha)/\mathrm{RMS}(\bar{r})]$. Means and alphas are percentages per year. Panel A uses monthly returns from January 1968 through December 2024; Panel B uses January 1996 through December 2024. Sig. is the percentage of test-portfolio alphas with a two-sided six-lag Newey–West p-value below 0.05. Discrete combines decile and quintile tail-breakpoint sorts, and Managed denotes characteristic-managed portfolios.

tail/crash factor, lower-tail CIQ factor, and residual-coskewness factor using only the rule class shown to the left of the arrow and evaluates all five ARM-family portfolios formed under the disjoint rule class shown to the right. The eleven configurations separately withhold weighting schemes, sorting rules, breakpoint universes, and minimum-price screens.

The basis lowers root mean squared pricing errors in every aggregate train–test comparison. Full-sample reductions range from 45.2% when the basis is formed from value-weighted portfolios and tested on equal-weighted portfolios to 70.7% when portfolios above the $1 screen form the basis and no-screen portfolios are tested. Over 1996–2024, the corresponding range is 40.3% to 60.9%. Median absolute alphas range from 0.40% to 1.48% in the full sample and from 0.46% to 1.93% after 1995. The cross-weighting tests are asymmetric across weighting schemes: an equal-weighted basis leaves no significant value-weighted alphas after 1995, while a value-weighted basis leaves 16.7% of equal-weighted test alphas significant. This heterogeneity is consistent with the subgroup-specific basis estimates in Table 6, while the pricing-error reductions establish that the three basis dimensions carry substantial common content across disjoint construction choices.

C.6 Joint Factor-Model Spanning Tests

C.6.1 Control-Factor Definitions

Table C.8 documents the exact factor vector behind each model label used in the main pricing tables, joint spanning tests, oracle exercise, and residual-PCA analysis.

Table C.8: **Definitions of the Control Factor Models**

Model	K	Factors included in the implemented regression	Primary reference
Benchmark traded-factor models			
CAPM	1	MKT.	Sharpe, 1964
FF3	3	MKT, SMB, and HML.	Fama and French, 1993
Carhart4	4	FF3 augmented by MOM.	Carhart, 1997
FF5	5	MKT, SMB, HML, RMW, and CMA.	Fama and French, 2015
FF6	6	FF5 augmented by MOM.	Fama and French, 2018
FF6+BAB+STR	8	FF6 augmented by betting-against-beta (BAB) and the short-term reversal factor (STR).	Frazzini and Pedersen, 2014; Jegadeesh, 1990
FF6+LIQ	7	FF6 augmented by the Pástor–Stambaugh traded liquidity factor.	Pastor and Stambaugh, 2003
q5	5	MKT, ME, IA, ROE, and EG.	Hou et al., 2021
BS6	6	MKT, SMB, MOM, IA, ROE, and the HML Devil value factor.	Barillas and Shanken, 2018; Asness and Frazzini, 2013
Behavioral, intermediary, and statistical models			
SY4	4	MKT, SMB, $MGMT$, and $PERF$.	Stambaugh and Yuan, 2017
DHS3	3	MKT, FIN, and $PEAD$.	Daniel et al., 2020
HKM2	2	MKT and the intermediary-capital risk factor.	He et al., 2017
PCA	5	The first five conventional PCA factors estimated from the external reference asset panel.	Lettau and Pelger, 2020
RP-PCA	5	The first five risk-premium PCA factors, which tilt statistical factor extraction toward expected-return information.	Lettau and Pelger, 2020
Higher-moment and asymmetric-risk controls			
Coskewness	2	MKT and MKT^2; this is a polynomial market-return specification rather than a separate traded coskewness factor.	Harvey and Siddique, 2000
Cokurtosis	3	MKT, MKT^2, and MKT^3, allowing the pricing kernel to load on market skewness and kurtosis.	Dittmar, 2002
FF6+PSS	7	FF6 augmented by the predicted systematic skewness factor (PSS).	Langlois, 2020
Downside2	2	r_w and $r_{w,D} = r_w D$.	Farago and Tédongap, 2018
GDA3	3	r_w, D, and $r_{w,D}$.	Farago and Tédongap, 2018
GDA5	5	r_w, rv, D, $r_{w,D}$, and $rv_D = rvD$.	Farago and Tédongap, 2018

Notes: This table defines the benchmark-model labels used in the factor-alpha, residual-PCA, joint-alpha, oracle, and state-dependence tables. Columns report the model label, number of factors K, implemented factor vector, and primary reference; row groups classify the models by type. MKT is the excess market return; SMB and HML are size and value; RMW and CMA are profitability and investment; and MOM is momentum. ME, IA, ROE, and EG denote the q-model size, investment, profitability, and expected-growth factors. $MGMT$ and $PERF$ are the Stambaugh–Yuan management and performance factors. FIN and $PEAD$ are the Daniel–Hirshleifer–Sun long- and short-horizon behavioral factors. r_w is the log aggregate wealth return, rv is the innovation in realized market variance, and D is the Farago–Tédongap downstate indicator; a subscript D denotes multiplication by that indicator. The empirical tables estimate each model over its maximum available monthly overlap and report the corresponding sample length T. PCA and RP-PCA denote external statistical factors and are distinct from the PCA estimated from the 150 ARM implementation returns.

C.6.2 Principal-Component Alpha Estimates

Table C.9 reports the factor-model alphas for PC1–PC5 that underlie the dimension-level comparison in Section 4.3. The estimates provide the PC-portfolio counterpart to the directly usable basis-factor results in Table 8.

Table C.9: **Means and Factor-Model Alphas of ARM Principal Components**

Model	K	T	PC1	PC2	PC3	PC4	PC5
Unadjusted							
Mean return	0	684	4.08***	3.86***	0.41	2.33***	0.73
			(3.67)	(3.10)	(0.45)	(2.89)	(1.23)
Benchmark traded factors							
CAPM	1	684	3.65***	5.27***	0.64	2.29***	0.83
			(3.19)	(4.67)	(0.69)	(2.88)	(1.34)
FF3	3	684	3.83***	3.81***	0.75	1.88**	0.97
			(4.24)	(3.88)	(0.78)	(2.41)	(1.49)
Carhart4	4	684	2.79***	4.55***	1.55	1.93***	0.78
			(3.13)	(4.32)	(1.42)	(2.58)	(1.22)
FF5	5	684	3.65***	3.57***	0.71	1.32*	1.38**
			(4.18)	(3.56)	(0.62)	(1.75)	(2.06)
FF6	6	684	2.73***	4.22***	1.39	1.36*	1.19*
			(3.18)	(3.96)	(1.20)	(1.83)	(1.77)
FF6+BAB+STR	8	684	3.59***	3.91***	0.84	1.31*	1.24*
			(4.00)	(3.13)	(0.73)	(1.65)	(1.77)
FF6+LIQ	7	684	2.83***	3.99***	1.54	1.30*	1.02
			(3.24)	(3.65)	(1.38)	(1.73)	(1.52)
q5	5	684	2.08**	4.32***	-0.33	1.54*	0.81
			(2.20)	(3.14)	(-0.23)	(1.78)	(1.21)
BS6	6	684	1.65*	3.28***	1.14	0.56	1.54**
			(1.70)	(2.84)	(0.90)	(0.67)	(2.04)
Behavioral / mispricing							
SY4	4	588	3.04***	3.80***	0.79	1.27	1.27*
			(2.76)	(3.44)	(0.64)	(1.44)	(1.71)
DHS3	3	558	0.01	4.08***	1.06	1.81*	0.94
			(0.01)	(3.37)	(0.76)	(1.94)	(1.18)
Intermediary capital							
HKM2	2	587	2.81**	4.90***	0.72	2.08**	0.63
			(2.44)	(4.40)	(0.73)	(2.42)	(0.98)
Statistical factors							
PCA	5	600	2.59***	3.25***	0.80	1.29	1.04*
			(2.72)	(3.20)	(0.78)	(1.49)	(1.66)
RP-PCA	5	600	2.61**	3.53***	-0.16	1.12	0.87
			(2.15)	(2.65)	(-0.13)	(0.84)	(1.25)
Higher moments / skewness							
Coskewness	2	684	1.80	3.39***	2.04**	1.71	0.77
			(1.29)	(2.71)	(1.96)	(1.64)	(1.03)
Cokurtosis	3	684	2.11	3.20**	2.37**	1.39	0.93
			(1.53)	(2.50)	(2.33)	(1.30)	(1.35)
FF6+PSS	7	600	2.44***	4.60***	1.38	1.03	1.03
			(2.69)	(4.46)	(1.07)	(1.24)	(1.46)
Downside / disappointment							
Downside2	2	684	0.54	5.03***	1.64	1.93*	1.05
			(0.35)	(3.61)	(1.49)	(1.71)	(1.38)
GDA3	3	684	1.30	4.83***	1.61	1.68	1.13
			(0.86)	(3.52)	(1.46)	(1.42)	(1.55)
GDA5	5	684	1.66	4.22***	1.61	1.43	1.21*
			(1.11)	(2.99)	(1.45)	(1.25)	(1.69)

Notes: This table reports unadjusted mean returns and factor-model alphas for PC1–PC5. The first row contains mean returns; subsequent row blocks contain intercepts from regressions on alternative benchmark models. The PC portfolios are obtained from covariance PCA of the 150 ARM implementation returns, normalized to unit gross exposure, and signed so that full-sample means are positive. Returns and alphas are annualized percentages. The underlying ARM returns are monthly from January 1968 through December 2024. Each benchmark regression uses the model's maximum available overlap, and T reports the resulting number of months. Newey–West t-statistics using six lags are reported in parentheses below the estimates. K is the number of benchmark factors. Appendix Table C.8 defines every model and its factor vector. PCA loadings and unadjusted means are estimated on the same sample. *, **, and *** denote conventional two-sided significance at the 10%, 5%, and 1% levels, respectively.

C.6.3 Joint Alpha Tests

The individual alpha estimates in Tables 8 and C.9 show which ARM dimensions remain unexplained by each factor model. Table C.10 first complements the direct basis-alpha estimates with a joint test of whether the three PCA-free basis alphas are all zero. The null is that the basis alphas for broad tail/crash, lower-tail CIQ, and residual coskewness are all zero for a given benchmark model. The test therefore asks whether the simple basis itself, rather than the estimated PC rotation, is jointly spanned by the standard factor models.

Table C.10: **Joint Tests of Factor-Model Alphas for the ARM Basis**

Model	K	T	HAC Wald	p-value
Benchmark traded factors				
CAPM	1	684	51.36	< 0.001
FF3	3	684	52.15	< 0.001
Carhart4	4	684	40.51	< 0.001
FF5	5	684	42.57	< 0.001
FF6	6	684	34.13	< 0.001
FF6+BAB+STR	8	684	27.55	< 0.001
FF6+LIQ	7	684	31.07	< 0.001
q5	5	684	19.97	< 0.001
BS6	6	684	13.53	0.004
Behavioral / mispricing				
SY4	4	588	25.23	< 0.001
DHS3	3	558	17.14	< 0.001
Intermediary capital				
HKM2	2	587	38.62	< 0.001
Statistical factors				
PCA	5	600	20.94	< 0.001
RP-PCA	5	600	12.05	0.007
Higher moments / skewness				
Coskewness	2	684	15.65	0.001
Cokurtosis	3	684	14.51	0.002
FF6+PSS	7	600	29.15	< 0.001
Downside / disappointment				
Downside2	2	684	20.25	< 0.001
GDA3	3	684	18.99	< 0.001
GDA5	5	684	15.04	0.002

Notes: This table reports joint tests that the three ARM-basis factor-model intercepts are zero for each benchmark model. The tested returns are the broad tail/crash, lower-tail CIQ, and residual-coskewness basis factors. Each is regressed on the factor vector of the indicated benchmark model. K is the number of benchmark factors, and T is the number of monthly observations. The ARM returns span January 1968–December 2024; each model uses its maximum available monthly overlap. Appendix Table C.8 defines the benchmark factor vectors. HAC Wald jointly tests the three intercept restrictions using the Bartlett-kernel covariance estimator of Newey and West (1987) with six lags and allows cross-equation residual covariance. Under the null, the statistic is asymptotically $\chi^2(3)$.

Table C.11 then repeats the same system test for PC1–PC5. The classical GRS test provides an exact finite-sample F-test under normal, temporally i.i.d. returns (Gibbons et al., 1989). We instead report its HAC-robust large-sample counterpart, consistent with the GMM approach to mean–variance-efficiency testing in MacKinlay and Richardson (1991). The covariance estimator follows Newey and West (1987), allowing for heteroskedasticity, serial dependence, and cross-equation residual covariance. Restricting the PC test to PC1–PC5 aligns the joint hypothesis with the PC portfolios emphasized in the main analysis and avoids allowing the small-eigenvalue later PCs to drive the conclusion.

Table C.11: **Joint Tests of Factor-Model Alphas for ARM Principal Components**

Model	K	T	HAC Wald	p-value
Benchmark traded factors				
CAPM	1	684	51.31	< 0.001
FF3	3	684	52.53	< 0.001
Carhart4	4	684	40.36	< 0.001
FF5	5	684	44.00	< 0.001
FF6	6	684	33.79	< 0.001
FF6+BAB+STR	8	684	29.88	< 0.001
FF6+LIQ	7	684	31.55	< 0.001
q5	5	684	18.64	0.002
BS6	6	684	15.32	0.009
Behavioral / mispricing				
SY4	4	588	24.43	< 0.001
DHS3	3	558	16.34	0.006
Intermediary capital				
HKM2	2	587	38.72	< 0.001
Statistical factors				
PCA	5	600	19.63	0.001
RP-PCA	5	600	13.56	0.019
Higher moments / skewness				
Coskewness	2	684	18.63	0.002
Cokurtosis	3	684	20.75	< 0.001
FF6+PSS	7	600	28.65	< 0.001
Downside / disappointment				
Downside2	2	684	21.78	< 0.001
GDA3	3	684	21.99	< 0.001
GDA5	5	684	18.59	0.002

Notes: This table reports joint tests that the factor-model intercepts of PC1–PC5 are zero for each benchmark model. The five PC portfolios are jointly regressed on the factor vector of the indicated benchmark model. K is the number of benchmark factors, and T is the number of monthly observations. The ARM returns span January 1968–December 2024; each model uses its maximum available monthly overlap. Appendix Table C.8 defines the benchmark factor vectors. HAC Wald jointly tests the five intercept restrictions using the Bartlett-kernel covariance estimator of Newey and West (1987) with six lags and allows cross-equation residual covariance. Under the null, the statistic is asymptotically $\chi^2(5)$.

Both joint-test designs reject all models at the 5% level. For the direct basis factors, the largest p-value is 0.007; for PC1–PC5, the largest is 0.019. The rejection applies to the benchmark traded-factor models as well as the behavioral, statistical, higher-moment, and downside/disappointment specifications. The result therefore confirms that no considered model jointly spans either the PCA-free ARM basis or the economically important portion of the ARM return space.

C.7 Residual Structure After ARM-Basis Adjustment

Table C.12 reports the PC-by-PC decomposition and family loading shares underlying the ARM-basis row of Table 9.

Table C.12: **Residual PCA Map after ARM-Basis Adjustment**

	Variance		Residual PC portfolio			ARM-family loading shares				
PC	Share	Cum.	α	t-stat.	Sharpe	HTCR	CIQ	COSK	MCRASH	TR β
PC1	35.7	35.7	0.19	0.22	0.03	30.2	3.7	1.0	17.4	**47.6**
						(+30.2)	(+0.0)	(+0.0)	(+17.4)	(-47.6)
PC2	13.0	48.7	0.50	0.73	0.11	35.6	5.1	6.7	**40.7**	11.9
						(-28.2)	(-0.0)	(+0.0)	(+39.6)	(-11.3)
PC3	10.7	59.5	0.01	0.02	0.00	**44.9**	19.8	1.5	29.0	4.9
						(+27.1)	(-0.0)	(-0.0)	(-29.0)	(+1.9)
PC4	6.0	65.5	0.10	0.32	0.04	9.1	21.5	24.6	10.3	**34.6**
						(+6.8)	(+0.0)	(-0.0)	(-3.1)	(-3.7)
PC5	4.9	70.4	0.77	2.63	0.37	22.2	8.6	27.0	**31.0**	11.2
						(-0.5)	(+0.0)	(+0.0)	(+6.0)	(-5.5)

Notes: This table reports the variance, alpha, and ARM-family loading composition of the first five residual PC portfolios after ARM-basis adjustment. Each of the 150 ARM implementation returns is regressed on the broad tail/crash, lower-tail CIQ, and residual-coskewness basis factors. Fitted factor exposures are subtracted while the intercept is retained, and covariance PCA is applied to the centered adjusted returns. PC weights are normalized to unit gross exposure and signed so that sample alphas are positive. Share and Cum. are percentages of adjusted-return variance; α is the annualized adjusted PC mean in percent; and Sharpe is annualized. Family columns report gross loading shares with signed net shares in parentheses; gross shares sum to 100 up to rounding. The monthly sample covers January 1968–December 2024. Alpha t-statistics use Newey–West standard errors with six lags. The largest gross family share in each row is bold. TR β denotes tail-risk beta, CIQ denotes lower-tail commonality in quantiles, and COSK denotes residual coskewness.

After adjustment for the three basis factors, PC1 remains the largest source of residual covariance but carries little average return. Its loadings primarily contrast TR β with HTCR and MCRASH. PC5 is the only residual PC portfolio priced at the 5% level and combines MCRASH, COSK, and HTCR exposure. Thus, the remaining priced component is economically small and mixes families already represented in the basis rather than identifying another recurring dimension.

C.8 Oracle Diagnostic

Table C.13 reports deliberately generous ex-post oracle benchmarks. Selection is performed separately for every ARM family–implementation cell after observing the alphas, so the exercise is a descriptive lower bound on pricing errors rather than a feasible model or valid post-selection test.

Table C.13: **Ex Post Oracle Pricing Diagnostics**

Oracle	Modal model	T	Median $\lvert\alpha\rvert$	90th pct. $\lvert\alpha\rvert$	Sig. 5%	Alpha reduction
Benchmark traded factors	BS6	684	1.42	5.24	24.0	53.0
Behavioral / mispricing	DHS3	558–588	1.64	4.78	32.0	41.9
Intermediary capital	HKM2	587	3.09	6.76	60.0	10.5
Statistical factors	RP-PCA	600	2.11	4.85	36.0	39.6
Higher moments / skewness	Coskewness	600–684	2.11	5.95	24.0	41.1
Downside / disappointment	GDA5	684	1.90	5.20	14.7	39.0
All models	**BS6**	**558–684**	**0.73**	**3.47**	**12.0**	**80.1**

Notes: This table reports ex post oracle pricing diagnostics by benchmark-model family and across all models. Each row summarizes the model selected separately for every ARM-family-by-implementation return. Within a model family, the oracle selects the specification with the smallest absolute alpha for each return; the All models row selects from the complete benchmark set. Median $|\alpha|$ and 90th pct. $|\alpha|$ summarize absolute annualized alphas in percent. Sig. 5% is the percentage of selected alphas whose model-specific t-statistic exceeds 1.96 in absolute value. Alpha reduction is the median percentage reduction in absolute alpha relative to CAPM, and Modal model is the most frequently selected specification. Models use their maximum available monthly overlap with ARM returns from January 1968 through December 2024. T reports the sample length or range of sample lengths among selected models. Selection is performed ex post. The reported significance frequencies use conventional model-specific tests and are not adjusted for model selection.

D Robustness of the Predictive Economic-State Evidence

D.1 Split-Sample Predictive State Dependence

Table D.1 applies PCA weights, PC matching, and signs estimated through December 1995 to predictive regressions for returns from January 1996 through December 2024.

Table D.1: **Out-of-Sample Predictive State Dependence of the Priced ARM Dimensions**

	(1) PC1 Broad tail/crash	(2) PC2 Lower-tail CIQ	(3) PC4 Residual coskewness
Matched split PC	+PC2 (0.99)	+PC1 (0.99)	+PC4 (0.82)
α / mean	1.82	4.13	3.67
COMMON	-0.74	4.03	1.25
	(-0.47)	(2.48)	(1.04)
RA−UNC	3.81	0.92	-0.68
	(2.19)	(0.51)	(-0.68)
R^2	0.024	0.031	0.006
Observations	348	348	348

Notes: This table reports predictive state-dependence regressions for three PC portfolios constructed using a pre-1996 training sample and evaluated after 1995. Columns align the split-sample portfolios with full-sample PC1, PC2, and PC4. PCA weights, one-to-one PC matching, and sign orientation are estimated through December 1995 and applied to subsequent ARM returns. The dependent variable is the matched PC return in month $t+1$. COMMON is the standardized average of standardized risk aversion and uncertainty, and RA−UNC is their standardized difference; both are observed in month t. Intercepts and slopes are annualized percentages. The evaluation sample covers January 1996–December 2024 and contains 348 monthly observations. Newey–West t-statistics using six lags are reported in parentheses below state coefficients. Matched split PC gives the training-sample PC used for each full-sample column; the leading sign is the orientation applied to its return, and the number in parentheses is the absolute alignment correlation. Because the state variables are centered, the intercept equals the sample mean.

D.2 Alternative RA/UC Series

Table D.2 repeats the predictive regressions using the alternative RA/UC series with coefficients estimated over 1990–2022. The table reports both the full-sample PC returns and the split-sample PC returns, using the overlap available for the alternative series.

Table D.2: **Predictive State Dependence Using the Alternative RA/UC Series**

	PC1		PC2		PC4		
State measure	COMMON	RA−UNC	COMMON	RA−UNC	COMMON	RA−UNC	Obs.
Panel A. Full-sample PC returns							
RA/UC, coeffs est. 1990–2022	1.93	6.72	4.47	1.02	-0.42	-2.06	385
	(1.15)	(3.45)	(2.57)	(0.43)	(-0.28)	(-1.36)	
Panel B. Chronological holdout							
RA/UC, coeffs est. 1990–2022	1.38	5.48	4.00	2.13	1.30	-0.36	314
	(0.86)	(2.89)	(2.57)	(0.93)	(1.03)	(-0.30)	

Notes: This table reports predictive state-dependence coefficients using the alternative RA/UC risk-aversion and uncertainty series based on Bekaert et al. (2013). Panel A uses the full-sample PC portfolios; Panel B uses PC weights, matching, and signs estimated through December 1995. The reported RA/UC series uses coefficients estimated over 1990–2022. COMMON is the standardized average of standardized risk aversion and uncertainty, and RA−UNC is their standardized difference. State variables are measured in month t, PC returns in month $t+1$, and coefficients are annualized percentages. Panel A contains 385 monthly observations available after return alignment. Panel B covers January 1996–February 2022 and contains 314 observations. Newey–West t-statistics using six lags are reported in parentheses below the coefficient estimates.

D.3 Additive Factor Controls in Predictive Regressions

Table D.3 includes benchmark factors as additive controls while holding factor betas fixed across economic states.

Table D.3: **Factor-Adjusted Predictive State Dependence of the Priced ARM Dimensions**

		FF6			BS6			GDA5		
PC	Mean	α	COMMON	RA−UNC	α	COMMON	RA−UNC	α	COMMON	RA−UNC
PC1	3.83	1.94*	1.50	3.23**	0.98	1.23	3.13**	-0.33	-1.52	3.05**
		(1.66)	(0.97)	(2.41)	(0.78)	(0.83)	(2.49)	(-0.16)	(-0.85)	(2.04)
PC2	4.47	4.16***	5.01***	0.40	3.54***	3.82***	0.08	3.14*	2.49	0.45
		(3.52)	(3.31)	(0.28)	(2.84)	(2.83)	(0.06)	(1.87)	(1.45)	(0.32)
PC4	2.32	2.27***	0.75	-0.56	1.77*	0.27	-0.81	2.94**	0.25	-0.63
		(2.74)	(0.64)	(-0.57)	(1.92)	(0.22)	(-0.77)	(2.36)	(0.18)	(-0.60)

Notes: This table reports predictive state-dependence regressions for PC1, PC2, and PC4 after adding FF6, BS6, or GDA5 factor returns as controls. Rows identify PC portfolios; model blocks report the factor-model intercept and the two state coefficients. For each PC portfolio, the month-$t+1$ return is regressed on month-t COMMON and RA−UNC and on the contemporaneous factor returns of the indicated benchmark model. Factor returns enter additively and are not interacted with the states. COMMON is the standardized average of the standardized risk-aversion and uncertainty series of Bekaert et al. (2022), and RA−UNC is their standardized difference. Mean is the unadjusted annualized PC return; intercepts and slopes are also annualized percentages. All regressions use 462 monthly observations from July 1986 through December 2024. Newey–West t-statistics using six lags are reported in parentheses below the estimates. FF6, BS6, and GDA5 are defined in Appendix Table C.8. Because factor returns are not centered, α is the factor-model intercept evaluated at the mean state. *, **, and *** denote conventional two-sided significance at the 10%, 5%, and 1% levels, respectively.